THESE DE DOCTORAT DE

L'ÉCOLE NATIONALE SUPERIEURE MINES-TELECOM ATLANTIQUE
BRETAGNE PAYS DE LA LOIRE - IMT ATLANTIQUE

in Joint PhD with

UNIVERSITY OF CAPE TOWN, Faculty of Health Sciences

ECOLE DOCTORALE N° 601
Mathématiques et Sciences et Technologies de l'Information et de la Communication
Spécialité : *Image, Signal et Vision*

Jean-Rassaire FOUEFACK

Towards a framework for multi class statistical modelling of shape, intensity, and kinematics in medical images

PhD defence on 18 march 2021

Research labs : Laboratoire de traitement de l'information médicale (LATIM INSERM UMR 1101)
Medical image inferencing and distributed diagnosis (Mi2d2)

Thèse N° 2021IMTA0240

University of Cape Town

Committee

President	Tien Tuan DAO, Full Professor, Ecole Centrale Lille, France
Reviewer	Ross T. WHITAKER, Full Professor, SCI Institute, University of Utah, USA
Reviewer	Philippe BUECHLER, Associate Professor, ARTORG, University of Bern, Switzerland
Examiner	Diana MATEUS, Full Professor, Ecole Centrale Nantes, France
Supervisor	Valérie BURDIN, Full Professor, IMT Atlantique, France
Supervisor	Tinashe MUTSVANGWA, Associate Professor, University of Cape Town, South Africa
Co-Supervisor	Bhushan BOROTIKAR, Associate Professor, SCMIA, Symbiosis International University, Pune, India

Invited members

Co-Supervisor	Tania DOUGLAS, Full Professor, University of Cape Town, South Africa
Examiner	Marcel LÜTHI, Associate Professor, University of Basel, Switzerland

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

Declaration of Authorship

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

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

• This research report is my own work.

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

Name: Jean-Rassaire Fouefack Date: 18th May 2021

Signed by candidate

Acknowledgements

I am heartily thankful to my supervisors A/Prof. Tinashe Mutsvangwa and Prof. Valérie Burdin, for selecting me for this project. I appreciate their energy, advice, patience and words of motivation that have helped me to understand the work and inspired me many times to do my best. I would like to express my sincere gratitude to my co-supervisors Prof. Tania Douglas and A/Prof. Bhushan Borotikar for their advice and suggestions, which were very helpful for the accomplishment of this project.

A special thanks to all my mates from the Medical Image-based Inferencing and Distributed Diagnostics (Mi2d2) research group and from the Laboratory of Medical Information Processing (LaTIM INSERM UMR1101). Their comments have enabled me to reflect upon this research from a more practical perspective. A sincere appreciation to the staff from the Division of Biomedical Engineering at the University of Cape Town and the Department of Image and Information Processing at IMT Atlantique for a great and motivating work atmosphere. I would also like to thank Dr. Marcel Luethi and the Gravis research team, University of Basel, for their inputs in implementing features needed for this project in Scalismo software.

This work is based on research supported by 1) the National Research Foundation (NRF) of South Africa (grant no's 105950 and 114393); the South African Research Chairs Initiative of the NRF and the Department of Science and Technology (grant no 98788), and the Medical Research Council of South Africa -2) the Brest Métropole, France (Grant MNB/MJM no 17-178).

Thanks to my family for their love and support in all possible ways. Finally, I am grateful to my wife for her unremitting encouragement, and unconditional love.

Abstract

Statistical modelling has become a ubiquitous tool for analysing of morphological variation of bone structures in medical images. For radiological images, the shape, relative pose between the bone structures and the intensity distribution are key features often modelled separately. A wide range of research has reported methods that incorporate these features as priors for machine learning purposes. Statistical shape, appearance (intensity profile in images) and pose models are popular priors to explain variability across a sample population of rigid structures. However, a principled and robust way to combine shape, pose and intensity features has been elusive for four main reasons: 1) heterogeneity of the data (data with linear and non-linear natural variation across features); 2) sub-optimal representation of three-dimensional Euclidean motion; 3) artificial discretization of the models; and 4) lack of an efficient transfer learning process to project observations into the latent space.

This work proposes a novel statistical modelling framework for multiple bone structures. The framework provides a latent space embedding shape, pose and intensity in a continuous domain allowing for new approaches to skeletal joint analysis from medical images. First, a robust registration method for multi-volumetric shapes is described. Both sampling and parametric based registration algorithms are proposed, which allow the establishment of dense correspondence across volumetric shapes (such as tetrahedral meshes) while preserving the spatial relationship between them. Next, the framework for developing statistical shape-kinematics models from in-correspondence multi-volumetric shapes embedding

image intensity distribution, is presented. The framework incorporates principal geodesic analysis and a non-linear metric for modelling the spatial orientation of the structures. More importantly, as all the features are in a joint statistical space and in a continuous domain; this permits on-demand marginalisation to a region or feature of interest without training separate models. Thereafter, an automated prediction of the structures in images is facilitated by a model-fitting method leveraging the models as priors in a Markov chain Monte Carlo approach. The framework is validated using controlled experimental data and the results demonstrate superior performance in comparison with state-of-the-art methods.

Finally, the application of the framework for analysing computed tomography images is presented. The analyses include estimation of shape, kinematic and intensity profiles of bone structures in the shoulder and hip joints. For both these datasets, the framework is demonstrated for segmentation, registration and reconstruction, including the recovery of patient-specific intensity profile.

The presented framework realises a new paradigm in modelling multi-object shape structures, allowing for probabilistic modelling of not only shape, but also relative pose and intensity as well as the correlations that exist between them. Future work will aim to optimise the framework for clinical use in medical image analysis.

Keywords : Shape-pose and intensity latent space, shoulder and hip human joint, medical image analysis, pattern recognition.

Version française - Résumé étendu

1. Introduction

1.1. Contexte et motivation

Les troubles musculo-squelettiques (TMS) produisent des pathologies liées à des atteintes douloureuses des os, muscles, tendons, ligaments ou nerfs, et leur prévalence mondiale est en augmentation constante (Main and de C Williams, 2002; Nikooyan et al., 2012). Ces maladies ont été déclarées comme étant la principale cause de handicap physique moteur à long terme (Pedoia et al., 2016). Ces pathologies musculo-squelettiques ont des causes multiples, notamment l'infirmité motrice cérébrale, la déformation des os, les traumatismes, et elles sont l'une des sources les plus fréquentes de douleur dans le corps humain. Elles ont ainsi un impact sur la vie quotidienne du patient en limitant le mouvement de son système locomoteur. Dans certains cas, les symptômes réduisent considérablement la capacité à poursuivre des activités régulières, ce qui peut conduire à un mode de vie sédentaire précoce.

Toutes les articulations du corps sont exposées aux TMS. Toutefois, le niveau de risque dépend de l'amplitude des mouvements et des degrés de liberté (DDL) de l'articulation. Ainsi, en raison de leur grande amplitude de mouvement et de leur importance pour les activités de la vie quotidienne, les articulations de l'épaule et de la hanche présentent généralement un risque plus élevé de TMS (Woon and Ham, 1991). La littérature en

rapport avec les troubles de la hanche (Tanzer and Noiseux, 2004; Clohisy et al., 2009; Anderson et al., 2010; Li et al., 2015) et de l'épaule (Ludewig and Cook, 2000; Ludewig and Braman, 2011; Van der Molen et al., 2017; Park et al., 2019) abonde. Par exemple, une anomalie anatomique du fémur a été associée à un risque accru de progression de l'arthrose de la hanche en raison d'une contrainte élevée sur le cartilage articulaire (Lespasio et al., 2018). Les facteurs morphologiques à l'origine du conflit fémoroacétabulaire impliqué dans l'étiopathogénèse sous-jacente de l'arthrose de la hanche et dans les lésions du labrum restent flous et font toujours l'objet de recherche sur les TMS (Murphy et al., 2016). De même pour l'articulation de l'épaule, la morphologie des os constitutifs contribue à la fonction de l'épaule ; par exemple, il a été signalé que la rétroversion humérale et la forme de l'omoplate déterminent l'amplitude de mouvement des athlètes de lancer et des joueurs de volley-ball (Ribeiro et al., 2012; Ribeiro and Pascoal, 2012).

L'imagerie tomodensitométrie (CT), résonance magnétique (RM) ou fluoroscopique à rayons X sont maintenant régulièrement utilisées pour étudier la morphologie et la fonction des articulations humaines en relation avec les blessures osseuses et articulaires, ainsi que pour évaluer la cinématique articulaire *in vivo* $3D$ (Kozono et al., 2017; Borotikar et al., 2017). En outre, les outils cliniques basés sur l'image pour l'assistance et le planning chirurgical, tels que BLUEPRINT™ (`https://www.shoulderblueprint.com/`) et MocapLab™ (`https://www.mocaplab.com/`), sont désormais largement acceptés. Les structures affichées dans ces images médicales sont caractérisées par trois informations principales :

- *la forme* : géométrie d'un objet, invariante par rapport à une transformation rigide et un changement d'échelle

- *la pose* : état statique d'un objet par rapport à un objet de référence et représentée par une transformation rigide entre les deux objets

- *l'intensité* : valeur numérique d'un pixel/voxel de l'image codant la réponse du tissu,

imagé par un système physique (CT, IRM, ...)

La quantification de la pose est appelée cinématique et décrit le mouvement de points, de corps (objets) et de systèmes de corps (groupes d'objets) dans un système de référence donné (Whittaker, 1937). Les traitements d'imagerie médicale tels que la segmentation, l'extraction et la reconstruction de la structure d'intérêt, dépendent essentiellement de ces caractéristiques clés. Le succès de ces traitements dans l'analyse morpho-fonctionnelle des articulations humaines dépend de la compréhension de la variabilité des caractéristiques de forme, d'intensité et de pose. La variabilité de ces caractéristiques comprend la variabilité de la morphologie complexe des articulations, la variabilité spatiale au sein de l'image, la variabilité de la forme entre les individus et la variabilité de l'intensité de l'image en ce qui concerne les caractéristiques de contraste (c'est-à-dire la délimitation entre les tissus voisins) (Seoud et al., 2019).

A ce jour, des algorithmes avancés de traitement d'images médicales ont été développés pour faciliter l'analyse des formes, des poses et de l'intensité (Rueckert and Schnabel, 2020). La stratégie commune consiste à estimer 3 sous-espaces homogènes indépendamment pour la forme, l'intensité et la cinématique au moyen de modèles construits à partir des données. D'après la littérature, la tendance actuelle pour trouver ces modèles homogènes est d'utiliser des modèles générateurs basés sur le deep learning (apprentissage profond) (Toshev and Szegedy, 2014; Ouyang et al., 2014). Cependant, ces modèles générateurs profonds ont tendance à avoir une spécificité élevée mais une faible généralité ; dans certains cas, ils expliquent les données exclues de l'apprentissage avec une grande confiance tout en étant complètement faux (Huang et al., 2017; Webster et al., 2019; Nalisnick et al., 2018). En outre, les approches deep learning nécessitent un grand nombre de données d'apprentissage pour être couronnées de succès.

Par ailleurs, des approches bien établies pour l'analyse de la forme, de l'intensité et de la cinématique utilisant des modèles statistiques ont été développées depuis 30 ans (Cootes

et al., 1992; Bossa and Olmos, 2006; Lüthi et al., 2017). Les modèles statistiques de forme (SSM), d'apparence active (AAM), de forme active (ASM), de forme et d'intensité (SSIM) et de forme-pose (SSPM), sont devenus omniprésents dans la littérature et visent à expliquer la variation naturelle de ces caractéristiques (forme, pose et intensité) dans les objets biologiques (Cootes et al., 1992; Sarkalkan et al., 2014; Bossa and Olmos, 2006; Lüthi et al., 2017). Ces modèles visent tous à trouver des espaces paramétrés de forme, d'intensité et de pose homogènes. Ces espaces sont ensuite utilisés comme connaissance a priori pour la prédiction automatisée à partir des corrélations observées apprises.

L'analyse morphologique des articulations humaines à l'aide des SSMs implique généralement une modélisation séparée des objets (os) constituant les articulations, tels que l'humérus et l'omoplate pour l'épaule. Alors que ces modèles analysent la forme et la pose séparément, des études montrent que la morphologie est déterminée par la cinématique en ce sens que la congruence de l'articulation, définie comme l'adéquation morphologique d'une surface articulaire à celle en vis-à-vis, est déterminée par sa cinématique ((Ibáñez et al., 2001; Gil et al., 2004). Inversement, la morphologie des os formant l'articulation peut contraindre leur position relative, c'est-à-dire leur cinématique (Jacq et al., 2008, 2009; Breton et al., 2014). Or, l'analyse séparée des structures d'une même articulation ne fournit aucune information sur leurs poses relatives et ne permet donc pas de tirer parti des corrélations, lorsqu'elles existent. L'utilisation de tels modèles pour la prédiction des structures anatomiques (par exemple, les articulations) sans intégrer les corrélations forme-pose conduirait en fin de compte à des prédictions inexactes, en particulier lorsque l'on dispose de peu de connaissances sur la structure à prédire. Cela limite l'utilité de ces approches dans les tâches d'analyse d'images médicales nécessitant la connaissance des variations de pose des objets d'intérêt. *Ainsi, disposer d'un espace d'analyse où la forme et la cinématique sont représentées ensemble, dans le même espace homogène, peut être intéressant pour la compréhension morpho-fonctionnelle des articulations humaines.*

Cela étant dit, plusieurs modèles intégrant la variation des formes et des poses en utili-

sant l'analyse en composantes principales (ACP) ont été proposés (Smoger et al., 2015; Fitzpatrick et al., 2011) pour étudier les changements intervenant au niveau des zones de contact entre les articulations. Cependant, ces modèles ont utilisé l'analyse en composante principale (ACP), qui n'est pas appropriée pour l'analyse statistique des transformations rigides qui ne sont pas dans un espace linéaire. Pour surmonter cette limitation, plusieurs études ont proposé des cadres d'analyse statistique pour la forme et la pose dans des variétés différentielles riemanniennes en utilisant l'analyse géodésique principale (AGP) (Pennec, 2006; Fletcher et al., 2004). Ces cadres ont été appliqués pour la modélisation des variations de forme et de pose des articulations dans différentes structures telles que les os du poignet (Anas et al., 2014) et les structures neuro-anatomiques (Bossa and Olmos, 2006). Cependant, ces modèles utilisent les angles d'Euler pour représenter les rotations décrivant la pose. *Moreau et al.* (Moreau et al., 2017) a indiqué qu'une telle représentation est limitée dans sa capacité à modéliser des transformations rigides d'objets non-compacts, comme les os long tels que l'humérus et le fémur. À la connaissance de l'auteur, aucun travail n'a été rapporté sur la modélisation de la forme et de la pose des os longs ayant une grande amplitude de mouvement. Cela pourrait être dû à l'inefficacité des approches actuelles pour atténuer l'effet d'élongation qui existe pour les os longs. Ce problème, également appelé "effet de non-compacité", se manifeste par une longueur d'arc de rotation erronée lorsque l'on considère le mouvement des os longs (non compacts) (voir figure 1.1). *L'absence de telles approches de modélisation a limité l'utilité des modèles statistiques pour le diagnostic clinique et le traitement des TMS chez l'humain pour les articulations concernées, comme l'épaule et la hanche.*

Si les modèles statistiques de forme et de pose se sont révélés prometteurs dans l'analyse d'images médicales, l'ajout d'intensité à ces modèles peut s'avérer inestimable pour fournir des connaissances sur les changements dans les contraintes et les déformations de contact entre les articulations. Une telle analyse, ainsi que la modélisation par éléments finis (FEM), peuvent être utilisées pour étudier l'interaction structurelle dans des conditions normales et pathologiques (Dao and Ho Ba Tho, 2015), et peuvent être utilisées pour

évaluer la répartition des contraintes internes sur la structure musculo-squelettique de l'articulation (Zheng et al., 2017). Cependant, les FEM fournissent des résultats génériques puisque les propriétés des matériaux fournies comme entrée des modèles proviennent généralement d'un seul patient ou bien sont des moyennes obtenues à partir de la littérature basée sur des cadavres. Alternativement, des SSMs combinés à une procédure de morphing ont été utilisés pour générer des FEM de forme-paramétrique des tissus et organes du foie et du fémur proximal (Lauzeral et al., 2019) ainsi que pour la colonne lombale (Campbell and Petrella, 2016). La caractérisation de la relation entre la cinématique et la forme ; et la répartition des contraintes (ou déformations) nécessite un grand nombre d'échantillons, ce qui, en pratique, est difficile. Des stratégies d'instanciation d'échantillons synthétiques plausibles ont été rencontrées dans la littérature sous la forme de SSIM (Yao and Taylor, 2001). Cependant, à la connaissance de l'auteur, la combinaison des SSIM avec des informations cinématiques n'a pas encore été signalée dans la littérature.

Enfin, les interventions chirurgicales nécessitent souvent la connaissance de la morphologie des articulations ainsi que de l'orientation relative de l'objet par rapport à d'autres structures. La morphologie des articulations et l'orientation de l'objet sont généralement estimées à partir de volumes $3D$ (images CT ou MR). Cependant, en raison du coût élevé de l'acquisition des images CT et RM, l'imagerie radiographique $2D$ reste l'alternative la plus abordable et la plus disponible, notammnent dans les pays émergents. Bien qu'elles soient plus facilement accessibles, les images radiographiques $2D$ ordinaires ne peuvent pas fournir la géométrie $3D$ requise pour la pratique clinique. Par exemple, il faut percer des trous dans l'os cortical de la diaphyse pour la fixation interne des fractures fémorales et l'opération peut nécessiter une vue $3D$ de la géométrie de l'os pour minimiser les échecs ou les récidives de l'opération (Fox et al., 2013). Pour contourner cette limitation, différents travaux (Yao and Taylor, 2001; Sadowsky et al., 2006; Ehlke et al., 2013; Klima et al., 2016; Reyneke et al., 2018a; Fotsin et al., 2019) ont étudié la morphologie et la pose des structures d'intérêt en estimant une reconstruction 3D à partir d'images radiographiques $2D$. Cependant, les méthodes rapportées n'utilisent que des connaissances a priori de forme,

ou une combinaison de connaisances a priori de forme et de pose, simultanément avec la référence d'intensité pour effectuer les reconstructions $3D$ à partir de $2D$. Ces approches n'intègrent pas toutes les caractéristiques (forme, pose, intensité) dans un seul modèle, et ne tirent donc pas parti de la corrélation qui peut exister entre elles pour des prévisions plus précises.

La discussion ci-dessus souligne l'utilité potentielle d'un cadre statistique générateur qui intègre des informations sur la forme, la pose et l'intensité et leurs corrélations ; le tout dans un modèle unifié. Cependant, avant les travaux présentés dans cette thèse, aucun cadre de modélisation de ce type n'avait été signalé dans la littérature.

1.2. Objectifs de la recherche

Motivés par les lacunes misent en évidence ci-dessus, et cherchant un cadre de principe pour les futures applications comme – i) prévision de la mécanique articulaire spécifique au sujet et du mouvement admissible pour la planification chirurgicale ii) amélioration de la conception des prothèses iii) estimation de la qualité de l'os pour quantifier le risque de fractures osseuses – le but et les objectifs suivants ont été développés pour ce travail de thèse :

> **But : Développement et validation d'un cadre pour la modélisation statistique multi-classes de la forme, de la cinématique et de l'intensité dans les images médicales et, en outre, démonstration de l'utilisation de ce cadre pour l'analyse d'images réelles de patients. Le but a été décliné en trois objectifs décrits ci-dessous.**
>
> ***Objectif 1*** *: Développer un cadre pour les modèles statistiques de la cinématique des formes des articulations humaines à partir d'images médicales qui intègre l'intensité.*
>
> Un nouveau cadre appelé "Dynamic Multi Feature-Class Gaussian Process Model"

(DMFC-GPM) a été développé. Ce cadre fournit un espace latent qui encode les caractéristiques de forme, de pose et d'intensité dans un domaine continu pour permettre la marginalisation à la demande d'une région ou d'une caractéristique d'intérêt sans avoir à apprendre des modèles séparés.

Objectif 2 *: Développer un cadre permettant d'adapter le DMFC-GPM aux données et de valider les performances des modèles.*

Une méthode d'ajustement des modèles pour l'utilisation du DMFC-GPM dans les tâches d'analyse a été mise au point. L'approche utilise le DMFC-GPM comme connaissance a priori dans une approche Monte Carlo par Chaînes de Markov (MCMC) pour la prédiction automatisée des structures, compte tenu de certaines observations dans l'image. Les prédictions peuvent être basées sur des observations partielles de maillages de surface, de contours ou de nuages de points pour des images à 2 ou 3 dimensions. La méthode exploite la corrélation des caractéristiques clés (ou d'un sous-ensemble de celles-ci) intégrées dans l'espace latent pour faire des prédictions sur les informations manquantes. En outre, le cadre d'ajustement peut être utilisé pour automatiser la segmentation, la reconstruction et le recalage des articulations humaines à partir d'images médicales.

Objectif 3 *: Démontrer l'utilisation des DMFC-GPM pour l'analyse des images médicales.*

Le cadre complet (modélisation et ajustement) a été démontré en l'appliquant pour l'analyse des articulations. Plus précisément, le cadre DMFC-GPM a été utilisé pour estimer la forme, la cinématique et le profil d'intensité des structures dans les images des articulations de l'épaule et de la hanche.

La structure de cette thèse est résumée dans le schéma de la figure 1.2. Après ce chapitre d'introduction, tous les chapitres, tels qu'ils sont présentés dans le manuscrit anglais de

manière détaillée, sont résumés ci-dessous. Les références citées ici sont celles du document principal en anglais.

2. Revue de littérature

Ce chapitre présente la synthèse de l'état actuel des connaissances en matière de modélisation statistique de la forme, de la pose et de l'intensité des images, ainsi que leurs applications et leurs limites, en particulier pour la modélisation des articulations humaines. La première section décrit les caractéristiques pertinentes de l'articulation humaine qui seront utilisées tout au long de cette thèse. Par la suite, des modèles statistiques de forme et d'intensité pour des objets individuels sont présentés et leurs limites sont mises en évidence. Ensuite, des modèles statistiques de forme et de pose sont décrits et les limites actuelles sont exposées. Enfin, les conclusions tirées de la littérature sont exposées dans le contexte de la motivation de cette recherche.

3. Mise en correspondance de forme, pose et intensité

Dans ce chapitre, deux méthodes de mise en correspondance multi-objets sont proposées, une paramétrique, l'autre basée sur l'échantillonnage. Ces méthodes utilisent des modèles deformables basés sur des processus gaussiens (Gaussian Process Morphable Models ou GPMMs, Lüthi et al. (2017)) comme espace latent et une approche basée sur les moindres carrés pour l'estimation de la transformation rigide donnant l'orientation spatiale des objets. La mise en correspondance des formes volumiques (intensités internes) est obtenue par interpolation des correspondances basées sur la surface. Enfin, pour valider les méthodes de mise en correspondance, des données synthétiques ont été générées. Elles consistent en des formes simplifiées de type "lollipop" décrites par leur surface maillée, complétée d'intensités internes. Un modèle statistique de forme a été construit sur une base de "lollipops" et utilisé pour effectuer le recalage. Puis un ensemble d'articulations test constituées de

plusieurs “lollipops” (deux ou trois par articulation) a été utilisé comme multi-structures cibles et le recalage paramétrique a été effectué pour établir une correspondance entre elles. Le résultat a montré que la méthode de recalage proposée peut établir correctement une correspondance forme-pose, et donc que chaque articulation peut être remplacée par son correspondant recalé pour l'analyse des variations et corrélations forme, pose et intensité.

4. Dynamic Multi Feature-Class Gaussian Process Model

Ce chapitre présente le cœur de la contribution théorique de cette thèse : un framework de modélisation dynamique à plusieurs classes de caractéristiques basé sur des processus gaussiens qui code la forme, la pose et l'intensité dans le même espace statistique. Le modèle qui en résulte est appelé “Dynamic Multi Feature-Class Gaussian Process Model” (DMFC-GPM). Le cadre DMFC-GPM se suffit à lui-même, c'est-à-dire que l'espace statistitique basé sur la forme-pose-intensité ne nécessite pas de modèle supplémentaire pour résoudre les tâches d'analyse d'images médicales. De plus, toutes les caractéristiques sont construites comme un processus gaussien sur les vecteurs de déformations, ce qui les rend indépendantes de la résolution des objets dans la base de donnée d'apprentissage et de leur discrétisation artificielle. Ce travail unifie les modèles basés sur les caractéristiques en étendant leur interprétation à un contexte plus large. Les corrélations des trois types de caractéristiques (forme, pose, intensité) intégrées dans le modèle devraient permettre de mieux comprendre l'information morpho-fonctionnelle spécifique d'un patient et de révéler les informations sous-jacentes. La figure 4.1 montre l'organigramme du DMFC-GPM proposé. Les verrous résolus pour construire ce modèle sont les suivants :

1. la linéarisation des transformations rigides et l'utilisation d'une métrique adaptée, basée sur l'énergie de déplacement,

2. l'homogénéisation des caractéristiques de forme, de pose et d'intensité dans un même

vecteur de déformations,

3. le calcul d'un espace latent continu basé sur un processus gaussien.

De plus, le calcul d'un noyau flexible permet d'adjuster le modèle si necessaire. Nous avons aussi proposé le calcul du modèle marginal, du modèle de permutation et le calcul du modèle de régression. En plus de ces modèles dérivés, le cadre DMFC-GPM va au-delà de la forme et permet la marginalisation de formes uniques, de formes multiples, d'intensité ou de caractéristiques multiples basées sur une région spéficique de l'objet, comme le résume la figure 4.5.

5. Ajustement du DMFC-GPM et validation

Dans ce chapitre, nous effectuons une validation complète pour évaluer la performance du DFMC-GPM developpé au chapitre 4. Avant cela, il a été nécessaire de mettre au point une méthode d'ajustement de type Monte Carlo par Chaînes de Markov (MCMC) pour l'utilisation des modèles dans différents scenari. Le cadre DFMC-GPM est validé en utilisant des données synthétiques construites à des fins spécifiques incluant des variations et des corrélations de caractéristiques entièrement contrôlées. Cette validation comporte deux sous-parties. Dans la première partie, la capacité du modèle à expliquer la forme et la pose est évaluée. Dans la deuxième partie, la capacité du modèle à expliquer la forme, la pose et l'intensité est évaluée.

1- **Analyse des formes et des poses**. Pour évaluer la performance de la méthode proposée dans l'explication des caractéristiques de la forme et de la pose, un ensemble de données synthétiques a été créé avec des modes de variation de la forme et de la pose rigoureusement connus. Des articulations artificielles ont été créées à l'aide de deux "lollipops" par scenario (figure 5.1). Chaque articulation était composée de deux "lollipops" (objets 1 et 2) dont la forme variait dans un seul sens. Pour

la variabilité de la pose, un mouvement d'articulation a été simulé. Pour chaque articulation générée, l'objet 2 a été tourné par rapport à l'objet 1 en utilisant la convention de l'angle d'Euler pour décrire les transformations rigides. L'objet a été déplacé dans le plan $yz-$ correspondant aux angles d'Euler définissant un mouvement (voir figure 5.2, droite). Dans cette partie, nous avons étudié plus spécifiquement le cadre "Dynamic Multi-Object Gaussian Process Model" (DMO-GPM) qui est un cadre derivé de DMFC-GPM expliquant uniquement les variations de forme et de pose. Le cadre DMO-GPM a été évalué pour sa capacité à extraire correctement la corrélation entre les caractéristiques de la forme de l'objet. Le résultat a montré que le modèle peut parfaitement capturer la corrélation si elle existe. Nous avons également évalué la capacité du cadre DMO-GPM à exploiter la corrélation pour améliorer les prédictions d'informations manquantes. Les objets ont été prédits sans qu'aucune caractéristique de l'objet à prédire ne soit observée. Les erreurs de prédiction étaient relativement importantes par rapport aux prédictions avec observation complète des caractéristiques. Cependant, les résultats obtenus montrent clairement que toute corrélation intégrée dans le modèle fournit des connaissances supplémentaires qui peuvent être utiles pour améliorer les prédictions. Nous avons également évalué les attributs du DMO-GPM, y compris les modèles de régression et de permutation. Le résultat a montré que le cadre pourrait parfaitement être marginalisé à toute caractéristique d'intérêt. La qualité du modèle proposé a été évaluée en comparant quantitativement ses mesures de spécificité et de généralisation à celles des SSPMs utilisant la représentation d'Euler pour des transformations rigides. Le résultat a montré que le modèle proposé est supérieur à celui avec la représentation d'Euler pour la spécificité et la généralisation.

2- **Analyse de la forme, de la pose et de l'intensité**. Le cadre DMFC-GPM avec ses trois classes de caractéristiques (forme, pose et intensité) a également été validé. La capacité de ce modèle à expliquer toutes les classes de caractéristiques par l'échantillonnage a été évaluée ainsi que sa qualité de prediction par ajustement. Des

volumes ont été générés à partir des surfaces maillées des “lollipops”. Les volumes binaires des “lollipops” ont été transformés pour intégrer l’intensité pour cette expérience. Une centaine d’échantillons de DMFC-GPM construits à partir des données des “lollipops” ont été générés de manière aléatoire. La corrélation entre les formes des objets, celle entre les formes et les intensités ainsi que celle entre les poses des objets des échantillons générés ont toutes été évaluées. Le résultat a été comparé à celui des exemples dans la base de données. Le résultat montre que le framework a correctement expliqué la corrélation à partir des objects dans la base de données, le pourcentage de corrélation étant approximativement le même dans les deux cas. Pour évaluer la performance du DMFC-GPM dans la prédiction de la forme-pose et de l’intensité, sa performance dans la prédiction à partir des observations a été comparée aux méthodes de l’état de l’art. Les erreurs RMS (root mean square) basées sur le profil d’intensité prédit et sur celui observé, ont montré que notre framework est plus performant que celui utilisant la représentation d’Euler. La figure 5.17 montre quelques exemples de prédiction à partir de volumes $3D$.

6. Application du cadre DMFC-GPM pour l’analyse forme-pose des articulations en imagerie médicale

Dans ce chapitre, le cadre Dynamic Multi-Object Gaussian Process Model (DMO-GPM) est illustré pour l’analyse des épaules à partir de données d’images CT. Le DMO-GPM est un sous-modèle du DMFC-GPM, il est particulièrement adapté à l’analyse couplée de la forme et de la pose. Les données relatives aux épaules étaient constituées de surfaces maillées de l’humérus et de l’omoplate segmentées à partir d’images CT des épaules bilatérales de cadavres frais. Le cadre DMO-GPM est appliqué pour expliquer la forme et un mouvement d’abduction simulé avec amplitude ample. Les données utilisées sont d’abord décrites, puis un aperçu de la manière dont la correspondance anatomique est établie dans les articulations de l’épaule, est fourni. Ensuite, un mouvement d’abduction est simulé pour

obtenir des données ayant une grande amplitude de mouvement. Ces données sont utilisées pour démontrer la performance du modèle pour l'analyse du mouvement. Le succès de la prédiction en utilisant le DMO-GPM dans divers scenari d'observation laisse entrevoir une utilité clinique pour la récupération des formes prémorbides, permettant l'analyse à partir d'images incomplètes, ou la prédiction d'un objet anatomique en utilisant la connaissance d'un objet anatomique voisin (voir figure 6.8 : ligne du haut). Il a été démontré que le cadre peut faire de bonnes estimations du mouvement de l'articulation de l'épaule (voir figure 6.8 : ligne du bas). Ainsi, ce nouveau cadre de modélisation pourrait servir de descripteur $3D$, couplé forme-pose, pour les articulations humaines à partir d'images médicales. Ce cadre peut être utilisé pour prédire la forme prémorbide pour la conception de prothèses et l'aide au contrôle chirurgical.

7. Application du cadre DMFC-GPM pour l'analyse de caractéristiques forme-pose-intensité en imagerie médicale

Ce chapitre présente l'utilité du cadre proposé grâce à des études de cas, dans lesquelles le cadre DFMC-GPM est utilisé pour reconstruire et prédire les articulations de l'épaule et de la hanche à partir de données d'images CT. Les données d'images utilisées pour les analyses ont été obtenues à partir de spécimens cadavériques et de sujets vivants. Les données relatives à l'épaule étaient constituées d'images CT bilatérales du complexe de l'épaule cadavérique et chaque volume de données d'image comprenait les clavicules, les humeri et les omoplates. Les données relatives à la hanche étaient constituées d'images de CT d'individus vivants et chaque volume d'images contenait l'os du bassin complet (os iliaque et sacrum) et les deux os du fémur.

1 - La première section présente l'analyse de l'épaule, où les données utilisées sont d'abord décrites, puis un aperçu de la manière dont la correspondance a été éta-

blie, est décrit. Une évaluation de la variabilité de la forme de l'os de l'épaule, de la pose et des caractéristiques d'intensité est ensuite présentée, suivie d'une démonstration de la prédiction des caractéristiques de forme, de pose et d'intensité à partir des images CT de l'épaule. La section sur l'analyse des données de l'épaule se termine par une illustration de la manière dont le DMFC-GPM de l'articulation de l'épaule peut être utilisé pour la génération de radiographies reconstruites numériquement. Les résultats ont montré que le cadre DMFC-GPM est une manière alternative d'analyser les images CT de l'articulation de l'épaule qui fournit plus d'informations sur la forme, la pose et les profils d'intensité simultanément. La figure 7.1 montre les échantillons volumiques de forme-intensité des complexes d'épaule des première, deuxième et troisième composantes principales (PCs) du DMFC-GPM de l'épaule en $3D$ et $2D$.

2 - La deuxième partie du chapitre se concentre sur l'analyse de la hanche effectuée à l'aide du cadre complet. Une description des données des images de la hanche est suivie d'un aperçu de la manière dont la correspondance a été établie pour les articulations de la hanche. Comme pour la section sur l'épaule, une évaluation de la variabilité de la forme de l'os de la hanche, de la pose et du profil d'intensité est présentée ; elle est suivie d'une démonstration de la prédiction des caractéristiques de la forme, de la pose et de l'intensité à partir des images CT de la hanche. En outre, des projections du DMFC-GPM de l'articulation de la hanche sur un espace $2D$, en utilisant un algorithme de radiographie reconstruite numériquement (DRR), sont présentées. La figure 7.7 montre des images de la prédiction de 4 sujets, où l'intensité prédite est indiquée à l'aide d'une échelle rouge-bleue. Les articulations de la hanche sont prédites avec le profil d'intensité et les contours du bassin, des fémurs gauche et droit.

8. Conclusion

Cette section présente les principaux résultats issus de cette thèse. Le travail de recherche a été motivé par trois aspects. Premièrement, il s'agissait de pallier l'absence d'outils informatiques permettant d'exploiter les informations sous-jacentes issues de la combinaison des principales caractéristiques des images médicales : la forme, la pose et l'intensité, utilisées cliniquement pour le diagnostic, la planification et l'orientation de la chirurgie. Les modèles statistiques actuels qui fournissent des espaces latents pour l'analyse de la forme, de la pose et de l'intensité de l'image ne tiennent pas compte d'une description continue des objets d'intérêt, ils n'intègrent pas non plus toutes les caractéristiques clés et leurs corrélations. Deuxièmement, la prédiction et l'évaluation des structures dans les images médicales à des fins de diagnostic ou de traitement (comme l'évaluation des troubles musculo-squelettiques) se fait encore par l'ajustement successif de modèles statistiques de forme, de pose ou d'intensité pour trouver la forme, l'orientation ou le profil optimal des objets. L'utilisation de modèles statistiques distincts ne permet pas de tirer parti des corrélations qui existent entre ces caractéristiques, ce qui augmente le risque de prévisions inexactes. Troisièmement, l'analyse des articulations humaines à partir d'images médicales à l'aide de modèles statistiques reste un domaine de recherche ouvert et représente un défi en raison du nombre de degrés de liberté (DDL) et de la grande amplitude de mouvement de certaines articulations, notamment celles de l'épaule et de la hanche présentées ici. Le défi est encore plus grand lorsqu'il est associé à une grande variation naturelle de forme et d'intensité. Ainsi, l'objectif global du projet était de développer et de valider un cadre pour la modélisation statistique multi-classes de la forme, de la cinématique et de l'intensité dans les images médicales. En outre, il était essentiel de démontrer l'utilisation de ce cadre pour l'analyse d'images réelles de patients. Cela a nécessité la généralisation, l'adaptation et l'application de certaines techniques statistiques existantes, ainsi que la proposition de nouvelles méthodologies. Les contributions techniques de ce travail sont classées en trois catégories, et résumées ci-dessous.

8.1. Première contribution

Une méthode a été mise au point pour estimer l'espace latent continu paramétrique à partir de structures dans les données d'images médicales avec variation selon les caractéristiques de forme, de pose et d'intensité. Des méthodes permettant d'établir la correspondance entre la forme, la pose et l'intensité pour de multiples objets ont été proposées au chapitre 3. Ces méthodes comprennent des approches paramétriques et basées sur l'échantillonnage. Ces approches d'établissement de correspondance ont été utilisées pour trouver des champs de déformation entre les objets de référence et les objets cibles. Les champs de déformation obtenus à partir des données d'entraînement ont ensuite été linéarisés et normalisés pour obtenir des descripteurs d'objets dans un espace vectoriel, c'est-à-dire en projetant les fonctions dans une Variété différentielle riemannienne vers l'espace tangent au champ de déformation moyen tel que présenté au chapitre 4. Le goulot d'étranglement de la non-linéarité associée à la transformation rigide a été abordé en utilisant une représentation qui conduit à une norme L^2 proportionnelle à l'énergie cinétique nécessaire pour déplacer l'objet d'une position à l'autre. Un espace latent basé sur un processus gaussien a été calculé sur ces champs de déformation codant les caractéristiques de forme, de pose et d'intensité dans le même espace. Un tel espace latent est paramétré par un ensemble de fonctions orthogonales codant non seulement la variabilité de chaque classe de caractéristiques dans un domaine continu, mais aussi les corrélations qui existent entre elles. Ce cadre de modélisation a été baptisé "Dynamic Multi-Feature Class Gaussian Process Models" (DMFC-GPM). Il est important de noter que, comme toutes les caractéristiques se trouvent dans un espace statistique commun et dans un domaine continu, ce cadre permet la marginalisation à la demande d'une région ou d'une caractéristique d'intérêt sans avoir à retravailler des modèles séparés, contrairement aux méthodes actuelles de la littérature (Cootes et al., 1992; Lüthi et al., 2017; Moreau et al., 2017; Anas et al., 2016). En outre, la continuité du modèle permet également de dériver des sous-modèles à partir du DMFC-GPM, tels que les modèles forme-cinématique "Dynamic Multi-Object Gaussian Process Model" (DMO-GPM) et les modèles de régression, qui ont également été présentés

et discutés. Le DMO-GPM est particulièrement adapté à l'analyse de la forme et de la pose. Le modèle de régression peut être limité à certaines observations dans l'image, ce qui améliore la précision des résultats des prévisions puisqu'il réduit l'espace des échantillons possibles. En outre, les modèles statistiques précédents (SSPM, SSM, SSIM, AAM ou SAM,) peuvent être considérés comme des attributs du cadre unifié développé (voir la figure 4.5). Pour résumer, les principales contributions au sein de cette première catégorie sont les suivantes :

- Nouvelles méthodes paramétriques basées sur l'échantillonnage pour établir la correspondance anatomique multi-objets.

- Modification d'une métrique non-linéaire pour comparer des transformations rigides dans un espace continu. Cela permet d'atténuer l'inconvénient de "non-compacité" pour les objets non-compacts.

- Processus de normalisation de différentes classes de caractéristiques (c'est-à-dire les caractéristiques de forme, de pose et d'intensité) afin d'obtenir un espace d'analyse homogène.

- Espace latent basé sur un processus gaussien pour la modélisation de plusieurs classes de caractéristiques avec ses propriétés associées caractérisées.

8.2. Deuxième contribution

Une méthode de processus d'adjustement a été développée dans le chapitre 5 pour établir un canal de transfert des caractéristiques de forme, de pose et d'intensité de l'espace latent vers les observations, y compris le nuage de points, les surfaces maillées, les images $3D$ (CT) ou $2D$ (radiographies). Le principal avantage du cadre présenté par rapport aux autres approches est qu'en plus des caractéristiques de forme intégrées dans l'espace latent, il tient également compte de la pose et des caractéristiques d'intensité, ainsi que de toutes leurs corrélations observées, le tout intégré dans un seul cadre. Cette intégration

de diverses caractéristiques dans l'espace latent est mise à profit lors de l'établissement du canal d'apprentissage. Un tel canal améliore la précision de prédiction des caractéristiques de forme, de pose et d'intensité qui sont intéressantes pour le diagnostic et le traitement des troubles musculo-squelettiques. L'approche proposée a été validée par des expériences contrôlées et approfondies, et les résultats ont montré qu'elle est plus performante que les approches existantes incluses dans l'évaluation comparative réalisée. Tout d'abord, en ce qui concerne la variabilité de la pose, le cadre DMFC-GPM était plus précis pour expliquer la transformation rigide sur de multiples objets. En particulier, il est plus spécifique aux données d'entraînement tout en généralisant bien sur de nouvelles données et de manière réaliste. Deuxièmement, il exploite la variation de la pose, encodée sous forme de champ de déplacement d'énergie dans l'espace latent, pour obtenir les paramètres optimaux de la pose à partir d'une observation. Enfin, il exploite le descripteur de la caractéristique d'intensité dans l'espace latent pour déduire le profil d'intensité de la structure interne des objets. Ici, les principales contributions techniques ont été :

- Adaptation de l'échantillonnage du MCMC et des algorithmes de Metropolis-Hastings pour la prédiction multi-objets.

- Incorporation de la projection de radiographie reconstruite numériquement (DRR) dans le processus d'échantillonnage du modèle pour permettre une projection de $2D$ à $3D$. Cela permet d'obtenir un espace latent de $2D$ par rapport à celui de $3D$.

- Développement de données synthétiques contrôlables pour valider le canal de transfert à partir de l'espace latent du modèle.

- Comparaison du modèle proposé avec d'autres approches couramment utilisées, y compris les modèles de forme active (ASM) basés sur les SSM, et les modèles de pose et de forme réalisés en utilisant une métrique de transformation rigide standard.

8.3. Troisième contribution

Le cadre DMFC-GPM a été appliqué pour analyser les données d'images d'articulations, épaule et hanche, dans les chapitres 6 et 7. En évaluant ses performances pour la reconstruction d'os partiels simulés de l'épaule, le cadre DMO-GPM a révélé un potentiel d'utilisation pour déduire la structure anatomique normale à partir de la partie restante de la structure ou d'autres structures dans une articulation ou un complexe. En outre, l'utilisation du DMFC-GPM basé sur les connaissances a priori pour la prédiction de la forme, de la pose et de l'intensité de différentes structures à partir d'images CT et d'images $2D$ de l'épaule et de la hanche, avec leurs os constitutifs (omoplate, clavicule et humérus pour l'épaule ; bassin, fémur gauche et droit pour la hanche) affichés aux orientations spatiales d'imagerie. Les résultats expérimentaux ont clairement indiqué que le cadre ne détecte pas seulement les contours d'un objet par segmentation des structures, comme la plupart des méthodes de pointe, mais qu'il fournit les profils d'intensité des os spécifiques du patient dans les images CT. Il a également été démontré que le DMFC-GPM des complexes de l'épaule et des articulations de la hanche peut être projeté dans un espace $2D$ en tant que DRR, ce qui pourrait être étendu pour l'analyse de n'importe quelle articulation en $2D$. Les principales contributions à l'analyse des images médicales dans cette partie ont été :

- Prétraitement des données de l'épaule et de la hanche à partir d'images CT pour obtenir les structures médicales souhaitées pour étudier le cadre proposé.

- Démonstration du cadre sur les complexes de l'épaule et de la hanche à partir d'images CT et de surfaces maillées pour évaluer son efficacité en utilisant des données d'images réelles.

Le cadre DMFC-GPM réalise un nouveau paradigme dans la modélisation et l'analyse de structures multiples présentes dans les images médicales, permettant une modélisation probabiliste et une analyse statistique non seulement de la forme mais aussi de la pose et de l'intensité. Les corrélations qui existent entre les caractéristiques et qui sont observées

pendant la création du modèle, peuvent maintenant être étudiées et exploitées dans des outils cliniques permettant l'évaluation des troubles musculo-squelettiques à partir des images médicales.

9. Financement

Ce travail de thèse a été réalisé dans le cadre d'un partenariat entre IMT-Atlantique, site de Brest, et l'Université de Cape Town en Afrique du Sud, pour l'obtention du diplôme de doctorat conjoint. Je remercie chaleureusement les soutiens financiers qui ont permis la réalisation de ces travaux pendant trois ans : pour la France, bourse de recherche de Brest Métropole (MNB/MJM no 17-178) et pour l'Afrique du Sud, bourse de la Fondation pour la Recherche Nationale (NRF, grant no 105950 et 114393), bourse de la "South African Research Chairs Initiative" de la NRF et du Département de Science et Technologie (grant no 98788), ainsi que le soutien du Conseil de la Recherche Médicale.

Contents

List of Figures

List of Tables

List of Acronyms

ASM	Active shape model
AAM	Active appearance model
BMD	Bone mineral density
CT	Computed tomography
CDML	Cauchy distribution based maximum likelihood
DMO-GPM	Dynamic multi-object Gaussian process model
DOF	Degrees of freedom
DMFC-GPM	Dynamic multi feature class Gaussian process model
DRR	Digitally reconstructed radiograh
FAI	Femoroacetabular impingement
FEM	Finite element model
FFD	Free-form deformation
FFDM	Free-form deformation model
GPA	Generalised Procrustes analysis
HD	Hausdorff distance
HU	Hounsfield unit

GPMM	Gaussian process morphable model
$\mathcal{GP}$	Gaussian process
IMCP-GMM	Iterative median closest point-Gaussian mixture model
ICP	Iterative closest point
MCMC	Markov chain Monte Carlo
MSD	Musculoskeletal disorder
MRI	Magnetic resonance imaging
RMS	Root mean square
RMSML	Root mean square maximum likelihood
OA	Osteoarthritis
PCA	Principal component analysis
PGA	Principal geodesic analysis
PC	Principal component
SSM	Statistical shape model
SE(3)	Lie group of orientation-preserving symmetries in $\mathbb{R}^3$
SAM	Statistical appearance model
SPM	Statistical pose model
SSPM	Statistical shape and pose model

Chapter One

Introduction

1.1 Motivation

Prevalent worldwide, musculoskeletal disorders (MSD) are often caused by an injury to muscles, bones, joints, tendons, ligaments or nerves (Main and de C Williams, 2002; Nikooyan et al., 2012). They have been reported as the leading cause of physical disability and are one of the most common causes of pain (Pedoia et al., 2016). They have multiple causes including cerebral palsy, bone deformity, trauma, and impact daily life by limiting movement, in some cases leading to early sedentary lifestyle.

All joints in the body are at risk of MSD. However, the level of the risk of MSD depends on the range of motion and the degrees of freedom (DOF) of the joint. Thus, due to their high range of motion and their importance for daily life activities, the shoulder and the hip joints are typically at higher risk of MSD (Woon and Ham, 1991). The literature abounds with reports on hip (Tanzer and Noiseux, 2004; Clohisy et al., 2009; Anderson et al., 2010; Li et al., 2015) and shoulder (Ludewig and Cook, 2000; Ludewig and Braman, 2011; Van der Molen et al., 2017; Park et al., 2019) disorders. Anatomical abnormality of the femur has been associated with increased risk of hip osteoarthritis (OA) progression as a result of elevated stress on the articular cartilage (Lespasio et al., 2018). The morphological factors resulting in femoroacetabular impingement (FAI) involved in the underlying etiopathogenesis of hip OA and in damage of the labrum remain unclear, and are still a focus of MSD and disease research (Murphy et al., 2016). Similarly for the

shoulder joint, the morphology of the constituent bones contributes to shoulder function; for example, it was reported that humeral retroversion and scapular shape determine the range-of-motion of throwing athletes and volleyball players (Ribeiro et al., 2012; Ribeiro and Pascoal, 2012).

Computed tomography (CT), magnetic resonance (MR) or X-ray fluoroscopic imaging are now regularly used for studying the morphology and function of human joints in relation to bone and joint injuries, as well as for the assessment of *in vivo* $3D$ joint kinematics (Kozono et al., 2017; Borotikar et al., 2017). Furthermore, image-based clinical tools for surgical guidance and planing such as BLUEPRINT™(`https://www.shoulderblueprint.com/`) and MocapLab™(`https://www.mocaplab.com/`) have gained mainstream acceptance. Structures displayed in these intra-modal medical images are characterised by three key features: shape (geometrical information of an object relative to another under rigid transformation and scaling invariance), pose (refers to a distinguishable static state of an object and represented by a rigid transformation) and intensity (a numerical value of an image pixel/voxel). Pose quantification is termed as kinematics which describes the motion of points, bodies (objects), and system of bodies (groups of objects) in a defined reference system (Whittaker, 1937). Medical image processes such as segmentation, reconstruction and prediction of the structure of interest from medical images essentially depend on these key features. The success of such processes in morpho-functional analysis of human joints depends on an understanding of the variability of shape, intensity and pose features. The variability in these features include complex joint morphology variability, spatial variability within the image, shape variability between individuals, and image quality variability in contrast characteristics (i.e. demarcation with surrounding tissues) (Seoud et al., 2019).

To date, advanced algorithms in medical image processing have been developed to facilitate shape, pose and intensity analysis (Rueckert and Schnabel, 2020). The common strategy consists of estimating a subspace of homogeneous patterns for shape, intensity and kinematics through data-driven models. From the literature, the current trend for

finding these homogeneous patterns is to employ deep learning-based generative models (Toshev and Szegedy, 2014; Ouyang et al., 2014). However, deep generative models tend to have high specificity but low generality; in some cases, they explain out-of-training data inputs with high confidence while being completely wrong (Huang et al., 2017; Webster et al., 2019; Nalisnick et al., 2018). In addition, deep learning approaches also require large numbers of training examples to be successful.

Alternatively, well established approaches for shape, intensity and kinematics analysis employing statistical models have been reported (Cootes et al., 1992; Bossa and Olmos, 2006; Lüthi et al., 2017). Statistical models of shape (SSM), intensity (SSIM) and pose (SPM), have become ubiquitous in the literature and aim to capture the natural variation of these features in biological objects (Cootes et al., 1992; Sarkalkan et al., 2014; Bossa and Olmos, 2006; Lüthi et al., 2017). These models all aim to find parameterized spaces of homogeneous shape, intensity and pose patterns. Such spaces are then used as priors for automated prediction from observations and interpretation of the learned patterns.

The morphological analysis of human joints using SSMs typically involves a separate consideration of the objects (bones) constituting the joints, such as the humerus and the scapula for the shoulder joints (Inyang et al., 2017; Yang et al., 2015; Salhi et al., 2017). While, these models analyse shape and pose separately it has been reported that joint motions are influenced by the shape of their articular geometries, which contributes to the regulation of their relative position (Ibáñez et al., 2001; Gil et al., 2004; Breton et al., 2014; Smoger et al., 2015; Fitzpatrick et al., 2011). Analysing the structures separately does not provide any pose information, and hence, does not leverage such a correlation when it exists. Using such models for anatomical structure (for example, joints) prediction without embedding shape-pose correlations would ultimately lead to inaccurate prediction, especially when there is little knowledge available of the structure to be predicted. This limits the utility of such approaches in medical image analysis tasks requiring knowledge of pose variations of the objects of interest.

That being said, several models embedding shape and pose variation using principal component analysis (PCA) have been proposed (Smoger et al., 2015; Fitzpatrick et al., 2011). However, the models used linear PCA, which is not appropriate for statistical analysis of rigid transformation. To overcome this limitation, several studies have proposed statistical analysis frameworks for shape and pose in Riemannian manifolds using principal geodesic analysis (PGA) (Pennec, 2006; Fletcher et al., 2004). These frameworks have been applied for the modelling of shape and pose variation of joints across different structures such as wrist bones (Anas et al., 2014) and neuroanatomical structures (Bossa and Olmos, 2006). However, such models used angles for modelling the rotations to describe the pose. Moreau et al. (2017) reported that such a representation is limited in its ability to model rigid transformations of non-compact objects. A non-compact object is an object with small area compared to the area of its bounding circle (for example, long bones such as humeri and femurs). To the author's best knowledge, no work has been reported on the shape and pose modelling of long bones with high range of motion. This could possibly be due to the inefficiency of current approaches in modelling high range of motion and the difficulty of accurate pose modelling of non-compact objects. The latter problem, also called the "non-compactness effect", manifests as erroneous rotation arc length when considering bones of different lengths (see figure 1.1) (Moreau et al., 2017). The lack of such modelling approaches has limited the utility of statistical models for clinical diagnosis and treatment of MSD of human joints with long bones such as shoulder and hip joints.

While statistical shape and pose models have shown promise in medical image analysis, the addition of intensity to these models may be invaluable in providing knowledge on the changes in joint contact stress and strain. Such analysis, along with finite element modelling (FEM), can be used to investigate the structural interaction under normal and pathological conditions (Dao and Ho Ba Tho, 2015), and can be used to assess the internal stress distribution on the musculoskeletal structure of the joint (Zheng et al., 2017). However, FEM provides generic results since the material properties provided as model input are typically from a single patient or are averages obtained from literature based on cada-

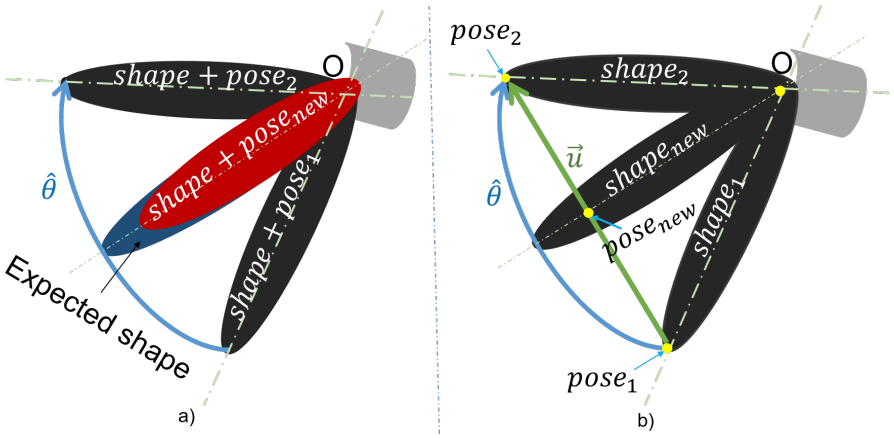

Figure 1.1 Illustration of the non-compactness problem. Left: a joint with a fixed object (grey) and an elongated moving one (black). The elongated object in red is the interpolation of the two positions of the moving object. The blue colour shows the error in the resultant shape. Right: The additional points (yellow) selected on the fixed object 1 and 2 can be used as position parameters. Interpolating the rotation ($\hat{\theta}$) curve is then the same as interpolating on the segment ($\vec{u}$) resulting in the correct shape.

vers. Alternatively, SSMs combined with a morphing procedure have been used to generate shape-parametric FEMs of tissues and organs of the liver and proximal femur (Lauzeral et al., 2019) as well as for the lumbar spine (Campbell and Petrella, 2016). Characterising the relationship between the shape-kinematics and the stress (or strain) distributions requires a large number of samples which, in practice, is challenging. Strategies for instantiating plausible synthetic samples have been encountered in the literature in the form of SSIMs (Yao and Taylor, 2001). However, to the author's best knowledge, the combination of SSIMs with kinematic information has thus far not been reported in the literature.

Finally, surgical procedures often require knowledge of joint morphology as well as the object's relative orientation to other structures. Joint morphology and the object's orientation are usually estimated from $3D$ volumes (CT or MR images). However, due to high cost of CT and MR image acquisition, $2D$ X-ray imaging remains the most affordable and available alternative. While more readily accessible, regular 2D X-ray images are unable to provide the $3D$ geometry required for clinical practices. For example, drilling holes through the cortical bone of the shaft of the femur is required for internal fixation

of femoral fractures and the operation may require $3D$ view of the bone geometry to minimise surgery failures or recurrences (Fox et al., 2013). To circumvent this limitation, $3D$ from $2D$ reconstruction has been reported as a possible means to estimate the morphology and pose of structures of interest from 2D X-ray images (Yao and Taylor, 2001; Sadowsky et al., 2006; Ehlke et al., 2013; Klima et al., 2016; Reyneke et al., 2018a; Fotsin et al., 2019). However, reported methods only use shape priors, or a combination of shape and pose priors, simultaneously with intensity distribution reference to perform $3D$ from $2D$ reconstructions. Such approaches do not embed all these features in a single prior and hence do not leverage the correlation that may exist between them for accurate predictions.

The above discussion points to the potential utility of a generative statistical framework that incorporates shape, pose and intensity information and their correlations; all in one unified model. However, prior to the work presented in this thesis, no such modelling framework had been reported in the literature.

1.2 Research aim and objectives

Motivated by the gap highlighted above, and seeking a principled framework for the following future applications: 1) prediction of subject-specific joint mechanics and allowable motion for surgical planning; 2) enhanced prosthesis design; and 3) bone quality estimation to quantify the risk of bone fractures; the following aim and objectives were developed for this thesis work.

Aim: Development and validation of a framework for multi-class statistical modelling of shape, kinematics and intensity in medical images and furthermore, demonstration of the use of the framework for analysing real patient images. The aim was factorised into three objectives outlined below.

Objective 1: To develop a framework for statistical shape-kinematics models of human joints from medical images that embeds the intensity.

A novel framework called Dynamic Multi Feature-Class Gaussian Process Modelling (DMFC-GPMs) was developed. The framework provides a latent space that encodes shape, pose and intensity features in a continuous domain to permit on-demand marginalisation to a region or feature of interest without training separate models.

Objective 2: To develop a framework for fitting DMFC-GPMs to data and validate model performance.

A model fitting method for using the DMFC-GPM in analysis tasks was developed. The approach uses a DMFC-GPM as a prior in a Markov chain Monte Carlo (MCMC) approach for automated prediction of structures, given some observations in the image. Predictions can be based on observations of partial surface meshes, contours or point clouds in both a $2D$ and $3D$ image setting. The method leverages the correlation of the key features (or a subset of them) embedded in the latent space in making predictions of missing information. Furthermore, the fitting framework can be used to automate the segmentation, reconstruction and registration of human joints from medical images.

Objective 3: To demonstrate the use of DMFC-GPMs for analysing medical images

The full framework (modelling and fitting) was demonstrated through applications for joint analysis. Specifically, the DMFC-GPM framework was used to estimate the shape, kinematics and intensity distribution of structures in images of shoulder and hip joints.

1.3 Ethical considerations

This research involved the use of medical images from patients as well as from fresh cadavers. Living subject CT images of the lower limb reviewed and approved by the ethics committee of the Ghent University Hospital (reference B670201111480) were obtained. Ethical approval was obtained from the Faculty of Health Sciences Human Research Ethics Committee (HREC) of the University of Cape Town for use of previously collected CT

images of the fresh frozen cadaver shoulders (reference HREC 283/2020). Additional CT image data of the torso were obtained, with permission, from the SICAS Medical Repository in Switzerland (SICAS:http://www.si-cas.com/).

1.4 Project outputs

International conference with Proceedings

- **Fouefack, J.R.**, Borotikar, B., Douglas, T.S., Burdin, V. and Mutsvangwa, T.E.M. 2020. Dynamic Multi-object Gaussian Process Models, In: Martel, A.L., Abolmaesumi, P., Stoyanov, D., Mateus, D., Zuluaga, M.A., Zhou, S.K., Racoceanu, D., Joskowicz, L. (Eds.), Lecture Notes in Computer Science Book Series 12264. Springer Nature Switzerland AG 2020, pp. 755-764.

- **Fouefack, J.R.**, Alemneh, T., Borotikar, B., Burdin, V., Douglas, T.S. and Mutsvangwa, T., 2019, July. Statistical shape-kinematics models of the skeletal joints: Application to the shoulder complex. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 2019, (pp. 4815-4818). IEEE.

- I presented our work on "Dynamic Multi-object Gaussian Process Models" at the 23^{rd} International Conference on Medical Image Computing and Computer Assisted Intervention MICCAI, 4-8 October 2020, Lima, Peru (virtual)

- I presented our work on "Statistical shape-kinematics models of the skeletal joints: Application to the shoulder complex" at the 41^{st} Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 23-27 July 2019, Berlin, Germany

Arxiv preprint

Fouefack, J.R., Borotikar, B., Douglas, T.S., Burdin, V. and Mutsvangwa, T.E., 2020. Dynamic multi-object Gaussian process models: A framework for data-driven functional modelling of human joints. arXiv:2001.07904.

Hackathon

I initiated the development of tetrahedral-mesh based shape modelling in Scalismo and organized, together with members from the Graphics and Vision Research Group, University of Base, a hackathon where this was integrated into the core part of the Scalismo software (https://scalismo.org/).

Data and Coding

The following research outputs are available on request to the author's supervisors for research use:

- The synthetic data developed in this project consisting of lollipop joints including surface meshes, tetrahedral meshes and 3D image volumes.
- Segmented Swiss shoulder data are available in the author's research group for research use. These data include processed 3D humeri, scapulae and clavicles.
- Well-commented project code implemented in Scalismo using the Scala programming language.

1.5 Organisation of the thesis

The structure of this thesis is summarised in the schematic in figure 1.2 and the overview is as follows: Chapter 2 presents a review of the literature relevant to the thesis. Starting with a general description of the notion of a human joint, the chapter includes a review

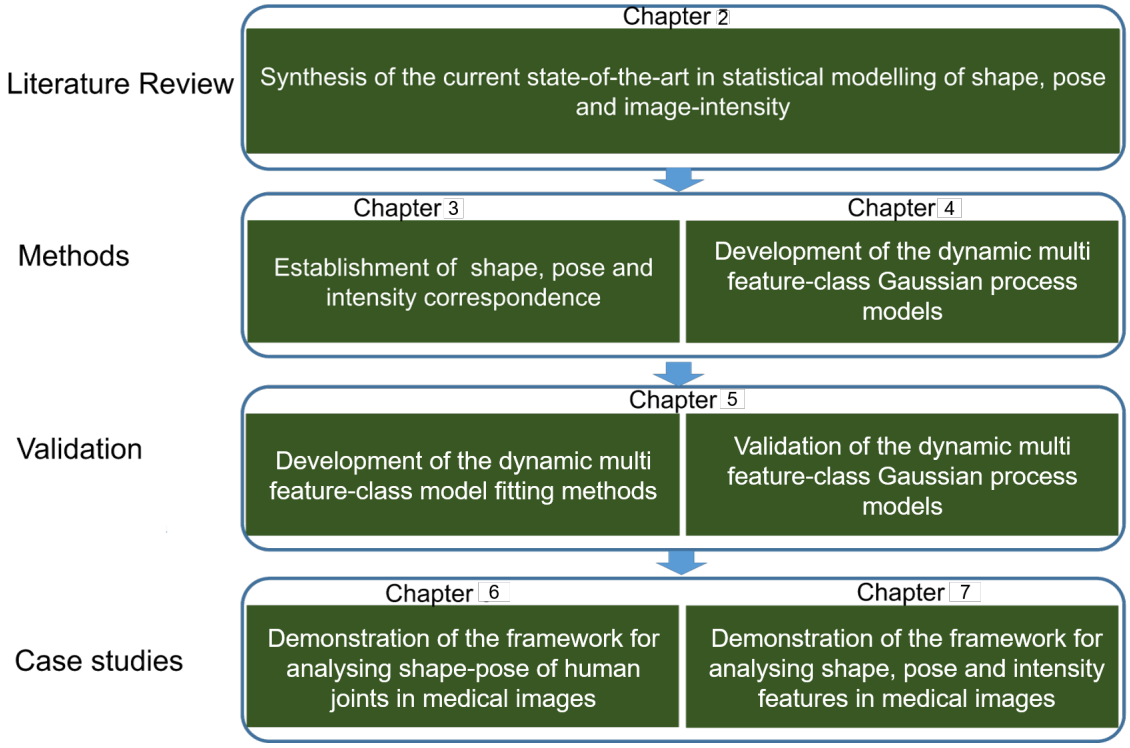

Figure 1.2 Summary of the project.

of the current state-of-the-art in statistical modelling of shape, pose and image-intensity; applications of such statistical modelling in image analysis; and limitations of statistical modelling, especially for modelling human joints. The chapter concludes with an identification of gaps and recommendations for identified research areas. Chapter 3 presents methods developed for parametric and sampling-based establishment of correspondence of multiple objects in a joint complex. Chapter 4 describes in full, the developed Gaussian process-based dynamic multi feature class modelling framework. Chapter 5 is a presentation of the MCMC-based fitting method for using the models in different scenarios, as well as a comprehensive validation of model performance. The development of synthetic data with controlled analytically defined feature variation is also described as these data form part of the validation protocol. In Chapter 6 and 7, the utility of the modelling framework is showcased using real imaging data of shoulder and hip joints. Finally, Chapter 8 outlines the main findings of the work, as well as limitations and recommendations for the future work. The chapter ends with a description of the main project outputs.

Chapter Two

Literature Review

This chapter presents the synthesis of the current state-of-the-art in statistical modelling of shape, pose and image-intensity along with their applications and limitations, especially when modelling human joints. The first section describes the relevant features of the human joint that will be used throughout this thesis. Thereafter, statistical shape and intensity models for single objects are presented and their limitations highlighted. Next, statistical shape and pose models are described and current limitations unpacked. Finally, conclusions drawn from the literature are outlined in the context of the motivation for this research.

2.1 Human joint and its features in medical images

Human joints are articulations between two bones on frictionless surfaces made up of cartilage and are stabilised by multiple soft tissue structures. They facilitate locomotion and help perform activities of daily life. They are part of the human musculoskeletal system (Martini et al., 2006). The human body consists of several joints of different levels of complexity (in terms of the range of motion and structure), each of which play a crucial role in mobility. Among these joints, the shoulder and hip joints are the most mobile, that is, they have more degrees of freedom (DOF) and allow higher ranges of motion. This makes these joints vulnerable to injury or damage. Hence, joint analysis methods demonstrated on these joints can be easily extended to other joints. A brief overview of the anatomical structure of these joints is presented below.

The shoulder complex consists of three bones: scapula, humerus and clavicle. It is a com-

bination of four joints including the glenohumeral joint (shoulder joint), acromioclavicular joint, sternoclavicular joint and the scapulothoracic joint (an imaginary floating joint) (Kendall, 1989; Good and MacGillivray, 2005). The glenohumeral and sternoclavicular joints are the main joints of the complex. The glenohumeral joint is a synovial ball-and-socket joint and the sternoclavicular joint a gliding style synovial joint. Bones, muscles and tendons involved in these joints provide a large ranges of motion to the upper extremity in all three planes of motion (Kendall, 1989). The shoulder complex is highly susceptible to dysfunction and instability due to its DOF coupled with its intricate structural and functional design.

The hip joint is the joint formed between the thighbone (femur) and the pelvis. Similar to the shoulder joint, the hip joint is a synovial ball-and-socket joint. The round head of the femur resting in a cavity (the acetabulum) together with the complex structure of muscles, bones, tendons and ligaments allow free rotation of the limb and results in a compromise between stability and mobility (Harty, 1984).

The study of human joints from medical images requires the analysis of the following three key features determined from the image: shape, pose and intensity.

- **The shape refers to the geometrical information of an object of interest relative to another under rigid transformations and scaling invariance** (Kendall, 1989). That is for example, any geometrical information of the scapula bone remaining after its alignment to a reference.

- **The pose refers to a distinguishable static state of an object** and it is represented by a rigid transformation (Brégier et al., 2018). Pose quantification is termed as kinematics which describes the motion of points, bodies (objects), and systems of bodies (groups of objects). An example is the transformations that move the humerus bone relative to the scapula during external rotation, extension and abduction motion as shown in figure 2.1.

- **Intensity refers to a numerical value of an image pixel/voxel**. An example is a voxel value in a computed tomographic (CT) image. Intensities depend on the contrast sensitivity of the imaging system, hence, typically require a preprocessing step for calibration.

The morphology of the bones forming the joint contribute to the regulation of their relative position and a simple way of understanding the correlation between shape variation and joint motion is essential for the management of multiple joint-related disorders (David et al., 2012; Ibáñez et al., 2001; Gil et al., 2004; Breton et al., 2014; Smoger et al., 2015). The challenge is due to the complexity in shape that highly varies across individuals. It is known that the biomechanics of joints are influenced by the shape of their articular geometries (Smoger et al., 2015). An understanding of bone deformity and kinematics, as well as the intensity profiles across the joints, may contribute to improved diagnosis and treatment of joints disorders. Statistical models have been shown to improve such understanding of shape and kinematic variability (Smoger et al., 2015). The following section presents the synthesis of modelling methods that have been proposed to address these challenges.

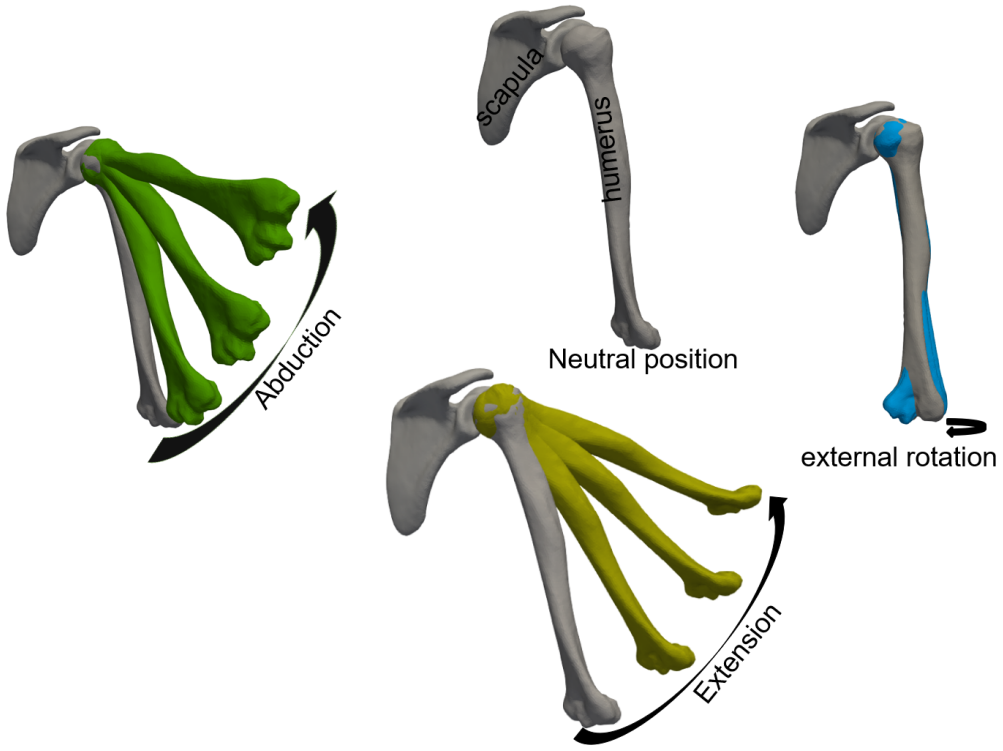

Figure 2.1 The shoulder joint and different motion types. The humerus in grey is at 0° relative to the scapula (in grey), humeri in green are at different abduction angles relative to the scapula, humeri in dark yellow are at different extension angles relative to the scapula and humeri in blue are at different internal rotation angles.

2.2 Statistical modelling of bone structures

Statistical modelling of objects (typically bone structures) explores naturally occurring variation across the object family, summarising this variation with a few parameters. These statistical models can be categorised into statistical shape models (SSMs), active appearance models (AAMs), active shape models (ASMs), statistical shape and intensity models (SSIMs) and statistical shape-pose models (SSPMs). All these models work by learning the variation from training examples to define the space of possible occurring shapes, intensities and poses (Cootes et al., 1992, 1995, 1998; Sarkalkan et al., 2014; Bossa and Olmos, 2006; Reyneke et al., 2018b). Strategies proposed to derive these models are complex; depending on the object representation (continuous or discrete representation),

as well as the number of objects per training example. Starting from single object models (SSMs, ASMs, AAMs, SSIMs) that only involve one object per training example to multi-object models (SSPMs) that incorporate multiple objects per training example, different models have been proposed in the literature. However, several challenges persist in all the outlined modelling approaches, limiting their utility as will be outlined in the following sections. The simplest modelling approach, single object models, provides a good starting point and is the topic of the next section.

2.2.1 Single object models

Single object models are built from datasets containing one object per training example. These modelling approaches started from SSMs (also called point distribution models (PDMs)) that have shown the potential of explaining variation across different structures represented as discretized objects (Cootes et al., 1992). The assumption, when developing and using single object statistical models, is that the shape of the same family of bones across individuals within a population can be obtained by adding major modes of variation to their respective averages. The models can be developed from objects' shape features obtained from $3D$ image volumes such as CT and magnetic resonance (MR) images. The SSMs are typically developed through the following sequence of steps: 1) Extraction of objects from $3D$ image volumes, often done through segmentation; 2) Shape correspondence establishment, often done through rigid and non-rigid registration as well as through the generalised Procrustes analysis (GPA) (Gower, 1975; Duchon, 1977; Kybic and Unser, 2003; Myronenko et al., 2007; Lüthi et al., 2013); and 3) Calculation of the average and the modes of variation using some form of ordination, typically principal component analysis (PCA) (Wold et al., 1987). Below is a mathematical formulation of the SSMs.

A- Statistical shape models

Let $\{X_S^i, i = 1, \ldots, n\}$ be the representation of the training dataset with n examples, where X_S^i is a vector representing the i^{th} object. Regardless of how the correspondence

is established, once obtained, the shape analysis pipeline follows the same process. An instance X_S from the resultant SSM for a given $\alpha = (\alpha_i)_i \sim N(0,1)$ is defined as:

$$X_S = \bar{X}_S + \sum_{i=1}^{n} \alpha_i \sqrt{\lambda_S^i} e_S^i \tag{2.1}$$

With the mean shape $\bar{X}$ and the covariance matrix Σ defined as

$$\bar{X}_S = n^{-1} \sum_{i=1}^{n} X_S^i$$

$$\Sigma_S = (n-1)^{-1} \sum_{i=1}^{n} (X_S^i - \bar{X}_S)(X_S^i - \bar{X}_S)^T$$

where e_S^i is the i^{th} eigenvector of the sample covariance matrix representing the i^{th} principal mode of variation. The parameters λ_S^i are eigenvalues of the covariance matrix and describe the contributions of the i^{th} principal mode of variation in shape.

These models are discrete models, meaning that they are based on a fixed discrete representation of the object. As such, the statistics learnt from the dataset are entirely based on the discrete points. It is then not possible to have the same statistic on different resolutions of the representation of interest without retraining the model. However, most bone structures are naturally continuous, that is, they are smooth and can be represented by an infinite number of points (Demisse et al., 2017). Hence, the optimal way to represent them is to preserve their continuity in the modelling process. Discretization can then be an on-demand operation customised to the application of the models. Recently this artificial discretization bottleneck has been addressed through curve-based shape deformation (Demisse et al., 2017) and Gaussian process morphable models (GPMMs) (Lüthi et al., 2017). The curve-based approach represents an object as a set of transformation matrices that map one point to another with respect to a fixed coordinate frame. Such a representation is computationally expensive since it requires an estimation of rigid transformations for each point of the object. Alternatively, GPMMs simply define probability distributions on deformation fields which are basically a set of vectors representing local variations from

the reference object to each object as formulated below.

Let us consider the same training dataset above $\{X_S^i, i = 1, \ldots, n\}$. The GPMM-based approach represents the i^{th} object X_i in the training data as a deformation field $\mathcal{F}_{iS}$ from the reference object $\Gamma_S \subset \Omega_S \subset \mathbb{R}^3$, with Ω_S being the reference domain. A Gaussian process $\mathcal{GP}(\bar{\mathcal{F}}_S, \mathcal{K}_S)$ may be defined over the set of deformation fields $\{\mathcal{F}_{iS}, i = 1, \ldots, n\}$. The mean ($\bar{\mathcal{F}}_S$) and the kernel ($\mathcal{K}_S$) functions, parameterising the $\mathcal{GP}$, are then defined as:

$$\begin{aligned} \bar{\mathcal{F}}_S &= n^{-1} \sum_{i=1}^{n} \mathcal{F}_{iS}(x), \\ \mathcal{K}_S(x, y) &= (n-1)^{-1} \sum_{i=1}^{n} (\mathcal{F}_{iS}(x) - \bar{\mathcal{F}}_S(x))(\mathcal{F}_{iS}(y) - \bar{\mathcal{F}}_S(y))^T, x, y \in \Omega_S \end{aligned} \tag{2.2}$$

To represent the model in a continuous domain, interpolation is used (Lüthi et al., 2017). Different interpolation approaches can be used including the barycentric coordinate-based interpolation (Meyer et al., 2002) where the statistic around a point in the domain is a function of its barycentric coordinates; and the nearest-neighbourhood interpolation where the statistic around a point in the domain is a function of its nearest neighbourhood points' statistics (Lüthi et al., 2017).

The probabilistic formulation of the mode in the continuous domain Ω_S is then defined as:

$$\mathcal{F}_S(x) = \bar{\mathcal{F}}_S(x) + \sum_{m=1}^{M} \alpha_m \sqrt{\lambda_m} \Phi_m(x), \text{ with } x \in \Omega_S, \tag{2.3}$$

where $\alpha = (\alpha_m)_m \sim \mathcal{N}(0, 1)$ is the unit normal distribution, (λ_m, Φ_m) are the eigenvalues/basis functions of the kernel obtained using Karhunen-Loève expansion and Nyström approximation, with M being the model's rank. Statistical shape models have been used for prediction of anatomical structures from incomplete ones (Blanc and Székely, 2012; Salhi et al., 2020). However, all the shape models described thus far, whether discrete or continuous, are limited to shape analysis tasks only. As such, these models have limited use in medical image analysis tasks requiring the knowledge of the intensity distribution in the images; tasks that include intensity-based segmentation or the analysis of the internal

structure of the bone.

B- Statistical intensity-based models

Intensity-based models have been proposed in the literature to address such limitations. Statistical shape and intensity models such as ASMs, AAMs and SSIMs are computed similarly to SSMs and GPMMs (Cootes et al., 1995, 1998; Sarkalkan et al., 2014; Reyneke et al., 2018b). However, the intensity correspondence establishment is more challenging compared to the shape case as will be discussed in chapter 3. For AAMs, vectors with coordinates representing shape features are replaced by intensities; and for ASMs by intensity profiles. For Gaussian process-based SSIMs (Reyneke et al., 2018b), deformation fields encoding shape features are replaced by deformation fields encoding intensities. These intensities are directly obtained from images such as those from CT or MR. Below is the formulation of the AAMs and ASMs based on the work of (Cootes et al., 1995, 1998).

Let $\{X_I^i, i = 1, \ldots, n\}$ be the representation of the training dataset, where X_I^i is a vector representing the intensities of the i^{th} object or the intensity profiles at each boundary point of the object. An instance X_I representing the intensity profile for a given $\alpha = (\alpha_i)_i \sim N(0, 1)$ is defined as:

$$X_I = \bar{X_I} + \sum_{i=1}^{n} \alpha_i \sqrt{\lambda_I^i} e_I^i \tag{2.4}$$

With the mean intensities $\bar{X}$ and the covariance matrix Σ defined as:

$$\bar{X_I} = n^{-1} \sum_{i=1}^{n} X_I^i$$

$$\Sigma_I = (n-1)^{-1} \sum_{i=1}^{n} (X_I^i - \bar{X_I})(X_I^i - \bar{X_I})^T$$

where e_I^i is the eigenvector of the sample covariance matrix representing the i^{th} principal mode of variations and $\bar{X_I}$ the mean vector representing the mean intensities. The eigenvalues λ_I^i describe the contributions of i^{th} the principal mode of variation in intensity. Active shape models are SSMs with the associated mean intensity profile ($\bar{X_I}$). They

work by simultaneously using the SSMs and the mean intensity profile for intensity-based task analysis (Cootes et al., 1995). Active appearance models are formalised as a combined PCA type model of shape and intensity resulting from the combined shape (b_S) and intensity (b_I) parameters defined as:

$$\gamma = [W_S b_S, b_I]$$

where b_S and b_I are functions of eigenvectors from shape and intensity models, respectively, as defined in (Cootes et al., 1998); and W_S is a diagonal matrix of weights for shape parameters. Active appearance models and ASMs have shown promise in the segmentation of human organs from $3D$ volumes (Ciller et al., 2015; Gaweł et al., 2018). However, AAMs and ASMs do not combine shape and intensity in the same covariance matrix, hence they do not capture all the correlations that may exist between both features. Using such models for prediction from medical images may lead to inaccurate results since correlations that could provide additional information are not embedded in the model. Moreover, existing AAMs are discrete models like PDMs. An alternative to them are Gaussian process ($\mathcal{GP}$) based SSIMs (Reyneke et al., 2018a) which are built in a continuous domain, as formulated below.

For GP-based SSIMs, models are defined as $\mathcal{GP}(\bar{\mathcal{F}}_I, K_I)$ over a set of intensity deformation fields $\{\mathcal{F}_{iI}, i = 1, \ldots, n\}$. With $\mathcal{F}_{iI}$ being a voxel value-based deformation field of the i^{th} image in the training dataset. The mean ($\bar{\mathcal{F}}_{iI}$) and the kernel (K_I) functions, parameterising the $\mathcal{GP}$, are then defined as:

$$\begin{aligned} \bar{\mathcal{F}}_I &= n^{-1} \sum_{i=1}^{n} \mathcal{F}_{iI}(x), \\ \mathcal{K}_I(x, y) &= (n-1)^{-1} \sum_{i=1}^{n} (\mathcal{F}_{iI}(x) - \bar{\mathcal{F}}_I(x))(\mathcal{F}_{iI}(y) - \bar{\mathcal{F}}_I(y))^T \end{aligned} \tag{2.5}$$

Applications of this family of models (SSMs, GPMMs, SSIMs, AAMs or ASMs) are limited to single-object modelling in the absence of any pose variation. With these models, artic-

ulated anatomical structures cannot be analysed. An articulated structure is composed of several substructures that can move relative to each other. Statistical pose models have been developed to model pose variations for multi-object structures. These are now presented in the next section.

2.2.2 Multi-object models

While single-object models provide useful information regarding the shape and intensity, they do not contain any information regarding pose distribution of objects in space. This means these models cannot simultaneously explain shape, intensity and pose variations across multi-objects. For the scope of this thesis, **a multi-object complex is considered to be a set of objects gathered in a common spatial frame, and moving coherently with respect to each other**. An example of a multi-object complex is the shoulder with three different bone structures, scapula, humerus and clavicle. Another example is the hip with two different bones, femur and pelvis. Several varying properties can be identified in a multi-object complex including shape, intensity and pose of every object (Gorczowski et al., 2010). Similar to SSMs, statistical shape and pose models (SSPMs) are parametric representations of a set of multi-objects and are defined as a variation around the mean of shape parameters and pose parameters. A validated SSPM is achieved through a correspondence establishment, model building and validation, respectively. As mentioned for single shape models, registration is a well established approach for correspondence establishment. However, in contrast to the SSMs, registration of several structures is more complex due to large pose variations between the objects to be modelled.

2.2.3 Multi-object registration

Performing the registration task using multi-object statistical model is more complex due to the pose variability that includes the physiological relationship between two or more bodies. The single object statistical modelling methods explained above (section

2.2.1) may be used for multi-object registration using single-object models for each of the multi-object structures. However, pose variability should be incorporated into the process because objects are not often imaged at the same pose. Rigid registration is a global transformation (that is a transformation that does not change the shape of the object being transformed) and thus it may not be sufficient for registration of complex structures. Rigid registration can simply be applied to each object separately, but the inter-object spatial relationship is not preserved. Several approaches to rigidly register multi-objects simultaneously while preserving the inter-object spatial relationship have been reported in the literature. Kalman filter-based registration was used for the rigid registration of several lumbar spines while preserving their relative spatial positions (Moghari and Abolmaesumi, 2007; Rasoulian, Abolmaesumi and Mousavi, 2012). The iterative closest point (ICP) algorithm which aims to find the closest point on a geometric entity to a given point in another entity, has also been used (Besl and McKay, 1992). Bindernagel et al. (2011) registered knee joints using ICP-based registration. The registration technique applied by Bindernagel et al. (2011) used different local transformations for the tibia and the femur, and a global transformation to limit the loss of the inter-object relationships. Agrawal et al. (2020) applied a similar approach for the registration of the hip joints. These techniques are computationally expensive because rigid transformations are estimated through an optimisation process. An alternative to this is the use of landmark-based techniques proposed by Umeyama (1991), which estimates rigid transformation parameters through explicit formulations. Such an approach is computationally more efficient. However, to the author's best knowledge this has only been used for shape-pose modelling. Hence, using this approach for multiple modelling including shape, pose and intensity would be of interest.

Non-rigid registration is a local registration technique (that transforms objects while changing the shape) and can handle complex structures. In order to non-rigidly register multi-object structures, a template can be selected and deformed to fit to remaining multi-object structures in the dataset. However, the pose difference considerably increases the degree

of variability of aligned specimens, making it challenging to perform the fitting of different objects, simultaneously. Many approaches have been proposed in the literature to deal with this challenge. Common approaches define deformation functions that transform the template to other specimens with three parameters: the first parameter (for e.g., shape) attributed to the first object; the second for the second object; and the third for the kinematics (inter-object relationship) as a parameter to model pose (Yeo et al., 2018; Kainmueller et al., 2009). Another approach is piece-wise registration in a rigid fashion using sliding surfaces for implicit kinematic restrictions, as used for a SSPM of the wrist joint (Van De Giessen et al., 2009). However, bias is often introduced in such models due to the choice of the reference. However, group-wise registration techniques have been introduced to mitigate bias in registration (Jacq et al., 2008; Rasoulian, Abolmaesumi and Mousavi, 2012; Bhatia et al., 2004; Anas et al., 2014).

While the above methods have shown promise in simultaneous registration of multiple structures; to the author's best knowledge, most of these approaches are applied for joints with a range of motion in only one rotational DOF such as the wrist joint. No protocol has been proposed for registering joints with larger ranges of motion for all rotational degrees of freedom while considering the full shape of the constituted bones such as shoulder or hip joints.

2.2.4 Computation of multi-object latent space

Even though simultaneous registration of multiple structures is challenging, the most difficult problem remains the computation of a compact space (space of compressed data descriptors computed from the training dataset) able to embed and explain shape, pose and intensity variability across these structures. This type of space, mapping the explicit information about $3D$ objects such as shape, intensity or pose, into compressed feature vectors is called as **latent space**. The space of possible shapes can be approximated by an Euclidean space ($\mathbb{R}^n$) (Moreau et al., 2017; Lüthi et al., 2017). Thus, for single object statistical models (SSMs, SAMs or SSIMs), PCA is commonly used to project the data

into a low-dimensional orthogonal latent space where linear statistics maybe performed (for e.g., calculating mean shape and its variance, etc.). For pose, the direct Euclidean isometries are considered. This set is called SE(3) and it is not a linear space. Thus, obtaining a mean for rigid transformations is challenging. Some studies have proposed pose transformation models using linear PCA with the assumption that a small range of pose variation can be approximated by a linear space. For example, Smoger et al. (2015) proposed an approach, based on the local shape and kinematic coordinate systems, to build an SSPM of the knee joint using standard PCA. The authors used the SSPM to characterise the relationship between shape and kinematics of components of the knee. However, a limitation to their approach was their use of standard PCA. Pose variations between different structures are represented by curved surfaces in space. Such curved surfaces are mathematically modelled in differential geometry using Riemannian manifolds (Fletcher et al., 2004). Linear PCA therefore, cannot be applied directly to such a curved space. In the manifold representation, a Riemannian metric allows one to measure the distance between transformations. The Riemannian metric's associated distance allows local assimilation of the manifold to an Euclidean space; called the tangent space (Fletcher et al., 2004).

Probabilities and statistical tools on a Riemannian manifold have been proposed using principal geodesic analysis (PGA) for statistical analysis of rigid transformations (Pennec, 2006; Fletcher et al., 2004). The PGA is a PCA in the tangent space of a Riemannian manifold. Principal geodesic analysis aims to find the geodesic subspace that minimises distances between elements in the manifold and their projections to the tangent space (Fletcher et al., 2003). However, due to higher computational costs associated with the estimation of the geodesic subspace, it is usually approximated by the tangent space using exponential and logarithmic mapping. The distribution around the mean in the tangent space is modelled by performing the PCA using the maximal explanation of the covariance matrix (Fletcher et al., 2003). Bossa and Olmos (2006) developed a multi-object pose model of brain structures using PGA. The model provided the location, orientation and

scale which were useful as a coarse initialisation for segmentation algorithms. Anas et al. (2014) proposed a shape-pose model using PGA to segment wrist bones from CT images. The formulation of PGA based SSPMs is developed below.

Principal geodesic analysis-based strategies consider a multi-object training dataset with n examples, $\{\{S_i^j\}_{j=1,\ldots,N}\}_{i=1,\ldots,n}$ where each training sample S_i is composed of N objects. The j^{th} object of the i^{th} training example can be represented by S_i^j, a vector of discrete landmarks with known correspondence between the examples $\{S_i^j\}_{i=1,\ldots,n}$ of the j^{th} object, the reference vector $\vec{\Gamma}_S^j$ is a vector of point coordinates of the object Γ_S^j. The reference is defined such that it is topologically unbiased relative to the dataset (Mutsvangwa et al., 2015) or it can be an object from the database that is close to the mean object. The rigid transformation that aligns the j^{th} object of the i^{th} example to its reference object Γ_S^j is defined by:

$$h_i^j = \underset{h \in SE(3)}{\arg\, min} \|h(S_i^j) - \vec{\Gamma}_S^j\|^2$$

where $SE(3)$ is the space of orientation-preserving rigid transformations with $\|.\|$ being the distance metric. To perform the PGA, the transformation of the j^{th} object in the i^{th} training example is projected onto the tangent space. This space is tangent to $SE(3)$ at the mean rigid transformation $\bar{h}_j$. The projection is done using log mapping as:

$$log_{\bar{h}_j} h_i^j = [p_1, \ldots, p_m]^T \tag{2.6}$$

where $(p_i)_{i=1,\ldots,m}$ are the parameters of the rigid transformation $\bar{h}_j^{-1} \circ h_i^j$ representing the Euler's, Rodrigues's or quaternion parameters (Bossa and Olmos, 2007).

These representations describe the $3D$ rotations with angles. Rodrigues and quaternion represent a 3D rotation as an axis vector and a rotation angle, while Euler represents it with three angles. **For the rest of this thesis, these representations will be called termed standard representation (SR).**

To map the i^{th} transformation back to the manifold, an exponential mapping is used:

$$h_i^j = \bar{h}_j exp(log_{\bar{h}_j} h_i^j)$$

To build the SSPMs, shape and pose features of the i^{th} example are concatenated in joint feature vectors (Bossa and Olmos, 2007) by:

$$X^i = [w^s X_S^i, w^p X_p^i]^T$$

where $X_S^i = [S_i^1, \ldots, S_i^N]^T$ and $X_p^i = [log_{\bar{h}_j} h_i^1, \ldots, log_{\bar{h}_j} h_i^N]^T$, w^s and w^p denote the shape and pose weightings, respectively, that balance the relative importance between shape and pose within the model. The resulted norm of the SR is defined as:

$$\|h_i^j\|_{SR} = t^T\, t + s\, r^T\, r \tag{2.7}$$

where r and t are rotation and translation vectors, respectively; and s the scaling factor.

To build the model, an additional assumption is that the joint feature residual dataset $\{X^i\}_{i=1}^n$ can be modelled by a normal distribution similar to PDMs (Cootes et al., 1992) as follows:

$$(S, P) \sim \mathcal{N}(\bar{X}, \Sigma)$$

where $\bar{X}$ is the joint mean and Σ, the covariance matrix, are estimated from the dataset $\{X^i\}_{i=1,\ldots,n}$ as:

$$\bar{X} = n^{-1} \sum_{i=1}^{n} X^i$$

$$\Sigma = n^{-1} \sum_{i=1}^{n} (X^i - \bar{X})(X^i - \bar{X})^T$$

The model is obtained by performing PCA which on using a probabilistic PCA interpret-

ation leads to:

$$X = \bar{X} + \sum_{i=1}^{n} \alpha_i \sqrt{\lambda_i} e^i$$

where $e^i = [e^i_S, e^i_p]^T$ is the combined i^{th} eigenvector of the covariance matrix (associated to the i^{th} eigenvalue λ_i), obtained from single value decomposition (SVD) and $\alpha = (\alpha_i)_i \sim \mathcal{N}(0, 1)$. Then, the SSPM is split into two parts, the shape component, typically a PDM, defined as follows:

$$X_S = \bar{X}_S + \sum_{i=1}^{n} \alpha_i \sqrt{\lambda^i}\, e^i_S \tag{2.8}$$

and the pose component, typically SPMs, defined as:

$$h = \bar{h}\, exp\left(\sum_{i=1}^{n} \alpha_i \sqrt{\lambda^i} e^i_p\right) \tag{2.9}$$

A new instance of a discrete SSPM is obtained as $S_p = h(X)$ which derives the shape at a relative spatial position. However, the associated norm to this SR encodes less information about the pose of the object (Moreau et al., 2017). This results in the so called "non-compactness effect" defined as the dis-proportionality between the area of the object being modelled and the area of its bounding circle. Objects at risk of exhibiting such an effect are called non-compact objects. Moreau et al. (2017) proposed a new metric for statistical analysis of rigid transformations that efficiently handles the "non-compactness effect". The metric proposed by (Moreau et al., 2017) was applied to build a SPM of the rib cage. The authors proposed a new distance for rigid transformations analysis, defined by:

$$d^2(h^j_i) = \delta^j (h^j_i)^T \delta^j (h^j_i)$$

where δ^j_i is the deformation field associated to the transformation h^j_i, defined as:

$$\delta^j_i (h^j_i) = h^j_i (\Gamma^j_S) - \Gamma^j_S \tag{2.10}$$

This metric is efficient in the sense that its resultant norm is equivalent to the L^2 norm which is related to energy and provides more information about the displacement of the object. Hence it is robust in capturing pose descriptors of non-compact objects. The model in (Moreau et al., 2017) was shown to be more robust in accurately estimating the similarity between rigid transformations compared to models using standard metrics. Using such a metric for a SSPM may provide a principled way to analyse the relationship between shape and pose. **However, to the author's best knowledge, this metric has not been used for building SSPMs** and embedding this new representation in linear models for shape-pose analysis remains challenging.

Furthermore, to the author's best knowledge, no SSPMs have been developed that embed statistical modelling of intensities. This limitation of SSPMs prevents their use in clinical applications that require intensity-based predictions. Three-dimensional reconstruction from $2D$ images such as $2D$ X-rays is one of the clinical applications that has benefited from shape and intensity-based models (Yao and Taylor, 2001; Sadowsky et al., 2006; Markelj et al., 2012; Ehlke et al., 2013; Klima et al., 2016; Reyneke et al., 2018a; Fotsin et al., 2019). The inclusion of intensity in SSPMs may lead to a modelling framework with a potential for even more clinical applications. These would include the prediction of subject-specific joint mechanics and permitted motion for surgical planning; segmentation of $3D$ volumes; prosthesis design; $3D$ reconstruction from $2D$ images to estimate pre-morbid shapes and joint contact spaces from single or biplanar X-ray images; and the efficient estimation of muscle forces and moment arms for pre-surgical and rehabilitation assessments. However, developing such a model remains challenging due to the lack of an efficient way to embed various feature types in a compact space while preserving the continuous property of the objects.

2.3 Conclusion

This chapter summarises the current state-of-the-art in statistical modelling of single or multiple anatomical objects that embed shape, pose, and intensity either together or separately. To the author's best knowledge, no statistical models have been reported in the literature that combine shape and kinematic parameters in an inherently continuous space. Furthermore, the few shape-kinematics models in the literature that do characterise the relationship between shape and pose do not efficiently use principal geodesic analysis (PGA); the appropriate method when operating on manifolds. In addition, existing shape and pose models do not use an appropriate metric that mitigates the "non-compactness effect". This means that current models are not capable of accurately modelling rigid transformations.

Some research has focused on developing intensity-based models (SSIMs or SAMs) but there have been no reports on embedding intensity in shape-pose models. Lacking an explicit modelling of intensity variation, current modelling paradigms are thus limited for certain medical image processing tasks such as the prediction of object's contours with their associated internal structure features from $3D$ volumes (CT scans or MRI). Embedding intensity in SSPMs may help generate $3D$ volumes (such as CT image) of joint mechanics while observing only the shape and the kinematics of the patient, that is to say, defining a posterior with shape and kinematics as prior knowledge. Such models may provide further knowledge for patient-specific joint mechanics and motion prediction. The literature abounds with reports on $3D$ reconstruction from $2D$ for image-guided surgical applications (Reyneke et al., 2018a). However, no work using statistical model prior embedding all key features (shape, pose and intensity) has been encountered thus far regardless of the potential for clinical applications (e.g clinical orthopaedics).

Since 2000, a range of work has been done to build several models as SSMs, SPMs, SSIMs and SAMs for achieving applications in medical image processing and analysis. Nevertheless, no work has been reported about a unified statistical model embedding all key

features (shape, pose and intensity) in order to study human skeletal joints (multi-objects) from medical images. Chapter 4 will be a presentation of a novel modelling framework, developed by the author, called Dynamic Multi Feature Class-Gaussian Process Model (DMFC-GPM), which overcomes the limitations described above. However, before studying the variability of the three key features to be embedded in the modelling framework, correspondence between the multi-object complexes of the training dataset must be established for each one of them. This is presented in the following chapter.

Chapter Three

Establishing shape, pose and intensity correspondence

3.1 Introduction

As described in the previous chapter, shape, pose and intensity are the key features that have clinical relevance in medical images of human joints. To develop statistical models embedding these features, dense correspondence needs to be established between examples across the training dataset. Dense shape-pose and intensity correspondence between two objects in the medical imaging domain means finding a one-to-one mapping between surface points defining the object as well as the intensity inside the objects while preserving their spatial orientation. In medical imaging, there is variability across these key features that we want to embed in our statistical model. The study of such variability requires correspondence which is often established through registration. Indeed, registration aims to find the best mapping between the features of two similar objects or among several instances (observations) of the same object, in order to bring object data into the same reference system as well as to provide anatomical similarity information across them. This chapter presents a method for dense correspondence establishment (surface and volumetric) across multi-object biological structures while preserving the spatial orientation of the objects.

The establishment of dense correspondence across the multi-object data is challenging, because objects are not only to be aligned but all the points defining the objects (on

the surface and inside) need to be in-correspondence as well. Dense correspondence differs from sparse anatomical correspondence which is often established through manual annotation of landmarks by experts. Dense correspondence requires an automated process as the points defining the objects are often in the thousands and thus can not be manually annotated. Regarding surface ($3D$ point cloud and mesh) registration, a survey by Tam et al. (2013) provides an overview of multiple approaches for dense correspondence. Regarding volumetric mesh registration (points inside the surface), there is a paucity of work in the literature with few reported studies (Klima et al., 2016; Yao and Taylor, 2001; Ehlke et al., 2013). These volumetric mesh registration approaches allow the establishment of intensity correspondence when analysing shape-intensity variabilities without requiring the use of image-based registration techniques (Preston et al., 2016; Sotiras et al., 2013; Wyawahare et al., 2009). For a single biological structure, different methods for registration exist including parametric (Duchon, 1977; Kybic and Unser, 2003; Myronenko et al., 2007; Lüthi et al., 2013) and sampling (Chen et al., 2014; Schönborn et al., 2017; Madsen et al., 2019) approaches.

The establishment of dense correspondence through registration starts with a rigid (global) transformation for removing rotation and translation that exist between objects. The second step is to find a non-rigid transformation that locally establishes correspondence across shape features. The first step is often done through the rigid registration algorithms such as iterative closest point (ICP) (Besl and McKay, 1992) and its numerous variants such as iterative median closest point (IMCP) (Jacq et al., 2008) or through other alignment methods such as Procrustes analysis (Gower, 1975). The second step can be done in two ways: 1) parametric-based process using non-rigid registration (Duchon, 1977; Kybic and Unser, 2003; Myronenko and Song, 2010; Lüthi et al., 2013) or 2) sampling-based process (Chen et al., 2014; Schönborn et al., 2017; Madsen et al., 2019). The parametric-based approach aims to optimise the cost function evaluating the similarity between two objects while the sampling-based approach aims to generate weighted samples (samples with an associated probabilities) and returns the one with the highest probability.

Let's briefly look at these non-rigid registration approaches. For parametric-based registration, correspondence can typically be achieved through fitting a template model. When available, the template can be a pre-trained statistical model (prior), otherwise a free-form deformation (FFD) model may be employed. Different FFD model formulations exist (Lüthi et al., 2017; Duchon, 1977; Myronenko et al., 2007; Mayya et al., 2013). However, the current state-of-the-art parametric method for establishing dense correspondence is the Gaussian process morphable models (GPMMs) based registration approach. The GPMM-based registration has the advantage of being able to combine pre-trained statistical and bespoke FFD models to improve model flexibility. Similarly, sampling-based dense correspondence methods may also employ prior models. These can be pre-trained models, FFD models, or a combination of both (in the GPMM paradigm). However, in contrast to parametric-based approaches, which are based on the optimisation of a cost function, sampling-based methods are based on the filtering of some distribution proposing parameters for the prior models. The Markov chain Monte Carlo (MCMC) method, a popular sampling-based method in numerical simulation and computational science, has been widely used for sample-based fitting (Chen et al., 2014; Schönborn et al., 2017; Madsen et al., 2019).

Despite the success of the above methods for establishing dense correspondence, thus far they have only been reported for single-object registration in point based surface models. Establishing such dense correspondence across several structures while keeping their anatomo-physiological relationship intact is not a straightforward task. Several works have been reported on multi-object registration (Moghari and Abolmaesumi, 2007; Jacq et al., 2009; Bindernagel et al., 2011; Rasoulian, Abolmaesumi and Mousavi, 2012; Agrawal et al., 2020) as reported in the literature review chapter (section 2.2.3). However, to the author's best knowledge no work has been reported on the use of parametric and MCMC-based registration using the unified GPMMs as priors for multi-object and intensity registration. Moreover, for all the above methods, the correspondences are surface-based. To establish correspondence across the intensities, correspondence needs to be established not only

across the surface of the objects but also inside the objects (Yao and Taylor, 2001). Thus, it remains challenging to establish correspondence across the geometric points defining the internal structures of these objects.

To overcome the limitations cited in the literature, we propose both parametric and sampling-based multi-object correspondence establishment methods in this chapter. First, GPMMs-based priors are used for registration to establish shape-pose correspondence across surfaces of the objects. Second, a volumetric shape correspondence approach (that is correspondence across points inside the volumetric mesh that allows the embedding of intensities) is derived through the interpolation of the deformations obtained from the surface-based correspondence. Finally, controllable data are generated for the validation experiments.

3.2 Establishment of shape-pose correspondence using GPMMs

In this section, dense correspondence across individual objects is achieved through alignment and subsequently, non-rigid registration performed using GPMMs-based priors (Lüthi et al., 2017). Alignment is performed by estimating the rigid transformation parameters using sets of anatomical landmarks selected on the objects. Here the registration method is derived for N ($N \in \mathbb{N}$) objects that extends the GPMM-based single object fitting.

Before deriving the proposed correspondence establishment process, let us formulate the GPMM-based priors used as FFD models for individual objects. To build the template model, a template object is often defined to make the topology unbiased across the training dataset (Stoll et al., 2006; Kraevoy and Sheffer, 2005). Iterative median closest point-Gaussian mixture model (IMCP-GMM) is a state-of-the-art approach for unbiased template generation (Mutsvangwa et al., 2015). The algorithm works by extracting a set of consensus points cloud (virtual shape) built from the median points of each local point

cloud across the dataset. The virtual shape acts as a canonical representation of all the objects of the same shape family in the training dataset. Given a template object defined in a domain Ω, a GPMM-based FFD model is defined as a deformation $\mathcal{F}_S$, as formulated below:

$$\mathcal{F}_S \sim \mathcal{GP}(\bar{\mathcal{F}}_S, \mathcal{K})$$

where $\bar{\mathcal{F}}_S$ is the mean deformation field and k the kernel functions, parameterising the Gaussian process ($\mathcal{GP}$). An example of kernel often used is the Gaussian kernel defined by

$$\mathcal{K}^{(s,\sigma)}(x, y) = s \exp(\|x - y\|^2/\sigma^2), \; x, y \in \Omega \tag{3.1}$$

where s is the scaling factor of the deformation vectors and σ the correlation range of the deformations. A large σ implies a large variability and the sensitivity of this parameter is discussed in (Lüthi et al., 2017).

To register N objects while keeping their anatomo-physiological relationship, one of the N objects is assumed to remain fixed while the others move relative to the fixed one. Individual FFD models (N models) can then be built for each object family and used to register target objects from their respective family. After the registration, each inverse of the rigid registration transformation can then be used to retrieve the original spatial position of moving objects. This operation retrieves the spatial relationship between moving objects. Rigid transformation of the fixed object is finally applied to all moving objects to obtain their spatial relationship with the fixed object. Dense correspondence (shape-pose) is established across the dataset by using the same FFD model as the reference for the registration. The mathematical formulations of algorithms for obtaining the best deformations for establishing correspondence are presented in the next two sections for both parametric and sampling-based approaches.

3.2.1 Obtaining parametric-based deformations

Here, the model's parameters that best represent a target object need to be obtained through optimisation.

Let us consider two joints (reference and target) consisting of N objects each. We can define these reference and target joints, respectively, as:

$$\{\Gamma_S^j, j = 1, \ldots, N\}, \{S^j, j = 1, \ldots, N\}$$

The object S^j is a surface mesh defined as $S^j = \{x_i^j, i = 1, \ldots, n_j\}$, with n_j being the number of vertices and x_i^j are $3D$ points.

Let Θ be the parameterised space of a FFD model, which in the case of GPMMs, is the space of possible deformation fields. Let T_j be the rigid transformation that transforms the j^{th} object to the reference Γ_S^j, and T the rigid transformation that transforms the fixed object (S^{j_0}) to its reference Γ^{j_0}. The j^{th} registered object, S_r^j, with its spatial position retrieved can then be obtained as:

$$S_r^j = T \circ T_j^{-1}(\underset{\mathcal{F}_S \in \Theta}{\arg min}\ \mathcal{SM}(\mathcal{F}_S(\Gamma_S^j), S^j) + \beta \|\mathcal{F}_S\|^2) \tag{3.2}$$

where DF is the space of deformation fields, $\|.\|$ is a norm in DF, $\mathcal{SM}$ is a similarity measure, and β a regularisation coefficient. The rigid transformations T_j and T can then be computed through least-squares based estimation (Umeyama, 1991) as:

$$T^j = \underset{h \in SE(3)}{\arg min} \|\overrightarrow{h(S^j)} - \vec{\Gamma}_S^j\|^2 \text{ and } T = \underset{h \in SE(3)}{\arg min} \|\overrightarrow{h(S^{j_0})} - \vec{\Gamma}_S^{j_0}\|^2 \tag{3.3}$$

Figure 3.1 illustrates the multi-object shape correspondence establishment process described above. It starts with rigid registration (step 1), followed by non-rigid registration

and ends with the retrieval of the spatial relationship between target objects (target 1 and 2).

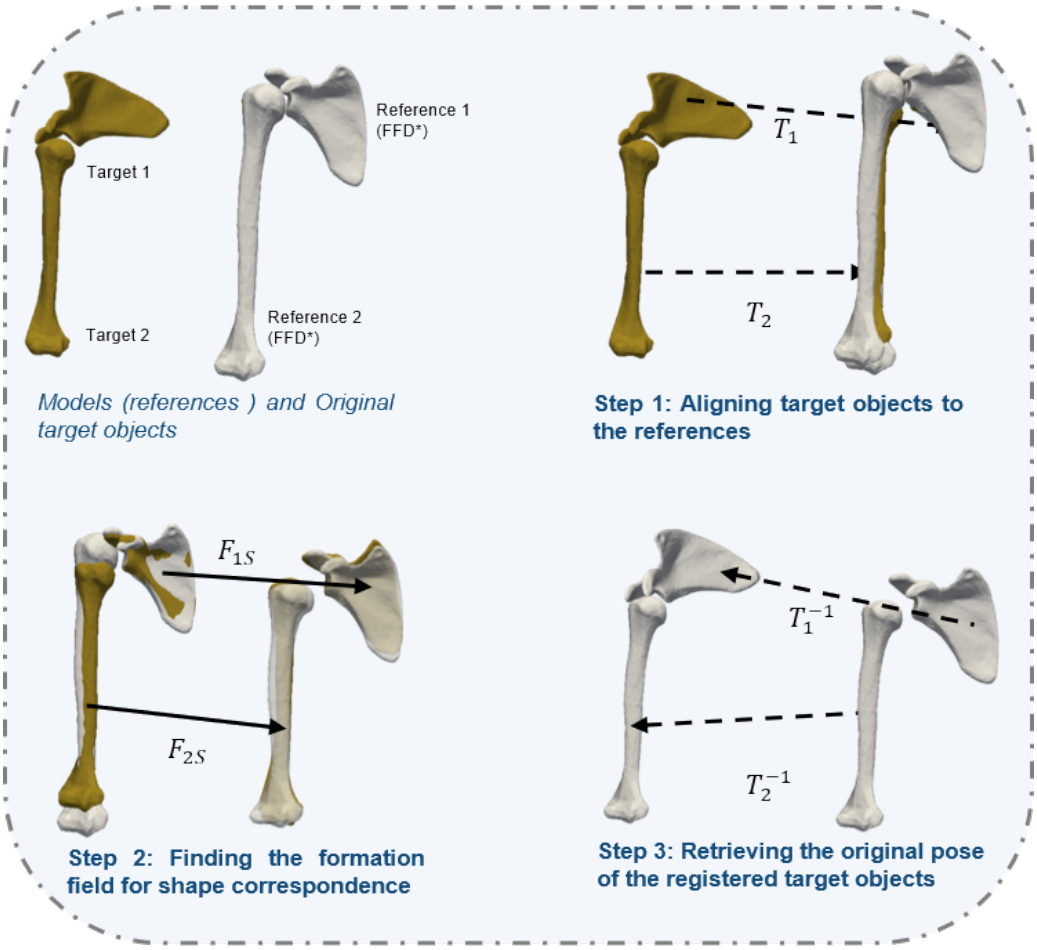

Figure 3.1 Multi-object registration schematic. Top-left: Reference and target joints consisting of two objects each. Step 1: T_1 and T_2 rigid transformations removing rotation and translation between individual reference and targets for the first and the second object, respectively. Step 2: $\mathcal{F}_{1S}$ and $\mathcal{F}_{2S}$ are non-rigid deformations for establishing shape correspondence between the reference and the target objects for objects 1 and 2, respectively. Step 3: The relative positioning of the target objects retrieves with the inverse of T_1 and T_2.

Considering a simple similarity measure $\mathcal{SM}$ such as the mean squared distance from the reference to the closest target point as described in (Lüthi et al., 2017), equation 3.3 becomes

$$S_r^j = T \circ T_j^{-1} \left(\underset{(\alpha_m)_m \in \Theta}{\arg min} \int_{\Gamma_S^j} d^2(x + \sum_{m=1}^{M} \alpha_m \sqrt{\lambda_m} \Phi_m(x)) dx + \beta \sum_{m}^{M} \alpha_m^2 \right) \tag{3.4}$$

Where $d(x) = \|x - \arg\min_{x' \in S^j}(\|x - x'\|)\|$, $\beta \in \mathbb{R}$ a weighing parameter, and $(\alpha_m)_m$ are the model parameters defining the deformation field $\mathcal{F}_S$. This formulation describes the computation of the parametric-based deformation $\mathcal{F}_S$ (non-rigid transformation), which establishes anatomical correspondence between the reference and the target object. It is also shown in the formulation how the associated rigid transformation describing the spatial orientation of the target object is retrieved. Solving the problem formulated in equation 3.4 provides model parameters that best represent the target. This establishes dense anatomical correspondence across the reference and the target objects while preserving their spatial relationship. However, robust optimisation strategies to avoid local minima remain challenging. The section below presents a sampling-based strategy as an alternative to parametric-based deformations.

3.2.2 Obtaining sampling-based deformation

Similar to the description of parametric-based registration above, below is the description of the sampling-based formulation approach developed to find the best deformation field and rigid transformation between the target object and the reference one.

For the sampling-based computation of the deformation fields u (eq 3.2), Markov chain Monte Carlo (MCMC) and Metropolis-Hastings algorithms presented in (Chen et al., 2014; Schönborn et al., 2017; Madsen et al., 2019), are adopted. For the formulation presented here, the goal of the MCMC fitting process is to find a model deformation field u that optimally represents the target object. The optimum deformation field u that minimises the difference between the model sample and the target can be estimated as the posterior model using Bayes rule:

$$p(\mathcal{F}_S|T_j) = \frac{p(\mathcal{F}_S)p(T_j|\mathcal{F}_S)}{\int p(\mathcal{F}_S)p(T_j|\mathcal{F}_S)d\mathcal{F}_S} \tag{3.5}$$

where $\int p(\mathcal{F}_S)p(T_j|\mathcal{F}_S)d\mathcal{F}_S$ is intractable, $p(\mathcal{F}_S)$ is the model, and $p(T_j|\mathcal{F}_S)$ is the likelihood. Finding the deformation $\mathcal{F}_S$ in equation 3.5 is equivalent to maximising the product

of the likelihood and the prior; hence equation 3.5 is equivalent to:

$$\underset{\mathcal{F}_S \in DF}{\arg min}\ p(T_j|\mathcal{F}_S)p(\mathcal{F}_S) \tag{3.6}$$

Thus, the j^{th} registered object, S_r^j, and associated spatial position can be obtained using:

$$S_r^j = T \circ T_j^{-1}(\underset{\mathcal{F}_S \in DF}{\arg min}\ p(T_j|\mathcal{F}_{iS})p(\mathcal{F}_S)) \tag{3.7}$$

Equation 3.7 provides a means to obtain the optimum deformation field that establishes dense correspondence between the reference and the target. However, both the parametric and sampling-based approaches proposed above only establish correspondence across the outer surfaces of the objects. Thus, the correspondence between volumetric meshes still needs to be obtained. The following section describes the manner in which to leverage the surface correspondence to establish correspondence across volumetric meshes.

3.3 Volumetric shape correspondence

Once dense correspondence across joint surface meshes is established, dense correspondence across volumetric meshes can also be established. Such a correspondence allows a one-to-one mapping of the intensities inside the objects. To do this, the discrete transformation that establishes correspondence across the reference and the target mesh surface can be interpolated to obtain a transformation over the entire continuous domain; through an adoption of the approach reported in (Yao and Taylor, 2001). For this task, the nearest neighbourhood interpolation can be used. It works by assuming that the coefficients of a point in the neighbourhood basis are the same to those of its corresponding point in the target neighbourhood basis. This is formalised below.

Let x be a reference volumetric mesh point, and $\mathcal{V}(x)$ is its neighbourhood point on the outer surface such that $x = \beta\mathcal{V}(x)$. In addition, let $T(T_j^{-1}(u(\mathcal{V}(x))))$ be the target surface

point of the neighbourhood $\mathcal{V}(x)$. Finally, let x' be the corresponding target point of x, such that $x' = \beta T(T_j^{-1}(u(\mathcal{V}(x))))$.

The continuous correspondence transformation over the volumetric domain can then be defined by:

$$u_{int} : \Omega \longrightarrow \mathbb{R}^3$$
$$x \mapsto \beta T(T_j^{-1}(u(\mathcal{V}(x))))$$

The j^{th} target volumetric mesh (which is in-correspondence with the reference volumetric mesh) can be obtained by applying the interpolated transformation to reference volume. The intensity correspondence is then obtained by aligning each volumetric mesh to its corresponding $3D$ image, and the voxel intensity value of the closest voxel to a mesh point is assigned to that point.

3.4 Validation of surface and volumetric correspondence establishment framework

In this section, the proposed method for multi-object registration is validated using synthetic data.

3.4.1 Generation of synthetic data

To validate the registration method, synthetic data were generated. These data were used for the experiment because the shape variation was designed to be controllable. Using such data avoids difficulties that could arise with objects having complex shape variability such as human anatomies. Moreover, with these data, it is possible to compare the registration results and the ground truth.

The synthetic data consists of surface meshes of a "lollipop" adopted from Gee and Treece

(Gee and Treece, 2014). In this project, the lollipop has a "shaft", a "neck" and a "head". All generated meshes had 6000 vertices and identical surface topology. The neck angle was uniformly distributed in the range $34 - 71$ degrees and the head length was uniformly distributed in the range $3.8 - 4.9$ cm while the shaft length was constant (9 cm), so the shape variation in the synthetic data was only along two degrees of freedom (DOF). The full dataset consisted of 100 synthetically generated lollipops.

3.4.2 Building of the prior for registration

A GPMM-based statistical shape model (SSM) of the lollipop meshes was built to act as a shape prior for the registration. Samples were drawn from GPMM built above from 100 lollipops. Samples from the first principal component (PC) are shown in figure 3.2. The first PC explains the variation of the angle of the lollipop's neck as was in the training data. Thus, the model captures the same variability as in the training dataset and therefore can be used to perform validation of the registration method across lollipops. This model was used to perform the registration in order to establish correspondence across the joints.

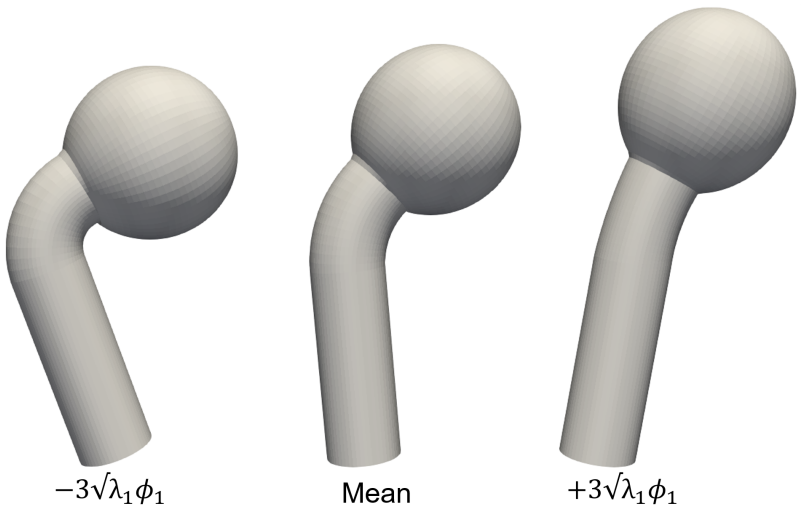

Figure 3.2 Samples from the first PC of the lollipop model. Samples from -3 and $+3$ standard deviation around the mean.

3.4.3 Generation of lollipop joints

The synthetic data described above consists of lollipops which are single objects. However, the proposed approach needs to be evaluated in registering multi-objects, which is the focus of this work. Thus, lollipop joints needed to be created which simulated pose variability. After that, the GPMM built above was used to establish shape correspondence while preserving the pose of the objects.

Of the 100 lollipops above, a subset 5 lollipop joints (two lollipops per joint) were used as multi-object targets on which the registration errors would be tested. To constitute a joint, shape and pose variability between lollipops was required. Lollipop joints were generated as couples of shape and pose through a random generation of shape and rotation parameters (examples of target joints are shown in figure 3.3 (top). Parametric registration was performed to establish correspondence across this subset of lollipop joints in order to evaluate the performance of the proposed approach for multi-object registration to establish correspondence.

3.4.4 Lollipop shape-pose correspondence establishment

A- Experimental method for testing quality of multi-object correspondence

The GPMM built above was used as prior to register all the lollipop joints in the test dataset using the parametric-based approach described in section 3.2.1. After the registration, surface to surface distances (Hausdorff distance [HD]) between the predicted objects and the original target objects at their relative spatial positions were computed. A t-test was performed at each vertex to evaluate the similarity between the correspondence established and the original one. To that end, a correspondence p-map (a map of p-value at each vertex) for the t-test was developed between the set of fitted vertices (corresponding vertices across the fitted dataset) and the set of target vertices (corresponding vertices across the original datasets).

B- Results of quality of multi-object correspondence

For the surface to surface distance errors, the results are shown in figure 3.3 (bottom), where the colormap indicates the resulting errors. The errors were low with the maximum HD surface to surface error of 0.5 mm. The shape similarity analysis was performed to evaluate the registration, which was done at each vertex across the datasets.

For the t-test results, figure 3.4 shows the resulting p-maps for all vertices in the mesh. The analysis reveals high precision in terms of shape, the test fails to find significant differences at any of the vertices. The small HD errors associated with the failure in finding significant differences suggest that each fit lollipop can replace its corresponding original one. This shows that the proposed registration method can suitably establish shape-pose correspondence, hence, each joint can be replaced by its fit for any subsequent analysis.

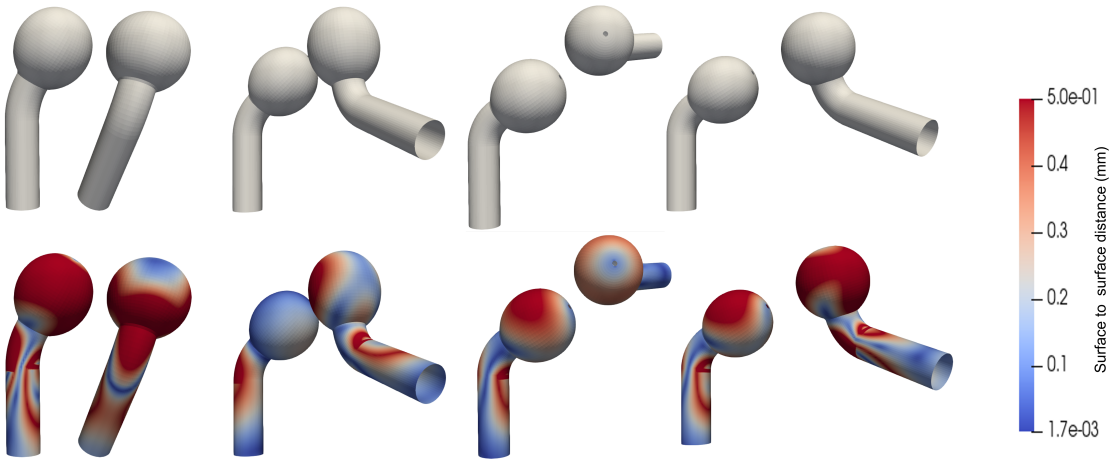

Figure 3.3 Target joint registration results. Top: target joints. Bottom: surface to surface distance between the registration output joints and the targets.

3.4.5 Lollipop volumetric-shape correspondence establishment

In this experiment, the same GPMM-based prior and the lollipop joint test dataset above were used to evaluate the quality of the volumetric mesh correspondence as described in section 3.3.

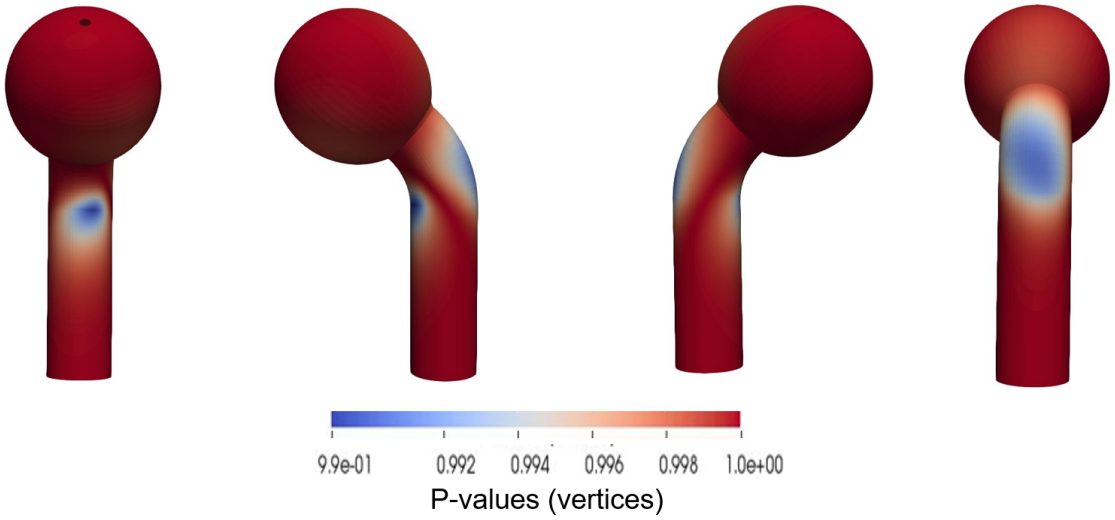

Figure 3.4 Correspondence test at each vertex. Lollipop meshes with associated vertex similarity value as a p-value resulting from the t-test.

A- Experimental method for evaluating of the correspondence

Six lollipops in-correspondence and one reference were selected. A reference volumetric-shape was constructed as a tetrahedral mesh from the reference using Amira software v6.2.0 (http://www.fei.com/). This software allows to create tetrahedral meshes from surface meshes. Since the lollipop meshes were not initially closed surfaces, they were first transformed into closed surfaces to generate tetrahedral meshes. To transform open lollipop surfaces into closed surfaces: first, they were rendered into binary images; and second, they were segmented and final closed surface meshes were reconstructed from the segmentation. Once closed surfaces were generated, the deformation fields establishing surface correspondence were interpolated. The interpolated deformations were applied to the reference tetrahedral mesh to generate in-correspondence tetrahedral meshes. To evaluate the correspondence, the outer surfaces were extracted from the obtained target volumetric meshes and compared to their ground truth surface meshes. It is important to note that the correspondence was only evaluated on the outer surface because there was no ground truth for tetrahedral meshes to validate the correspondence inside the structure.

B- Results of quality of the volumetric-shape correspondence

Figure 3.5 shows the HD surface to surface distance maps of the objects. It can be observed that both surfaces (the one extracted from the volumetric and the original one) are similar with the maximum error being 0.8 mm.

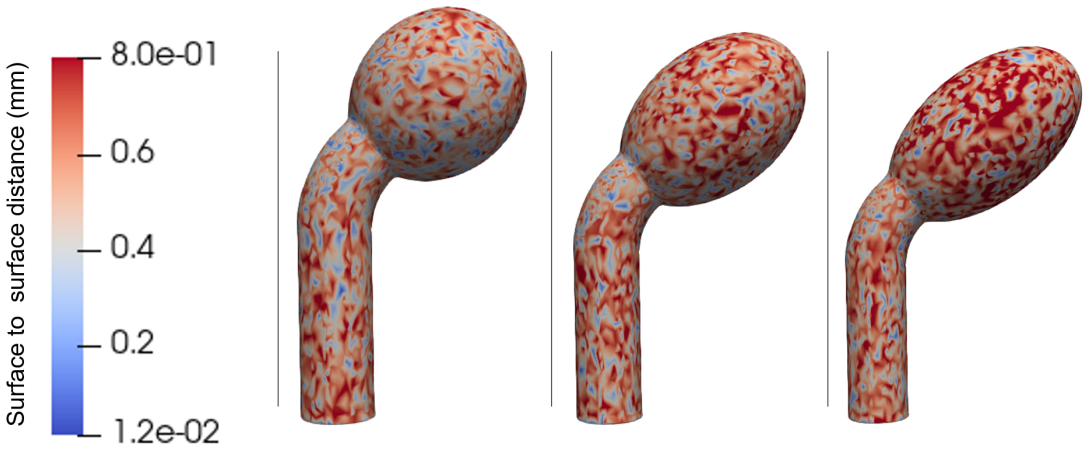

Figure 3.5 Lollipop volumetric-shape registration results. Surface to surface distance between outer surfaces of the obtained volumetric after registration and the target surfaces.

3.5 Conclusion

Establishing correspondence across objects remains challenging with the fundamental issue being the explicit characterisation of the similarity/correspondence between objects. Correspondence establishment becomes more challenging when confronted with multi-object and volumetric mesh scenarios.

In this chapter, a registration approach for establishing dense correspondence across multiple objects in a joint complex was described. In the presented approach, the aim is to find a composition of a rigid and non-rigid deformation that explains object feature transformation, in this case, $3D$ coordinates. Rigid transformations are computed through a deterministic optimisation. However, non-rigid deformations are more complex and are computed through a probabilistic optimisation. Both parametric and sampling-based optimisation approaches for estimation of a non-rigid deformation were developed and

presented in this chapter.

Correspondence for the internal structure of the object is more complex to explicitly characterise as compared to the outer surface. Using similar approaches used in surface registration for volumetric registration would lead to various challenges including the difficulty in computing the similarity between two volumes, and the difficulty in incorporating prior assumptions such as landmarks. Instead, the method presented here used the attribute of the GPMM which assumes that the surrounding of a surface deforms with the surface, hence, internal structure correspondence can be deduced from the surface one. While this approach shows great potential, the limitations typical of most registration techniques persist. These include the difference in complexity of the shape of interest relative to the template used, the optimisation algorithm adopted, and the characterisation of the correspondence itself. Moreover, it is important to note that the correspondence of the internal structure relies entirely on the correspondence of the outer surface. A limitation to this approach is that there may not be high enough correlation between points of the internal structure that are far away from the outer surface, and the outer surface, to accurately guide correspondence establishment. Furthermore, there was no ground truth for tetrahedral meshes to validate the correspondence inside the structure. Future work could include an investigation of other intensity-based correspondence establishment techniques which have shown promise for intensity modelling (Bonaretti et al., 2014; Taghizadeh et al., 2017).

Establishing correspondence across multiple-objects was the first step towards finding a compact space that can efficiently capture shape, pose, and internal structure variability (intensity). This is the topic of the next chapter.

Chapter Four

Dynamic multi feature-class Gaussian process models

4.1 Introduction

The previous chapter described how to establish dense correspondence across multiple objects in a joint complex using unique Gaussian process morphable model (GPMM)-based free-form deformation models (FFDMs). The transformation establishing the correspondence between the reference and each individual training object of the joint is composed of two parts: a rigid transformation and a shape deformation field. The challenge is to find a minimum subset $\mathbf{D}_0$ that can generally and specifically represents the set of correspondence transformations $\mathbf{D}$. Such a problem is currently an active research area (Li and Maguire, 2010). The common strategy consists of estimating a subspace of homogeneous patterns (for example, geometric shapes) through models. This inductive learning approach leads to models that allow for automated prediction from observations and interpretation of the learned patterns. In the medical image analysis domain, pattern analysis has to contend with two existing problems. First, patterns may be present across non-homogeneous feature classes. To give an example, in morpho-functional analysis, determining the changes in relative spatial dynamics of different anatomical structures adds an additional pattern class on top of the morphological (geometric) pattern class (Gorczowski et al., 2010). Second, patterns in some feature classes may reside in non-linear spaces. From the literature, the current trend in finding compact spaces for linear and non-linear patterns is to formalise the problem as an inverse problem and use deep learning-based generative

models to estimate a subspace of patterns (Toshev and Szegedy, 2014; Ouyang et al., 2014). However, deep generative models tend to have high specificity but low generality; in some cases, they explain out-of-training data inputs with high confidence while being wrong (Huang et al., 2017; Webster et al., 2019; Nalisnick et al., 2018). Furthermore, deep generative models require large numbers of training examples to be successful. Finally, the interpretability of deep generative models remains complex. Thus, non-parametric analysis of non-linear variations of non-homogeneous features for different medical image analysis tasks (shape completion, morphometric analysis, segmentation) with the same model, remains desirable.

With regard to the analysis of shape, a well established and understood formalism for analysing $3D$ geometric variation in a linearized statistical space exists in the form of statistical shape modelling (SSM). Statistical shape models typically model the data in a Euclidean vector space using principal component analysis (PCA) that treats feature variation as a linear combination of local displacements only (Cootes et al., 1992; Blanc and Székely, 2012; Mutsvangwa et al., 2015). While limited in capturing the non-linearity in shape space, the validity of this linearization for rigid shapes has been codified in the literature for single anatomical structures (Cootes et al., 1992; Blanc and Székely, 2012). Efforts for faithfully representing the non-linearity of shape space have been reported but have not become mainstream due to computational inefficiency and a lack of robustness (von Tycowicz et al., 2018). Recently, Lüthi et al. (2017) introduced a generalisation of SSMs, referred to as Gaussian process morphable models (GPMMs). In this framework, the parametric low-dimensional model was represented as a Gaussian process over deformation fields obtained from training examples. In contrast to discrete models (for example, point distribution models) that are dependent on artificial discretization, GPMMs are inherently continuous, that is, permitting of the arbitrary discretization of the domain on which the model is defined. However, GPMMs do not embed inter-object shape correlation, nor the anatomo-physiological relationship between articulating objects. This is because current SSMs and GPMMs are unable to represent, in a statistically robust and

intuitive way, an articulating anatomical complex composed of several rigid substructures which can move relative to each other.

Similar to the shape case, pose variation using linear descriptions has been reported (Smoger et al., 2015; Klinder et al., 2008; Fitzpatrick et al., 2011). A straightforward approach to model shape and pose together is to model joint flexibility implicitly by incorporating joint motion in statistical space (Klinder et al., 2008). Fitzpatrick et al. (2011) reported on a characterisation of the relationship between shape and contact mechanics of the patella and femur using an SSM of the articulating surfaces between both bones. The authors considered the patello-femoral contact space as a new object and modelled it as a standard SSM. Another approach is to concatenate pose and anatomical features into a joint vector and apply linear PCA to obtain joint principal components (PCs) (Smoger et al., 2015). A similar approach was reported for shape and pose modelling of the hip, PCs were obtained from a PCA computed over joint shape vectors (Agrawal et al., 2020). However, these methods model pose transformations as elements of an Euclidean vector space (linear space), contrary to reality, since they belong to a manifold and need to be described using non-linear statistical methods (Bossa and Olmos, 2007).

Reports abound in the literature on efforts to model shape and pose features together using methods more suited for managing the non-linearity of pose descriptors (Anas et al., 2014; Bossa and Olmos, 2007; Moreau et al., 2017; Fletcher et al., 2004). Anas et al. (2014) reported the embedding of relative pose variation of articulating bones into SSM frameworks, in the form of statistical shape and pose models (SSPM) using Rodrigues's parameters and spherical description for pose descriptors. Other authors have reported statistical pose models (SPM) only, that is to say, the variations modelled across spatial orientations without inclusion of their shapes (Bossa and Olmos, 2006). Bossa and Olmos (2007) proposed parametric low-dimensional models of several subjects in different poses with the pose variation representation obtained through principal geodesic analysis (PGA), a non-linear extension of PCA (Fletcher et al., 2004) as described in the literature review

(chapter 2, section 2.2.4). Shape and pose features of the neuroanatomical structures were concatenated in a long vector and standard multivariate statistics were extracted (Bossa and Olmos, 2007). Chen et al. (2014) proposed an articulated SSM for segmentation of the wrist from computed tomographic (CT) images without concatenating the shape and pose features; their model was a combination of an SSM and an SPM developed using PGA (Chen et al., 2014). All the above shape and pose PGA based models used a standard representation (Rodrigues, quaternion and Euler representation) for modelling the rotations describing the pose. Moreau et al. (2017) reported that the standard representation (SR) is limited in its ability to model rigid transformations of non-compact objects. A non-compact object is an object with small area compared to the area of its bounding circle. They instead proposed a new norm for statistical analysis of rigid transformations which is robust for analysing non-compact objects. However, they only demonstrated the norm for SPMs without considering shape and pose variability analysis. A non-linear shape modelling approach based on differential coordinates (avoiding global rigid transformation through local rotation and stretching) has been reported (Ambellan et al., 2019; von Tycowicz et al., 2018). While this approach captures the inherent non-linearity in shape variation, it still suffers from artificial discretization which prevents marginalisation of the resultant probabilistic model, and its capability of prediction. Additionally, it does not allow for modelling multiple objects/structures in a complex.

Despite the importance of shape and pose features in the analysis of multi-object complexes, an important feature class from the raw images themselves, namely intensity or grey-level distribution, needs to be considered. Typically, the shape and pose features are inferred from the intensity profile in medical images through pre-processing steps including segmentation and landmark annotation. Any errors in this pre-processing are propagated in subsequent analysis tasks. Leveraging the intensity profile as observations in images may lead to better analysis outcomes since it is a priori known that different tissue structures have unique intensity distributions. For example, in CT images, bones are brighter and soft tissue such as muscles are darker. Hence, shape and appearance-based models

have been considered as a good choice for the prediction of all organs in various medical image synthesis and applications (see chapter 2, section 2.2.1). Starting from Cootes et al. (2001, 1992), statistical models of appearance have been applied to images using a set of model parameters for both shape and grey-level variations learned from a training dataset. The same idea has been used on various reported SSPMs (as detailed in chapter 2, section 2.2.4) in the literature to predict structures from the observed images (Anas et al., 2016; Sebastian et al., 2003; Klima et al., 2016; Fotsin et al., 2019). Typically, the approach is to simultaneously use shape-pose priors and controllable grey-level models in segmentation/registration algorithms. However, none of these models embeds the grey-level features in the shape-pose priors. Additionally, these methods are based on inherent discretization of the structures of interest, structures which are continuous by nature. To summarise, the idea of embedding all features (shape, pose, intensity) in a unique model is not yet explored in the literature, likely due to some difficult theoretical considerations.

This chapter presents a Gaussian process-based dynamic multi-feature class modelling framework which encodes shape, pose and intensity in the same statistical space. The resultant models are called dynamic multi feature class Gaussian process models (DMFC-GPMs). A DMFC-GPM is self-contained, that is to say, this shape-pose-intensity based prior should not require any additional model to solve medical image analysis tasks. Furthermore, all features are built as a Gaussian process over deformations which makes them independent of the resolution of the training images and their artificial discretization. This work unifies feature-based models extending their interpretation to a larger context. The driving hypothesis is that the correlations of the three feature types (shape, pose, intensity) embedded in one model may provide more insight into patient-specific anatomo-physiological relationship signals and reveal underlying information. Figure 4.1 shows the flowchart of the proposed DMFC-GPM. Bottlenecks resolved to compute these models include: the linearization of rigid transformations in section 4.2; the homogenisation of shape, pose and intensity features in section 4.3; the computation of a Gaussian process-based latent space in section 4.4.1; the computation of a flexible kernel in section

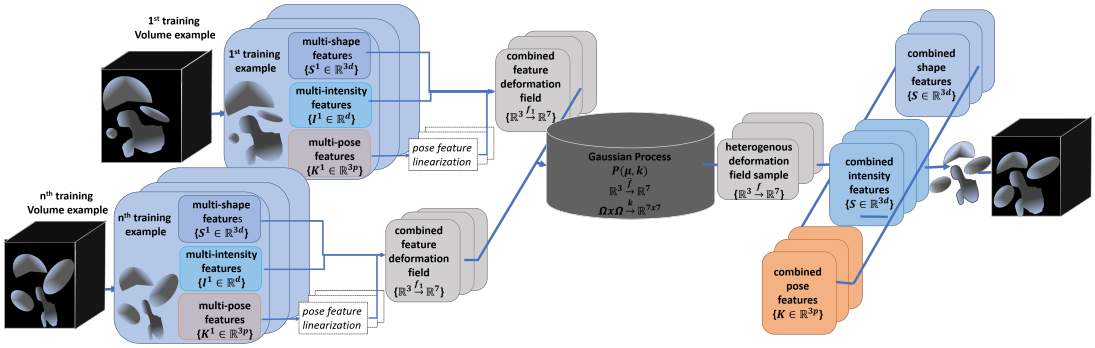

Figure 4.1 Proposed multi feature-class model architecture. Multiple volumetric shape-intensity and pose features are obtained after establishing dense correspondence. The pose feature linearization is performed through tangent space projection using log mapping. A Gaussian process is defined over deformation fields obtained from combined volumetric shape-intensity and pose features. Multi feature-class samples are obtained by 1) using the shape deformation field component to instantiate volumetric shape instances; 2) using the intensity distribution deformation field component to instantiate volumetric shape-intensity instances; and 3) retrieval of their spatial orientation using rigid transformations associated to the pose deformation field component

4.5.2; the computation of the marginal model in section 4.5.1; the computation of the permutation model in section 4.5.3; and the computation of the regression model in section 4.5.4.

4.2 Linearising rigid transformations

This section describes how to define the relative pose of an object as a deformation field from an atlas containing a reference object. Before deriving the linearization approach, terms and notations are defined to avoid confusion.

4.2.1 Definitions and notations

Volumetric shape also called object shape refers to a volumetric mesh such as tetrahedral mesh. **A volumetric shape-intensity refers to a volumetric shape with a scalar value associated with each vertex**. A volumetric shape-intensity joint refers to an ensemble constituted of volumetric shape-intensities. **A feature is a distinctive attribute of shape, pose or intensity**. **A feature class refers to a type of feature**,

that could be shape, pose or intensity.

A volumetric shape-intensity (V) is represented as a triplet: an object shape class (S), a pose class (P) and an intensity class (I). Such a representation is noted as:

$$V = (S, P, I)$$

The object shape class (S) is a volumetric mesh defined as a set $3D$ geometric points noted $S = \{s_i\}_{i=1}^{n_s}$, with s_i being a $3D$ point and n_s the number of vertices of mesh. Pose class (P) is also represented as a set points capturing the spatial orientation of the object noted as $P = \{p_i\}_{i=1}^{n_p}$, with p_i being a $3D$ point and n_p the number of pose points. The intensity class (I) is represented as set of scalar values denoted as $I = \{i_i\}_{i=1}^{n_s}$, with i_i being a scalar value.

A volumetric shape-intensity joint is a set defined as:

$$\{V^j = (S^j, P^j, I^j), j = 1, \ldots, N\} \tag{4.1}$$

where S^j, P^j, I^j represent the shape, pose and intensity feature classes of the j^{th} volumetric shape-intensity (V^j).

Assuming n joints of N volumetric shape-intensity each, as training examples, these can be used to describe a training dataset represented as a set:

$$\left\{V_i^N = \{(S_i^j, P_i^j, I_i^j), j = 1, \ldots N\}, i = 1, \ldots, n\right\}$$

With the terms to be used now defined, lets start by developing an efficient representation for rigid transformations which is necessary for a description of pose as element of the a linear space.

4.2.2 Developing a representation for rigid transformations

The literature on SSPMs usually emphasises a common shape-pose space for medical image processing (Anas et al., 2014; Bossa and Olmos, 2007). However, the standard representation (SR) of rigid transformations (described in section 2.2.4) may not always encapsulate the kinetic energy necessary to move points on an object from one position to another, particularly for non-compact shapes. This reduces the efficiency of this parameterisation for statistical analysis of rigid transformations. Moreau et al. (2017) proposed a kinetic energy-based representation that describes rigid transformations between two reference objects and a target object as a set of deformation fields from the reference object to the target. This is referred to in equation 2.10 (chapter 2). By representing pose as deformation fields, the resulting norm is proportional to the kinetic energy necessary for the displacement of the object. However, the size of the displacement is equivalent to the number of vertices of the reference mesh. The number of vertices can be large in high fidelity image analysis tasks. This, coupled with the use of an optimisation process for computing the Fréchet mean, can lead to high computational cost.

Below is a description of how this representation can now be extended to a continuous space allowing it to encode rigid transformations in DMFC-GPMs. This new transformation representation, which the author calls the energy displacement representation, will be referred to as EDR for the remainder of the thesis. With this representation, the number of points to be used for the estimation of the rigid transformation is customised. The Fréchet mean is directly obtained from the mean deformation fields, hence, its computation is more efficient.

Assuming there are n volumetric shape-intensity joints from the training dataset that are in-correspondence, we define the EDR of the j^{th} volumetric shape-intensity of the i^{th} joint example on its reference spatial position domain $\Omega_K^j \subset \mathbb{R}^3$ with their respective reference shape Γ_S^j and their reference pose points Γ_K^j. The EDR $\mathcal{F}_{iP}^j$ and its associated metric d are defined as:

$$\begin{aligned}
&\mathcal{F}_{iP}^{j}(x) = h_i^j(x) - Id(x), x \in \Omega_K^J \supset \Gamma_K^j, i = 1, \ldots, n \\
&h_i^j = \underset{h \in SE(3)}{\arg min} \|\overrightarrow{h(S_i^j)} - \overrightarrow{\Gamma}^j_S\|^2 \\
&d^2(h_i^j, Id) = \overrightarrow{\mathcal{F}_{iP}^{j}(\Gamma_K^j)}^T \overrightarrow{\mathcal{F}_{iP}^{j}(\Gamma_K^j)}
\end{aligned} \tag{4.2}$$

where $SE(3)$ (special Euclidian group) is the Lie group of orientation-preserving symmetries (translations, rotations) in $\mathbb{R}^3$. The metric d defined on the Riemannian manifold is efficient in differentiating and comparing rigid transformations (Moreau et al., 2017); P refers to the pose feature. Rigid transformations can be estimated using a least squares strategy (Umeyama, 1991), which is computationally inexpensive compared to the gradient descent strategy used in (Moreau et al., 2017).

The EDR metric has the property we desire for an efficient representation of rigid transformations (the resulting norm is proportional to the kinetic energy necessary for the displacement of the object's points). The above optimisation problem for rigid transformation parameters estimation uses point set based estimation, where the resulting translation vector highly depends on the spatial distribution of those points. The following section discusses the effect of the point set distribution on the translation vector estimation.

4.2.3 Point set-based estimation effect on the translation vector

Shape objects are typically represented as a set of points in $3D$ space. These can be a sparse set providing a crude approximation of the shape or a dense set of points for higher definition of the shape. Regardless, it is important to define an appropriate translation vector for rigid transformation of point sets. Typically, the estimation of rigid transformation parameters (Umeyama, 1991; Li et al., 2018) returns rotation R and translation parameters between two sets of points $\{x_i\}_i$ and $\{y_i\}_i$. The resulting translation vector is a function of the position of points in the point set. In a least squares sense, the translation is a function of the μ_x and μ_y, the average positions of the set $\{x_i\}_i$ and $\{y_i\}_i$, respectively. The average of a point set depends on the positional distribution used to describe it. As

such, two different point sets can give two different translation vectors for the same two objects. However, for an object moving in space $\mathbb{R}$, a translation vector depends on the displacement of the rotation centre. Figure 4.2 shows two different sets of landmarks (figure 4.2 a and b) used for the estimation of rigid transformation parameters and different resulting alignments. This point was not discussed in the work reported on the displacement field based non-linear metric for statistical analysis of rigid transformation (Moreau et al., 2017).

There are two ways to obtain a translation, invariant to positional distribution of points. In one approach, the rotation centre of the shape objects is computed and used to retrieve the translation vectors after the rigid transformation estimation (as illustrated in figure 4.2 a). Several biomechanical processes have been investigated for the computation of the rotation centre or the rotation axis including sphere and quadratic fitting methods (Eberly, 2000; Leboucher et al., 2010; Lempereur et al., 2010). In the second approach, only the points around (points that their centre of mass should be close to the rotation centre or on the rotation axis) the rotation centre or rotation axis are used to estimate the rigid transform parameters; the translation is subsequently a function of the average point (figure 4.2 b). However, while the first approach is more desirable, it does require an optimisation process for the computation of the rotation centre or rotation axis. The second approach is certainly less expensive in terms of computation, but it is also limited because different points may give different results or their centre of mass may not be close to the centre of rotation or the axis of rotation.

Now that we have an efficient representation of rigid transformations that ultimately leads to linear pose features, it remains for this representation to be harmonised with shape and intensity features to obtain a unified representation.

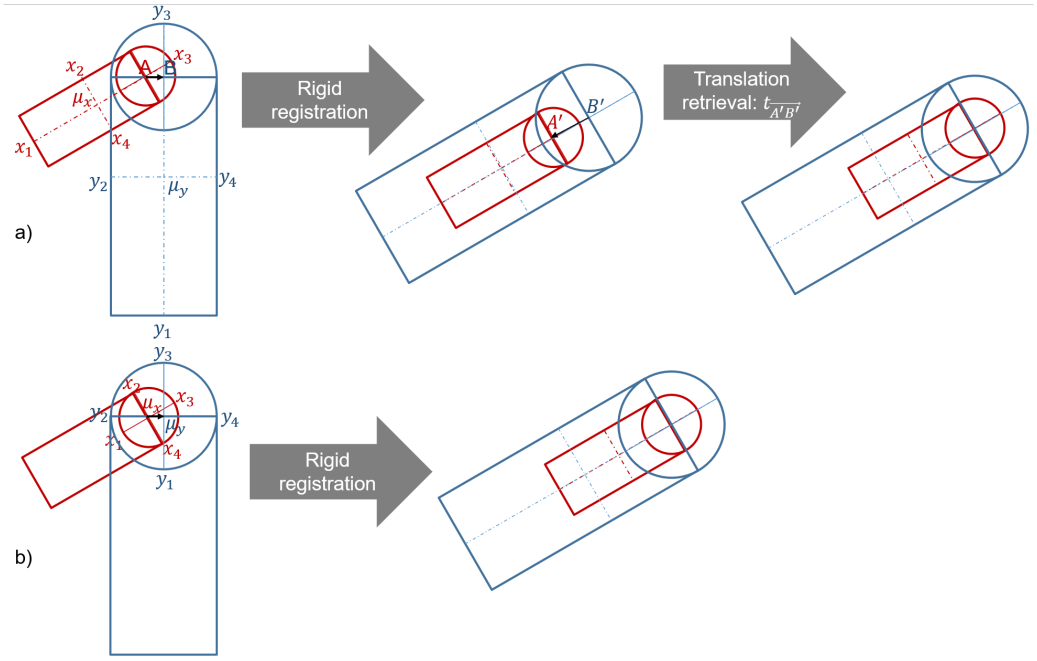

Figure 4.2 Illustration of how the distribution of points affects a translation vector. Note that only a sparse set of points is indicated for illustrative purposes. a) The average (μ_x and μ_y) of the points sets ($\{x_i\}_1^4, \{y_i\}_1^4$)) used for rigid transformation estimation are far from the rotation centre (A and B). The translation retrieval is necessary to obtain the translation after the registration. b) The average (μ_x and μ_y) of the used points sets approximate the rotation centres (i.e $\mu_x = A$ and $\mu_y = B$). The expected alignment is obtained after rigid registration.

4.3 Development of multi feature-class deformation fields

This section proposes a unified shape, pose and intensity deformation over which a Gaussian process ($\mathcal{GP}$) will be defined (section 4.4.1). **Shape, pose and intensity features are represented as deformation fields**. Representing features as deformation fields would: first **allow their interpolation to obtain continuous models** which would overcome the artificial discretization bottleneck encountered in previous models; and second, **allow marginalisation over any region of interest in the domain as well as over any feature class**. Since each feature class (shape, pose or intensity) belongs to its own space, different by nature, units and range (the amplitude of the variability across the dataset), **each requires normalisation prior to defining a unified representation of multi feature-classes**.

To have all feature classes embedded in the same latent space, a single representation combining all of them for each example in the training dataset is needed. In the above section (section 4.2.2), pose deformation field components ($\mathcal{F}_{iP}$) have been defined. Shape deformation field components ($\mathcal{F}_{iS}$)) obtained after applying the generalised Procrustes analysis (GPA) are defined as in GPMMs (see section 2.2.1) but over a volumetric domain for multiple objects. Intensity deformation field components ($\mathcal{F}_{iI}$) are defined as functions that map intensities of the reference domain to the ones of the target joint (see section 2.2.1).

To define shared deformation fields across the training dataset, a reference multi volumetric shape-intensity $\Gamma = \cup_{j=1}^{N}(\Gamma^j) \subset \mathbb{R}^3$ can be chosen from the training examples, with $\Gamma^j \subset \Omega^j$ (j^{th} reference domain). The deformation field of the i^{th} example in the training dataset

can then be defined as:

$$\mathcal{F}_i \colon \Omega \longrightarrow \mathbb{R}^7 \tag{4.3}$$

$$x \mapsto \mathcal{F}_i(x) = (\mathcal{F}_{iS}(x), \ldots, \mathcal{F}_{iP}(x), \ldots, \mathcal{F}_{iI}(x)) \tag{4.4}$$

with

$$\begin{cases} \mathcal{F}_{iS}(x) = \mathcal{F}^j_{iS}(x) \quad if\ x \in \Omega^j \\ \mathcal{F}_{iP}(x) = \mathcal{F}^j_{iP}(x) \quad if\ x \in \Omega^j \\ \mathcal{F}_{iI}(x) = \mathcal{F}^j_{iI}(x) \quad if\ x \in \Omega^j. \end{cases} \tag{4.5}$$

where Ω is the reference domain, and $\mathcal{F}_{iS}$ is the shape deformation field component, $\mathcal{F}_{iP}$ is the pose deformation field (see section 4.2), and $\mathcal{F}_{iI}$ is the intensity deformation field. The deformation field with a superscript j indicates the j^{th} object component.

Then, before obtaining residual features (around the unified mean) that represent normalised features, the joint mean function (unifying the mean shape, pose and density transformations) has to be defined. We estimate this mean function $\bar{\mathcal{F}}$ as:

$$\bar{\mathcal{F}} : (\bar{\mathcal{F}}_S, \bar{\mathcal{F}}_P, \bar{\mathcal{F}}_I) \colon \Omega^j \longrightarrow \mathbb{R}^7, j = 1 \ldots, N$$
$$x \mapsto (\frac{1}{n}\sum_{i=1}^{n} \mathcal{F}^j_{iS}(x), \bar{h}^j(x), \frac{1}{n}\sum_{i=1}^{n} \mathcal{F}^j_{iI}(x)) \tag{4.6}$$

with $\bar{h}^j = \underset{h \in SE(3)}{\arg min} \left\| h\left[\frac{1}{n}\sum_{i=1}^{n} \mathcal{F}^j_{iP}(\vec{\Gamma}^j_K)\right] - \vec{\Gamma}^j_K \right\|^2$.

In order to obtain the unified space of shape, pose and intensity feature classes, the pose transformation needs to be projected onto a linear space. We project rigid transformations (this is the linearisation of relative spatial transformations) onto a tangent space to the Riemannian manifold at the mean rigid transformation using the EDR (eq 4.2). Relative spatial transformations are linearised through exp / log bijective mapping.

To do so, let us define the exp / log function associated to multi feature-class transformations. The log mapping is defined as:

$$\begin{aligned}
&\log[(\mathcal{F}^j_{iS}, h^j_i, \mathcal{F}^j_{iI})] = (\mathcal{F}^j_{iS}, \mathcal{F}^j_{iP}, \mathcal{F}^j_{iI}), i = 1, \dots n, \qquad j = 1, \dots, N, \\
&\log_{(\mathcal{F}^j_{i_1S}, h^j_{i_1}, \mathcal{F}^j_{i_1I})}[(\mathcal{F}^j_{i_2S}, h^j_{i_2}, \mathcal{F}^j_{i_2I})] = \log[((\mathcal{F}^j_{i_2S})^{-1} \circ \mathcal{F}^j_{i_1S}, (h^j_{i_2})^{-1} \circ h^j_{i_1}, (\mathcal{F}^j_{i_2I})^{-1} \circ \mathcal{F}^j_{i_1I})]
\end{aligned} \tag{4.7}$$

And the exp mapping (inverse of the log) is defined as:

$$\begin{aligned}
&\exp[\mathcal{F}^j_{iP}] = \underset{h \in SE(3)}{\arg min} \|h[\mathcal{F}^j_{iP}(\vec{\Gamma}^j_K)] - \vec{\Gamma}^j_K\|^2 \\
&\exp_{(\mathcal{F}^j_{iS}, h^j_i, \mathcal{F}^j_{iI})}[(\mathcal{F}^j_{iS}, \mathcal{F}^j_{iP}, \mathcal{F}^j_{iI})] = (\mathcal{F}^j_{i_2S} \circ \mathcal{F}^j_{i_1S}, h^j_{i_1} \circ \exp[\mathcal{F}^j_{i_2P}], \mathcal{F}^j_{i_2I} \circ \mathcal{F}^j_{i_1I})
\end{aligned} \tag{4.8}$$

where $(\mathcal{F}^j_{i_1S}, h^j_{i_1}, \mathcal{F}^j_{i_1I})$ and $(\mathcal{F}^j_{i_2S}, h^j_{i_2}, \mathcal{F}^j_{i_2I})$ are two unified transformations fields.

The joint deformation fields projected onto the vector space (figure 4.3) over which the $\mathcal{GP}$ will be defined are $\{\mathcal{F}_i\}_{i=1}^n$

$$\begin{aligned}
\mathcal{F}_i &: \Omega \longrightarrow \mathbb{R}^7 \\
&x \mapsto \log_{\bar{\mathcal{F}}}[(\mathcal{F}_{iS}, \mathcal{F}_{iP}, \mathcal{F}_{iI})](x)
\end{aligned}$$

Once joint deformation fields are defined, a $\mathcal{GP}$ can be built over them to estimate a latent space.

4.4 Building and sampling continuous multi feature-class latent space

In this section the DMFC-GPM is built through the latent space computation as a set of orthogonal functions. The model sampling process is proposed as well as the attributes of

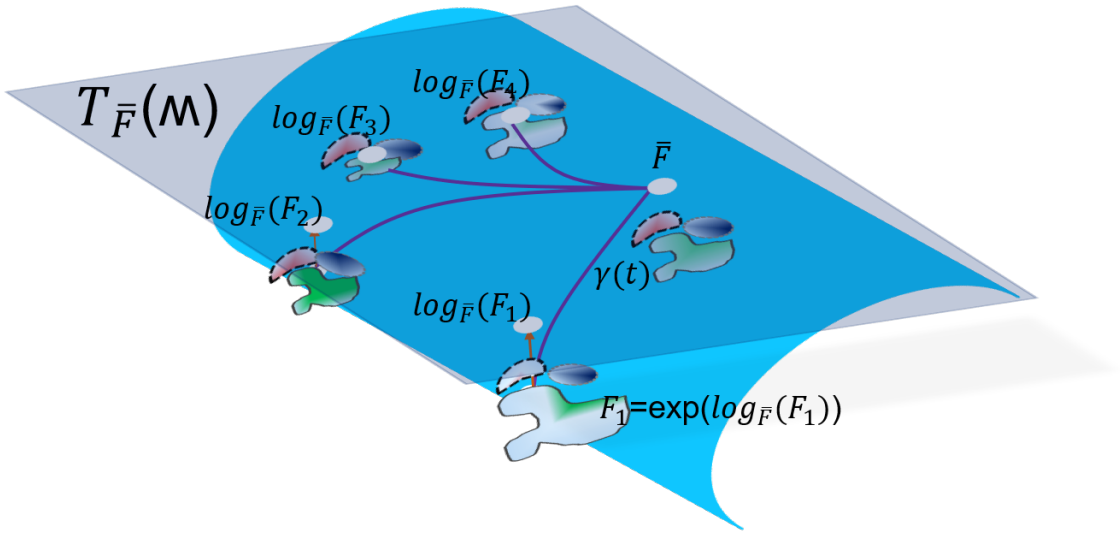

Figure 4.3 Training volumetric shape-intensity joint examples are projected into the tangent space at the mean deformation ($T_{\bar{F}}(M)$). The manifold of volume shape-intensity examples is represented in blue and the tangent at the mean deformation is represented in grey.

the models.

4.4.1 Establishing a latent space

From a training dataset of n examples with N volumetric shape-intensities each, the deformation fields $\{\mathcal{F}_i\}_{i=1}^{n}$ defined on the continuous domain $\left(\Omega = \cup_{j=1}^{N}\Omega^j\right)$ can now be modelled by a Gaussian process defined as:

$$\mathcal{F} \sim \mathcal{GP}(\bar{\mathcal{F}}, \mathcal{K}) \tag{4.9}$$

where $\bar{\mathcal{F}}$ is defined in Eq. 4.6, and $\mathcal{K}$ is the kernel function defined as:

$$\mathcal{K}(x,y) = \frac{1}{n-1}\sum_{i=1}^{n}\mathcal{F}_i(x)\mathcal{F}_i(y)^T, \quad (x,y) \in \Omega \times \Omega \tag{4.10}$$

Through the $\mathcal{GP}$, the multi feature-class deformation fields $\mathcal{F}$ are represented by an orthogonal set of basis functions denoted by $\{\Phi_m\}_{m=1}^{M}$, $M \leq n$ (number of examples) and

defined as:

$$\Phi_m(x) = \left(\Phi_m^S(x), \Phi_m^P(x), \Phi_m^I(x)\right) \tag{4.11}$$

where Φ_m^S, Φ_m^P and Φ_m^I represent the shape, the pose and the intensity component of the m^{th} basis function, respectively. After computing these latent space functions, the GP is made continuous through interpolation. Different interpolation methods are described in section 2.2.1. The functions $\{\Phi_m\}_m$ are now continuous, leading to the desired continuous latent space.

4.4.2 Sample generation

After the parameterisation of the latent space above, we can make the model generative through a probabilistic formulation (Gaussian noise) to permit sampling from the model. The j^{th} volumetric shape-intensity is thus modelled as a combination of shape deformation field, a rigid transformation (obtained using the exponential mapping described in section 4.3), and intensity deformation field, as defined below:

$$\mathcal{F}_\alpha \sim \exp_{\bar{\mathcal{F}}}[(\mathcal{F}_{\alpha S}, \mathcal{F}_{\alpha P}, \mathcal{F}_{\alpha I})], \qquad \alpha = (\alpha_m)_m \sim \mathcal{N}(0, \mathbf{I}) \tag{4.12}$$

with

$$\begin{cases} \mathcal{F}_{\alpha S}(x) = \sum_{m=1}^{M} \alpha_m \sqrt{\Lambda_m}\, \Phi_m^{S_j}(x) \\ \mathcal{F}_{\alpha P}(x) = \sum_{m=1}^{n} \alpha_m \sqrt{\Lambda_i}\, \Phi_m^{P_j}(x) \\ \mathcal{F}_{\alpha I}(x) = \sum_{m=1}^{M} \alpha_m \sqrt{\Lambda_m}\, \Phi_m^{I_j}(x). \end{cases} \tag{4.13}$$

where $x \in \Omega^j$ is a point in the domain, the pair $(\Lambda_m, \Phi_m)_{m=1,\dots,M}$ represents the m^{th} eigenvalue and associated eigenfunction; and $\mathbf{I}$ being the $M \times M$ identity matrix. These couples are obtained using Karhunen-Loève expansion and Nyström approximation (Berlinet and Thomas-Agnan, 2011; Rasmussen, 2003) as shown in equation 2.2. The m^{th} eigenfunction Φ_m is then a continuous multi feature-class deformation field defined on Ω, in the latent space.

An example in figure 4.4 shows how a j^{th} volumetric shape-intensity is sampled from the model. An instance of the j^{th} volumetric shape-intensity and its associated spatial position can be generated from our probabilistic model for $\alpha = (\alpha_m)_m$ as:

$$V(S^j, h^j, I^j) = V(\mathcal{F}_\alpha(\Gamma^j)) \tag{4.14}$$

where $\Gamma^j \subset \Omega^j$ is the reference shape of the j^{th} object in the joint. The DMFC-GPM model

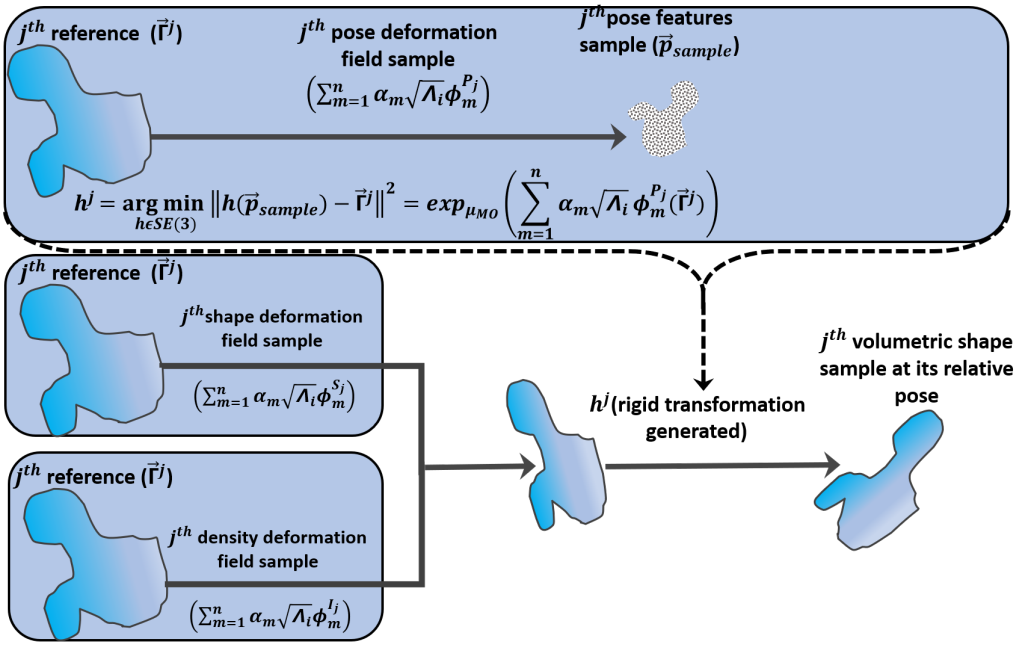

Figure 4.4 Obtaining volumetric shape-intensity sample in its associated spatial orientation. Top: the rigid transformation is obtained from pose component deformation field sample. Middle-bottom: the sampled volumetric shape-intensity is obtained from shape and intensity deformation fields, and its spatial orientation is retrieved using the generated rigid transformation.

can be used to derive single or multi-feature class models on demand without the need for retraining. The following section presents how different sub-models can be derived from the DMFC-GPMs.

4.5 Sub-model attributes of the DMFC-GPM framework

Since Gaussian process-based models operate in a continuous domain, the DMFC-GPM framework inherits properties intrinsic to GPMMs (Lüthi et al., 2017). These attributes include flexibility adjustment, marginalisation and the regression. However, in addition to the region of interest shape marginalisation allowed by GPMMs, the DMFC-GPM framework goes beyond shape and allows single shape, multiple shape-spose, intensity or region based multiple feature marginalisation as summarised in figure 4.5. Similarly, the framework also allows regression to any of the features included in the domain. Moreover, a permutation model can be retrained to allow more flexible poses in the model as can also be seen in figure 4.5. All these attributes are discussed in this section.

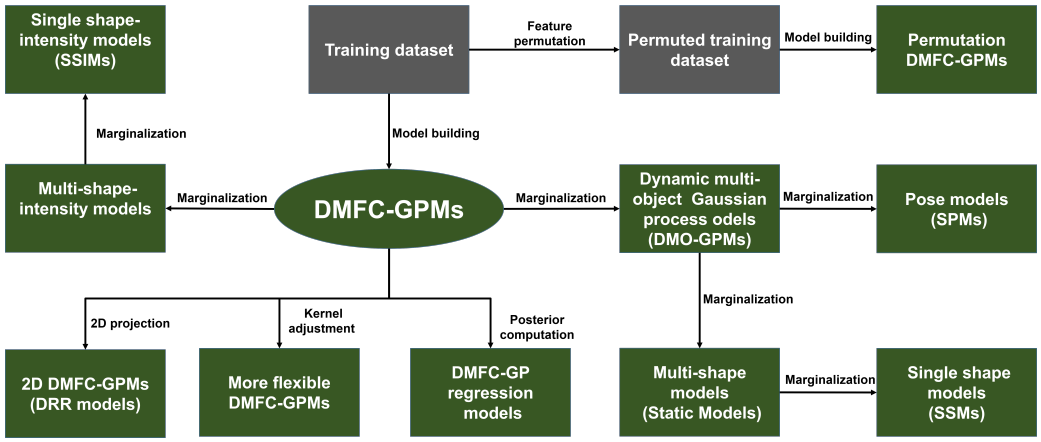

Figure 4.5 Different attributes of the DMFC-GPMs. Different models obtained through marginalisation, kernel adjustment or regression of the DMFC-GPMs. The permutation model is obtained through a retraining using dataset with pose permuted.

4.5.1 Marginalising the model

As in GPMMs, our DMFC-GPM framework allows for the estimation of missing parts of a given knowledge of a specific part of the domain; thanks to the marginal probability properties.

Let us assume that the DMFC-GPM is a joint probabilistic model denoted by $p(\Omega)$. The model could be marginalised over a subdomain as inherited in GPMMs, which could be features: from a specific class, from a specific structure or from different classes. For example, the DMFC-GPM of the shoulder complex (including scapula and humerus and clavicle) could be marginalised into three different individual object models. Furthermore, DMFC-GPM model includes static multi-object models, that is to say, setting the pose parameters constant to obtain a classical SSM of the joint. Conversely, it also includes a motion model, that is to say, setting the shape parameters constant would lead to a pose-only statistical model. Given a subdomain A, its marginalised model is defined as

$$p(A) := \mathcal{GP}(\bar{\mathcal{F}}, \mathcal{K})|_{A \subset \Omega}. \tag{4.15}$$

The DMFC-GPM marginalised to shape and pose or the DMFC-GPM with only shape and pose (that is the DMFC-GPM with intensity constant) is termed a Dynamic Multi-Object Gaussian Process Model (DMO-GPM). Such a model extends SSPMs through not only making them inherently continuous, but also by embedding a robust metric (Moreau et al. (2017)) for rigid transformation comparisons (see chapter 2, section 2.2.4).

Below is the formulation of the DMO-GPM built as a DMFC-GPM with only shape-pose variability (Fouefack et al., 2020). The DMO-GPM is defined as:

$$\mathcal{F} \sim \mathcal{GP}(\bar{\mathcal{F}}, \mathcal{K}) \tag{4.16}$$

where $\bar{\mathcal{F}}$ is defined in Eq. 4.6, and $\mathcal{K}$ is the kernel function defined as:

$$\mathcal{K}(x, y) = \frac{1}{n-1} \sum_{i=1}^{n} \mathcal{F}_i(x) \mathcal{F}_i(y)^T, \quad (x, y) \in \Omega \times \Omega \tag{4.17}$$

with a joint deformation field $\mathcal{F}_i = (\mathcal{F}_{iS}, \mathcal{F}_{iP}, \mathcal{F}_{iI})$, having no intensity variation, that is to say, the intensity deformation field component $\mathcal{F}_{iI}$ is constant.

4.5.2 Modelling with kernels

With the DMFC-GPM it is possible to increase the variability around the mean without retraining the model; a property inherited from GPMMs. Often, the number of examples for training a model is insufficient to provide a subspace of shapes to represent all possible shape variations within the family, which introduces bias. The extension of the variability using a bias kernel can overcome this limitation (Lüthi et al., 2017). The framework proposed above allows for the understanding of such variability and defining an appropriate kernel that encodes the variability, not only for shape as in GPMMs but also for pose. A new kernel can be added to the trained model kernel without a need of retraining examples.

An example of a kernel that can increase smoothness in a model is the Gaussian kernel as used in (Lüthi et al., 2017). It can be defined in a specific subdomain of $\Omega^j, j = 1, \ldots, N$ by:

$$\mathcal{K}_{\Omega}^{(s,\sigma)}(x, y) = s \exp(\|x - y\|^2/\sigma^2), \qquad (x, y) \in \Omega \times \Omega \tag{4.18}$$

where the parameter s defines the scale of the average error and σ the range over which deformations are correlated (Lüthi et al. (2017)). The new DMFC-GPM kernel can be a combination of the pre-trained kernel with different Gaussian kernels in each domain defined by

$$\mathcal{K}_{new} = \mathcal{K} + \sum_{j=1}^{N} I_{7\times7}\mathcal{K}_{A}^{(s_j,\sigma_j)}, \ A \subset \Omega_j \tag{4.19}$$

with $I_{7\times7}$ is the 7×7 identity matrix.

Another possibility of increasing motion variability could be through extending training data pose. This aims to enable a training joint example to extend beyond its original pose (pose from the acquisition), that is to say, matching a volumetric shape-intensity deformation from one example with other joint pose deformations. By doing that, the correlation between shape-intensity and pose is reduced or completely removed. This increases the degrees of freedom or the range of motion of training examples.

4.5.3 Pose permutation models

The basis functions $\{\Phi_m\}$ over which the $\mathcal{GP}$ is built have pose and shape components, and it is a priori known that an object could have multiple possible poses that define its motion, relative to neighbouring objects. In practice, there are various constraints for obtaining possible poses of a given object. These constraints include limitation of patient's movement in image acquisition instruments (scanners) and high cost associated with imaging. Building a model with different motion ranges for different objects could result in shape-pose correlations that may not exist. To overcome this difficulty, a pose permutation-based modelling approach can be leveraged from the DMFC-GPM framework. The permutation approach consists of associating poses from other objects in training dataset to each object.

Given the training dataset deformation fields $\{(\mathcal{F}_{iS}, \mathcal{F}_{iP}, \mathcal{F}_{iI}), i = 1, \ldots n\}$, the permutation training is defined as;

$$\{(\mathcal{F}_{iS}, \mathcal{F}_{jP}, \mathcal{F}_{iI}), j = 1, \ldots, n, i = 1, \ldots, n\}$$

The resulting model is more generalised, however, the price of this generalisation may be the encoding of motion that is unrealistic for biological structures in the model since allowable motion is typically constrained by the shape. To mitigate this drawback, the permutation defined above can be constrained using expert prior knowledge about a shape-pose relationship, for example, only associating the pose of one shape to another one if both shapes are similar. While these permutation models could represent alternatives, the best scenario for modelling pose variation remains the use of specific poses from real subjects which necessitates acquisition of data in different positions of constituents bones.

4.5.4 Posterior models

Statistical models usually leverage the correlation that exists within features to perform the analysis while only having limited knowledge. The correlation may not exist in some scenarios but one may still want to constraint models to limited observations. An alternative to correlation-based analysis is regression analysis. Let us consider that one does not have knowledge about the correlation, but can observe part of the multi feature-class deformation fields, that is to say, for instance, knowing some corresponding feature of the reference domain on the observation. Let us consider that we have an observed joint $O = \{V^1, \ldots, V^N\}$, with N volumetric shape-intensity at various poses. Let us further assume that we know some points on the reference domain and their correspondence on the observed joint, noted as $\{x_O, \mathcal{F}_O, \sigma^2\}$, where x_O are observed features on the reference domain constituting the set Ω_O, and $\mathcal{F}_O$ the deformation field mapping them to the target ones; σ^2 is the Gaussian noise level of the observation. We can formulate the dynamic multi feature-class Gaussian process (DMFC-GP) regression models (Williams and Rasmussen, 2006; Lüthi et al., 2017) as:

$$\mathcal{F}|x_O : \mathcal{F}_O \sim \mathcal{GP}(\bar{\mathcal{F}}_O, k_O) \tag{4.20}$$

by conditioning the mean:

$$\bar{\mathcal{F}}_O(x) = \bar{\mathcal{F}}(x) + k(x, x_O)^T (k(x_O, x_O) + \sigma^2 I)^{-1} (x_O - \mu(x_O)), \quad x_O \in \Omega_O, x \in \Omega \tag{4.21}$$

and conditioning the kernel:

$$k_O(x, y) = k(x, y) - k(x, x_O)^T (k(x_O, x_O) + \sigma^2 I)^{-1} k(x_O, y), \quad x_O \in \Omega_O, y \in \Omega \tag{4.22}$$

where

$$\mathcal{K}(x_O, x_O) = \frac{1}{n-1} \sum_{i=1}^{n} \mathcal{F}_O(x_O) \mathcal{F}_O(x_O)^T \tag{4.23}$$

and

$$\mathcal{K}(x_O, y) = \frac{1}{n-1}\sum_{i=1}^{n}\mathcal{F}_O(x_O)\mathcal{F}(y)^T, x_O \in \Omega_O, y \in \Omega \tag{4.24}$$

One of the practical applications of this regression is to constrain a model at the desired relative pose. An example case is where x_O could be the reference pose landmarks, and $\mathcal{F}_O$ would be a rigid transformation mapping those landmarks to the desired relative spatial orientation.

4.6 Computational consideration

This section describes computational challenges associated to the DMFC-GPM framework. As a GPMM-based framework, the DMFC-GPM framework is formulated in such a way that it is computationally independent of the representation of the objects. The computational time required to obtain the latent space does relies on the size of the domain (which is a function of the number of features), but only on the number of the training examples (n) (see equation 4.10). The GPA algorithm embedded in the framework (see section 4.3) requires an optimization process and may be computationally expensive (Gower, 1975). However, the GPA removes remaining rotation, translation and scaling variations after rigid registration, hence the process often requires a small number of iterations. Moreover, as in the GPMM framework, the least squares-based estimation of the rigid transformation parameters is used for the GPA which is not computationally expensive as described in section 4.2.2.

4.7 Conclusion

This chapter presented a framework for combining statistical representations of multiple feature classes, namely shape and pose and intensity features, thus providing a principled way for building data-driven generative models for medical image analysis. Previous statistical models in the literature (SSPMs, SPMs, SSMs or SAMs) can be instantiated using the unified framework (as summarised in the figure 4.5). A new shape-pose modelling ap-

proach DMO-GPMs that extends GPMMs (Lüthi et al., 2013), SSPMs (Bossa and Olmos, 2007) and SPMs (Moreau et al., 2017) was also presented as an attribute of the DMFC-GPM framework. The formulation employs interpolated deformation fields for shape, pose and intensity feature classes that incorporate the natural representation of the objects in terms of smoothness, proposing a continuous dynamic multi-feature latent space.

The presented framework not only uses a homogeneous and linear statistical representation of both shape and pose features, but it also embeds the intensity distribution within the internal structure of the object of interest. Additionally, it embeds a Gaussian noise thereby making the model generative, as well as able to estimate the latent space in a continuous domain. This type of framework maybe useful in the medical imaging domain where structures of interest can be articulating joints. Human joint structures in images represent non-linear and non-homogeneous feature classes in terms of their bone shapes, poses, and grey-level representation. The optimal and homogeneous statistical representation of the variation across multi feature classes create a new paradigm that ultimately provides latent information necessary for more accurate prediction in medical images.

More importantly, the model's attributes and sub-categorical extension of the models were proposed for different analyses. The presented modelling framework also allows for the incorporation of extension methods that are very useful when dealing with limited data. For the extension property, first, the GPMMs kernel-based extension was adopted to allow more flexibility to be added to the model based on an operator's prior knowledge of the morphology of interest and expected variation. Second, a pose variability extension was proposed using permutation properties that reduce shape-pose correlation and increases pose variability which is very important when dealing with joints with high range of motion. The illustrated marginalisation property of the model allows the statistical analysis of a region of interest while conserving the global statistics. Moreover, it was shown that a posterior model can be derived for regression analysis which can be regarded as a viable alternative to correlation-based analysis.

Chapter Five

Dynamic multi feature-class model fitting and validation

5.1 Introduction

This chapter presents the validation of the dynamic multi feature-class Gaussian process model (DMFC-GPM) framework. In particular, the prediction ability of models built using the framework is evaluated through fitting the models to data. The goal of the model fitting is to determine the optimal parameters of the model that explain a given observation ($3D$ mesh surface or image). The model fitting corresponds to a projection of the observation into a latent space of few descriptors. In 1994, Cootes et al introduced shape model fitting, where the model parameters serve as compact descriptors of shapes (Cootes et al., 1994). This forms a good starting point for understanding the use of statistical shape models (SSMs) as priors for the analysis of given observations. The model prior should have the following property: it needs to truly represent all the variation of the object's family that it describes. As such, the fitting outcome should be restricted to the correct features and should not include noise or artefacts. This intrinsic attribute puts statistical models (shape, pose, appearance) into the restrictive popular tools for medical image analysis tasks including image segmentation, and prediction of pre-morbid anatomy (Salhi et al., 2020; Lüthi et al., 2013; Cootes et al., 2001). Through fitting SSMs (shape), ASMs (shape and appearance) and SSPMs (shape and pose) onto observations, a representative sample of the observation is obtained for the shape, intensity or pose that is modelled. Any annotation given on the model can be transferred to the observation via its representation in the

model space, allowing for transfer learning; for example, the identification of pathological regions of a bone. This shows the big difference between model registration (chapter 3) and model fitting. Registration does not establish a transfer learning channel because the model used for registration often encodes generic deformations and generic deformations may be unrealistic.

For the DMFC-GPM, the desired fitting process should establish a means for transferring not only shape information from the model to the observation as previous models but also provide a manner to determine the correct relative position and intensity profile. The fitting may not be able to perfectly explain the observations. However, it should explain it at least to the level of validity of the priors, which depends on the number and the variability of the training dataset, as well as the registration quality. Various fitting methods have been reported in the literature. Starting from Cootes et al. (1994) who introduced a parametric method that uses shape models to localise structures in medical images, other parametric fitting methods reported and applied in medical image analysis include methods for fitting onto mesh surfaces (point clouds) (Brunton et al., 2016) or images (Heimann and Meinzer, 2009; Zhou and Xu, 2016) etc. A unified parametric approach in a continuous domain using the mean square metric to fit models onto surfaces, points or images has been reported (Lüthi et al., 2017). While single object model-fitting approaches have shown great potential, fitting multi-structure models remains challenging. The expectation-maximisation algorithm (Dempster et al., 1977) and parametric non-rigid registration (see survey for different approaches (Myronenko and Song, 2010)) have been adopted to fit shape and pose models into computed tomography (CT) images (Rasoulian, Rohling and Abolmaesumi, 2012; Anas et al., 2016; Rasoulian et al., 2013). These parametric-based approaches are faster because they depend on the optimisation of a cost function explicitly defined by metrics. Finding the global minimum/maximum of a cost function, however, remains challenging because of the existence of local minima/maxima. As such, sampling-based methods for approximating the global minimum/maximum have been suggested (Morel-Forster et al., 2018; Madsen et al., 2019; Schönborn et al.,

2017).

Sampling-based methods have provided the computational basis for the rapid expansion of statistical machine learning in recent years (Ma et al., 2019). For example, the Markov chain methods and the popular Metropolis-Hastings algorithms are intensively used to simulate multivariate distributions (Schönborn et al., 2017). In SSM-based fitting applications, some work using sampling-based methods has been reported, in particular, the probabilistic morphable model fitting framework (Morel-Forster et al., 2018; Madsen et al., 2019; Schönborn et al., 2017). This approach works by finding optimal shape parameters that best represent the observation in the latent space. However, to the best of the author's knowledge, the approach has not been extended for use on models embedding combined variations of shape, pose and intensity features. The validation of the fitting with complex models such as DMCF-GPMs that include three feature-class types (shape, pose and intensity) has to be carefully conducted to be able to analyse inter-correlations and feature-class type differences.

In this chapter, a sampling-based Markov chain Monte Carlo (MCMC) fitting method for DMFC-GPMs is described. In addition, adaptations developed for different observation scenarios (point cloud-based mesh surfaces, $3D$ and $2D$ images) are also presented. Finally, model validation using proposed built synthetic data with controllable feature variation and correlations is presented. Some of the experiments and results are parts of the content presented at the Medical Image Computing and Computer-Assisted Intervention (MICCAI2020) conference and published in the MICCAI2020 conference proceedings (Lecture Notes in Computer Science 12264) and on Arxiv.

5.2 DMFC-GPM fitting

The Markov chain Monte Carlo (MCMC) and Metropolis-Hastings algorithms (Morel-Forster et al., 2018; Madsen et al., 2019; Schönborn et al., 2017), as introduced in chapter 3 for registration, can be adopted as a sampling-based framework for model fitting. Let us

consider observations to be objects with arbitrary relative positions to each other. In the case of observations from image data (CT, magnetic resonance (MR), X-rays) for instance, each displayed structure of interest is in its relative position at the time of image acquisition with respect to other structures. The goal of the MCMC fitting process is to find the best model parameters θ that optimally represent the target example. The uncertainties related to the pose and the intensity estimation are embedded in the DMFC-GPM and therefore we should do not need a separate model. The observation examples may consist of partial objects: bones with slight cracks caused by osteomalacia or any incomplete anatomical structure due to pathology or due to image artefacts (cropping etc.). The advantage of the DMFC-GPM is that on encountering a missing part on one object, the other objects in the target example can be used as additional observations since the correlation between the objects is embedded in the model. Thus, the DMFC-GPM with MCMC fitting can be used to obtain a posterior model of one object while another is missing. Below is a presentation of the formulation for MCMC fitting using DMFC-GPM as prior.

The MCMC fitting presented now aims to find the deformation field $\mathcal{F}_\theta$ that maximises the product of the likelihood and the prior. The resulting Bayes's formula for the sampling-based fitting is thus:

$$\arg\min_{\theta\in\Theta} \; p(T|\mathcal{F}_\theta)p(\mathcal{F}_\theta) \tag{5.1}$$

where Θ is the latent space which is a space of possible deformation field parameters and T the observation. Full shape, pose and intensity estimations require a likelihood which captures within-object and between-object features. A multi-feature class comparison likelihood, taking into account both global and local components, can thus be proposed. The global component ensures the prediction of all feature classes while each local component ensures the predominance of its feature class in the global likelihood. But before finding the similarity between predicted and observed objects, the fitting process requires an initialisation.

5.2.1 Fitting initialisation

An important step in model fitting is to build a process where both observation and model remain invariant under irrelevant rigid transformations. These irrelevant transformations are rigid transformations existing between two images due to the acquisition protocol or the type of the machine. For example, one can have two CT images of the shoulder of the same patient with the humerus being at 20 degree abduction relatively to the scapula, the irrelevant rigid transformation will be any other rigid transformation between the scapula and the humerus than the 20 degree abduction between them. Various ways of initialising model fitting have been reported. Anas et al. (2016) used a rough segmentation of the object's boundaries in the image followed by the alignment of principal axes of the model mean and the segmented object. Lüthi et al. (2013, 2017) selected landmarks on the object in the image and on the model mean followed by the alignment through least-squares methods. Compared to the landmarks-based method which only needs the selection of few landmarks, the rough segmentation of multiple structures can be time-consuming and lead to the propagation of the segmentation errors through the fitting process. However, when dealing with multiple structures, it may be hard to identify landmarks and associate them to the corresponding objects. The DMFC-GPM developed in this project requires the annotation of one of the objects in the model for initialisation. This object is the fixed object relative to which the other objects in the DMFC-GPM model are considered to be moving. This is explained in detail in the registration method described in chapter 3.

5.2.2 Local likelihood

To locally ensure a good fit, predicted and observed features should be locally compared. To do so, a local likelihood should be associated to each object in the model to independently compute the local similarity. Such a local process is very important especially when structures do not have the same number of features, for example, the number of clavicle features may be lower than that of the scapula.

The j^{th} local likelihood l_j compares the observation of the object with an $\mathcal{F}_\theta(\Omega^j)$ generated from the model using independent feature-wise comparison. This can be considered as a probability distribution of possible shape, pose and intensity samples evaluated for the target. It locally fits the DMFC-GPM by maximum likelihood. T^j could represent a subset of the j^{th} target object space.

$$l_j(\mathcal{F}_\theta, T^j) = p(T^j|\mathcal{F}_\theta(\Omega^j)) = \prod_{x_j \in T^j} p(x_j|x_j(\mathcal{F}_\theta)) \tag{5.2}$$

where Ω^j is the domain of the j^{th} object, p the probability density function (PDF) and $x_j(\mathcal{F}_\theta)$ is the most probable correspondent of the points x_j on the model instance $\mathcal{F}_\theta(\Omega^j)$.

5.2.3 Global likelihood

In contrast to the local likelihood, the global likelihood maximises the ensemble of the objects of interest. It may not take into account the specificity of each individual structure, however it can drive the model fitting away from local maxima/minima. The global likelihood L compares the observation of target T with the model instance $\theta(\Omega)$ and it is defined as:

$$L(\mathcal{F}_\theta, T) = p(T|\mathcal{F}_\theta(\Omega)) = \prod_{x_j \in T} p(x_j|x_j(\mathcal{F}_\theta)) \tag{5.3}$$

where Ω is the global domain of the ensemble, and $x_j(\mathcal{F}_\theta)$ is the most probable correspondent of the points x_j on the model's instance $\mathcal{F}_\theta(\Omega)$.

5.2.4 Obtaining the optimal parameter set

A sample $\theta^{'}$ is accepted as a new state θ for the j^{th} object with a Metropolis decision probability β_j, and the target T with a probability β, as below:

$$\beta_j = \min\left\{\frac{l_j(\mathcal{F}_\theta, T^j)p(\mathcal{F}_\theta)}{l_j(\mathcal{F}_{\theta'}, T^j)p(\mathcal{F}_{\theta'})}\frac{Q(\theta^{'}|\theta)}{Q(\theta^{'}|\theta)}, 1\right\} \tag{5.4}$$

$$\beta = \min\left\{\frac{L(\mathcal{F}_\theta, T)p(\mathcal{F}_\theta)}{L(\mathcal{F}_{\theta'}, T)p(\mathcal{F}_{\theta'})}\frac{Q(\theta'|\theta)}{Q(\theta'|\theta)}, 1\right\} \tag{5.5}$$

where $Q(\theta'|\theta) = \mathcal{N}(\theta'|\theta, \Sigma_\theta)$ is the proposal generator. Finally, the proposal Q is fed through a sequence of Metropolis acceptance decisions:

$$\{\beta_j, \beta, j = 1, \ldots, N\} \tag{5.6}$$

The predicted target is the instance of the posterior model with the highest probability β.

5.3 Different observation scenarios

Observations refer to what is being measured during a study or experiment. In medical image processing, observations can vary and the prediction strategy may be specific to the type of observation. These observations could be point clouds (which include structured and unstructured points) or images ($2D/3D$ images or a sequence of images). The prediction quality depends on the amount of knowledge provided by the observation as well as the model fitting algorithms. In this section, prediction strategies associated to different types of observations are presented.

5.3.1 Point cloud observation

For $3D$ surface meshes represented as point clouds, a likelihood is required to estimate the similarity between the proposed set of points with the target ones. For the observed point x under a Gaussian noise σ, the likelihood value can be estimated as:

$$p(x|x(\mathcal{F}_\theta), \sigma) = \mathcal{N}(x|\mathcal{V}(x, \mathcal{F}_\theta(\Omega)), \sigma^2) \tag{5.7}$$

where $\mathcal{V}(x, \mathcal{F}_\theta(\Omega))$ is the nearest neighbour point of the point x in current model's instance $\mathcal{F}_\theta(\Omega)$.

5.3.2 3D image observation

For $3D$ images, a likelihood is required, which computes the similarity between the generated intensity distribution and the observed distribution in the image. Cauchy and Gaussian distributions which are the most used distribution for image intensity have been reported for approximating CT and MR image intensity distributions. For example, the MRI intensity has been reported to follow a normal distribution (Gudbjartsson and Patz, 1995), while could also follow Cauchy distribution (Janani et al., 2015). Hence, Cauchy and Gaussian noises are used as prior likelihoods for similarity evaluation between generated intensity distribution and observed image intensity distribution.

Let x be a point from observed image I, the Cauchy likelihood for the point x is then given by:

$$p(x|\mathcal{F}_\theta(x)), \sigma) = \frac{1}{\pi\sigma} \frac{\sigma^2}{(I(x) - \mathcal{F}_{\theta_I}(x))^2 + \sigma^2} \tag{5.8}$$

The Gaussian distribution prior likelihood at x is given for target image I by:

$$p(x|\mathcal{F}_\theta(x), \sigma) = \mathcal{N}(I(x)|\mathcal{F}_{\theta_I}(x), \sigma^2) \tag{5.9}$$

where σ is the noise and $\mathcal{F}_{\theta_I}$ the intensity component of the deformation field $\mathcal{F}_\theta$. These likelihoods can be used in equations 5.2, 5.3.

5.3.3 2D image observation

To fit the DMFC-GPMs onto $2D$ images such as those from X-ray, a model projection transformation is required; that is to say, a transformation that maps a volume render onto image pixels. The model sample image pixel (I_θ) can then compared to the observed image pixel (I) in 2D image space. The projection transformation, $\mathcal{P}$ is defined as:

$$\mathcal{P} \colon \Theta \to \mathbb{R}^2, \quad \theta \mapsto \mathcal{P}(\theta) = \mathcal{P} \circ \mathcal{F}_\theta$$

where $\mathcal{P} \circ \mathcal{F}_\theta$ is a 2D image representation of the model deformation fields $\mathcal{F}_\theta$ projected (the projector $\mathcal{P}$) and it is defined as:

$$\mathcal{P} \circ \mathcal{F}_\theta : x \in \mathbb{R}^2 \mapsto I_\theta(x)$$

This function takes a parameter θ that represents a model instance (volume shape-intensity) and returns a $2D$ image. The same prediction process for $3D$ then holds for $2D$ images through replacing $\mathcal{F}_\theta$ with I_θ. For a model based on CT image data, the projector ($\mathcal{P}$)can be a digitally reconstructed radiograph (DRR).

Several DRR rendering approaches have been reported and a survey by Reyneke et al. (2018a) overviews many such approaches. The well-established rendering technique is implemented as the Beer-Lambert law using "direct volume ray-castin", which models the propagation of X-rays through matter (Max, 1995). This is accomplished using a volume rendering technique (from the field of computer graphics), known as "direct volume ray-castin" (Weiskopf, 2007). The intensity value of any given pixel of the simulated X-ray image is computed by aggregating sample points from the CT volume taken along a straight path from a predefined "ray-origin" to the current pixel. Thus, the volume integral describing the Beer-Lambert law is approximated by numerical integration (a Riemann-sum approximation in the case of equidistant samples). For DRR generation throughout this thesis, a GPU-based implementation of the direct volume ray-casting technique available in the author's laboratory was used to obtain the simulated X-ray images.

5.4 Model validation: Shape-pose evaluation

This section focuses on the validation of the framework in explaining shape-pose features using a marginalised shape-pose model. To evaluate the performance of the proposed method, synthetic datasets were created with precisely defined modes of shape and pose

variation. The preliminary evaluation of model performance avoided the use of real biological data because such data typically exhibit complex shape and pose variation; and embeds prepossessing errors making it difficult to assess the true modelling framework characteristics. The use of synthetic data also avoids the need for acquisition of ground truth image data for bench marking as it uses analytically defined and precise ground truths.

Synthetic data were generated and consisted of surface mesh data of a "lollipop" as defined in (Gee and Treece, 2014). The lollipop had a "shaft", a "neck" and "head". All meshes had 6000 vertices with identical surface topology. The neck angle was uniformly distributed in the range $34 - 71$ degrees and the head length was uniformly distributed in the range $3.8 - 4.9$ cm. The shaft length was constant, so the variation expressed had two degrees of freedom (DOF). A dataset of 100 lollipops was created.

5.4.1 Shape-pose correlation evaluation

The proposed framework was evaluated for its capacity to explain shape and pose variability, that is, its capacity in explaining shape correlation between objects as well as shape-pose correlation, when they exist. This evaluation was done by comparing the correlation captured by the model to the one existing in the training. In addition, the leveraging of such a correlation for improving the prediction accuracy, is demonstrated.

To evaluate the potential of the proposed framework in modelling shape correlation between different objects, synthetic data with two objects per training examples were created. Only two objects were used for this experiment so that the correlation would be interpreted without ambiguity. Artificial joints were created using two lollipops per scenario. Each joint was composed of two lollipops with major axes for corresponding pairs of lollipops of r_1 and r_2, for object 1 and object 2, respectively. The span of the dataset of joints was created by varying r_1 and r_2 such that $(r_1, r_2) = (i/10, j/10), i = 16, \ldots, 30, j = 15, \ldots, 1$; creating a negative correlation between objects 1 and 2 (figure 5.1).

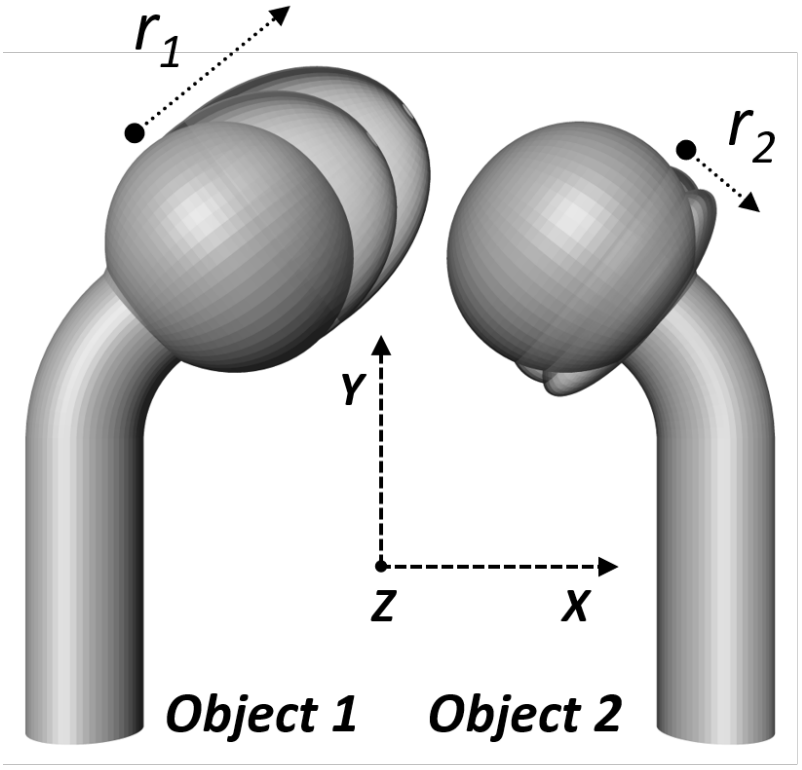

Figure 5.1 Illustration of shape variation of the lollipop data. Shape variation of the first and the second objects, the first object parameter r_1 increasing in the direction of the arrow and the second object parameter r_2 decreasing in the direction of the arrow.

For pose variability, joint motion was simulated. For each joint generated above, the second lollipop (object 2) was rotated relative to the first one (object 1) using the Euler's angle convention for describing rigid transformations (φ, θ, ψ). The second lollipop was moved in the xy-plane corresponding to Euler's angles $(0, \theta, 0)$; where only θ was varied by four angles ($\theta = \frac{1}{5}\pi, \frac{2}{5}\pi, \frac{3}{5}\pi, \frac{4}{5}\pi$) defining a motion (figure 5.2). Defining the same motion for each joint created a dataset where the shape was uncorrelated with motion because each pair of shapes performed the same movement. For shape-pose analysis, a Dynamic

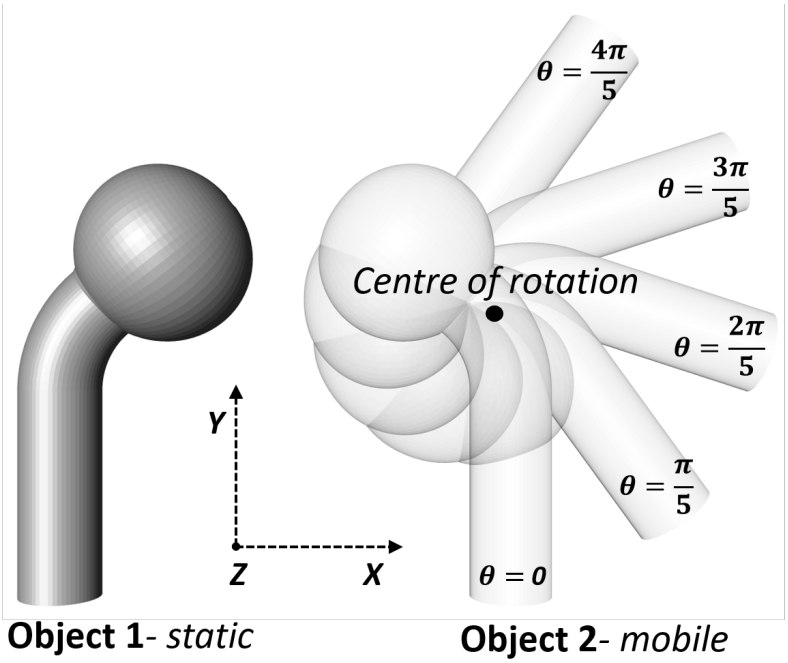

Figure 5.2 Illustration of pose variation of the lollipop data. An example motion of a lollipop joint, the fixed object (grey) and the moving object (white) at a different angles θ.

Multi-Object Gaussian Process Model (DMO-GPM) which is a shape-pose only model was built as a DMFC-GPM with no intensity variability, as described in the previous chapter (section 4.5.1).

A- Empirical validation of correlation between objects

The DMO-GPM framework was evaluated for its ability to correctly capture the correlation between object shape features when such a correlation is explicitly defined in the training

data. One would like to know whether the model's samples would also have the same correlation since it is hypothesised that any correlation existing in the training data should be captured by the model.

The lollipop joint dataset had a negative correlation between the first object shape parameter (r_1) and the second object shape parameter (r_2), that is, these parameters were inversely varying. This dataset was used to build a DMO-GPM. The correlation coefficient was -0.99.

Samples were drawn from the DMO-GPM and their correlation with the synthetically created examples was compared. The correlation was evaluated between r_1 and r_2 which are the shape quantities of the first and the second objects, respectively. The regression analysis of the samples from the DMO-GPM showed negative high correlation with a non-zero slope at 0.0001 significance level and a correlation coefficient of -0.99. This correlation is exactly what was prescribed in the training dataset. The samples from the model and those from training data showed the same regression curve (see Figure 5.3). This demonstrated that the DMO-GPM is indeed able to capture the correlation between objects of the joints with highest precision.

B- Prediction using static model

Prediction of premorbid shape is essential in clinical practices; statistical models are often used for premorbid estimation from observations (Fox et al., 2013). The use of statistical models for inference requires observations. When no observations are available, inference in medical image analysis becomes challenging. The DMO-GPM embeds many objects and may capture latent information that provides additional knowledge of one object from observation of another. Such knowledge can be leveraged to improve the prediction of an object with limited or partial observations (i.e with partial features available) or in the extreme, only from observations from other objects in the joint. However, for the latter case, the assumption is that the objects being modelled are correlated. In this section, the

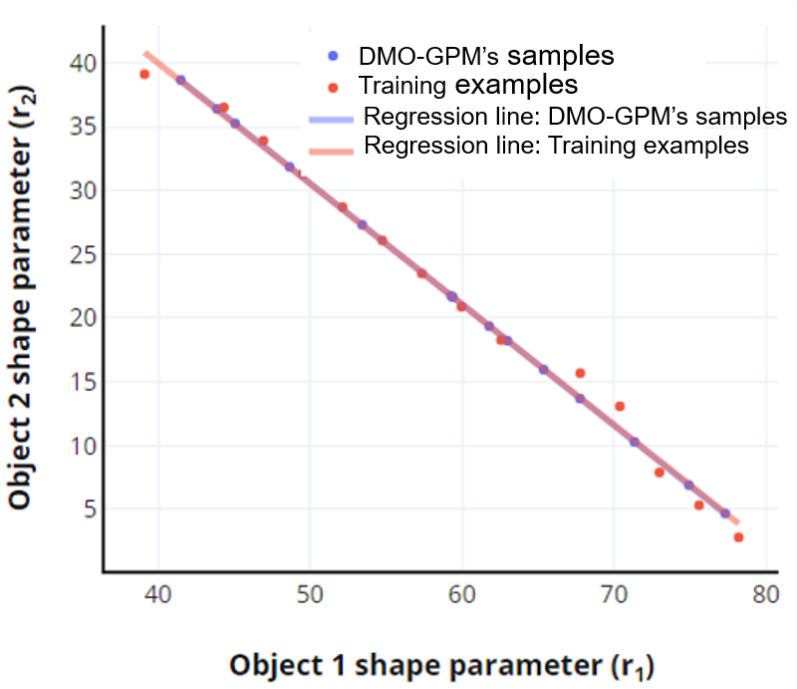

Figure 5.3 Regression analysis comparison of the DMO-GPM's samples and the training examples. The regression curve between r_1 (the first object's shape quantity) and r_2 (the second object's quantity).

capacity of the DMO-GPM framework in performing such a task is evaluated.

The DMO-GPM used in the previous section was marginalised to shape only (without pose variation) leading to what can be considered as a static model (as described in chapter 4, section 2.9 and figure 4.5). The marginalised model works by constraining the Gaussian process model to a specific region of the reference domain. By denoting A the shape region of the shape-pose domain (Ω) (as described in chapter 4, eq. 4.15), the marginalised shape only (static model) can be defined as:

$$p(A) := \mathcal{GP}(\bar{\mathcal{F}}, \mathcal{K})|_{A \subset \Omega}. \tag{5.10}$$

To evaluate these inference properties, four types of observations of the target structures (objects) were defined and used for predictions in the rest of the experiments (see table 5.1).

Table 5.1 Taxonomy of types of observations of target objects used for evaluating model prediction capacity

Type I	Full object to be predicted is available
Type II	Partial object available
Type III	Sparse anatomical features (e.g. landmarks) of the object available
Type IV	No direct information of the object is available

To test the hypothesis, five values of the parameters r_1 and r_2 were generated to obtain lollipop joint training examples. These parameters were generated between the minimum and the maximum values fixed for the training dataset. For each lollipop joint, the first object was predicted from observing the second object only. Conversely, prediction of the second object from observing the first object only was performed; that is to say, type IV observations for both predictions. Thus, the evaluation centred on ascertaining how well the DMO-GPM framework can leverage the existing correlation between a two object ensemble to predict one object given that the object itself was not observable but its counterpart was.

The results obtained for each object were compared to the ones obtained with type I observation (full observation of the object to be predicted). The Hausdorff distance (HD) and the root mean square (RMS) distance errors of the predictions are shown in table 5.2. For the prediction of the first and the second object with type I observation, the average RMS errors were 0.4 ± 0.06 and 0.22 ± 0.02 mm, respectively. For the prediction of the first and the second object with type IV observation, the average HD errors were 0.9 ± 0.22 and 0.82 ± 0.2 mm, respectively. The errors of the type I prediction are relatively small compared to the type IV prediction errors. However, these results obtained for type IV observations suggest show that any correlation embedded in the model provides additional knowledge that can be useful for the prediction. With such additional knowledge, the fitting process may lead to more accurate prediction.

It can be concluded that the DMO-GPM can capture any correlation existing between objects in the training dataset and leverage it to improve prediction.

Table 5.2 Prediction errors of the lollipop joint objects with Type I and IV observations using MCMC and DMO-GPM

	RMS (mm)	HD (mm)
Object 1 with Type I observation	0.002±2.2e-6	0.014±1.7e-3
Object 1 with Type IV observation	**0.4±0.06**	**0.9±0.22**
Object 2 with Type I observation	0.001±2.9e-6	0.01±0.001
Object 2 with Type IV observation	**0.22±0.02**	**0.82±0.2**

C- Uncorrelated and correlated shape-pose variation

So far only the shape correlation scenario has been presented, that is, the correlation that may exist between objects from a shape only perspective. However, as shown in the literature, joint motions are influenced by the shape of their articular geometries, which contributes to the regulation of their relative position (Breton et al., 2014; Jacq et al., 2008, 2009; Ibáñez et al., 2001; Jerbi et al., 2012; Gil et al., 2004). In this experiment, the ability of the DMO-GPM framework in capturing correlation that exists between shape-pose in the training data was tested. In addition, the model's capacity to minimise spurious shape-pose correlation was also evaluated; that is, testing whether the model is robust against exhibiting correlations that do not exist in the training data. To this end, the framework was first evaluated for its ability of capturing any correlation that exists in the training dataset. Second, the capacity of the framework in capturing existing correlation was evaluated.

To evaluate the potential of the framework in modelling shape and pose together without creating artificial correlation, the lollipop dataset with two objects per training example was used. The dataset did not have shape pose correlation since each joint had the same motion. A DMO-GPM was built for this dataset.

The DMO-GPM with no shape correlation was sampled. About 98.2% of the total variation was accounted for the first two principal components (PCs) which was the expected behaviour. Shape-only variation was explained by the first principal component (PC) without any pose variation and accounted for 87.7% while pose-only only was explained

by the second and accounted for 10.5% of the total variation. The variation explained by the fist and second PCs is shown in figure 5.4, where samples from from -3σ to $+3\sigma$ standard deviations around the mean are shown. For the first PC, the pose was fixed at -12 *rad* and the shape descriptors (i.e. r_1 and r_2) varied from from 41.5 to 77.3 *mm*, and from 38.7 to 4.7 *mm* for object 1 and object 2, respectively. This variation corresponds to the correlation existing in the training dataset. The second lollipop's head length (r_2) decreases in the first PC, it decreases up to its "neck" where it starts being unrealistic. This unrealistic behaviour is likely due to the large sampling range (up to $+3\sigma$) as indicated by * in figure 5.4. Shape parameters are fixed at the mean shape in the second PC, the mean shape parameters are $r_1 = 59.4$ *mm* and $r_2 = 21.6$ *mm* for object 1 and object 2, respectively. The pose variation represents the abduction motion in the training dataset, as can be observed that the pose parameter θ varies from -0.2 to -1.5 *rad*. The fact that two different PCs explain the shape and pose variation mirrors the shape and the pose variation across the training dataset, hence, confirms that the model represents the training dataset.

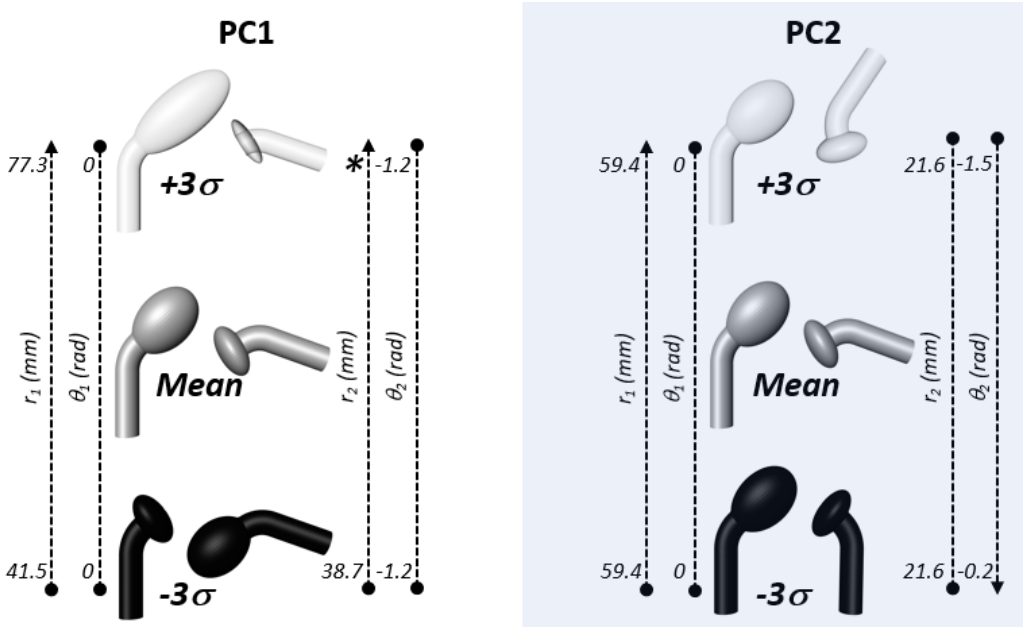

Figure 5.4 First and second PCs of the DMO-GPM with no correlation between shape and pose variation. The first PC accounts only for shape variation while the second PC accounts for pose variation only.

In the experiment above, the capacity of the framework to minimise spurious shape-pose correlation was evaluated. In contrast to the above experiment, the next experiment evaluated the capacity of the framework in capturing existing shape-pose correlation. Synthetic training data in which pose variation was constrained by shape was used to build the model. The experiment evaluated how well the model follows the training dataset's behaviour in shape-pose correlation.

A high shape-pose correlation was induced in the training dataset, creating a dataset where the motion is constrained by the shape. This was done by simulating the same movement as above but the maximum angle was inversely proportional to r_2 ($\theta_{max} = -\frac{3}{4r_2}\pi$), thus creating a correlation between shape and motion. A DMO-GPM with the pose constrained by the shape was built and sampled to evaluate shape-pose correlations.

About 99% of the total variation is accounted for by the first two PCs. Of this, 92.3% is accounted for by the first PC, which explains shape and pose variations. The second PC only accounts for 6.7% of the total variation and explains pose-only variation as shown in Figure 5.5. The first PC indicates that shape parameters r_1 and r_2 vary from 44.2 to 80.1 mm for object 1 and from 36.2 to 1.6 mm for object 2. The second shape parameter variation is inversely proportional to the pose parameter; representing the correlation in the training dataset. The second PC explains only pose variation, where the pose parameter θ varies from -0.01 to -1.4 rad, the shape parameters are fixed at the mean shape parameters which are 59.4 mm for object 1 and 21.6 for object 2. The shape-pose correlation in the first PC corresponds to the correlation between shape and pose variation in the training dataset.

As expected, with the shape correlation between objects and the correlation between shape and pose, the DMO-GPM framework showed its capacity to capture correlations prescribed in the synthetic data. It is important to note that such correlations in practice may not exist for some training datasets or for some anatomical structures. However, to the degree that correlations exist, the framework would be able to capture this and leverage them for

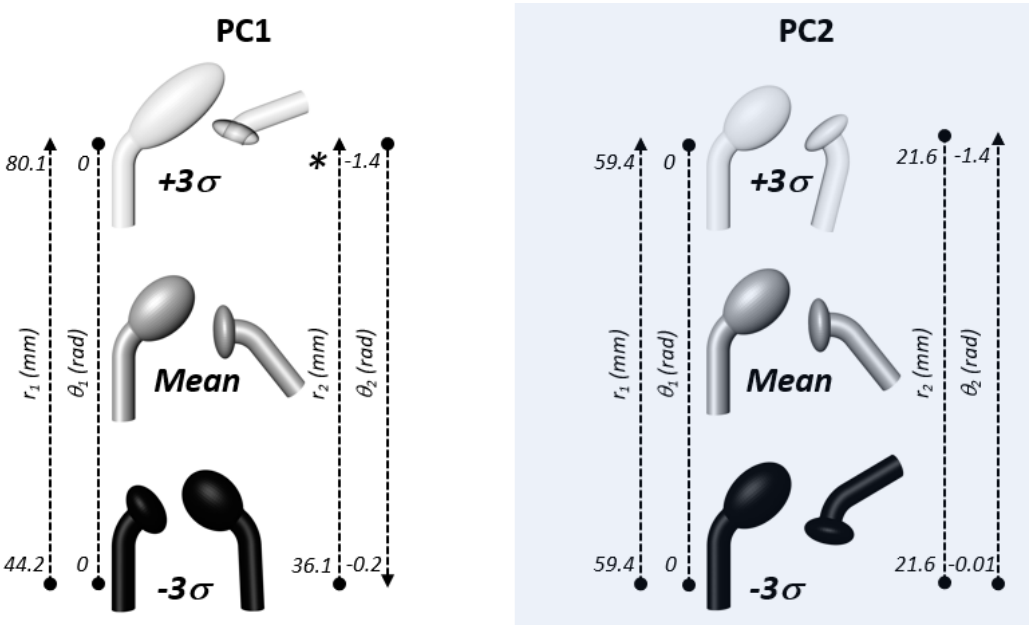

Figure 5.5 First and second PCs of the DMO-GPM. The first PC accounts for both shape and pose variation explaining the correlation between both feature sets. The second PC accounts for pose variation only.

prediction tasks.

5.4.2 Shape-pose model attributes and their validation

In the previous experiments, only shape correlation was evaluated which does not quantitatively tell us how good the framework is in representing the training dataset nor its capacity in generating valid unseen data. Moreover, only two objects were used in the above evaluations while the framework may be applied to any number of objects.

In this section, the DMO-GPM regression and permutation modelling attributes are demonstrated. The regression model constrains the DMO-GPM to specific observations and the permutation model is derived from the training dataset extended by permutating training poses (as detailed in chapter 4, section 4.5.3 and 4.5.4). The DMO-GPM framework's performance in generalising the training dataset as well as in explaining the training dataset, is also presented.

For these experiments, three objects were used to generate the training dataset. The

lollipop joints were modified such that each joint was composed of three lollipops with major axes for corresponding pairs of lollipops of r_1, r_2 and r_3, for object 1, object 2 and object 3, respectively. The span of the dataset of joints was created by varying r_i as $\{(r_1, r_2, r_3) = (r_i, 31 - r_i, 17 - r_i), i = 1, \ldots, 15\}$ creating a shape correlation between them as shown in figure 5.6 (left). For joint motion, for each joint generated above, the second lollipop (object 2) was rotated relative to the first one (object 1), and the third relative to object 2 using the Euler's angle convention for describing rigid transformations (φ, θ, ψ). The second lollipop was moved in the yz-plane through four angles $(\theta = \frac{1}{5}\pi, \frac{2}{5}\pi, \frac{3}{5}\pi, \frac{4}{5}\pi)$. The third object was moved in the xy-plane with four angles $(\theta = \frac{1}{2}\pi, \frac{1}{3}\pi, \frac{2}{9}\pi, \frac{1}{9}\pi)$ defining a Euclidean motion correlation with respect to the second one as in figure 5.6 (right).

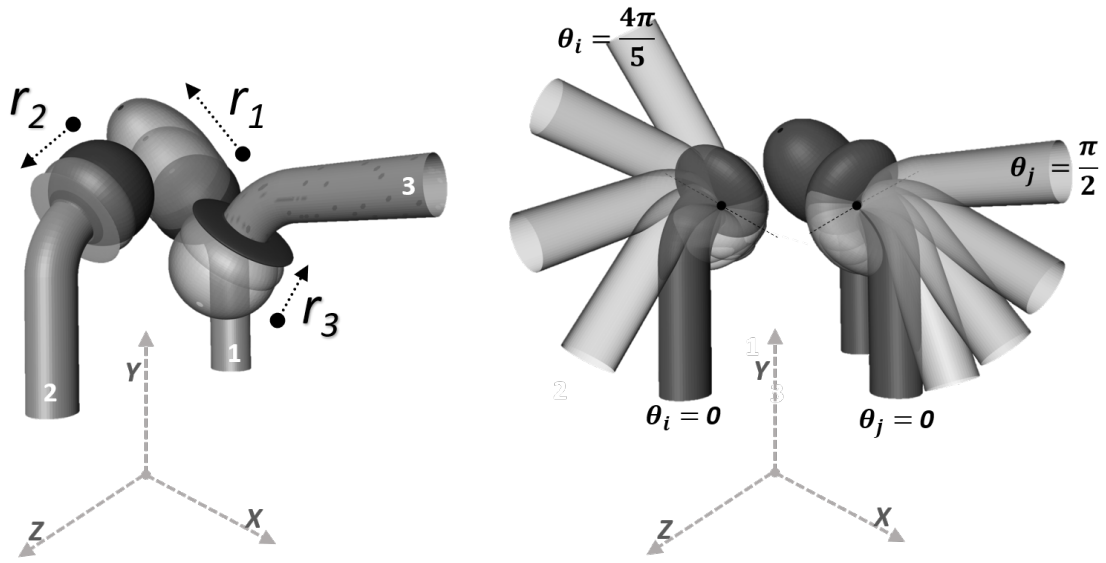

Figure 5.6 Shape-pose model training data. Left: Defined shape variability. Right: Defined pose variability.

A- Regression analysis of the lollipop joint

Assuming that prior knowledge of the shape is available, then a useful task could be to estimate the pose or the motion of the joint, which will lead to a model of the motion. For example, a patient-specific motion model can be estimated from the observation of its shape. To demonstrate this property of the DMO-GPM framework for shape-pose

analysis, different shape features were used to compute the associated posterior model (pose only model) from the model built with no shape-pose correlation.

After computing the latent space from the data, and making inference on observations, a posterior model (dynamic multi-object Gaussian process regression model) was obtained. The regression model was built from the DMO-GPM of the lollipop data with three objects and it was defined using:

$$\mathcal{F}|x_O : \mathcal{F}_O \sim \mathcal{GP}(\mu_O, k_O) \tag{5.11}$$

where μ_O and k_O are conditioning mean and kernel respectively; x_O is the observation and $\mathcal{F}_O$ the deformation field mapping them to the target ones as described in chapter 4 (section 4.5.4).

There were two posterior models, first, the observations were the objects with the shape parameters $(r_1 = 16, r_2 = 15, r_3 = 2)$; and second, the observations were the objects with shape parameters $(r_1 = 30, r_2 = 2, r_3 = 16)$.

Figure 5.7 shows the samples from the DMO-GPM and the posterior DMO-GPM. The samples were generated around the mean (black) from -2 to $+2$ standard deviation for the first, the second and the third PC. It can be seen that the posterior model variation corresponds to the pose variation of the DMO-GPM, that is to say, the variability explained by the first and the second PC of the posterior models corresponds to that explained by the second and the third PC of the DMO-GPM. The Gaussian noise of the observation was 1.0 and the same variability was obtained with two joint shapes $(r_1 = 16, r_2 = 15, r_3 = 2)$ and $(r_1 = 30, r_2 = 2, r_3 = 16)$, both at the reference pose. This shows that our regression model works as expected, allowing the DMO-GPM to be constrained to observation-specific model for getting more insight into desired features.

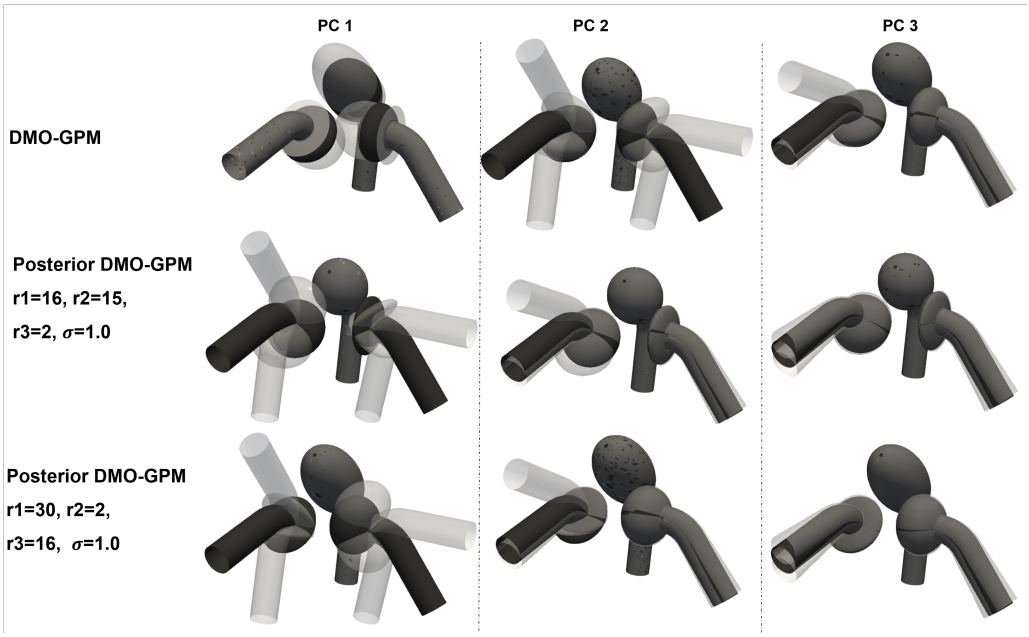

Figure 5.7 DMO-GPM and posterior models samples from the first, the second and the third PC. Top: samples from the DMO-GPM, Middle and bottom: Posterior model samples with fixed shape parameters (r_1, r_2, r_3) and Gaussian noise (σ).

B- Permutation analysis of the lollipop joint

One of the difficulties often encountered in medical imaging is the lack of availability of data for training models. This is especially true for studies exploring pose variability, where available datasets may have limited poses. To overcome such a limitation permutation models from the framework may be used. As described in chapter 4 (section 4.5.3), permutation models works by linking poses of all objects in training data to each object. Said another way, each training example is made aware of the poses of other training examples.

To validate the permutation model, a training dataset was created with 6 lollipop joint examples. Each example had a specific pose, that created a high correlation between shape and pose. The shape and poses were selected to cover the shape and the pose range (shape and pose range of the synthetic data described above).

Samples from the DMO-GPM model were drawn. It can be seen that for the first and

second object, the pose angle increases as the shape parameter increases as illustrated by the first PC in figure 5.8 (top). It clearly shows that only specific shapes can be at some specific poses, which may not be true and could be due to the lack of data. As an alternative, the permutation models were proposed as presented in chapter 4 (section 4.5.3). They allow a shape to be moved by other example rigid transformations (poses). The permutation model of the same training dataset was built, and the figure 5.8 (bottom) shows samples from that model. It can be observed that the permutation model successfully recovered the same pose variability as for the model with full training motion as shown in figure 5.7 (top). This proves that there is a potential of permutation models explaining pose variation of a subject beyond its observed pose (pose in the image), however, application to real data would require some knowledge about poses of similar shapes to be permuted.

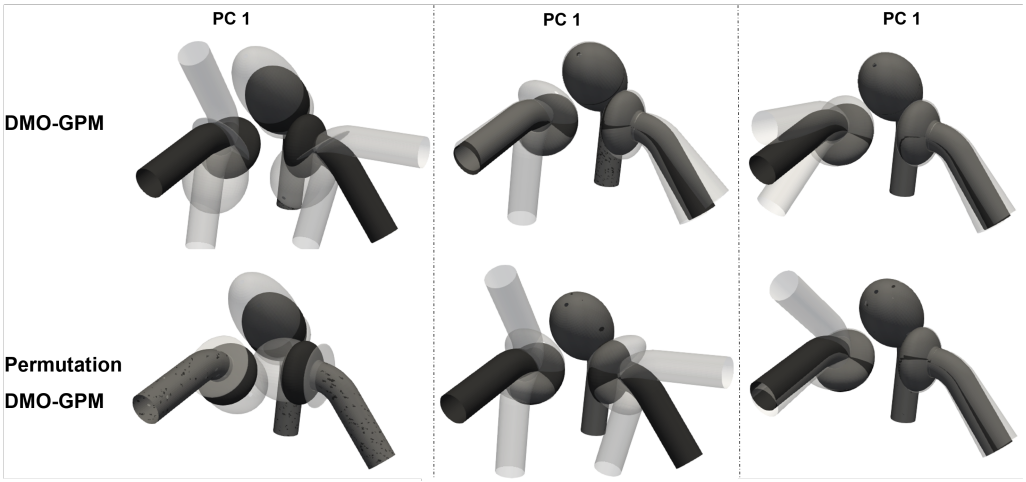

Figure 5.8 DMO-GPM and permutation model samples from the first, the second and the third PC. Top: samples from the DMO-GPM. Bottom: permutation model samples.

C- Performance comparison with the state-of-the-art methods

The standard validation of statistical models requires evaluation of specificity and generalisation capacity. Thus, both the specificity and generalisation capabilities of models built for the framework were compared to those of models built using current state-of-the-art methods. Prior to the standard validation of the framework, however, its capacity for

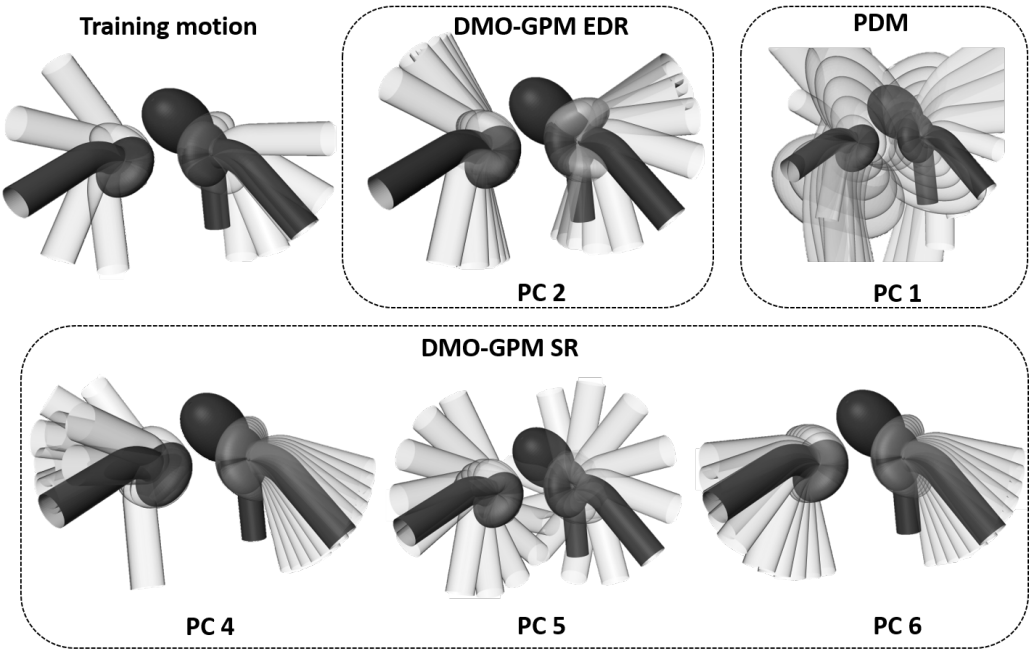

Figure 5.9 Comparison of pose variations with the training data. Top-left: Pose training data. Top-middle:DMFC-GPM with EDR (pose regression) Top-right: PDM. Bottom: DMO-GPM with SR (pose regression).

capturing the correct pose from training examples was compared to that of the state-of-the-art methods. The previously described dataset of lollipop joints with no shape-pose correlation was used for these experiments.

C.1- Pose capturing performance

The performance of the models in explaining the pose variability was quantitatively compared through pose regression using EDR (our model), or SR and a point distribution model (PDM).

Figure 5.9, from the top-left to the top- right, shows the DOF of the training data, the pose regression model with EDR and the PDM, respectively. At the bottom, the pose regression model with SR is shown. Our model (DMO-GPM with EDR) captures the training motion with one PC, while the model with SR needs up to 3 PCs to explain the motion (although motion in only in one axis in the training data). As expected, PDM led to unrealistic shape variation.

C.2- Specificity and Generality

The robustness of the proposed model using EDR was evaluated by quantitatively comparing its specificity and generalisation measures to the standard representation (SR) models. First, the specificity was evaluated by measuring the capability of the model to explain variation within the training data, in terms of the number of PCs used. The DMO-GPM was fitted onto 10 training examples, which included shape and pose variation. The fitting was performed using the Markov chain Monte Carlo (MCMC) method as in (Schönborn et al., 2017).

Figure 5.10 and 5.11 shows the box plot of the generality and the specificity results using the Hausdorff and RMS (figure 5.11) distances, respectively. The surface-to-surface HD (figure 5.10) and the RMS errors (left column) are shown in terms of the number of PCs for the first, the second and the third objects. The proposed model (DMFC-GPM with EDR) outperformed the one with SR for specificity. Moreover, the generalisation ability was evaluated by measuring the capability of the model in explaining examples outside the training data. Again, DMFC-GPM with EDR outperformed the one with SR as observed from figure 5.10 and 5.11. This illustrates the robustness of the proposed method in representing $3D$ Euclidean motion as well as in explaining associated shape variations.

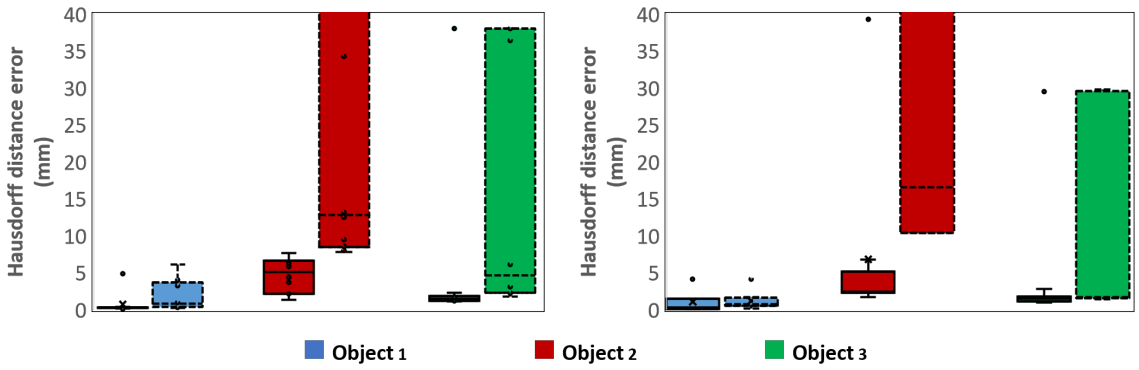

Figure 5.10 Quantitative comparison of the DMFC-GPM with EDR (our model) and DMFC-GPM with SR using HD. Left: Specificity performance using DMFC-GPM with EDR (solid line) and SR (dashed line). Right: Generalisation performance using DMFC-GPM with EDR (solid line) and SR (dashed line). Dots represent outliers.

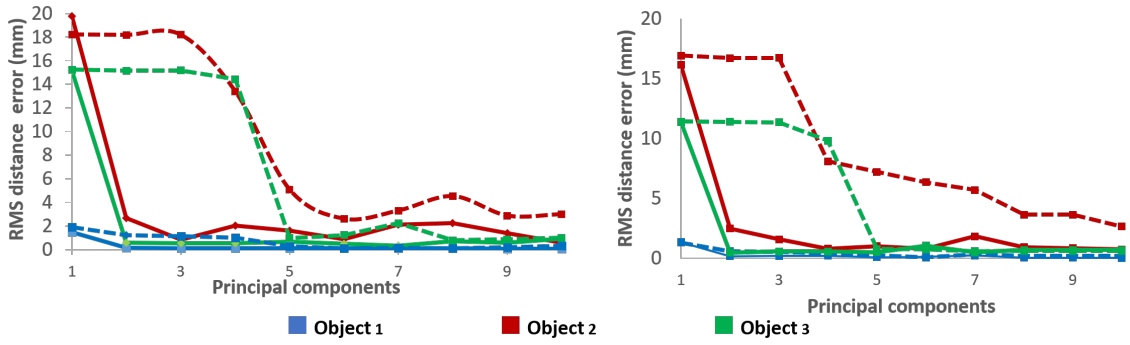

Figure 5.11 Quantitative comparison of the DMFC-GPM with EDR (our model) and DMFC-GPM with SR using RMS distance. Left: Specificity performance using DMFC-GPM with EDR (solid line) and SR (dashed line). Right: Generalisation performance using DMFC-GPM with EDR (solid line) and SR (dashed line).

5.5 Model validation: Shape, pose and intensity evaluation

In contrast to the experiments in the previous section where only the marginalised model (DMO-GPM) was used, the DMFC-GPM framework's ability to explain all feature classes (shape, pose, and intensity) through sampling was evaluated along through model fitting. Its ability for leveraging all embedded feature classes and their correlations to improve the prediction was also evaluated.

The synthetic dataset used for this evaluation consisted of three lollipop joints containing in-correspondence joint lollipop surfaces with known prescribed shape and pose variation. From these in-correspondence lollipops, several lollipop volumetric mesh joints were obtained through the interpolation of the surface mesh correspondence deformation fields as described in chapter 3 (section 3.3).

The lollipop volumetric meshes obtained were rendered to embed intensity. This was done by defining an image domain using a bounding box of the joint with a pixel value assigned

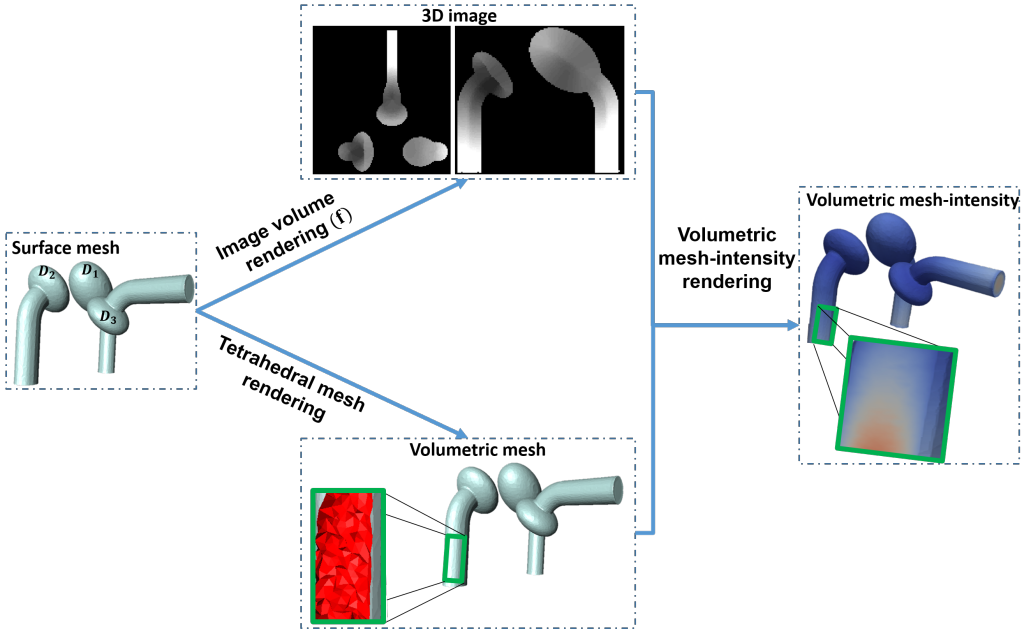

Figure 5.12 $3D$ image render from a joint lollipop and the joint volumetric shape-intensity generation. First: the lollipop joint is rendered to $3D$ image and also transformed into joint volumetric shape (tetrahedral meshes in this case, shown in red). Finally, the intensity distribution of the image rendered is associated with the joint volumetric shape to obtain the joint volumetric shape-intensity.

to each point using the function defined below: 5.12.

$$f(x) = \begin{cases} \|x - x_1\| & \text{if } x \in D_1 \\ \|x - x_2\| & \text{if } x \in D_2 \\ \|x - x_3\| & \text{if } x \in D_3 \\ 0, & \text{otherwise.} \end{cases} \tag{5.12}$$

where D_1, D_2, D_3 are the shape domain of the first, the second and the third object, respectively; and x_1, x_2 and x_2 are the selected reference for the first, the second and the third lollipops, respectively. Figure 5.12 shows an example of image rendering from a lollipop joint and the resulting volumetric shape-intensity joint.

5.5.1 Model sampling and correlation capturing

This section presents the evaluation of the DMFC-GPM framework in explaining within and between feature-class (including intensity) correlation. As previously done for shape-pose evaluation, the cross-correlation is evaluated here, where the correlation is evaluated between feature-classes.

One hundred samples of the DMFC-GPM built from the lollipop data were randomly generated. Between object shape correlation, shape-intensity correlation as well as between object pose correlation of generated samples were evaluated. The results were compared to the known correlations from the training examples.

The results show that the DMFC-GPM framework properly explains the correlation from the training data with the correlation percentage being approximately the same in both cases (training examples and model samples). Samples from -2 and $+2$ standard deviations around the mean were generated for the first and the second modes of variation of the model. The first PC explains 86% of the total variation which accounts for shape and intensity variation. The second PC explains 9.3% of the total variation, accounting for the joint motion. This was expected because there was a correlation between shape and intensity distribution, hence, both accounted for high variability across the training data.

Furthermore, joint shape samples were generated along with volumetric shape-intensity joints as observed in figure 5.13. A volume renderer was also added in the framework to generate $3D$ images as well as one $3D$ to $2D$ projector to generate DRRs as observed in figure 5.13.

To evaluate the cross-correlation between different feature types and objects, 100 samples were randomly generated from DMFC-GPM of the lollipop data. Since the intensity features were assigned as function of shape (r_1, r_2, r_3), the intensity quantities (d_1,d_2, $d+3$) were computed using the same points used to compute shape quantities. For the pose quantities ($\theta_1, \theta_2, \theta_3$), the first Euler parameters were used which describe the motion sim-

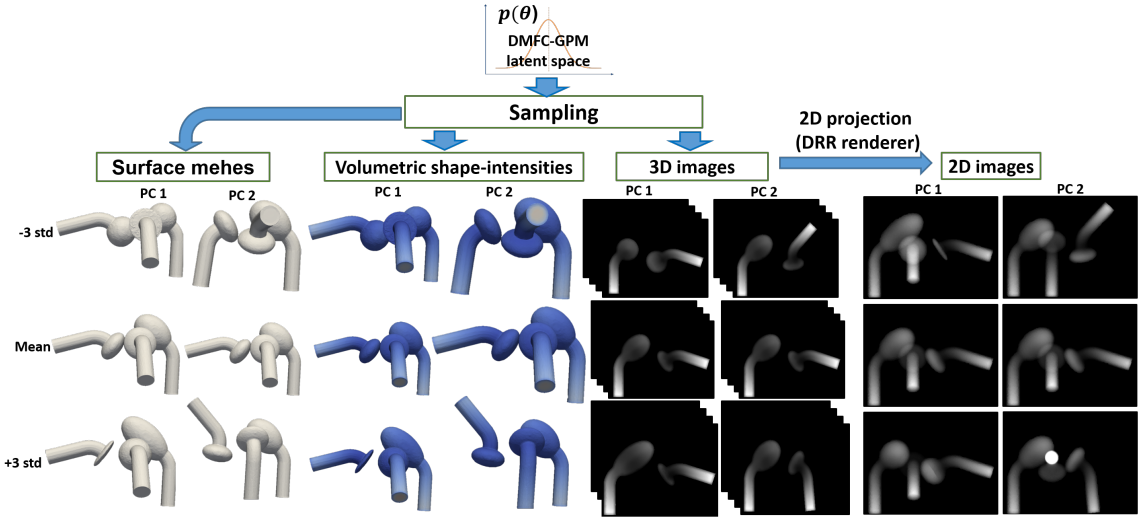

Figure 5.13 Sampling of the lollipop DMFC-GPM. Top: the DMF-GPMs ($p(\theta)$ is the probabilistic formulation of the model). Bottom: shape, volumetric shape intensity and image modality sampling. From left to right the first and the second PC sampling of shape joint, volumetric shape-intensity joint, $3D$ image rendered and digitally reconstructed radiographs. Samples from -3 and $+3$ std around the mean.

ulated in the training dataset. Between object shape correlation ($r_1 vs. r_2$ and $r_2 vs. r_3$), shape-intensity correlation ($r_1 vs. d_1$, $r_2 vs. d_2$ and $r_3 vs. d_3$), as well as between object pose correlation ($\theta_2 vs. \theta_3$) of the DMFC-GPM samples, were evaluated. The correlation coefficients from the DMFC-GPM's samples were compared to the ones from the training examples.

The results are shown in table 5.3. It can be observed that the DMFC-GPM framework properly explains the correlations from the training data with the correlation percentage being approximately the same in both cases. The slight difference between the shape and intensity for the first object could be attributed to random sampling of the model.

Table 5.3 Correlation coefficients between and within feature classes. Comparison of the model samples and the training examples correlations. d_1, d_2, d_3 are the intensity information for the first, the second and third object, respectively. θ_2 and θ_3 are the motion parameters of the second and the third objects, respectively.

	r_1 vs d_1	r_2 vs d_2	r_3 vs d_3	θ_2 vs θ_3	r_1 vs r_2	r_2 vs r_3
Training	46%	92%	96%	60%	99%	93%
DMFC-GPM	**56%**	**92%**	**93%**	**53%**	**98%**	**91%**

5.5.2 Shape, pose and intensity modelling validation

To assess the DMFC-GPM framework for predicting shape-pose and intensity, its performance in prediction from observations was compared with state-of-art methods using common synthetic data and a common model-fitting approach (MCMC sample-based fitting). On one hand, the DMFC-GPM framework was compared with the GPMM-based active shape model (ASM) (Cootes et al., 2001) and on the other hand, the MFC-GPM framework (DMFC-GPM with EDR) was compared the DMFC-GPM with SR.

A- EDR vs SR

Similar to the DMO-GPM framework specificity and generalisation quality evaluated in the shape-pose analysis section, here, the specificity and the generalisation performance of the DMFC-GPMs with EDR was evaluated. The quality of specificity and the generalisation of the DMFC-GPM with EDR was compared to DMFC-GPM with SR.

For testing, a synthetically created dynamic $3D$ image dataset (joint volumetric shape-intensity at various poses) was used. The images that were included in the training dataset were predicted with both models (DMFC-GPM EDR and DMFC-GPM SR) to compare the specificity performance. The images had six poses, that is, each image of the six images displayed object 2 and object 3 at specific poses. To implement the DMO-GPM with SR, rigid transformations within the DMO-GPM setting were represented by Euler's parameters, that is, rotations were represented by Euler's angles and the translation by $3D$ Euclidean vectors.

Figure 5.14 (bottom) shows the intensity-based RMS errors of the predicted and observed intensity distribution, where the model with EDR outperforms the one with SR.

For generalisation performance, a dynamic $3D$ image dataset with joint poses not included in the training was generated and predicted using both models. Figure 5.14 (top) shows the RMS errors and for each object, the model with EDR outperforms the one with SR.

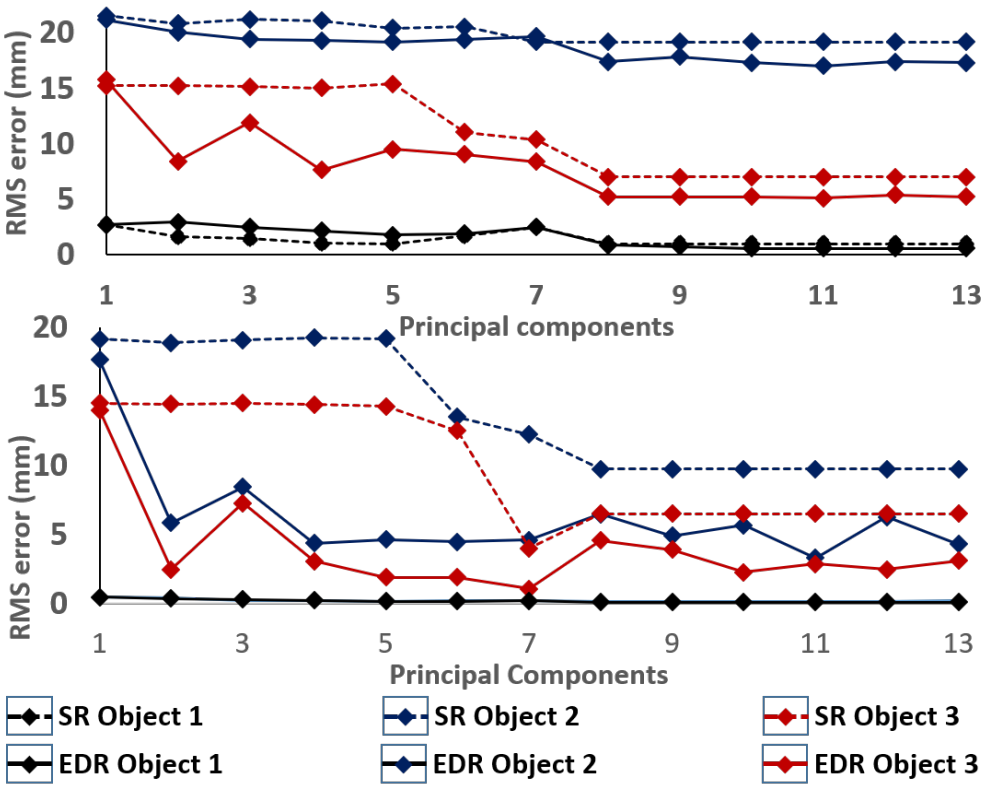

Figure 5.14 Generalisation (top) and specificity (bottom) of the DMFC-GPM with EDR and the DMFC-GPM with SR. The intensity distribution based RMS errors of the prediction using EDR (solid line) and the one using SR (dashed line) for first (black), second (red) and third (blue) shape-intensity.

B- DMFC-GPMs vs ASM

Statistical shape models are often used to perform segmentation through ASMs. Objects are segmented by using the SSM simultaneously with the average intensity distribution of the internal structure of the object. The average intensity may not be sufficient to delineate the boundary of the object and also such a process does not provide any information about the internal structure. The DMFC-GPM framework based latent space embeds the intensity features as well as their correlation with object shapes. Having such information embedded in the DMFC-GPM framework could be leveraged to improve the prediction accuracy of the object of interest. Both the GPMM-based ASM and DMFC-GPM were used to predict an individual object and the results were compared.

A test dataset was created consisting of 6 lollipop joints, for each joint. A $3D$ image was rendered with a missing part on the object to be predicted leading to a set images with missing structures. The DMFC-GPM and individual ASM were fitted onto the test $3D$ images to predict complete objects including the missing parts. The RMS and the HD errors of the shapes predicted using DMFC-GPM and those predicted using GPMM-based ASM were compared.

To implement the GPMM-based ASM, first, the GPMM-based SSM was developed from in-correspondence mesh surfaces for each object. Second, objects were aligned with their corresponding $3D$ image volumes and the intensity profile was generated at each vertex of the mesh surface. Finally, the average intensity profile was computed. For each individual object, the SSM and the average intensity profile were simultaneously used within the fitting setting for the prediction.

Figure 5.15 and 5.16 show the surface to surface HD and RMS distance errors of the prediction with DMFC-GPM and ASM. It can be seen that the prediction errors of the DMFC-GPM framework (solid line) are smaller compared to the GPMM-based ASM ones (dashed line). This shows that the DMFC-GPM outperforms ASM on this prediction task. This performance gains from using the DMFC-GPM could be attributed to the additional knowledge obtained from the correlation between objects embedded in that type of model.

5.5.3 Intensity-based motion prediction of the lollipop joint

In this experiment, the ability of the DMFC-GPM framework to leverage shape-intensity correlations to improve the prediction accuracy was evaluated. As shown in the feature class correlation, any correlation existing between shape and intensity can be captured by the DMFC-GPM framework. The framework's ability to be used for predicting an object's interior intensity profile from the object's shape was evaluated.

To this end, lollipop joints from $3D$ images were predicted. These images consisted of

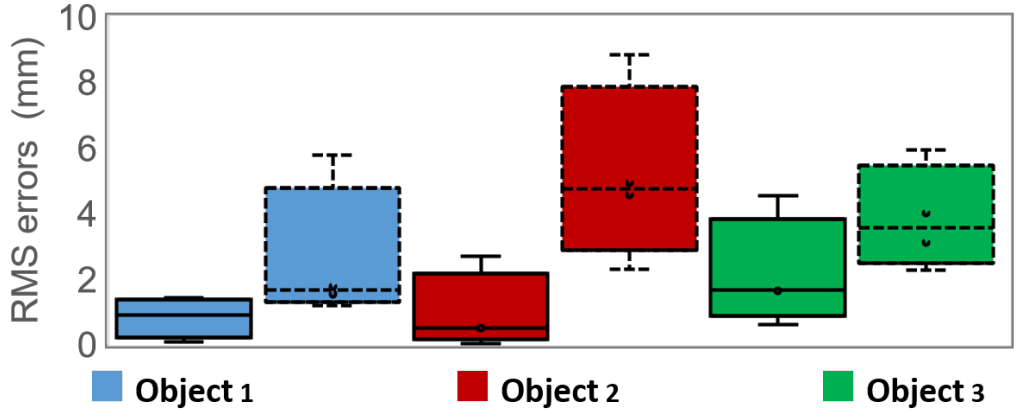

Figure 5.15 RMS distance errors for the prediction from $3D$ images using DMF-GPMs (solid line) and ASMs (dashed line). The error for the DFCM-GPM and the ASM for the first (blue), second (red) and third (green) objects.

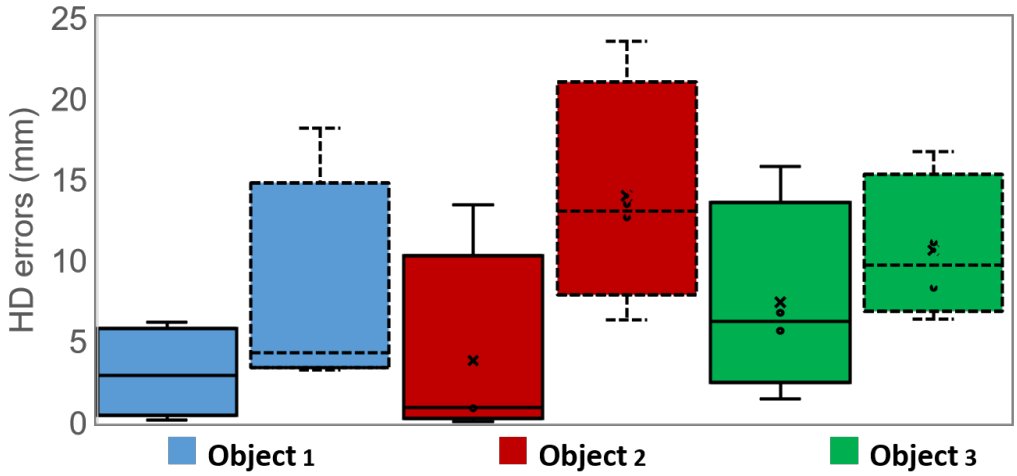

Figure 5.16 HD errors for the prediction from $3D$ images using DMF-GPMs (solid line) and ASMs (dashed line). The error for the DFCM-GPM and the ASM for the first (blue), second (red) and third (green) objects.

lollipop joints at various relative orientations to each other that also varied in shape. To evaluate the efficiency of the feature correlation accuracy in prediction, the prediction was done with three categories of observations. First, it was assumed that one observes the intensity distribution (that is the $3D$ image) which is the best-case scenario in practice. Second, it was assumed that one observes only the outer surface of the objects of interest (surface mesh); this can also be landmark points. Third, the observation of the volumetric mesh (outer and internal structure of the lollipops) was considered.

Table 5.4 shows the intensity-based normalised RMS of the predicted and the ground truths lollipop intensity distributions. The prediction errors were approximately the same in the three scenarios. This performance is attributed to the shape-intensity correlation that is embedded in the model as shown in table 5.3 above and meets the expected outcome. Figure 5.17 shows two successfully predicted lollipop joint examples from image volumes where displayed lollipops had different feature-class parameters.

Table 5.4 Normalised RMS error (mm) of the image volume estimation from different observations.

Observation	object 1	object 2	object 3
$3D$ Image	$\mathbf{0.005 \pm 0.002}$	$\mathbf{0.01 \pm 0.01}$	$\mathbf{0.01 \pm 0.005}$
Surface mesh	0.006 ± 0.002	0.01 ± 0.01	0.01 ± 0.003
Volumetric mesh	0.005 ± 0.005	0.01 ± 0.01	0.01 ± 0.002

5.6 Conclusion

In this chapter, model fitting to obtain the model parameters that optimally represent an observation was presented. The MCMC-based fitting method was adopted for multi feature-class fitting that is robust in finding global maxima/minima. Likelihoods were explicitly derived for different types of observations including point clouds and $3D$ images as well as a fitting initialisation process that was more efficient and reliable. The DMFC-GPM was critically evaluated and validated against multiple tasks through controlled experiments using the presented model fitting methods.

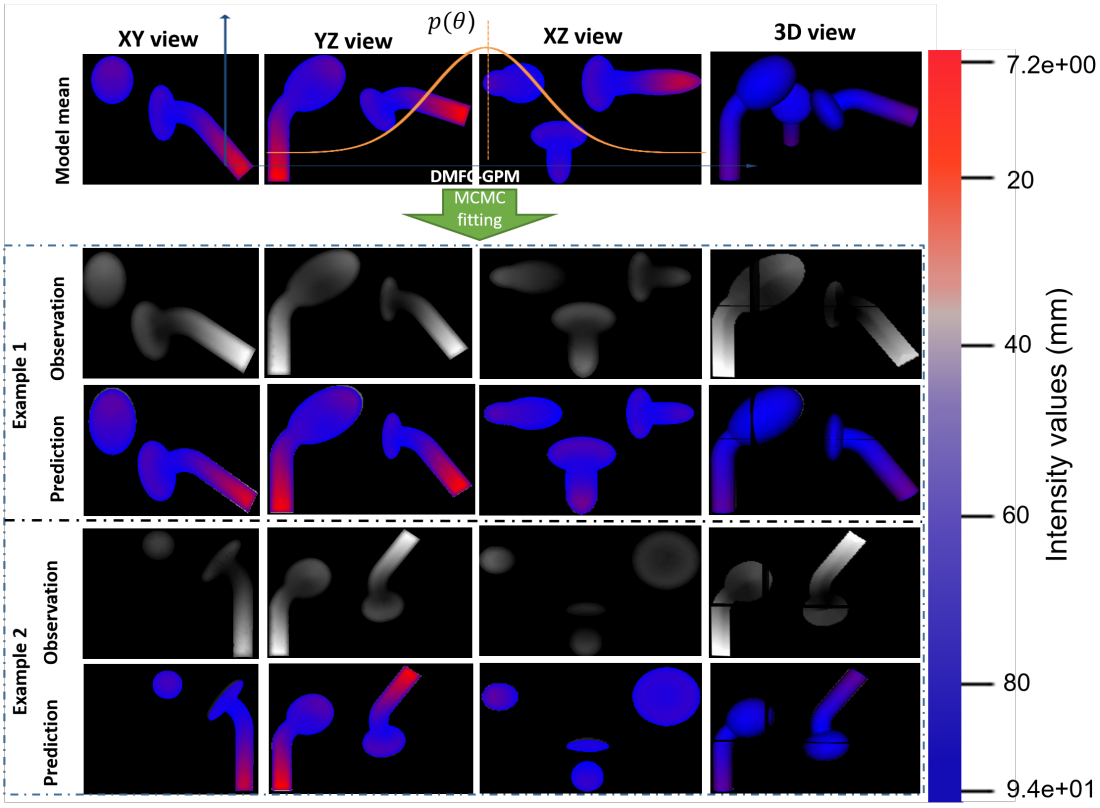

Figure 5.17 Prediction of the lollipop volumetric shape-intensity joint from image volumes. Model mean: represents the initial model sample in the fitting process. Observation: represents the image volume displaying the lollipops with various shape, pose and intensity distributions to be predicted (Observation 1 for the first example and observation 2 for the second one). Prediction: the image volume (observation) with the predicted structures (prediction 1 for the first example and prediction 2 for the second one) in red-blue colour (colour map representing the predicted intensity distribution).

It was demonstrated that the DMFC-GPM framework can incorporate more than two objects in a compact representation using Gaussian process. This means that the developed models provide a continuous analysis space with variations of shape, intensity and spatial dynamics. Different model attributes were evaluated including regression analysis, marginalisation and permutation analysis. Marginalising and regressing DMFC-GPMs allow us to analyse and/or predict shape-only or pose-only or combined extractions/generations. The capacity for the framework for modelling shape-pose variations was evaluated. The results with synthetic data revealed the efficiency of the DMFC-GPM framework for explaining shape-pose correlations across a given population of dynamic multi-objects. The DMFC-GPM framework was able accurately explain the shape correlation between objects and correlation between shape and spatial variation. This presents an advantage of multi-object modelling over single structure SSMs. Furthermore, DMFC-GPM also preserves the optimal representation of $3D$ motion using the EDR approach as illustrated on the non-compact lollipop dataset.

In addition, the capacity of the framework for explaining shape, pose and intensity distribution was experimentally evaluated. Its efficiency in explaining within and between feature class correlation with respect to the training example correlation was demonstrated. The DMO-GPM with EDR was compared with the one using standard metric (SR) and its robustness during prediction from $3D$ point clouds with respect to the object's dynamics was demonstrated. The DMFC-GPM framework further was compared to the ASM and DMFC-GPM showed its superiority for predicting single object contours. Its robustness in predicting multi-structures from image volumes displaying such multi-structures at various relative spatial orientations was also demonstrated. These important characteristics and attributes of the DMFC-GPM framework indicate potential for multiple applications ranging from pre-surgical planning to population-based implant designs.

Chapter Six

Demonstration of the framework for analysing shape-pose of human joints in medical images

6.1 Introduction

So far a framework for the analysis of multiple structures using a latent space has been presented. This latent space encodes the variability of feature-classes of a given training dataset, including shape, pose and intensity. Moreover, unlike statistical shape-pose models (SSPMs) (see section 2.2.4) which operate in a discrete domain, the presented framework works in a continuous domain. The potential of the framework has been illustrated using controlled synthetic data. However, there is a difference between synthetic data and real data. Medical image data are more complex in terms of shape-pose variability, leading to more challenges with regard to automated analysis. The complexity of clinical data comes from data that are often noisy and incomplete. This chapter describes how the presented framework can be applied to medical image data for shape-pose analysis of human joints. The global aim is to allow clinicians to load a set of observations such as $3D/2D$ landmark points or $3D$ surface meshes for automated prediction or analysis of the structures of interest.

In this chapter, the Dynamic Multi-Object Gaussian Process Model (DMO-GPM) framework is demonstrated for shoulder analysis. Recall that the DMO-GPM is a sub-model

type of the Dynamic Multi Feature-class Gaussian Process model (DMFC-GPM) framework and it is particularly suited for the analysis of shape and pose (see section 4.5.1). The shoulder joint is one of the most mobile joints of the human body. In addition to the high anatomical shape variability of the shoulder, it has the highest range of motion in the human body. Shoulder shape analysis using statistical shape models (SSMs) (Cootes et al., 2001) have been successful in literature (Salhi et al., 2020; Abler et al., 2018; Pitocchi et al., 2020). The analysis includes the prediction of missing scapular bone (Salhi et al., 2020), or pre-morbid glenoid cavity (Abler et al., 2018), or the scapula cortical bone (Pitocchi et al., 2020). Upper limb bones have been reconstructed using SSMs (Lebailly et al., 2012) and shape modelling has been demonstrated for the restoration of the patient-specific distal and proximal regions of the humerus (Vlachopoulos et al., 2018). However, the prediction of the pre-morbid shapes of shoulder bones using separate SSMs may suffer from some inherent drawbacks. The main drawback is that of impingement, since the relative position between the scapula and the humerus is not accounted for when using single object SSMs. Moreover, single object SSMs are not able to leverage any shape correlation that may exist between adjacent bones, to improve the prediction accuracy. Due to its high range of motion, the shoulder joint model requires a latent space embedding all key features that can provide additional knowledge for the estimation of both major bones (scapula and humerus), at any spatial orientation. As indicated in Section 2.2.4 previous methods employed for shape-pose analysis in the literature have not been applied to analyse joints with high range of motion.

For the shoulder analysis presented here, the DMO-GPM framework was applied for explaining shape and a simulated high range of abduction. The data used are first described, followed by an outline of how anatomical correspondence was established in the shoulder joints through methods described in section 3.2.1. The simulation of an abduction motion to obtain data across a high range of motion, is described. These data are used to demonstrate the performance of the model for pose analysis. This is followed by a demonstration of shape-completion performed on both scapula and humerus data. Virtual surgery of

both anatomies was used to develop incomplete shapes, while retaining their ground truth shapes. Finally, the prediction of shapes in different poses is described to evaluate the model framework's capacity for motion analysis. Some of the experiments and results of this chapter are part of the content presented by the author at the International Medical Image Computing and Computer-Assisted Intervention conference (Fouefack et al., 2020); and at the IEEE Engineering in Medicine and Biology Society conference (Fouefack et al., 2019), as well as in a pre-print on Arxiv. In all cases the author is the primary author and researcher.

6.2 Description of shoulder data

For the case study presented here, shoulder data consisted of $3D$ mesh surfaces segmented from computed tomography (CT) images of the bilateral shoulders of fresh cadavers. The cadaveric specimens were collected from the Division of Clinical Anatomy and Biological Anthropology, Faculty of Health Sciences, University of Cape Town, South Africa. Institutional ethics approval was granted for including the image data in this study (Approval No: HREC 283/2020). Eighteen bilateral shoulders joints were obtained. Of these data, 8 were from female, and 10 from male, decedents. Left shoulders were mirrored to the right to get a total of 36 shoulder joints. The age of the decedents ranged from 21 to 90 years. These data were imaged after dissection of the shoulder, where some surrounding tissues were removed and most of the images only display the humerus and the scapula bones.

6.3 Preprocessing the shoulder image data

As presented in chapter 3, anatomical correspondence establishment is a preprocessing step for the development of the DMO-GPM. The reconstructed surface meshes from the CT images were not in correspondence. That is to say, there was no anatomical one-to-one mapping between mesh vertices of the same bone family. To build a DMO-GPM, such a correspondence needs to be established first. Secondly, the shoulder dataset consisted

of humeri and scapulae, which were obtained from cadaveric CT image post-dissection. This means there existed a non-physiological relative positioning of the humeri to the relevant scapulae. Using such data for the analysis would lead to an unrealistic motion study. Furthermore, the simulated supine imaging position meant a narrow range of pose variation across specimens. However, the main idea to be presented here was to evaluate the ability of the developed framework for analysing high range of motion in a quasi-realistic context. To overcome these limitations, abduction motion was simulated to obtain in-correspondence shoulder data with a higher range of motion for the analysis.

6.3.1 Establishing anatomical correspondence in shoulder data

In contrast to the operator designed in-correspondence synthetic data used for validating the framework in chapter 5 (Section 5.4), real data were not in dense anatomical correspondence. To establish such correspondence across mesh surfaces from the shoulder dataset, a median virtual shape (MVS) was created for the humerus and scapula separately, using the iterative median closest point Gaussian mixture model (IMCP-GMM) (Jacq et al., 2008; Mutsvangwa et al., 2015) algorithm. The algorithm tries to find the "consensus shape" or "canonical" from a dataset which is based on median points defined from the information redundancy in multiple examples of the same biological objects. The MVSs were used as morphologically unbiased templates to built separate Gaussian process morphable models (GPMMs) for scapula and humerus. The MVS (one for each bone family) were then used as templates to register all the samples across the datasets using the parametric registration method described in chapter 3 (section 3.2). More precisely, GPMM-based free-form deformation models (FFDMs) were built from the templates for each bone separately (scapula and humerus). The registration involved fitting each of these single object GPMMs to bones of their respective family for all the training dataset.

After the registration, each fitted object (object corresponding to the best model parameters representing the target after the fitting) was compared to the original target object using Hausdorff (HD) and root mean square (RMS) surface to surface distances. The

Table 6.1 Registration errors of the glenohumeral training dataset. The average RMS and the HD errors of fitted scapulae and humeri vs. the original targets.

	RMS (mm)	HD (mm)
Humeri	**0.73±0.18**	**2.62±0.84**
Scapulae	**1.11±0.20**	**3.01±0.80**

fitted object would then replace the original target object for the model building process provided the distance errors between this fitted object and the original target were relatively small. The registration of the template to all the training examples allows the one to one mapping between them.

Table 6.1 shows the average (averaged over all training data) surface to surface RMS and HD error between the original meshes and their estimations after the registration process for humeri and scapulae for both shoulder datasets. The average RMS error was 0.73 *mm* and 1.11 *mm* for the humerus and the scapula, respectively. The average HD error was 2.62 *mm* for the humerus, and 3.01 *mm* for the scapula. The registration errors were low compared to the ones reported in (Inyang et al., 2017) where the maximum HD error was approximately 12.0 *mm*. This indicates that adequate anatomical correspondence was established across examples in the training datasets, hence target objects can be replaced by their fit in the training dataset. Moreover, the humerus registration errors are relatively low compared to the scapula registration errors. This could be attributed to the greater shape complexity of the scapula relative to the humerus (Mutsvangwa et al., 2015).

6.3.2 Simulation of shoulder joint motion

To overcome the issue of unrealistic pose coming from cadaveric imaging, an artificial motion was simulated using the biomechanics protocol proposed in (Wu et al., 2005). For abduction motion, Wu et al. (2005) notes that within a range of 0 to 120 degrees, the humerus can be considered a moving body while the scapula remains fixed. Beyond this range the assumption of a static scapula is no longer valid. In this experiment, a random shoulder joint from the in-correspondence dataset was used as a template and artificial

pose variations were obtained by: first, computing the rotation centre of the humeral head using a sphere fitting algorithm (Eberly, 2000); second, determining the International Society of Biomechanics (ISB) recommended coordinates for both humerus and scapula bones (Wu et al., 2005); third, aligning the ISB coordinate systems of the scapula and the humerus; fourth, moving the scapula ISB origin to the glenoid centre to obtain the neutral position of the humerus; and finally, obtaining ten poses by rotating the humerus around the glenoid centre in the yz-plane (simulating abduction). Five poses were generated for each joint $(0, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3})$. Figure 6.1 illustrates the simulated motion.

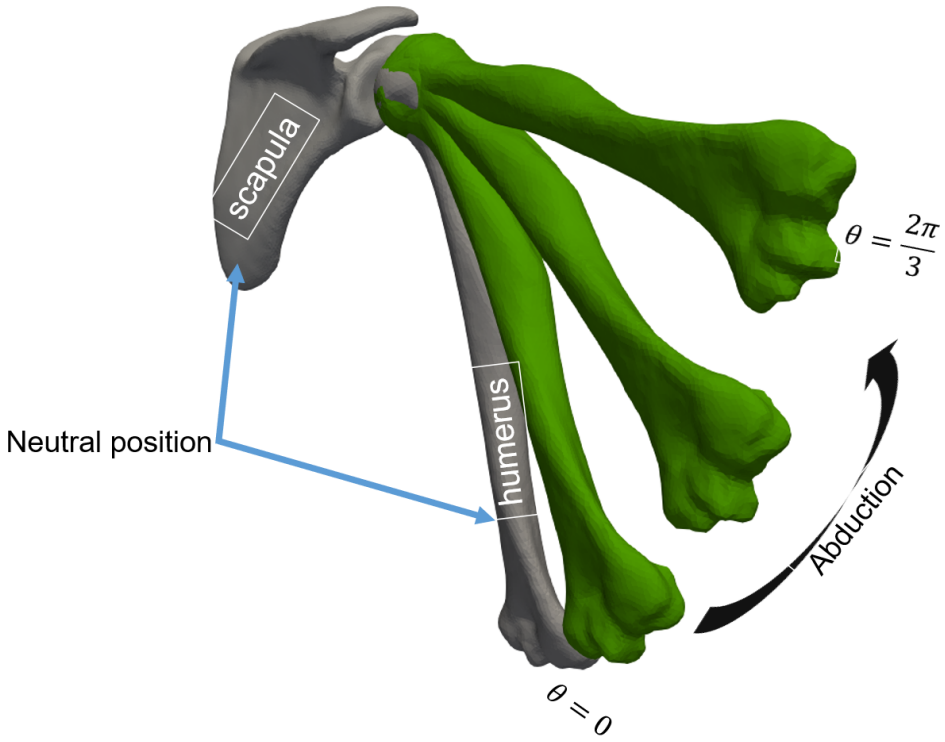

Figure 6.1 Illustration of simulated abduction motion. The bones in grey are at the neutral position. The green bones are the spatial positioning of the humerus relative to the scapula at various abduction angles ($\theta \in \{0, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}\}$).

After generating the abduction motion, the shoulder joints in the training dataset consisted of in-correspondence surface meshes in various abduction poses (defining a high range of motion). Since only shape and pose were being analysed, the shape-pose derivative of the DMFC-GPM (namely DMO-GPM) of the shoulder joint was developed for further

analysis. The method used for developing the shoulder DMO-GPM followed the one that is described in chapter 4 (section 4.5.1 and figure 4.5).

6.4 Building and validating a shoulder joint model

The DMO-GPM of the shoulder built from the dataset with the simulated abduction motion was intrinsically validated. This validation aimed to understand the shape-pose variability of the shoulder joint revealed by the DMO-GPM, as well as the shape-pose correlation. Samples were drawn along the primary principal components (PC) of the shoulder DMO-GPM in order to evaluate the model dynamics. The primary PCs were the first, the second and the third components, and samples were generated from -3 to $+3$ standard deviations about the mean along each component.

The model sampling showed that the first PC accounts for 68% of the total variation, explaining the change in elongation of the humeral shaft, and the width of the scapula along the medial and lateral borders. This result implies a clear correlation between the humeral shaft length and the scapula width as can be seen in figure 6.2. It can be observed that the sample at -3σ for the first PC shows impingement between the glenoid and the scapula head, which is likely due to the large sampling bandwidth (-3σ to $+3\sigma$). Thus, this information could be leveraged when using the model for shape-pose analysis. For example, one could constrain the first PC's parameters to only generate valid samples (samples with no impingement). The second PC accounts for 13% of the total variation and explains the variation in pose. The variation in PC 2 corresponds to the abduction in the training data. The third PC accounts for 7% the total variation and explains the change of the humeral shaft's length and width as well as the lateral border and the inferior angle of the scapula; this indicates correlation between these shape features.

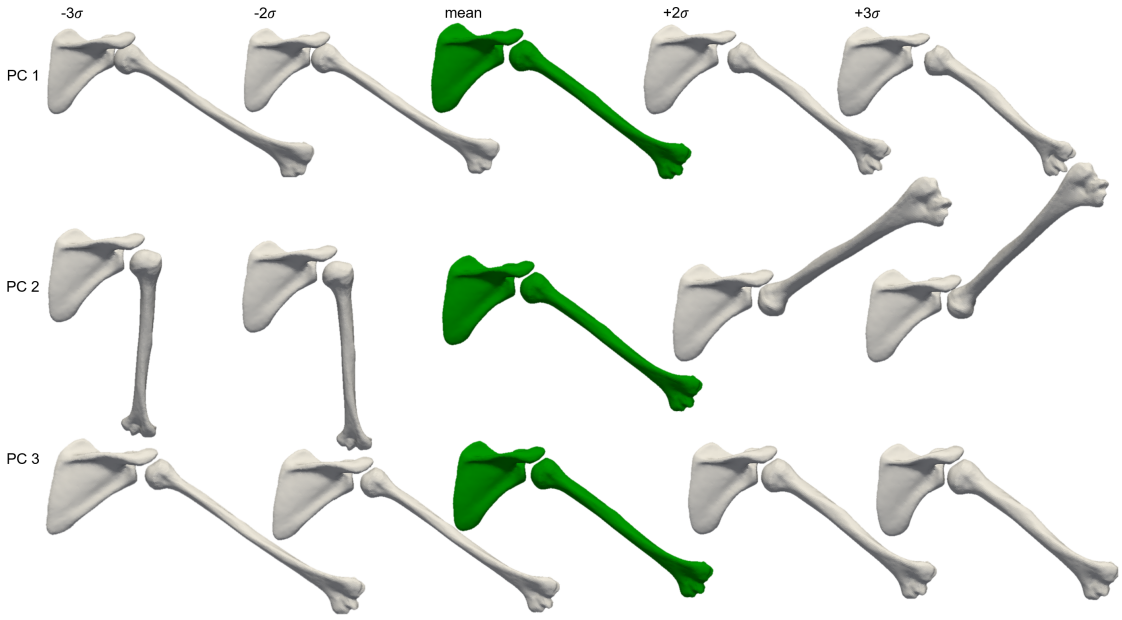

Figure 6.2 A sampling of the DMO-GPM of the shoulder. From the top to the bottom, an illustration of the first, second and third PC variability. From the left to the right, samples from -3 to $+3$ standard deviation around the mean (green).

6.5 Predicting shoulder shape features using shoulder joint model

Prediction of pre-morbid anatomy is a standard application of SSMs (Salhi et al., 2020; Abler et al., 2018; Pitocchi et al., 2020). As such, evaluation of the developed framework's performance on this clinically relevant task could be useful. The objective of this experiment was to use the shoulder DMO-GPM model for predicting pre-morbid shapes in simulated scenarios. The scenarios range from the object to be predicted being available, to no single feature of the structure to be predicted being available as described in table 6.2 [1]. This mimics the scenarios with increasing complexity, encountered in typical clinical workflows. For each scenario, the framework was demonstrated on the ability to reconstruct the normal anatomy of the joint through prediction.

[1]The table is identical to that presented in Chapter 5

Table 6.2 Taxonomy of types of observations of target objects used for evaluating model prediction capacity

Type I	Full object to be predicted is available
Type II	Partial object available
Type III	Sparse anatomical features (e.g. landmarks) of the object available
Type IV	No direct information of the object is available

6.5.1 Modifications to preprocessed data relevant to this experiment

The prediction method for this experiment was the sampling-based model-fitting approach described in Chapter 5. The idea was to use the DMO-GPM built above as a prior within the Markov chain Monte Carlo (MCMC) framework; i.e projecting the observations into the model's latent space. Thereafter, a prediction of the humerus from the observation of the scapula in different scenarios (see table 6.2) could be demonstrated. Thus, a test dataset was created, consisting of 10 shoulder joints (5 from the training data and 5 outside the training set) for shoulder joint shape feature prediction.

6.5.2 Methods specific to this experiment

To demonstrate realistic scenarios for prediction of pre-morbid shapes, shape-completion was performed for both scapulae and humeri from data which underwent virtual surgeries.

The first scenario was the prediction with Type IV observation (no knowledge available with regard to the humerus to be predicted); that is, only the scapula of the patient was observable. This scenario may happen for example, when trauma to a patient completely destroys the humerus; or when the image is cropped and the humerus is not visible at all.

The second scenario was the prediction with Type III observation (that is, anatomical landmarks of the humerus to be predicted, are available). For this experiment, two anatomical landmarks (tips of the medial and lateral epicondyles) were selected on the distal part of the targeted humerus. The aim here was to predict the humerus when presented

with a scapula and the two humerus landmarks using the DMO-GPM.

The third scenario was a bone resection simulation, that is, prediction with Type II observation. Partial humeri were obtained through virtual surgery of test set humeri surfaces. This scenario can occur for example, through trauma causing a humeral fracture while the scapula remains unaffected.

The last scenario was the prediction of the pre-morbid scapula. The same analysis was done for the scapula, where normal scapulae were reconstructed from partial scapulae using DMO-GPM and GPMM. Virtual surgeries to the distal scapula were performed to simulate resected scapulae.

To evaluate the prediction result for each scenario, the prediction errors using the DMO-GPM were compared to the ones obtained using a single structure GPMM-based SSM of the humerus for the second, third and the last scenarios. For the first scenario, a direct comparison of DMO-GPM performance with that of single object SSM was not possible as single object SSMs are incapable of predicting one structure given observations of another. In each case the error metrics used were the surface to surface distances between the predicted objects and the original objects using HD and RMS metrics.

6.5.3 Results of shoulder shape feature prediction

Table 6.3 shows the errors in predicting humeri without any observations at all of the respective humerus (with Type IV observation). It can be hypothesised that where there exists a strong correlation between the scapula and humerus shapes, then it should be possible to obtain a meaningful prediction of one given the other. Such a hypothesis is supported by the results and experiments on synthetic data in chapter 5 (section 5.4.1 and table 5.2). Furthermore, as mentioned above, single structure models would not be able to leverage such a correlation for prediction.

Figure 6.3 shows the errors in predicting the humerus (Type III observation) using the

Table 6.3 Prediction errors of the humerus with Type IV observations; average RMS and HD

DMO-GPM		
Humerus with Type IV observation	RMS (mm)	HD (mm)
	2.9±1.18	**14.0±5.23**
Single structure GPMM		
Humerus with Type IV observation	RMS (mm)	HD (mm)
	Not possible	

GPMM and those using the DMO-GPM framework. Prediction errors using the shoulder GPMM were large compared to the prediction errors using the shoulder DMO-GPM. The superiority of the DMO-GPM relative to the single structure GPMM could be attributed to the fact that the DMO-GPM framework leverages any additional knowledge of the humerus from its corresponding scapula through any naturally occurring correlation of features.

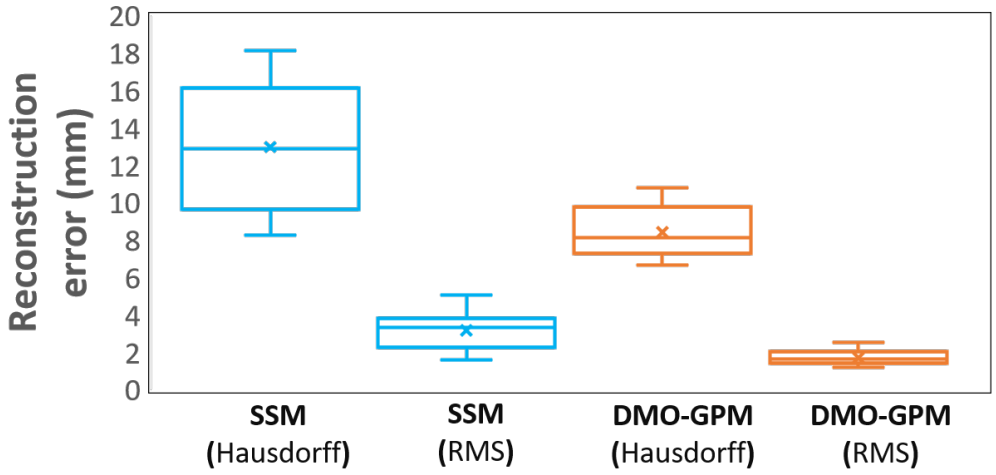

Figure 6.3 Comparison of the prediction of the humerus with Type III observation: shoulder GPMM vs. shoulder DMO-GPM. Blue: the HD (left) and the RMS distance (right) of the prediction using GPMM. orange: HD (left) and the RMS distance (right) of the prediction using DMO-GPM.

Figure 6.4 shows the box-plots of the RMS and the HD prediction errors of the humerus (with Type II observation) using shoulder DMO-GP and GPM models. The DMO-GPM outperforms the single structure GPMM in terms of reconstruction accuracy. Figure 6.5 shows a reconstructed scapula with DMO-GPM and individuals SSMs. Once again, this

performance could be due to the shape correlation that exists between the humerus and the scapula that can be considered to carry underlying information about the humerus.

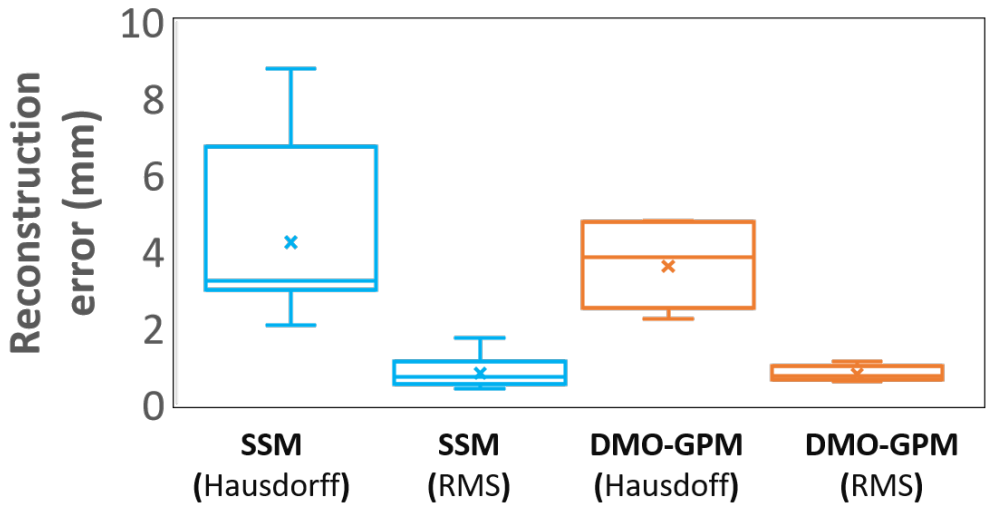

Figure 6.4 Comparison of the prediction of the humerus with Type II observation: shoulder GPMM vs. shoulder DMO-GPM. Blue: the HD (left) and the RMS distance (right) of the prediction using SSM. Orange: the HD (left) and the RMS distance (right) of the prediction using DMO-GPM.

Figure 6.6 shows the prediction errors of the scapula (with Type II observation), where the shoulder DMO-GPM outperforms the scapula GPMM. As expected, latent information of the scapula embedded in the humerus was used as additional acknowledge to outperform the single structure model.

6.5.4 Conclusion of DMO-GPM shape feature prediction capability

The above analysis suggests that the DMO-GPM framework improved upon the use of statistical shape models for pre-morbid shape estimation. Having scapula and humerus shape descriptors in the same latent space provided not only individual bone feature correlation but also between bone shape correlations that can be leveraged to improve the prediction accuracy. However, an evaluation of the performance of the framework in predicting shape features while preserving the relative position between the two bones

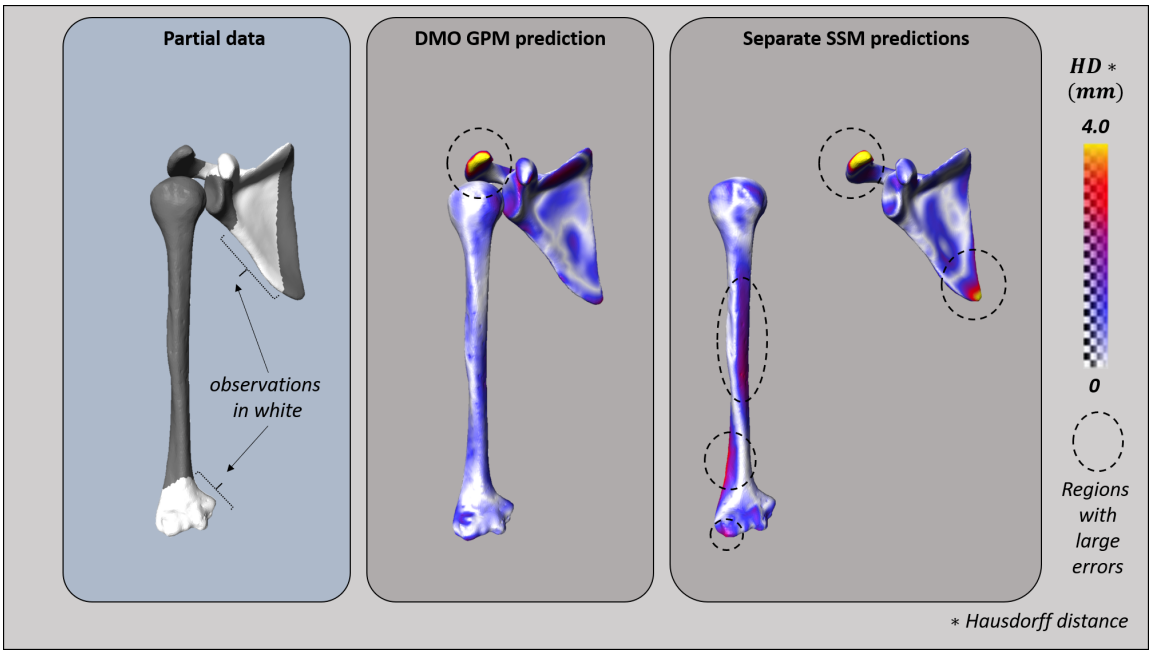

Figure 6.5 Reconstruction of a partial target scapula and humerus. Observed part of the shoulder for prediction (white). Middle: the Hausdorff distance (HD) reconstruction errors with shoulder DMO-GPM. Right: the HD reconstruction errors using separate GPMM of the humerus and the scapula.

was also necessary. Such a capability would be useful in mitigating the possibility of impingement when making predictions. The next experiment evaluated the DMO-GPM framework's capacity for predicting both shape and pose for the shoulder joint.

6.6 Shoulder shape-pose feature prediction

This section reports on the use of the DMO-GPM for predicting poses embedded in training data, in this case, abduction motion. This means predicting not only the shapes of the bones in the joint from images, but also retrieving their actual relative spatial positioning existing in the training data. Hence, the model can be used to predict acquired poses in the training data. **To the author's best knowledge there has not been any reported work on the simultaneous prediction of the scapula and humerus shapes, together with their relative spatial positioning; particularly for a larger pose variation.**

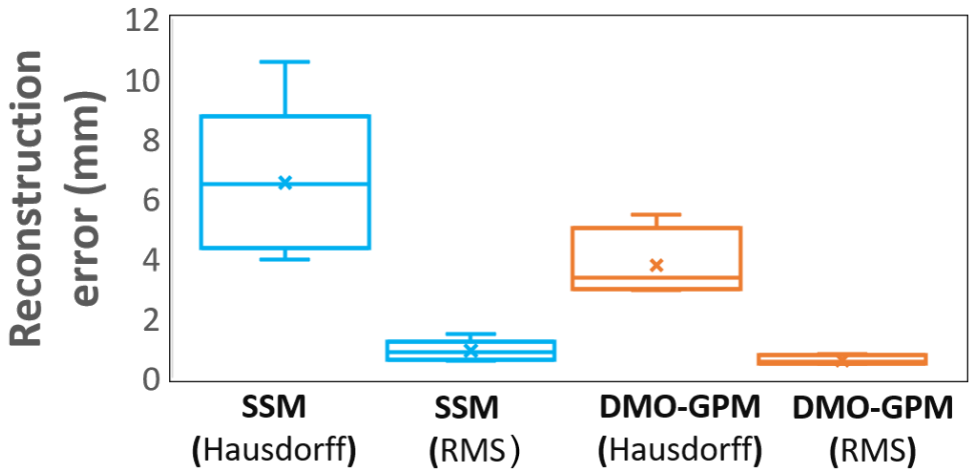

Figure 6.6 Comparison of the prediction of the scapula with Type II observation: shoulder GPMM vs. shoulder DMO-GPM. Blue: the HD (left) and the RMS distance (right) of the prediction using SSM. Red: the HD (left) and the RMS distance (right) of the prediction using the shoulder DMO-GPM.

6.6.1 Modifications to preprocessed data relevant to this experiment

For the experiment, five poses were generated within the abduction motion range $(0, \frac{3}{10}\pi, \frac{1}{2}\pi, \frac{7}{10}\pi, \frac{4}{5}\pi)$. The poses were selected to be different from the ones used in the training dataset but remaining in the same plane. This was to make sure that the motion described by testing dataset was not different from the abduction motion in the training. A shoulder joint DMO-GPM built with shape-pose variability was used for this experiment.

6.6.2 Methods specific to this experiment

For the evaluation, a DMO-GPM using the standard representation (SR) (see section 2.2.4) was built from the same shoulder data; that is, using the same simulated abduction motion. This second model was used for comparison with the DMO-GPM with energy displacement representation (EDR) (see Chapter 5) which follows a more principled manner to manage rigid transformations. To implement the DMO-GPM with SR, rigid transformations within the DMO-GPM setting were represented by Euler's parameters, that is, rotations were represented by Euler's angles, while translations were represented by 3D

Euclidean vectors. Thus, abduction poses were represented with Euler's parameters. Each example in the test dataset was predicted using both models (DMO-GPM with EDR) and DMO-GP with SR. The predictions were compared between the two constructs to evaluate their performance. The task in both cases was to predict the shapes of the shoulder joint and also their relative positions. The metric used was the RMS surface to surface distance. The average errors are the average over the surface to surface distance errors between the predicted and the original bones.

6.6.3 Results of shoulder shape-pose prediction

Figure 6.7 shows the box-plots of the average prediction errors obtained using pose marginalised DMO-GPM with EDR and also the DMO-GPM with SR. The average error of the prediction of the scapula was 0.63 mm for the DMO-GPM with SR and 0.55 mm for the DMO-GPM with EDR. For the humerus, the average error was 1.4 mm for the DMO-GPM with SR, and 0.9 mm for the DMO-GPM with EDR. The DMO-GPM with the EDR was superior to that with SR. For both scapula and humerus, the median error (cross line in the box) of the prediction with the SR model was higher than the one with EDR.

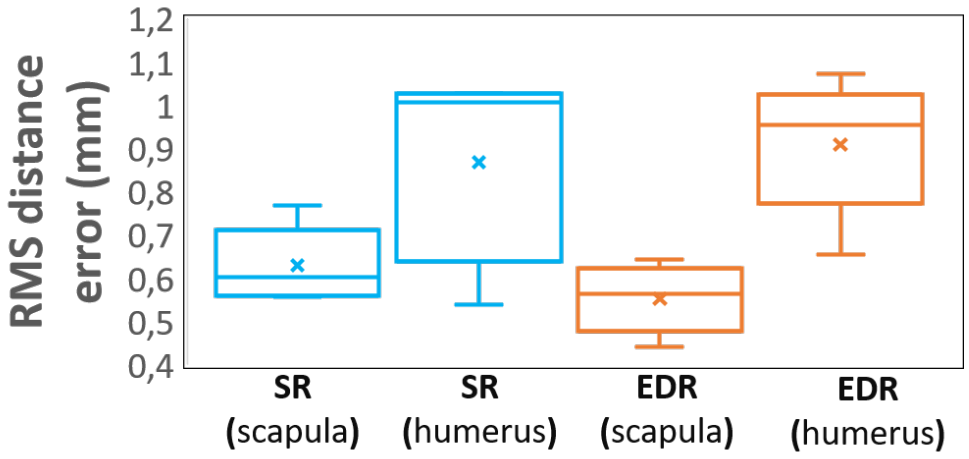

Figure 6.7 Prediction of the shoulder joint motion. DMO-GPM with EDR fitting compared to DMO with SR fitting. RMS errors between the predicted bone and the test targets for the scapula and the humerus.

Figure 6.8 shows the reconstructed abduction motion using the DMO-GPM with EDR, where the initial position of the model is shown on the left, and the target joints are shown to the right. The Hausdorff surface to surface distance between the predicted and the target joints are shown (right). It can be observed that joints were reconstructed with good shape and pose accuracy; 3.30 mm HD error and 0.003 rad angular error, respectively.

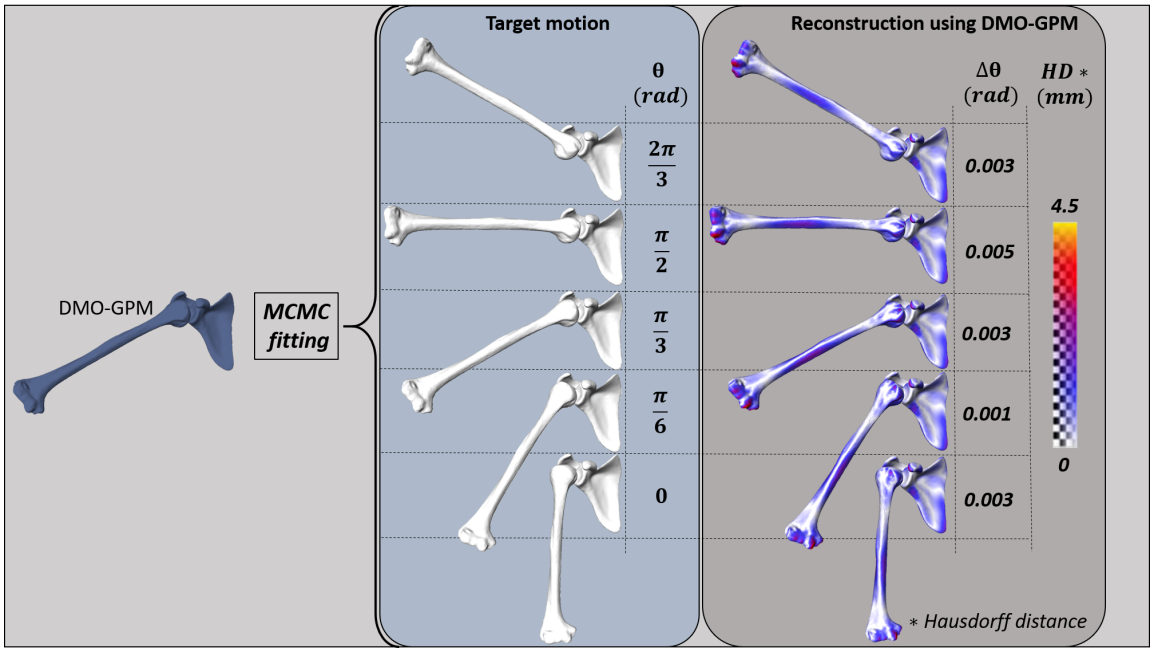

Figure 6.8 Reconstruction of target shoulders. Left: The DMO-GPM. Middle: target shoulder at various poses (from 0 to $\frac{2\pi}{3}$ rad) to be predicted. Right: Hausdorff distance and angle errors for the predicted shoulder using DMO-GPM with EDR.

6.6.4 Conclusion of DMO-GPM shape-pose prediction ability

The DMO-GPM framework performed well in predicting not only the shapes of the scapula and humerus, but also their relative positions. The test set motion range was relatively high ($0 - 120$ degrees) suggesting that the DMO-GPM framework can be useful for predicting high range motion. This may be clinically useful, for example, predicting joint motion from dynamic magnetic resonance (MR) imaging. However, the demonstration here was using bone segmentations, which are not always readily available. To improve

the utility of the framework would require the modelling of the image intensity profile so that this feature could also be used as an observed cue for prediction purposes.

6.7 Conclusion

This chapter describes the application of the DMO-GPM framework for analysing shape and pose features in medical images. The method was applied for shoulder joint analysis. It was shown that the use of a unified prior embedding shape and pose information is promising, particularly when analysing the complex anatomo-physiological relationships typical of human joints; and also where shape variation is large. The success of the framework for prediction in various observation scenarios hints at clinical utility for pre-morbid shape recovery, enabling analysis from incomplete images, or prediction of one anatomical object using the knowledge of a neighbouring anatomical object.

Furthermore, it was shown that the framework can make good estimates of the motion of the shoulder joint, at least in the case of controlled training data. Hence, this novel modelling framework could serve as a descriptor of $3D$ shape-pose of human joints in medical images. The framework can be used to predict pre-morbid shape for prosthesis design and surgical guidance. However, while the shoulder abduction motion embedded in the model was simulated through a biomechanics protocol, it remains a quasi-realistic, hence, may not reflect life patient motion. Moreover, simulated fractures used in the experiments may not reflect fractures that typically occur in real life.

For the analysis with $2D/3D$ images where $3D$ points representing bones are not available, the intensity feature should be included in the framework for the analysis, as will be demonstrated in the next chapter for the shoulder and hip joints.

Chapter Seven

Demonstration of the framework for analysing shape, pose and intensity features in medical images

7.1 Introduction

The Dynamic Multi Feature-Class Gaussian Process Model (DMFC-GPM) framework extends shape and pose modelling by incorporating intensity information. In the previous chapter, a sub-model of the full DMFC-GPM framework (the Dynamic Multi-Object Gaussian Process Model) was applied for analysing shape and pose in medical images of joints. This chapter describes the use of the full DMFC-GPM framework for analysing shape, pose and intensity in medical images. While the demonstration below is similar to the analysis on synthetic data in chapter 5 (section 5.5.2), using real images leads to some additional challenges. Some challenges encountered by clinicians when diagnosing or treating patients using medical images include noisy, incomplete data, or data with artefacts. Additionally, because of the high dimensionality of image volumes, the use of intensity-based numerical methods such as statistical shape and intensity models (SSIMs) (Reyneke et al., 2018b) leads to increased computational cost. It is the aim of the developed DMFC-GPM framework to allow clinicians to load a set of observations such as computed tomographic (CT), X-ray, or magnetic resonance (MR) images, and automatically predict structures of interest.

In this project, the DMFC-GPM framework was applied for shoulder and hip joint analysis; both of these are highly mobile joints. However, in contrast to the shoulder, the hip joint is a more stable joint, in part because it contains the most curved cavity of all ball-socket joints of the body. Active shape models (ASMs) (Cootes et al., 2001) have been successfully used for the segmentation of the hip bones by combining statistical shape models (SSMs) and the mean intensity profile to obtain the bony contours in the hip joint (Schmid et al., 2011; Xia et al., 2013; Park et al., 2012; Bredbenner et al., 2014), as well as humeral head contours from CT images (Mahdi et al., 2020). Automatic segmentation of the humerus and scapula bone using individual SSMs for each bone combined with another model incorporating the spatial relationship between them has been reported (Yang et al., 2015). Statistical shape models and the intensity models were also applied to investigate demographic variation in shoulder morphology (Soltanmohammadi et al., 2020). Regardless, segmentation of both the shoulder and hip joints remains challenging. For example, it is hard to selectively locate the articular cartilage of the hip joint (Li et al., 2008). Neural networks (NN) have been reported for hip segmentation based on bootstrap aggregating strategies (Xia et al., 2017). Shoulder bones were also segmented using NN with patch-based adjusted strategies (Yunpeng et al., 2018) and using adversarial networks with combined regularisation from shape priors (Boutillon et al., 2020). However, the NN approaches still suffer from various drawbacks including high variability in the relative spatial orientation of the segmented structures of interest (Seoud et al., 2019). Moreover, most of the NN methods cannot be used to segment objects that have different structures from those used for training the models. For example, a model trained on CT image data of the shoulder using NN can not be used to perform a prediction from a mere observation of anatomical landmarks of the shoulder bones. Typically, most of the NN segmentation methods do not retrieve the intensity profile representing the internal structure of bones of interest; a feature which may encode clinically relevant information. These drawbacks have introduced some limitations to the clinical utility of NN approaches. The bottlenecks also prevent the use of NN-based models trained on CT images from estimating $3D$ geometries

of bones from $2D$ images (X-ray images) without the need of paired training data ($3D$ and $2D$ images). The latter application has been dominated by shape modelling approaches through iterative digitally reconstructed radiograph (DRR) rendering and optimisation (Sadowsky et al., 2006; Fotsin et al., 2019; Klima et al., 2016; Ehlke et al., 2013).

Digitally reconstructed radiographs (DRRs) play an important role in the management of musculoskeletal disorders (MSDs) for diagnostic and orthopaedic surgical guidance. Due to high cost of imaging in $3D$, tedium of manual segmentation of $3D$ image volumes, and potential of exposing patients to relatively higher ionising radiation with CT imaging, $2D$ radiological images are routinely used for intra-interventional imaging and or follow-up studies in orthopaedics. However, the musculoskeletal system is $3D$ by nature and as such, assessment of the system using $2D$ images is error prone and introduces subjective bias on the part of the clinician. This is because the clinician has to mentally resolve the $2D$ images in $3D$; a process which relies on the clinician's experience. Alternatively, $3D$ representations of the object can be estimated from $2D$ image views. This may be especially useful when high costs, radiation exposure (for CT only) or lack of availability precludes the use of traditional $3D$ imaging. The literature abounds with research on the use of statistical model-based methods for $3D$ reconstruction from $2D$ images e.g. for $3D$ hip bone reconstruction from X-rays images (Klima et al., 2016; Yao and Taylor, 2001; Ehlke et al., 2013); and through use of the same techniques, but with associated pose and region-based intensity retrieval, for the knee joint (Fotsin et al., 2019). However, all the above methods work by associating a secondary, and separate, intensity distribution model in combination with the shape and pose models. Consequently, the ability to generate human joint DRRs for $3D$ from $2D$ registration, and reconstruction, using a statistical prior embedding shape, pose and intensity features in one unified model, remains unexplored.

The image data used for the analyses that follow were obtained from both cadaveric specimens and living subjects. The first section presents the shoulder analysis, where the

data used are first described, followed by an outline of how correspondence was established through methods described in section 3.3. An evaluation of shoulder bone shape, pose and intensity feature variability is then presented, followed by a demonstration of predicting shape, pose and intensity features from CT images of the shoulder. The shoulder data analysis section ends with a demonstration of how the DMFC-GPM of the shoulder joint can be used for DRR generation.

The second part of the chapter focuses on hip analysis conducted using the full framework. A description of the hip image data is followed by an outline of how correspondence was established for the hip joints. Similar to the shoulder section, an evaluation of hip bone shape, pose and intensity profile variability is presented; this is followed by a demonstration of predicting shape, pose and intensity features from CT images of the hip. In addition, projections of the DMFC-GPM of the hip joint to a $2D$ space using a DRR renderer are presented.

7.2 Shoulder joint analyses using the DMFC-GPM

In this experiment, images of the shoulder joint complex (consisting of the glenohumeral and the acromioclavicular joints) were analysed. In contrast to the shoulder data used in chapter 6, where shape and high range of motion analysis were the focus, here the main concern was the analysis and the prediction of all the key features that can be embedded in the modelling framework (shape, pose and intensity).

7.2.1 Description of the shoulder data

For the case study presented in this section, shoulder data consisted of 12 bilateral cadaveric shoulder complex CT images (total of 24 shoulders) obtained from the SICAS Medical Repository (SICAS: http://www.si-cas.com/). These data are freely available for research use. The voxel size of the images range from $0.97 \times 0.9 \times 0.5$ to $1.27 \times 1.27 \times 0.8$ mm. These image data were obtained from cadavers and each image data volume includes the

clavicles, humeri, and scapulae. Data with shoulder implants were excluded from this study. The recorded age at the time of death for cadaveric specimens ranges from 31 to 75, with 4 females and 8 males.

In contrast to the shoulder data used in chapter 6, each CT image data volume in this chapter includes the clavicle, humerus and scapula; and spatial orientations between them are physiologically plausible as they were obtained from intact cadaveric specimens, unlike the previous cadavers which were post dissection.

7.2.2 Establishing anatomical correspondence in shoulder data

In contrast to the shape-pose model building process in chapter 6 which only required surface meshes in-correspondence, here the correspondence needed to be established across volumetric shape-intensity of the shoulder joints. As described in chapter 3, correspondence across volumetric shape-intensity was induced from correspondence across outer surfaces of the objects. Thus, correspondence across the outer surfaces of each of the bones (scapula, humerus and clavicle) was established first. Also, similar to the previous chapter, a median virtual shape (MVS) was created for each of the humeri, scapulae and clavicles training data, separately, using an iterative median closest point-Gaussian mixture model (IMCP-GMM). Mean virtual shapes were isometric meshes and had 20000, 9820 and 4485 mesh points for the MVS of the scapula, humerus and clavicle, respectively. Surface registration was then performed to establish correspondence across the training data using the MVS of each bone, as unbiased templates.

Table 7.1 shows the average surface to surface root mean square (RMS) and Hausdorff (HD) distances between the original meshes and their post-registration approximations for clavicle, humerus and scapula bones. The registration errors were similar to the registration errors in the previous chapter (section 6.3.1) indicating that adequate anatomical correspondence was established across examples in the training dataset.

Table 7.1 Shoulder joint training dataset registration errors. The average RMS and the HD of the clavicle, scapula and humerus.

	RMS (mm)	HD (mm)
Clavicles	**0.18±0.10**	**2.12±0.50**
Humeri	**0.73±0.18**	**2.80±0.84**
Scapulae	**1.11±0.20**	**3.01±0.80**

From the in-correspondence surface meshes, the template surface mesh (MVS) was used to create tetrahedral mesh for each bone. These tetrahedral meshes were created using AMIRA software (http://www.fei.com/) as in chapter 3 (section 3.4.5). They were used as template volumetric shapes to establish volumetric shape-intensity correspondence. Each template tetrahedral mesh is isometric and had 21326, 11317, 4484 tetrahedrons for the scapula, humerus and clavicle, respectively. The correspondence across tetrahedral mesh was induced from the correspondence across surface meshes established above through an interpolation of the deformation fields as described in chapter 3 (section 3.3). The correspondence across volumetric shape-intensity of the shoulder complexes was then obtained by aligning each tetrahedral mesh to its corresponding CT image, and the voxel intensity value of the closest voxel to a mesh vertex was assigned to that vertex. Once intensity values are assigned to tetrahedral mesh points for each volumetric mesh (already in correspondence), it creates correspondence across volumetric shape-intensity. The in-correspondence volumetric shape-intensity of the shoulder complexes were used to create a DMFC-GPM of the shoulder.

7.2.3 Building and intrinsic validation of DMCF-FPM of the shoulder

The DMFC-GPM of the shoulder was intrinsically validated through sampling. The aim of this process was to evaluate the shape, pose and intensity variability of the shoulder complex revealed by the DMFC-GPM, as well as the correlation between different features.

A- Shape, pose and intensity variability of the shoulder complex

Figure 7.1A shows the volumetric shape-intensity samples of the shoulder complexes of the first, second and third principal components (PCs). The samples were generated from -3 to $+3$ standard deviations about the mean. Changes in shape, pose and intensity can be observed, where the colours indicate the intensity values in Hounsfield units (HU). For PC 1, the intensity at the humeral shaft changes from dark red (-3σ) to light red ($+3\sigma$) as well as the angle between the medial border of the scapula and the humerus increases from -3σ to $+3\sigma$. For the same PC, the width of the humeral shaft and the subscapular fossa increases from -3σ to $+3\sigma$. For PC 2, the clavicle seems to change its orientation relative to the scapula; the deltoid attachment area becomes more curved from $+3\sigma$ to -3σ about the mean. The humeral pose relative to the scapula changes as well. Humeral shaft length and width changes are the most prominent changes for PC 3 from $+3\sigma$ to -3σ about the mean.

The intensity variability is more difficult to interpret through visualisation alone (see figure 7.1A). To provide a more intuitive understanding of the change in intensity from the DMFC-GPM shoulder model, the intensity values were plotted against their location index for samples from -3 to $+3$ standard deviations about the mean as shown in figure 7.1B. The first PC shows high variability in intensity, with a relative decrease in variability for the second and the third PCs for all three bones: clavicle, humerus and scapula. However, the second PC indicates higher variability in intensity for the clavicle and the humerus compared to the scapula.

B- Correlation of shape and intensity in the shoulder complex

In this experiment, the hypothesis that combining shape and intensity in the same latent space may reveal underlying correlations between them, was evaluated. That is, the correlation between shape and intensity features for the scapula, the humerus and clavicle were evaluated for DMFC-GPM of the shoulder joint.

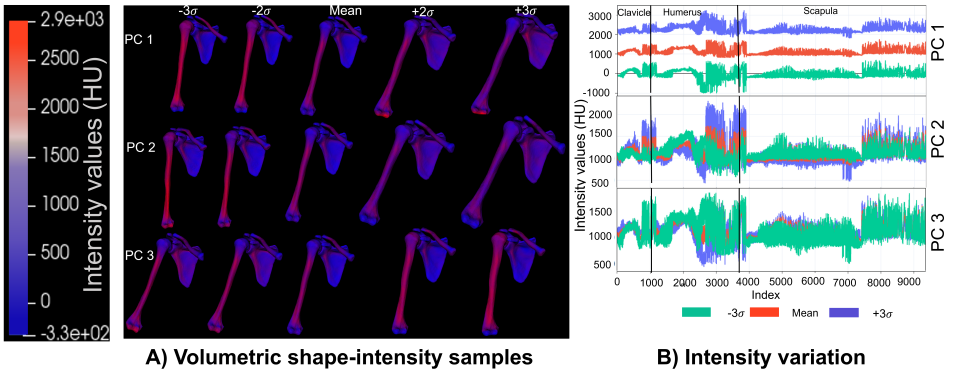

Figure 7.1 Sampling from the DMFC-GPM of the shoulder. (A) From the top to the bottom: Samples from the first, second and third PCs. From the left to the right: Samples from -3 to $+3$ standard deviation about the mean. (B) From the top to the bottom: Variation of intensity for the first, second, and third PCs of the clavicle, the humerus and the scapula, respectively.

B.1- Method for shape-intensity correlation experiment

To evaluate the correlation, Pearson's correlation coefficients between shape and intensity features were computed. The following was performed to calculate the correlation coefficients for a PC: firstly, samples were generated along the PC; and secondly, the correlation coefficients were computed between the set of corresponding points and their associated intensity values. The shape was measured at a point as the norm of the vector from the origin (point with all coordinates being zeros) to that point. The formalism of the process is below.

Let $\{(p_i^j, I_i^j)_{j=1}^{n_P}\}_{i=1}^{n_S}$ be the samples generated along a PC, where p_i^j is the i^{th} point on the j^{th} sample, I_i^j its associated intensity value, n_S, the number of samples, and n_p, the number of points. The Pearson's correlation coefficient of the i^{th} point can then be computed as:

$$r_i = \frac{n(\sum_j \|p_i^j\| I_i^j) - (\sum_j \|p_i^j\|)(\sum_j I_i^j)}{\sqrt{[n\sum_j \|p_i^j\|^2 - (\sum_j \|p_i^j\|)^2)][n\sum_j (I_i^j)^2 - (\sum_j I_i^j)^2)]}} \tag{7.1}$$

The indices of the volumetric meshes of the three bones were concatenated so that they became one volumetric mesh. This was done to obtain all the points and the intensities of the constituent bones (scapula, humerus and clavicle) in one graph.

B.2- Results of the shape-pose correlation in the shoulder complex

Figure 7.2 shows the histogram of the Pearson's correlation coefficient for shoulder bones for the first, the second and the third PCs. The results show high correlations ($|r_i| > 0.5$) for more than 95% of the points constituting the volumetric meshes, which supports the hypothesis, that shape and intensity for the shoulder bones are correlated. However, such results can only be conclusive if the training dataset is representative of the given population which depends on the size of the dataset as well as on the range of variability within the training dataset. Hence, for a large dataset (in terms of size and feature variability), such knowledge could be of interest in predicting intensity from observations of shape features, and vice-versa.

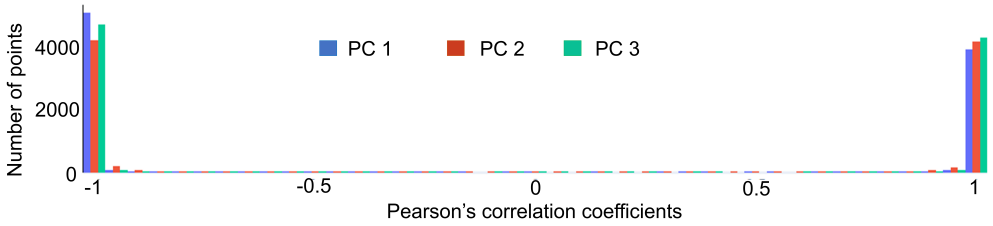

Figure 7.2 Histogram of shape and intensity correlation of the DMFC-GPM of the shoulder. Histogram of the correlation coefficients of the first (blue), the second (red) and the third (green) PCs.

7.2.4 Predicting shape, pose and intensity features using the shoulder complex model

The ability of the DMFC-GPM of the shoulder for predicting all the key features (shape, pose and intensity) was evaluated. Statistical models such as ASMs (Cootes et al., 2001) have been used in the literature for the segmentation of shoulder bones (Boutillon et al., 2020; Schmid et al., 2011; Xia et al., 2013; Park et al., 2012; Bredbenner et al., 2014). Segmentation plays an important role in medical imaging applications by transforming medical image representations into semantically meaningful objects. However, a meaning-

ful object for clinicians goes beyond the contour of the object. Clinicians would also like to know the position of the object relative to other objects in the joint, as well as to estimate the quality of the internal structure of the object. A primary motivation for developing the DMFC-GPM was to allow for the extraction the contours of bones in a joint, their relative spatial positioning, and their internal structural information; all in one process.

A- Method for the shoulder shape, pose and intensity prediction

To automatically fit the DMFC-GPM onto a CT image data for the prediction, the scapula was assumed to be a fixed object. The observations were full CT image volumes, that is, all the slices of the CT data. An initialisation step was required to remove any non-physiological relationship that may exist between the model and the CT image; that is, making that relationship invariant under non-physiological rigid transformations as described in chapter 5 (section 5.2.1). One of the attributes of the DMFC-GPM framework is that having some knowledge about an object is enough for such initialisation. For this demonstration, shoulder complexes were predicted from CT images using two different likelihoods for the MCMC-based fitting described in section 5.3.2. One was the Cauchy distribution based maximum likelihood (CDML), and the other was RMS-based maximum likelihood (RMSML). A test dataset of 6 CT images was used for the predictions. Each shoulder complex was predicted from a CT image in the test dataset using both the likelihoods and the errors were computed in Hounsfield units (HU).

B- Results for the shoulder shape, pose and intensity prediction

Table 7.2 shows the average errors, in HU, of the prediction from all the CT images in the test dataset. The errors are the normalised RMS similarity measures between the best model's sample intensity obtained and the observed shoulder intensity for both RMSML and CDML. The prediction errors are relatively small in both cases. The average error using the RMSML is 0.053 HU for the clavicle, 0.065 HU for humerus and 0.040 HU for the scapula. The average errors using CDML are 0.050, 0.060 and 0.040 for the clavicle,

humerus and scapula, respectively. The scapulae were predicted with the smallest errors although their shapes are more complex compared to those of the humerus and the clavicle. This could be attributed to the fact that scapula bones were fixed objects and were used to initialize the model, hence they did not have pose variability.

Table 7.2 Normalised root mean square errors of the prediction of the shoulder joint from the observation of CT images in the test dataset. Comparison of the prediction with RMSML to the one with CDML. The average RMSML and the CDML of the clavicle, humerus and scapula.

	Clavicle	Humerus	Scapula
RMSML	0.053 ± 0.017	0.065 ± 0.026	0.040 ± 0.041
CDML	0.050 ± 0.008	0.060 ± 0.020	0.040 ± 0.010

Figure 7.3 shows sample images of the prediction of 3 subjects, where the predicted intensity distribution is shown in a red-blue scale. The shoulder complex is predicted with intensity profiles that are close to those of the original images for each structure as indicated by the mean errors in table 7.2 . This is novel in medical image analysis where most of the machine learning segmentation algorithms only generate binary images of the target meaning they are unable to provide any information beyond target shape.

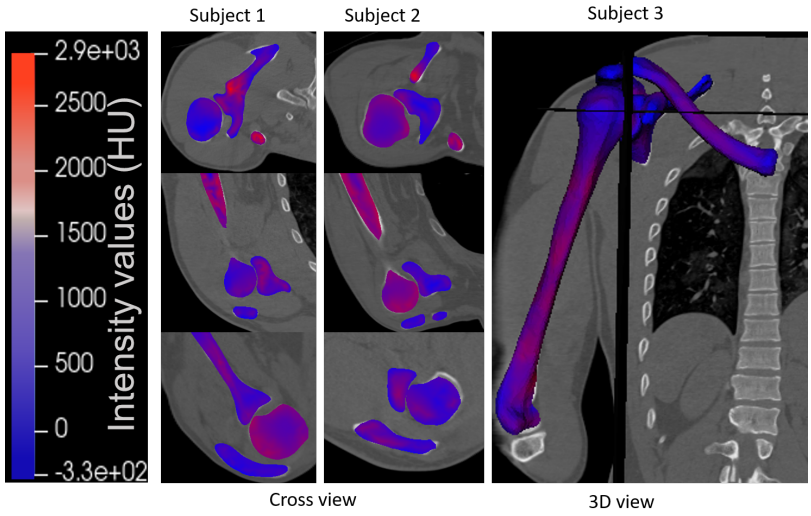

Figure 7.3 Shoulder prediction from CT images. The predicted shoulder intensity distribution and shapes (colour intensity map) for three subjects.

7.2.5 Two-dimensional model sampling and validation

It was demonstrated in chapter 5 (section 5.3.3) that the DMFC-GPM can be projected into a $2D$ space to obtain a digitally reconstructed radiograph (DRR) model. This experiment extends the $3D$ DMFC-GPM of the shoulder complex to a $2D$ model to analyse shape, pose and intensity variability in the resulting DRRs. Despite the fact that $2D$ representations of the joints are more compact because bones overlap, the changes in intensity should be more visible and interpretable in $2D$ image than in the volumetric shape-intensity representation. To demonstrate this, samples from the shoulder DMFC-GPM built above were projected into $2D$ using the volume rendering technique described in section 5.3.3.

The $2D$ projected samples of the shoulder complexes are shown in figure 7.4. For the intensity variation, it can be observed in PC 1 and 2 that the scapula, the humerus and clavicle are getting brighter from -3σ to $+3\sigma$. Similar variation is observed in PC 3 but more visible on the scapula compared to the clavicle and humerus. The clavicle pose changes are also more visible in PC 1 and the scapula shape changes are visible in PC 3 for the same standard deviations about the mean in their respective directions. The humeral shaft changes are more clearly observable in PC 2, as well as the changes on the scapula's medial border.

7.2.6 Conclusion for the shoulder analysis

The DMFC-GPM framework shows an alternative manner to analyze CT images of the shoulder joint which provide more insight into shape, pose and intensity profiles simultaneously. When the shoulder DMFC-GPM was projected into $2D$ space, the variability in shape, pose and intensity could also be visualised. This demonstrated the potential for the DMFC-GPM framework to be used in the analysis of $2D$ radiological images relevant to the shoulder complex, and also its potential to be a pertinent tool for image-based surgical guidance.

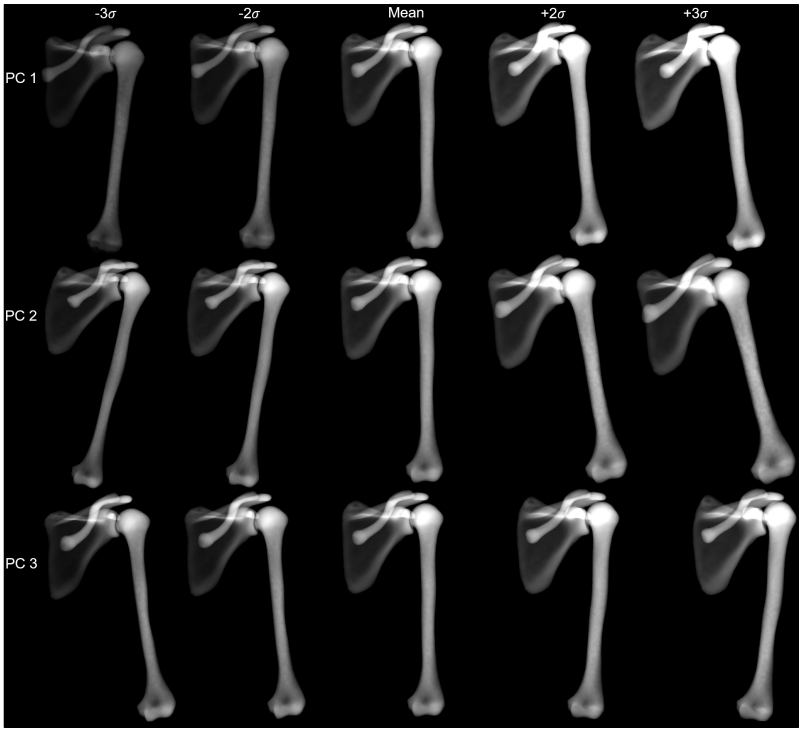

Figure 7.4 DMFC-GPM's samples of the shoulder complex projected as DRRs. From the top to the bottom: Samples from the first, second and third PCs. From the left to the right: Samples from -3 to $+3$ standard deviations about the mean for each of the PCs.

7.3 Hip joint analyses using the DMFC-GPM

In this experiment, images of the hip joints were analysed. In contrast to the shoulder data used in the previous section, where image data were acquired from cadaveric specimens, here the focus was to perform the analysis and the prediction of the key features of the image data from living subjects. Applying the framework to analyse these data is more likely to reflect clinical contexts.

7.3.1 Hip data description and preprocessing

The hip dataset consisted of 30 CT images from individuals with associated hip impairments. These data were part of the dataset reported in (Audenaert et al., 2020).Each image volume consisted of an average of 1864 slices with trans-axial pixel size ranging between 0.575 mm to 0.975 mm. The voxel size of the images range from $0.62 \times 0.62 \times 0.60$ to

$0.94 \times 0.94 \times 5.00\ mm$. Each image volume contained the full pelvis bone and both the femur bones. These data were approved for research use by the ethics committee of the Ghent University Hospital (Reference no: B670201111480). Surface meshes of the pelvis, left and right femurs were already segmented from these image data, were in correspondence, and were all at their original spatial orientations as described in (Audenaert et al., 2019).

For this project, a template tetrahedral mesh was created and used to established volumetric shape-intensity correspondence across the hips joints in the dataset as described for the shoulder data in section 7.2.2. Each template tetrahedral mesh was isometric and had 23378, 17905, 17905 tetrahedrons for the pelvis, left and right femurs, respectively. The volumetric shape-intensity of the hip joints in correspondence were used to build a DMFC-GPM of the hip joint, which was used for the following experiments.

7.3.2 Hip model building and intrinsic validation

A DMFC-GPM was built from the in-correspondence volumetric shape-intensity of the hip joints and was intrinsically validated. The shape, pose and intensity variability across hip joints were evaluated through sampling of the DMFC-GPM. The correlation between shape and intensity was also evaluated using Pearson's correlation coefficient, similar to the correlation analysis performed in section 7.2.3 for the shoulder.

A- Hip shape, pose and intensity variability

Figure 7.5A shows the hip volumetric shape-intensity samples drawn from the first, the second and the third PCs. For each PC, the variance ranged from -3σ to $+3\sigma$ about the mean. In terms of variation encoded in the first PC, the intensity values of the left and right femur shafts change from lower values (blue) to larger ones (red) with change from -3 to $+3\sigma$, respectively. For PC 2, the space between the left and the right femur increases from -3 to $+3\sigma$, indicating the relative spatial orientation variability between both bones captured by the DMFC-GPM. For PC 3 femur lengths increase from -3 to

$+3\sigma$, indicating shape changes encoded in the model.

The change in intensity is more visible in figure 7.5B, where the intensity value is plotted against its location index. The plots show how the intensity values of each bone vary when sampling from -3 to $+3\sigma$ for each PC (1, 2 and 3). The intensity varies the most in the first PC, approximately ranging from -500 HU to 2000 HU. In the second and the third PCs, the intensity variability ranges from -500 HU to 1500 HU and -200 HU to 1400 HU, respectively.

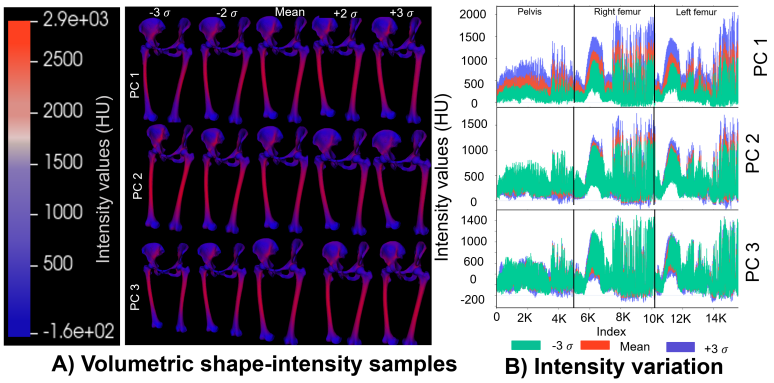

Figure 7.5 Sampling from the DMFC-GPM of the hip. A) From the top to the bottom: Samples from the first, second and third PCs . From the left to the right: Samples from -3 to $+3$ standard deviations about the mean for each of the PCs. B) From the top to the bottom: Variation of intensity for the first, second, and third PCs of the pelvis, the right and the left femurs, respectively.

B- Hip shape-intensity correlation

The correlation between shape and intensity for the pelvis, right and left femurs were evaluated using Person's correlation coefficients as described section 7.2.3. Figure 7.5 shows the histogram of the correlation coefficient between the shape and intensity of the hip joints in the images. The results show high correlations ($|r_i| > 0.5$) for more than 90% of the points constituting the volumetric mesh. Again this result supports the hypothesis that the DMFC-GPM framework can capture any correlation that may exist between shape and intensity features. This findings open doors for leveraging such knowledge in clinical practice to improve prediction accuracy. But as mentioned for the shoulder data,

this result depends on the variability and the size of the training data. Hence, the finding is not conclusive without ensuring that the training dataset is representative of the target population.

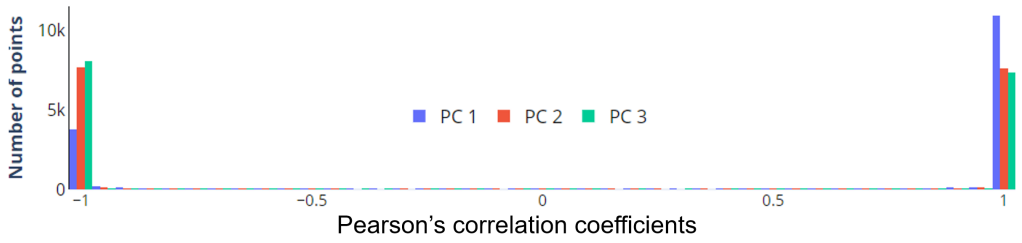

Figure 7.6 Histogram of shape and intensity correlation of the hip from the DMFC-GPM. Histogram of the correlation coefficients of the first (blue), the second (red) and the third (green) PCs.

7.3.3 Shape-pose and intensity prediction of the hip from image volumes

Similarly to the prediction done for the shoulder complex, shape, pose and intensity features of hip were predicted from CT image data using DMFC-GPM of the hip as prior. A test dataset of 6 CT images was used and hip joints were predicted from each CT image in the test dataset using CDML and RMSML. The observations were full CT image volumes, that is, all the slides of the CT data.

A- Results for the prediction of hip joints from CT images

Table 7.3 shows the normalised RMS similarity measures between the predicted intensity and the observed hip intensity for both RMSML and CDML. The mean error for the pelvis was 0.043 ± 0.010 HU for the RMSM compared to 0.044 ± 0.008 HU for the CDML. For the left and right femur, the mean errors were 0.028 ± 0.007 HU and 0.027 ± 0.005 HU for the RMSML compared to 0.028 ± 0.005 HU and 0.027 ± 0.004 HU for the CDML. Both femurs were predicted with similar errors, which was expected because both bones have similar shape and intensity profiles. The right and left femurs were predicted with

smaller errors compared to the pelvis; this could be attributed to the complexity of the pelvis shape compared to that of the femurs.

Figure 7.7 shows images of the prediction of 4 of the 6 subjects, where the predicted intensity is shown using red-blue scale. The hip joints are predicted with the intensity profile and the contours of the pelvis, the left and the right femurs. The prediction preserved the relative positioning between bones. This demonstrates that the DMFC-GPM may overcome impingements that often occur when performing segmentation using a separate SSM of each bone.

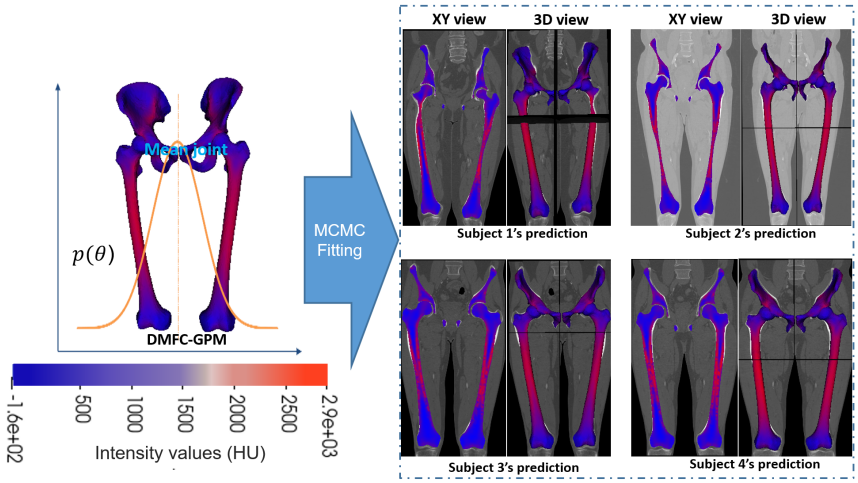

Figure 7.7 Hip joint prediction from CT images. The predicted hip intensity and shape (colour intensity map) for four subjects. Left: the mean of the hip DMFC-GPM. Right: the predicted joints (red-blue) and the target volumes (grey-scale).

Table 7.3 The normalised root means square errors of the prediction of the hip joints from the observation of CT image data. Comparison of the prediction with RMSML to the one with CDML. The average RMSML and the CDML of the pelvis, right and left femurs.

	Pelvis	Right femur	Left femur
RMSML	0.043 ± 0.010	0.028 ± 0.007	0.027 ± 0.005
CDML	0.044 ± 0.008	0.028 ± 0.005	0.027 ± 0.04

7.3.4 Two dimensional model of the hip from 3D model

The DMFC-GPM of the hip described above was projected onto a $2D$ space and samples drawn from the DMFC-GPM projected as DRRs using the same renderer described above

for the shoulder data.

Samples from the first, the second and the third PCs are shown in figure 7.8. The DRRs were generated as projected samples of the DMFC-GPM from -3 to $+3$ standard deviations from the mean for each PC. The shape, pose and intensity variability of the hip can be observed in all PCs. However, to aid in visualising the shape and pose differences, contours of highly varying regions are superimposed in the images in red. For PC 1, the red contours indicate shape changes around the greater trochanter of the femur and the iliac notch of the pelvis. For PC 2, as is also visible in figure 7.5, the space between the left and the right femurs increases with changes from -3σ to $+3\sigma$, as well as the variability around the pelvis inlet. In PC 2 and 3, the intensity variation is observed in all bones from -3 to $+3$ standard deviations with the grey scale changes from high to lower intensity. PC 3 also indicates changes around the pelvis inlet.

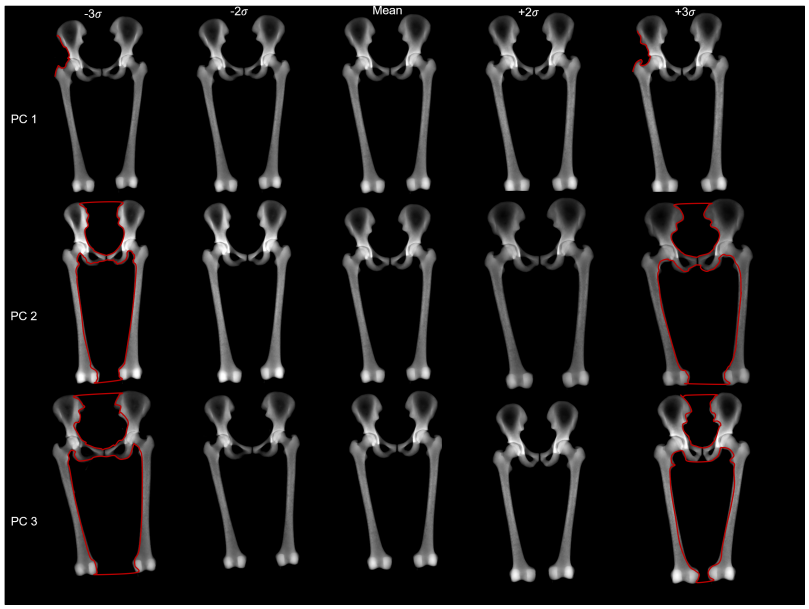

Figure 7.8 DMFC-GPM DRR samples of the hip joint. From the top to the bottom, samples for the first, the second and the third PCs. From the left to the right, samples from -3 to $+3$ standard deviation about the mean for each of the PCs. Red contours indicate areas of variation along each PC.

7.3.5 Conclusion for the hip joint analysis

Similar to shoulder data, the DMFC-GPM framework was demonstrated for hip joint analysis. The DMFC-GPM framework also shows that shape and intensity profile of the hip bones can be predicted from CT images while preserving their positioning with respect to each other. The DMFC-GPM was also projected in a $2D$ space and the encoded variation of the features was observed. This demonstrated that the DMFC-GPM framework can be used to overcome the positioning and intensity profile issues that occur when using ASMs and SSMs.

7.4 Software and computational time

Aside from demonstrating the potential of the DMFC-GPM framework through different experiments, one would also want to know whether they are computational challenges associated with using the model framework. Computation time plays an important role in the utility of a model for clinical applications. A modelling framework requiring supercomputers and long computational time may be limited in some applications. In this section, the software used for the experiments as well as the computational times are presented.

All the experiments throughout this thesis were run on central processing unit (CPU) software. There were run on a Lenovo E580 Laptop (IntelCore i7 -7500U, 16 GB RAM, 64 bit) in IMT Atlantique and, in the University of Cape Town (UCT), on UCT's information and communication technology services (ICTS) High-Performance Computing (HPC) cluster (CPU, Maxi nodes: 40) (http://www.icts.uct.ac.za/). Note that the UCT's cluster was used only to be able to run several jobs at the same time, not because the computation requires supercomputers.

As described in chapter 4 (section 4.6), the computation of the DMFC-GPM based latent space is not time-consuming, but the generalised Procrustes analysis (GPA) associated with aligning shapes may increase the training time. Table 7.4 shows the training time

Table 7.4 Training time for different of the shoulder and hip models used in this chapter and in chapter 6

Model	**Training time (s)**	
	with GPA	**without GPA**
Shoulder DMO-GPM	35	20
Shoulder DMFC-GPM	633	105
Hip DMFC-GPM	1677	262

for shoulder and hip models used in this chapter and in chapter 6. The training times were relatively short. The training times for the model with GPA are longer than the ones of the model without GPA. This was expected because the GPA requires an optimisation process. All the models used in our experiments included the GPA. The training times of the shoulder and hip DMFC-GPM (included shape, pose and intensity) are more longer compared to the one of the DMO-GPM (included shape and pose only). This is because the DMFC-GPM includes more structures and more feature classes, hence, requires some extra processing. To summarise, the training times for these models are short compared to models such as deep learning that may require hours or days for the training. After the training, the model is used for learning through a fitting process that may be computationally expensive as describe below.

The learning process used in the DFDMC-GPM framework is the sampling-based model fitting as described in chapter 5. As one of the sampling-based methods, the MCMC-based fitting is useful for finding the best model sample representing the observation. This is an advantage over parametric-based methods which can get stuck in local maxima/minima of the cost function; hence miss the global solution. However, sampling based approaches are computationally expensive. The MCMC-based fitting took approximately the same time for each model. Most of the experiments were done with an average number of iterations of 8000 with a running time of 2694 seconds. Table 7.5 shows example numbers of iterations with their associated running time.

Table 7.5 Model fitting times with corresponding number of iterations

Number of iteration	2000	5000	8000	12000
Running time (s)	720	1599	2694	3900

7.5 Conclusion

This chapter describes the application of the DMFC-GPM framework for analysing joints in medical images. The framework was applied for shoulder complex and hip joint analyses. The experiments demonstrate the use of a unified prior embedding shape, pose and intensity features towards a holistic analysis of joints in medical images. The experiments showed that the framework not only detects bone contours through segmentation, registration and reconstruction of structures, as most of the state-of-art methods, but it provides an estimate of the intensity profile inside the bones. It was also demonstrated that the DMFC-GPM of the shoulder complex and hip joints can be projected in a $2D$ space, a process which can be used for the analysis of other joints in $2D$.

It is important to note that, as any data-driven modelling, the performance of the model relies on the variability in terms of shape, pose and intensity distribution within the training data. Moreover as the presented framework is correspondence based, model performance also relies on the correspondence quality. Furthermore, there was an unexpected high correlation observed between the intensity and shape. Such a correction could be due to the intensity correspondence establishment approach used. The internal surface correspondence used for intensity correspondence was based on the surface one, which may lead to a high correlation between shape and intensity. In addition, the shoulder and hip data used did not have full range of motion, thus the framework remains to be evaluated for predicting shape, pose and intensity in such cases. Future work will aim to develop more sophisticated models which encode a large range of shape and pose variability of the shoulder and hip joints from living subjects. It would also be of interest to consider using other approaches for intensity correspondence establishment that may lead to more accurate correlations between shape and intensity as discussed in chapter 3.

Chapter Eight

Conclusion

8.1 Contributions of this work

This chapter describes the main findings and suggests possible extensions to the research presented in this thesis. The motivation for this work was three-fold. First, to overcome the lack of a computational tools for leveraging the underlying information emergent from a combination of shape, pose and image intensity, the key features used clinically for diagnosis, surgical planning and guidance. Current statistical models that provide latent spaces for shape, pose and image intensity analysis do not consider a continuous description of the objects of interest, nor do they embed all the key features and their correlations. Second, the prediction and evaluation of structures in medical images for learning purposes (such as musculoskeletal disorder evaluation) is still done through simultaneous fitting of shape, pose and intensity priors to find the optimal shape, orientation, as well as the intensity profile of the objects. The use of separate priors does not leverage correlations that exist between these features, hence increases the risk of inaccurate predictions. Third, human joint analysis from medical images using statistical models is still an open research area and is challenging due to the number of degrees of freedom (DOF), as well as high range of motion of some joints including the shoulder and hip joints presented here. The challenge is even more acute when coupled with large naturally occurring shape and intensity variation, the latter being a proxy measure of bone structure. Thus, the global aim of the project was to develop and validate a framework for multi-class statistical modelling of shape, kinematics and intensity in medical images. Furthermore, it was critical to demonstrate the use of the framework for analysing real patient images. This required the

generalisation, adaptation and application of some existing statistical techniques, as well as proposing novel methodologies. The technical contributions of this work are categorised in three items as summarised below.

First item: A method has been developed for estimating the parametric continuous latent space from structures in medical image data with variation across shape, pose and intensity features. Methods for establishing shape, pose and intensity correspondence for multiple objects have been proposed in chapter 3. These methods include parametric and sampling-based approaches. These correspondence establishment approaches were used to find deformation fields between reference and target objects. The deformation fields obtained from the training data were then linearized and normalised to obtain object descriptors in a vector space, that is, projecting functions from the Riemannian manifold to the tangent space at the mean deformation field as presented in chapter 4. The non-linearity bottleneck associated with rigid transformation was tackled by using a representation that leads to a non-linear norm proportion to the kinetic energy necessary to move the object from one position to the other. A Gaussian process-based latent space was computed over these deformation fields encoding shape, pose and intensity features in the tangent space. Such a latent space is parameterised by a set of orthogonal functions encoding not only the variability of each feature class in a continuous domain, but also the correlations that exist between them. This modelling framework was termed Dynamic Multi-Feature Class Gaussian Process Models (DMFC-GPM). Importantly, as all the features are in a joint statistical space, and in a continuous domain, this framework permits on-demand marginalisation to a region or feature of interest without re-training separate models; in contrast to current methods in the literature (Cootes et al., 1992; Lüthi et al., 2017; Moreau et al., 2017; Anas et al., 2016). Moreover, the continuity of the model also allows one to derive sub-model types from the DMFC-GPM such as Dynamic Multi-Object Gaussian Process Models (DMO-GPMs) and regression models, which were were also introduced and discussed. The DMO-GPM is particularly suited for use in analysing shape and pose. The regression model can be constrained to some observations in the image,

which improves the accuracy of the prediction outcomes since it reduces the space of possible accruing samples. In addition, previous statistical models (SSPMs, SPMs, SSMs or SAMs) can be seen as attributes of the developed unified framework (as summarised in see figure 4.5). To sum up, the principal contributions within this first category are:

- Novel parametric and sampling-based methods for establishing multi-object anatomical correspondence.

- Modification of a non linear-metric for comparing rigid transformations in a continuous space. This mitigates the non-compactness drawback for non-compact objects.

- A process for normalising different feature classes (i.e shape, pose and intensity features) in order to obtain an homogeneous analysis space.

- A Gaussian process based latent space for multi feature-class modelling with its associated properties characterised.

Second item: A learning process method was developed in chapter 5 to establish a transfer channel of shape, pose and intensity features from the latent space to the observations including point cloud, mesh surfaces, $3D$ images (computed tomography (CT)) or $2D$ images (X-rays). The main advantage of the presented framework compared to other approaches is that in addition to shape features embedded in the latent space, it also accounts for the pose and the intensity features, as well as all their correlations; all embedded in one framework. This embedding of various features in the latent space is leveraged when establishing the learning channel (see the fitting method in chapter 5). Such a learning channel improves the prediction accuracy of the shape, pose and intensity features that are of interest in the diagnosis and treatment of musculoskeletal disorders. The proposed approach was validated through extensive experiments and the results showed that it outperforms existing approaches included in the benchmarking performed. Firstly, with regard to pose variability, the DMFC-GPM framework was more accurate in explaining

rigid transformation across multiple objects. In particular, it is more specific to the training data and also generalises outside of the training data, in a realistic fashion (see section 5.4.2). Secondly, it leverages the pose variation, encoded as an energy displacement field in the latent space, to obtain the optimal pose parameters from an observation. Finally, it leverages the intensity feature descriptor in the latent space to infer the intensity profile of the internal structure of the objects. The main technical contributions here were:

- The adaptation of MCMC sampling and the Metropolis-Hastings algorithms for multi-object prediction.

- The incorporation of digitally reconstructed radiograph (DRR) projection in the model sampling process to allow $3D$ to $2D$ projection. This provides a $2D$ latent space from the $3D$ one.

- The development of controllable synthetic data to validate transfer learning from the model latent space.

- The benchmarking of the proposed model against other commonly used approaches including active shape model (ASM) based SSMs, and shape pose models made using a standard rigid transformation metric.

Third item: The DMFC-GPM framework was applied to analyse shoulder and hip image data in chapters 6 and 7. By evaluating its performance for the reconstruction of simulated partial shoulder bones, the DMO-GPM framework revealed potential for use in inferring the normal anatomical structure from the remaining part of the structure or from other structures in a joint or complex. Furthermore, the use of DMFC-GPM-based prior for predictions of shape, pose and intensity of different structures from CT images and $2D$ images of the shoulder and hip, with their constituent bones (scapula, clavicle and humerus for the shoulder; pelvis, left and right femur for the hip) displayed at imaging spatial orientations. The experimental results clearly indicated that the framework not only

detects an object's contours through segmentation of structures, as most of the state-of-art methods, but provides the patient-specific intensity profiles of the bones in the CT images. It was also demonstrated that the DMFC-GPM of the shoulder complexes and hip joints can be projected in a $2D$ space as DRRs, which could be extended for the analysis of any joints in $2D$. The main contributions in medical image analysis within this part were:

- Preprocessing of the shoulder and hip data from CT images to obtain desired data structures for the framework.

- Demonstration of the framework on shoulder complexes and hip joints from CT images and mesh surfaces to evaluate its efficiency using real image data.

The DMFC-GPM framework realises a new paradigm in modelling multiple structures in medical images, allowing for probabilistic modelling of not only the shape but also pose and intensity. The correlations that may exist between features can now be leveraged to develop clinical tools for the evaluation of musculoskeltal disorders in medical images. This will form part of future work.

8.2 Limitations and future research directions

Despite the findings and problems tackled in this thesis having added to the body of research on statistical modelling, the framework developed is subject to some limitations. In addition, it also creates a research foundation for further investigations. These limitations and future directions are outlined below.

8.2.1 Correspondence establishment

The use of large datasets for statistical model building increases the capacity of the model to represent a family of objects. A large dataset may contain a high range of feature variability. The model prior-based registration used in the DMFC-GPM framework for

shape correspondence establishment requires a template that is representative of the dataset (see section 3.2). However, with large variability, it becomes difficult to find a template for the training dataset. Hence, establishing correspondence using a template that does not represent the training dataset becomes more challenging. Moreover, volumetric shape correspondence (from which the intensity correspondence is defined) in the DMFC-GPM framework is induced from surface correspondence through the nearest neighbourhood interpolation (see section 3.3). With such an approach, the quality of the intensity correspondence highly depends on how the nearest neighbour of a point is estimated as well as on the surface correspondence quality. All these challenges may reduce the capacity of the model in covering the space of possible occurring shapes and intensities within the population. Furthermore, the method proposed for internal structure correspondence was only evaluated on the outer surface points because there were no ground-truth volumetric meshes. Some works on the volume registration can help. Recent deep learning techniques on correspondence (Agrawal et al., 2017; Sun et al., 2017; Boscaini et al., 2016; Fan and Wang, 2020; Lang et al., 2020) have shown promise in establishing shape correspondence across data with high range of variability. It would be interesting to investigate these state-of-the-art approaches to improve shape and intensity correspondence across large datasets. Moreover, one would also consider generating tetrahedral mesh data similar to the lollipop ones to evaluate the quality of the correspondence inside the structure of the objects.

For pose retrieval, the translation vector in the approach presented in this work is estimated using point clouds. This may lead to a transformation where the invariant point (point that does move after the transformation) is not the rotation centre when dealing with points covering non-compact objects as described in section 4.2.3. Future work may consider the use of conformal mapping based estimation of the rigid transformation. This approach considers a $3D$ object as a diffeomorphism of the unit sphere to itself (Narra et al., 2018).

8.2.2 Prediction from 2D/3D images

The establishment of a learning channel between the latent space and the observation through model fitting leads to various applications such as segmentation of $3D$ image volumes. Prediction from image volumes using DMFC-GPM requires a fitting method to retrieve the most relevant sample to be used. While sampling-based model-fitting approaches used in the DMFC-GPM framework inherently avoid convergence to a local minim/maximum, this process is time-consuming. Future work may consider investigating a parametric-based method for DMFC-GPM fitting similar to the shape model-fitting approach reported in (Lüthi et al., 2017). This could lead to a faster model-based segmentation compared to the sampling-based approach described in chapter 5.

Another challenge when making predictions from images using a DMFC-GPM prior is the necessity of the model to be flexible. For example, a way to increase the pose variability in the training data was proposed by using pose permutation models to obtain more generalised models (see section 4.5.3). However, the cost of this generalisation may be the encoding of unrealistic motion in the model for biological structures since biologically allowable motion is typically constrained by the shape. These challenges limited the real data analysis presented in this work. Moreover, while the shoulder abduction motion used in chapter 6 was simulated through a validated biomechanics protocol (Wu et al., 2005) which may not reflect reality. Furthermore, simulated fractures used in the experiments remain artificial and may not reflect commonly occurring real fractures.

Further work may consider building models with more data with large shape, pose and intensity variabilities. Such models would include high ranges of realistic motion and would be used to address biological and clinical questions. Moreover, it would be of interest to consider modelling humans joints with a large number of structures per training example. Such a model which may embed more correlations, which would ultimately improve prediction accuracy.

8.2.3 Modelling multi feature-class using deep learning

The core of the developed framework is based on principal component analysis (PCA). To address the non-linearity of pose variations, pose features were factorised from shapes and principal geodesic analysis (PGA) was applied to linearise the pose. However, to factorise pose and pose features, required the structures of interest to be segmented separately. For example, a multiple-structure joint (such as shoulder bones) should not be considered to be a single triangle mesh. One may want to model multi-objects as single structures. Moreover, shape and intensity variations are assumed linear (Cootes et al., 1998, 1992; Lüthi et al., 2017). While models based on this assumption have been reported to capture true shape and intensity, alternative methods based on deep learning have been reported. Variational autoencoders (VAEs) have shown to be efficient in modelling shape and pose as non-linear variations (Tan et al., 2018; Jiang et al., 2020; Xu et al., 2020). It would be of interest to develop shape, pose and intensity models using VAE and compare them with the DMFC-GPMs.

8.2.4 Synthesising high resolution image volumes

The DMFC-GPM framework estimates a latent space for shape, pose and the intensity profile of the interior of structures in the training data. One could use this model for image volume generation such as the estimation of super-resolution image volumes from low-resolution ones. However, the observed image volume may embed additional structures to the ones embedded in the models, for example, muscles surrounding bones. This limits a straightforward application of the DMFC-GPM framework for such a task. Thus, the model should not only include rigid structures being modelled but other rigid structures and soft tissues in the image.

Some work has been reported on the generation of a sequence of high resolution dynamic magnetic resonance images (MRI) from a single high-resolution image and a sequence of low-resolution frames (Makki et al., 2019). However, such a reconstruction needs patient-

specific labelled data, that is, a sequence of low-resolution image volumes and a single high-resolution one of the same patient. Obtaining labelled data is challenging due to the cost and imaging protocol constraints. Taking dynamic MRI of human joints with high ranges of motion such as the shoulder and the hip is extremely difficult. Embedding structures surrounding human joints in the DMFC-GPM could lead to a model that can be used to generate the frame sequence (volumetric shape-intensity at various poses). With such a model, a single high resolution volume would be enough to obtain the full intensity distribution of the image volumes. While the DMFC-GPM opens a new paradigm for statistical modelling, investigating above future directions would ultimately let to more advanced medical image analysis techniques.

8.3 Other research works and collaborations

This section presents outputs from other research projects where the author is a primary author or a collaborator.

Proceedings

- Tegang N.H.N., **Fouefack J.R.**, Borotikar B., Burdin V., Douglas T.S. and Mutsvangwa T.E., 2020. A Gaussian Process Model Based Generative Framework for Data Augmentation of Multi-modal 3D Image Volumes. In International Workshop on Simulation and Synthesis in Medical Imaging. Springer, Cham. pp. 90-100

- Thusini, X.O., Reyneke, C.J., Aellen, J., Forster, A., **Fouefack, J.R.**, Tegang, N.H.N., Vetter, T., Douglas, T.S. and Mutsvangwa, T.E., 2020, October. Uncertainty Reduction in Contour-Based 3D/2D Registration of Bone Surfaces. In International Workshop on Shape in Medical Imaging. Springer, Cham. pp. 18-29

- Inyang A.O., **Fouefack J.R.**, Sivarasu S., Roche S., Borotikar B., Burdin V. and Mutsvangwa T.E., 2017. Assessment of 3D morphological characteristics of the

shoulder bones using statistical shape modeling: Prospective application to handedness, 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC17), Seogwipo, Korea. pp. 1629-1632

- Narra N., **Fouefack J.R.**, Douglas T.S. and Mutsvangwa T.E., 2018. Conformal mapping of the human scapula to generate dense landmark feature. In IEEE 3rd Biennial South African Biomedical Engineering Conference (SAIBMEC). pp. 1-4

Presentations

- **Fouefack J.R**, Inyang A.O., Luethi, M., Borotikar B., Burdin V., Douglas, T.S. and Mutsvangwa T.E. Statistical analysis of dense correspondence and low sample size surface meshes. 7th World Congress of Biomechanics, July 2018, Dublin, Ireland

- **Fouefack J. R**., Borotikar B., Burdin V., Douglas T. S. and Mutsvangwa T. E. Automatic segmentation of the humerus from CT volumes using statistical shape models and Markov chain Monte Carlo methods. IEEE EMBS Summer School in Biomedical Engineering, June 2018, St Jacut, France.

- **Fouefack, J.R**., Borotikar, B., Burdin, V., Douglas, T.S. and Mutsvangwa, T.E., 2018, July. Automatic segmentation of the humerus from CT volumes using Statistical shape models and Markov Chain Monte Carlo methods. Poster at the Medical Image Summer Shool (MISS) on deep learning , Favignana, Italy.

References

Abler, D., Berger, S., Terrier, A., Becce, F., Farron, A. and Büchler, P. 2018, 'A statistical shape model to predict the premorbid glenoid cavity', *Journal of shoulder and elbow surgery* **27**(10), 1800–1808.

Agrawal, P., Mozingo, J. D., Elhabian, S. Y., Anderson, A. E. and Whitaker, R. T. 2020, Combined estimation of shape and pose for statistical analysis of articulating joints, *in* 'International Workshop on Shape in Medical Imaging', Springer, pp. 111–121.

Agrawal, P., Whitaker, R. T. and Elhabian, S. Y. 2017, Learning deep features for automated placement of correspondence points on ensembles of complex shapes, *in* 'In International Conference on Medical Image Computing and Computer-Assisted Intervention', Springer, pp. 185–193.

Ambellan, F., Zachow, S. and von Tycowicz, C. 2019, A surface-theoretic approach for statistical shape modeling, *in* 'International Conference on Medical Image Computing and Computer-Assisted Intervention', Springer, pp. 21–29.

Anas, E. M. A., Rasoulian, A., John, P. S., Pichora, D., Rohling, R. and Abolmaesumi, P. 2014, A statistical shape+pose model for segmentation of wrist ct images, *in* 'Medical Imaging 2014: Image Processing', p. 90340T.

Anas, E. M. A., Rasoulian, A., Seitel, A., Darras, K., Wilson, D., John, P. S., Pichora, D., Mousavi, P., Rohling, R. and Abolmaesumi, P. 2016, 'Automatic segmentation of

wrist bones in ct using a statistical wrist shape + pose model', *IEEE transactions on medical imaging* **35**(8), 1789–1801.

Anderson, S. E., Siebenrock, K. A. and Tannast, M. 2010, 'Femoroacetabular impingement: evidence of an established hip abnormality', *Radiology* **257**(1), 8–13.

Audenaert, E. A., Van Houcke, J., Almeida, D. F., Paelinck, L., Peiffer, M., Steenackers, G. and Vandermeulen, D. 2019, 'Cascaded statistical shape model based segmentation of the full lower limb in ct', *Computer methods in biomechanics and biomedical engineering* **22**(6), 644–657.

Audenaert, E., Van den Eynde, J., de Almeida, D., Steenackers, G., Vandermeulen, D. and Claes, P. 2020, 'Separating positional noise from neutral alignment in multicomponent statistical shape models', *Bone Reports* **12**, 100243.

Berlinet, A. and Thomas-Agnan, C. 2011, *Reproducing kernel Hilbert spaces in probability and statistics*, Springer Science & Business Media.

Besl, P. J. and McKay, N. D. 1992, Method for registration of 3-d shapes, *in* 'Sensor Fusion IV: Control Paradigms and Data Structures', Vol. 1611, International Society for Optics and Photonics, pp. 586–607.

Bhatia, K. K., Hajnal, J. V., Puri, B. K., Edwards, A. D. and Rueckert, D. 2004, Consistent groupwise non-rigid registration for atlas construction, *in* '2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821)', IEEE, pp. 908–911.

Bindernagel, M., Kainmueller, D., Seim, H., Lamecker, H., Zachow, S. and Hege, H.-C. 2011, An articulated statistical shape model of the human knee, *in* 'Bildverarbeitung für die Medizin 2011', Springer, pp. 59–63.

Blanc, R. and Székely, G. 2012, 'Confidence regions for statistical model based shape prediction from sparse observations', *IEEE Transactions on Medical Imaging* **31**(6), 1300–1310.

Bonaretti, S., Seiler, C., Boichon, C., Reyes, M. and Büchler, P. 2014, 'Image-based vs. mesh-based statistical appearance models of the human femur: implications for finite element simulations', *Medical engineering & physics* **36**(12), 1626–1635.

Borotikar, B., Lempereur, M., Lelievre, M., Burdin, V., Ben Salem, D. and Brochard, S. 2017, 'Dynamic mri to quantify musculoskeletal motion: A systematic review of concurrent validity and reliability, and perspectives for evaluation of musculoskeletal disorders', *PloS one* **12**(12), e0189587.

Boscaini, D., Masci, J., Rodolà, E. and Bronstein, M. 2016, Learning shape correspondence with anisotropic convolutional neural networks, *in* 'Advances in neural information processing systems', pp. 3189–3197.

Bossa, M. N. and Olmos, S. 2006, Statistical model of similarity transformations: Building a multi-object pose, *in* '2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06)', IEEE, pp. 59–59.

Bossa, M. N. and Olmos, S. 2007, Multi-object statistical pose+ shape models, *in* '2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro', IEEE, pp. 1204–1207.

Boutillon, A., Borotikar, B., Burdin, V. and Conze, P.-H. 2020, 'Multi-structure bone segmentation in pediatric mr images with combined regularization from shape priors and adversarial network', *arXiv preprint arXiv:2009.07092* .

Bredbenner, T. L., Mason, R. L., Havill, L. M., Orwoll, E. S., Nicolella, D. P. and in Men (MrOS) Study, O. F. 2014, 'Fracture risk predictions based on statistical shape and density modeling of the proximal femur', *Journal of Bone and Mineral Research* **29**(9), 2090–2100.

Brégier, R., Devernay, F., Leyrit, L. and Crowley, J. L. 2018, 'Defining the pose of any 3d rigid object and an associated distance', *International Journal of Computer Vision* **126**(6), 571–596.

Breton, D., Burdin, V., Leboucher, J. and Rémy-Néris, O. 2014, 'Study of the joint configuration of the knee using a morpho-functional analysis', *Innovation and Research in BioMedical engineering(IRBM)* **35**(1), 53–57.

Brunton, A., Salazar, A., Bolkart, T. and Wuhrer, S. 2016, Statistical shape spaces for 3d data: a review, *in* 'Handbook of Pattern Recognition and Computer Vision', World Scientific, pp. 217–238.

Campbell, J. and Petrella, A. 2016, 'Automated finite element modeling of the lumbar spine: Using a statistical shape model to generate a virtual population of models', *Journal of biomechanics* **49**(13), 2593–2599.

Chen, X., Graham, J., Hutchinson, C. and Muir, L. 2014, 'Automatic generation of statistical pose and shape models for articulated joints', *IEEE Transactions on Medical Imaging* **33**(2), 372–383.

Ciller, C., De Zanet, S. I., Rüegsegger, M. B., Pica, A., Sznitman, R., Thiran, J.-P., Maeder, P., Munier, F. L., Kowal, J. H. and Cuadra, M. B. 2015, 'Automatic segmentation of the eye in 3d magnetic resonance imaging: a novel statistical shape model for treatment planning of retinoblastoma', *International Journal of Radiation Oncology* Biology* Physics* **92**(4), 794–802.

Clohisy, J. C., Knaus, E. R., Hunt, D. M., Lesher, J. M., Harris-Hayes, M. and Prather, H. 2009, 'Clinical presentation of patients with symptomatic anterior hip impingement', *Clinical orthopaedics and related research* **467**(3), 638–644.

Cootes, T. F., Edwards, G. J. and Taylor, C. J. 1998, Active appearance models, *in* 'European conference on computer vision', Springer, pp. 484–498.

Cootes, T. F., Edwards, G. J. and Taylor, C. J. 2001, 'Active appearance models', *IEEE Transactions on pattern analysis and machine intelligence* **23**(6), 681–685.

Cootes, T. F., Hill, A., Taylor, C. J. and Haslam, J. 1994, 'Use of active shape models for locating structures in medical images', *Image and vision computing* **12**(6), 355–365.

Cootes, T. F., Taylor, C. J., Cooper, D. H. and Graham, J. 1992, Training models of shape from sets of examples, *in* 'BMVC92', Springer, pp. 9–18.

Cootes, T. F., Taylor, C. J., Cooper, D. H. and Graham, J. 1995, 'Active shape models-their training and application', *Computer vision and image understanding* **61**(1), 38–59.

Dao, T. T. and Ho Ba Tho, M.-C. 2015, 'Assessment of parameter uncertainty in rigid musculoskeletal simulation using a probabilistic approach', *Journal of Musculoskeletal Research* **18**(03), 1550013.

David, I., Philippe, M., Alexandre, T., Pralong, E. and Farron, A. 2012, Dynamical biomechanical model of the shoulder for muscle-force estimation, *in* 'Biomedical Robotics and Biomechatronics (BioRob), 2012 4th IEEE RAS & EMBS International Conference on', Ieee, pp. 407–412.

Demisse, G. G., Aouada, D. and Ottersten, B. 2017, 'Deformation based curved shape representation', *IEEE transactions on pattern analysis and machine intelligence* **40**(6), 1338–1351.

Dempster, A. P., Laird, N. M. and Rubin, D. B. 1977, 'Maximum likelihood from incomplete data via the em algorithm', *Journal of the Royal Statistical Society: Series B (Methodological)* **39**(1), 1–22.

Duchon, J. 1977, Splines minimizing rotation-invariant semi-norms in sobolev spaces, *in* 'Constructive theory of functions of several variables', Springer, pp. 85–100.

Eberly, D. 2000, *3D game engine design: a practical approach to real-time computer graphics*, Morgan Kaufmann Publishers Inc, San Francisco, CA, USA.

Ehlke, M., Ramm, H., Lamecker, H., Hege, H.-C. and Zachow, S. 2013, 'Fast generation of virtual x-ray images for reconstruction of 3d anatomy', *IEEE transactions on visualization and computer graphics* **19**(12), 2673–2682.

Fan, Y. and Wang, Y. 2020, Convolutional bayesian models for anatomical landmarking on multi-dimensional shapes, *in* 'International Conference on Medical Image Computing and Computer-Assisted Intervention', Springer, pp. 786–796.

Fitzpatrick, C. K., Baldwin, M. A., Laz, P. J., FitzPatrick, D. P., Lerner, A. L. and Rullkoetter, P. J. 2011, 'Development of a statistical shape model of the patellofemoral joint for investigating relationships between shape and function', *Journal of Biomechanics* **44**(13), 2446–2452.

Fletcher, P. T., Lu, C. and Joshi, S. 2003, Statistics of shape via principal geodesic analysis on lie groups, *in* '2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.', Vol. 1, IEEE, pp. I–I.

Fletcher, P. T., Lu, C., Pizer, S. M. and Joshi, S. 2004, 'Principal geodesic analysis for the study of nonlinear statistics of shape', *IEEE Transactions on Medical Imaging* **23**(8), 995–1005.

Fotsin, T. J. T., Vázquez, C., Cresson, T. and De Guise, J. 2019, Shape, pose and density statistical model for 3d reconstruction of articulated structures from x-ray images, *in* '2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)', IEEE, pp. 2748–2751.

Fouefack, J.-R., Alemneh, T., Borotikar, B., Burdin, V., Douglas, T. S. and Mutsvangwa, T. 2019, Statistical shape-kinematics models of the skeletal joints: Application to the shoulder complex, *in* '2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)', IEEE, pp. 4815–4818.

Fouefack, J.-R., Borotikar, B., Douglas, T. S., Burdin, V. and Mutsvangwa, T. 2020, Dynamic multi-object gaussian process models, *in* 'MICCAI, Lecture Notes in Computer Science Book Series 12264', Springer Nature Switzerland, pp. 755–764.

Fox, M. J., Scarvell, J. M., Smith, P. N., Kalyanasundaram, S. and Stachurski, Z. H. 2013, 'Lateral drill holes decrease strength of the femur: an observational study using fi-

nite element and experimental analyses', *Journal of Orthopaedic Surgery and Research* **8**(1), 1–8.

Gaweł, D., Główka, P., Kotwicki, T. and Nowak, M. 2018, 'Automatic spine tissue segmentation from mri data based on cascade of boosted classifiers and active appearance model', *BioMed Research International* **2018**.

Gee, A. H. and Treece, G. M. 2014, 'Systematic misregistration and the statistical analysis of surface data', *Medical image analysis* **18**(2), 385–393.

Gil, J. D., Stindel, E., Guillard, G., Hamitouche, C. and Roux, C. 2004, Detection of the functional knee center using the mean helical axis: Application in computer assisted high tibial osteotomy, *in* '2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821)', IEEE, pp. 1557–1560.

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

Gorczowski, K., Styner, M., Jeong, J. Y., Marron, J., Piven, J., Hazlett, H. C., Pizer, S. M. and Gerig, G. 2010, 'Multi-object analysis of volume, pose, and shape using statistical discrimination', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **32**(4), 652–661.

Gower, J. C. 1975, 'Generalized procrustes analysis', *Psychometrika* **40**(1), 33–51.

Gudbjartsson, H. and Patz, S. 1995, 'The rician distribution of noisy mri data', *Magnetic resonance in medicine* **34**(6), 910–914.

Harty, M. 1984, The anatomy of the hip joint, *in* 'Surgery of the hip joint', Springer, pp. 45–74.

Heimann, T. and Meinzer, H.-P. 2009, 'Statistical shape models for 3d medical image segmentation: a review', *Medical image analysis* **13**(4), 543–563.

Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K. Q. 2017, Densely connected convolutional networks, *in* 'Proceedings of the IEEE conference on computer vision and pattern recognition', pp. 4700–4708.

Ibáñez, L., Hamitouche, C., Boniou, M. and Roux, C. 2001, Morphogenesis-based deformable models application to 3d medical image segmentation and analysis, *in* 'International Conference on Medical Image Computing and Computer-Assisted Intervention', Springer, pp. 1369–1370.

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

Jacq, J.-J., Cresson, T., Burdin, V. and Roux, C. 2008, 'Performing accurate joint kinematics from 3-d in vivo image sequences through consensus-driven simultaneous registration', *IEEE Transactions on Biomedical Engineering* **55**(5), 1620–1633.

Jacq, J.-J., Schwartz, C., Burdin, V., Gérard, R., Lefevre, C., Roux, C. and Rémy-Néris, O. 2009, 'Building and tracking root shapes', *IEEE transactions on BioMedical Engineering* **57**(3), 696–707.

Janani, P., Premaladha, J. and Ravichandran, K. 2015, 'Image enhancement techniques: A study', *Indian Journal of Science and Technology* **8**(22), 1–12.

Jerbi, T., Burdin, V., Leboucher, J., Stindel, E. and Roux, C. 2012, '2-d–3-d frequency registration using a low-dose radiographic system for knee motion estimation', *IEEE Transactions on Biomedical Engineering* **60**(3), 813–820.

Jiang, B., Zhang, J., Cai, J. and Zheng, J. 2020, 'Disentangled human body embedding based on deep hierarchical neural network', *IEEE Transactions on Visualization and Computer Graphics* .

Kainmueller, D., Lamecker, H., Zachow, S. and Hege, H.-C. 2009, An articulated statistical shape model for accurate hip joint segmentation, *in* 'Conf Proc IEEE Eng Med Biol Soc', Vol. 2009, pp. 6345–6351.

Kendall, D. G. 1989, 'A survey of the statistical theory of shape', *Statistical Science* pp. 87–99.

Klima, O., Kleparnik, P., Spanel, M. and Zemcik, P. 2016, Intensity-based femoral atlas 2d/3d registration using levenberg-marquardt optimisation, *in* 'Medical Imaging 2016: Biomedical Applications in Molecular, Structural, and Functional Imaging', International Society for Optics and Photonics, p. 97880F.

Klinder, T., Wolz, R., Lorenz, C., Franz, A. and Ostermann, J. 2008, Spine segmentation using articulated shape models, *in* 'International conference on medical image computing and computer-assisted intervention', Springer, pp. 227–234.

Kozono, N., Okada, T., Takeuchi, N., Hamai, S., Higaki, H., Ikebe, S., Shimoto, T., Miake, G., Nakanishi, Y. and Iwamoto, Y. 2017, 'In vivo kinematic analysis of the glenohumeral joint during dynamic full axial rotation and scapular plane full abduction in healthy shoulders', *Knee Surgery, Sports Traumatology, Arthroscopy* **25**(7), 2032–2040.

Kraevoy, V. and Sheffer, A. 2005, Template-based mesh completion., *in* 'Symposium on Geometry Processing', Vol. 385, Citeseer, pp. 13–22.

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

Lang, Y., Lian, C., Xiao, D., Deng, H., Yuan, P., Gateno, J., Shen, S. G., Alfi, D. M., Yap, P.-T., Xia, J. J. et al. 2020, Automatic localization of landmarks in craniomaxillofacial cbct images using a local attention-based graph convolution network, *in* 'Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd Interna-

tional Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part IV 23', Springer International Publishing, pp. 817–826.

Lauzeral, N., Borzacchiello, D., Kugler, M., George, D., Rémond, Y., Hostettler, A. and Chinesta, F. 2019, 'Shape parametrization of bio-mechanical finite element models based on medical images', *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* **7**(5-6), 480–489.

Lebailly, F., Lima, L., Clairemidi, A., Aubert, B., Guerard, S., Chaïbi, Y., De Guise, J., Fontaine, C. and Skalli, W. 2012, 'Semi-automated stereoradiographic upper limb 3d reconstructions using a combined parametric and statistical model: a preliminary study', *Surgical and radiologic anatomy* **34**(8), 757–765.

Leboucher, J., Lempereur, M., Savéan, J., Burdin, V. and Rémy-Néris, O. 2010, Location of the ulno-humeral joint axis based on anatomical landmarks, *in* 'CAOS 2012: 10th Conference on Computer Assisted Orthopaedic Surgery'.

Lempereur, M., Brochard, S., Burdin, V. and Rémy-Néris, O. 2010, 'Difference between palpation and optoelectronics recording of scapular motion', *Computer Methods in Biomechanics and Biomedical Engineering* **13**(1), 49–57.

Lespasio, M. J., Sultan, A. A., Piuzzi, N. S., Khlopas, A., Husni, M. E., Muschler, G. F. and Mont, M. A. 2018, 'Hip osteoarthritis: a primer', *The Permanente Journal* **22**, 17–84.

Li, J., McWilliams, A. B., Jin, Z., Fisher, J., Stone, M. H., Redmond, A. C. and Stewart, T. D. 2015, 'Unilateral total hip replacement patients with symptomatic leg length inequality have abnormal hip biomechanics during walking', *Clinical Biomechanics* **30**(5), 513–519.

Li, W., Abram, F., Beaudoin, G., Berthiaume, M.-J., Pelletier, J.-P. and Martel-Pelletier, J. 2008, 'Human hip joint cartilage: Mri quantitative thickness and volume measure-

ments discriminating acetabulum and femoral head', *IEEE Transactions on Biomedical Engineering* **55**(12), 2731–2740.

Li, X., Liu, Y., Wang, Y., Wang, C., Wang, M. and Song, Z. 2018, 'Fast and globally optimal rigid registration of 3d point sets by transformation decomposition', *arXiv preprint arXiv:1812.11307* .

Li, Y. and Maguire, L. 2010, 'Selecting critical patterns based on local geometrical and statistical information', *IEEE transactions on pattern analysis and machine intelligence* **33**(6), 1189–1201.

Ludewig, P. M. and Braman, J. P. 2011, 'Shoulder impingement: biomechanical considerations in rehabilitation', *Manual therapy* **16**(1), 33–39.

Ludewig, P. M. and Cook, T. M. 2000, 'Alterations in shoulder kinematics and associated muscle activity in people with symptoms of shoulder impingement', *Physical therapy* **80**(3), 276–291.

Lüthi, M., Gerig, T., Jud, C. and Vetter, T. 2017, 'Gaussian process morphable models', *IEEE transactions on pattern analysis and machine intelligence* **40**(8), 1860–1873.

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

Ma, Y.-A., Chen, Y., Jin, C., Flammarion, N. and Jordan, M. I. 2019, 'Sampling can be faster than optimization', *Proceedings of the National Academy of Sciences* **116**(42), 20881–20885.

Madsen, D., Morel-Forster, A., Kahr, P., Rahbani, D., Vetter, T. and Lüthi, M. 2019, 'A closest point proposal for mcmc-based probabilistic surface registration', *arXiv preprint arXiv:1907.01414* .

Mahdi, F. P., Muto, T., Tanaka, H., Inui, H., Nobuhara, K. and Kobashi, S. 2020, 'Construction of 3-d humeral head statistical shape model in ct images', *Applied Sciences* **10**(16), 5591.

Main, C. J. and de C Williams, A. C. 2002, 'Musculoskeletal pain', *British Medical Journal* **325**(7363), 534–537.

Makki, K., Borotikar, B., Garetier, M., Brochard, S., Salem, D. B. and Rousseau, F. 2019, 'In vivo ankle joint kinematics from dynamic magnetic resonance imaging using a registration-based framework', *Journal of biomechanics* **86**, 193–203.

Markelj, P., Tomaževič, D., Likar, B. and Pernuš, F. 2012, 'A review of 3d/2d registration methods for image-guided interventions', *Medical image analysis* **16**(3), 642–661.

Martini, F., Timmons, M. J., Tallitsch, R. B., Ober, W. C., Garrison, C. W., Welch, K. B. and Hutchings, R. T. 2006, *Human anatomy*, Pearson/Benjamin Cummings San Francisco, CA.

Max, N. 1995, 'Optical models for direct volume rendering', *IEEE Transactions on Visualization and Computer Graphics* **1**(2), 99–108.

Mayya, M., Poltaretskyi, S., Hamitouche, C. and Chaoui, J. 2013, Scapula statistical shape model construction based on watershed segmentation and elastic registration., *in* 'ISBI', Citeseer, pp. 101–104.

Meyer, M., Barr, A., Lee, H. and Desbrun, M. 2002, 'Generalized barycentric coordinates on irregular polygons', *Journal of graphics tools* **7**(1), 13–22.

Moghari, M. H. and Abolmaesumi, P. 2007, 'Point-based rigid-body registration using an unscented kalman filter', *IEEE Transactions on Medical Imaging* **26**(12), 1708–1728.

Moreau, B., Gilles, B., Jolivet, E., Petit, P. and Subsol, G. 2017, A new metric for statistical analysis of rigid transformations: Application to the rib cage, *in* 'Graphs in

Biomedical Image Analysis, Computational Anatomy and Imaging Genetics', Springer, pp. 114–124.

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', Springer, pp. 137–146.

Murphy, N. J., Eyles, J. P. and Hunter, D. J. 2016, 'Hip osteoarthritis: etiopathogenesis and implications for management', *Advances in therapy* **33**(11), 1921–1946.

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

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

Myronenko, A., Song, X. and Carreira-Perpinán, M. A. 2007, Non-rigid point set registration: Coherent point drift, *in* 'Advances in Neural Information Processing Systems', pp. 1009–1016.

Nalisnick, E., Matsukawa, A., Teh, Y. W., Gorur, D. and Lakshminarayanan, B. 2018, 'Do deep generative models know what they don't know?', *arXiv preprint arXiv:1810.09136* .

Narra, N., Fouefack, J.-R., Douglas, T. and Mutsvangwa, T. 2018, Conformal mapping of the human scapula to generate dense landmark feature, *in* '2018 3rd Biennial South African Biomedical Engineering Conference (SAIBMEC)', IEEE, pp. 1–4.

Nikooyan, A., Veeger, H., Westerhoff, P., Bolsterlee, B., Graichen, F., Bergmann, G. and Van der Helm, F. 2012, 'An emg-driven musculoskeletal model of the shoulder', *Human Movement Science* **31**(2), 429–447.

Ouyang, W., Chu, X. and Wang, X. 2014, Multi-source deep learning for human pose estimation, *in* 'Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition', pp. 2329–2336.

Park, J., Choi, Y.-H., Chai, J. W., Cha, S. W., Lim, J. H., Jo, C. H. and Kim, D. H. 2019, 'Anterior capsular abnormality: another important mri finding for the diagnosis of adhesive capsulitis of the shoulder', *Skeletal radiology* **48**(4), 543–552.

Park, S. H., Lee, S., Yun, I. D. and Lee, S. U. 2012, Automatic bone segmentation in knee mr images using a coarse-to-fine strategy, *in* 'Medical Imaging 2012: Image Processing', International Society for Optics and Photonics, p. 831405.

Pedoia, V., Majumdar, S. and Link, T. M. 2016, 'Segmentation of joint and musculoskeletal tissue in the study of arthritis', *Magnetic Resonance Materials in Physics, Biology and Medicine* **29**(2), 207–221.

Pennec, X. 2006, 'Intrinsic statistics on riemannian manifolds: Basic tools for geometric measurements', *Journal of Mathematical Imaging and Vision* **25**(1), 127.

Pitocchi, J., Wirix-Speetjens, R., Van Lenthe, G. H. and Pérez, M. Á. 2020, 'Integration of cortical thickness data in a statistical shape model of the scapula', *Computer Methods in Biomechanics and Biomedical Engineering* pp. 1–7.

Preston, J. S., Joshi, S. and Whitaker, R. 2016, Deformation estimation with automatic sliding boundary computation, *in* 'International Conference on Medical Image Computing and Computer-Assisted Intervention', Springer, pp. 72–80.

Rasmussen, C. E. 2003, Gaussian processes in machine learning, *in* 'Summer School on Machine Learning', Springer, pp. 63–71.

Rasoulian, A., Abolmaesumi, P. and Mousavi, P. 2012, 'Feature-based multibody rigid registration of ct and ultrasound images of lumbar spine', *Medical Physics* **39**(6), 3154–3166.

Rasoulian, A., Rohling, R. and Abolmaesumi, P. 2012, 'Group-wise registration of point sets for statistical shape models', *IEEE Transactions on Medical Imaging* **31**(11), 2025–2034.

Rasoulian, A., Rohling, R. and Abolmaesumi, P. 2013, 'Lumbar spine segmentation using a statistical multi-vertebrae anatomical shape+ pose model', *IEEE transactions on medical imaging* **32**(10), 1890–1900.

Reyneke, C. J. F., Lüthi, M., Burdin, V., Douglas, T. S., Vetter, T. and Mutsvangwa, T. E. 2018a, 'Review of 2-d/3-d reconstruction using statistical shape and intensity models and x-ray image synthesis: Toward a unified framework', *IEEE reviews in biomedical engineering* **12**, 269–286.

Reyneke, C., Thusini, X., Douglas, T., Vetter, T. and Mutsvangwa, T. 2018b, Construction and validation of image-based statistical shape and intensity models of bone, *in* '2018 3rd Biennial South African Biomedical Engineering Conference (SAIBMEC)', IEEE, pp. 1–4.

Ribeiro, A. and Pascoal, A. G. 2012, 'The relationship between the humeral retroversion angle and the active shoulder rotation in volleyball players', *Journal of Biomechanics* **45**, S626.

Ribeiro, A., Pascoal, A. G. and Morais, N. 2012, The scapular contribution to the amplitude of shoulder external rotation on throwing athletes, *in* 'Technologies for Medical Sciences', Springer, pp. 227–242.

Rueckert, D. and Schnabel, J. A. 2020, 'Model-based and data-driven strategies in medical image computing', *Proceedings of the IEEE* **108**(1), 110–124.

Sadowsky, O., Ramamurthi, K., Ellingsen, L. M., Chintalapani, G., Prince, J. L. and Taylor, R. H. 2006, Atlas-assisted tomography: registration of a deformable atlas to compensate for limited-angle cone-beam trajectory, *in* '3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006.', IEEE, pp. 1244–1247.

Salhi, A., Burdin, V., Boutillon, A., Brochard, S., Mutsvangwa, T. and Borotikar, B. 2020, 'Statistical shape modeling approach to predict missing scapular bone', *Annals of Biomedical Engineering* **48**(1), 367–379.

Salhi, A., Burdin, V., Mutsvangwa, T., Sivarasu, S., Brochard, S. and Borotikar, B. 2017, Subject-specific shoulder muscle attachment region prediction using statistical shape models: A validity study, *in* '2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)', IEEE, pp. 1640–1643.

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

Schmid, J., Kim, J. and Magnenat-Thalmann, N. 2011, 'Robust statistical shape models for mri bone segmentation in presence of small field of view', *Medical image analysis* **15**(1), 155–168.

Schönborn, S., Egger, B., Morel-Forster, A. and Vetter, T. 2017, 'Markov chain monte carlo for automated face image analysis', *International Journal of Computer Vision* **123**(2), 160–183.

Sebastian, T. B., Tek, H., Crisco, J. J. and Kimia, B. B. 2003, 'Segmentation of carpal bones from ct images using skeletally coupled deformable models', *Medical Image Analysis* **7**(1), 21–45.

Seoud, L., Boisvert, J., Drouin, M.-A., Picard, M. and Godin, G. 2019, Increasing the robustness of cnn-based human body segmentation in range images by modeling sensor-specific artifacts, *in* 'Computer Vision – ECCV 2018 Workshops', pp. 729–743.

Smoger, L. M., Fitzpatrick, C. K., Clary, C. W., Cyr, A. J., Maletsky, L. P., Rullkoetter, P. J. and Laz, P. J. 2015, 'Statistical modeling to characterize relationships between knee anatomy and kinematics', *Journal of Orthopaedic Research* **33**(11), 1620–1630.

Soltanmohammadi, P., Elwell, J., Veeraraghavan, V., Athwal, G. S. and Willing, R. 2020, 'Investigating the effects of demographics on shoulder morphology and density us-

ing statistical shape and density modeling', *Journal of Biomechanical Engineering* **142**(12), 121005.

Sotiras, A., Davatzikos, C. and Paragios, N. 2013, 'Deformable medical image registration: A survey', *IEEE transactions on medical imaging* **32**(7), 1153–1190.

Stoll, C., Karni, Z., Rössl, C., Yamauchi, H. and Seidel, H.-P. 2006, Template deformation for point cloud fitting., *in* 'Symposium on Point-Based Graphics', pp. 27–35.

Sun, Z., He, Y., Gritsenko, A., Lendasse, A. and Baek, S. 2017, 'Deep spectral descriptors: Learning the point-wise correspondence metric via siamese deep neural networks', *arXiv preprint arXiv:1710.06368* .

Taghizadeh, E., Chandran, V., Reyes, M., Zysset, P. and Büchler, P. 2017, 'Statistical analysis of the inter-individual variations of the bone shape, volume fraction and fabric and their correlations in the proximal femur', *Bone* **103**, 252–261.

Tam, G. K., Cheng, Z.-Q., Lai, Y.-K., Langbein, F. C., Liu, Y., Marshall, D., Martin, R. R., Sun, X.-F. and Rosin, P. L. 2013, 'Registration of 3d point clouds and meshes: A survey from rigid to nonrigid', *IEEE transactions on visualization and computer graphics* **19**(7), 1199–1217.

Tan, Q., Gao, L., Lai, Y.-K. and Xia, S. 2018, Variational autoencoders for deforming 3d mesh models, *in* 'Proceedings of the IEEE conference on computer vision and pattern recognition', pp. 5841–5850.

Tanzer, M. and Noiseux, N. 2004, 'Osseous abnormalities and early osteoarthritis: the role of hip impingement', *Clinical Orthopaedics and Related Research®* **429**, 170–177.

Toshev, A. and Szegedy, C. 2014, Deeppose: Human pose estimation via deep neural networks, *in* 'Proceedings of the IEEE conference on computer vision and pattern recognition', pp. 1653–1660.

Umeyama, S. 1991, 'Least-squares estimation of transformation parameters between two point patterns', *IEEE Transactions on Pattern Analysis & Machine Intelligence* **13**(04), 376–380.

Van De Giessen, M., Streekstra, G. J., Strackee, S. D., Maas, M., Grimbergen, K. A., Van Vliet, L. J. and Vos, F. M. 2009, 'Constrained registration of the wrist joint', *IEEE Transactions on Medical Imaging* **28**(12), 1861–1869.

Van der Molen, H. F., Foresti, C., Daams, J. G., Frings-Dresen, M. H. and Kuijer, P. P. F. 2017, 'Work-related risk factors for specific shoulder disorders: a systematic review and meta-analysis', *Occupational and environmental medicine* **74**(10), 745–755.

Vlachopoulos, L., Lüthi, M., Carrillo, F., Gerber, C., Székely, G. and Fürnstahl, P. 2018, 'Restoration of the patient-specific anatomy of the proximal and distal parts of the humerus: statistical shape modeling versus contralateral registration method', *Journal of Bone and Joint Surgery Am.* **100**(8), e50.

von Tycowicz, C., Ambellan, F., Mukhopadhyay, A. and Zachow, S. 2018, 'An efficient riemannian statistical shape model using differential coordinates: with application to the classification of data from the osteoarthritis initiative', *Medical image analysis* **43**, 1–9.

Webster, R., Rabin, J., Simon, L. and Jurie, F. 2019, Detecting overfitting of deep generative networks via latent recovery, *in* 'Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition', pp. 11273–11282.

Weiskopf, D. 2007, *GPU-based interactive visualization techniques*, Springer.

Whittaker, E. T. 1937, *A treatise on the analytical dynamics of particles and rigid bodies*, CUP Archive.

Williams, C. K. and Rasmussen, C. E. 2006, *Gaussian processes for machine learning*, Vol. 2 of *3*, MIT press Cambridge, MA.

Wold, S., Esbensen, K. and Geladi, P. 1987, 'Principal component analysis', *Chemometrics and intelligent laboratory systems* **2**(1-3), 37–52.

Woon, Y. and Ham, R. 1991, 'Shoulder and hip joint range of motion in normal adults', *Journal of Korean Physical Therapy* **3**(1), 97–108.

Wu, G., Van der Helm, F. C., Veeger, H. D., Makhsous, M., Van Roy, P., Anglin, C., Nagels, J., Karduna, A. R., McQuade, K., Wang, X. et al. 2005, 'Isb recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion - part ii: shoulder, elbow, wrist and hand', *Journal of Biomechanics* **38**(5), 981–992.

Wyawahare, M. V., Patil, P. M., Abhyankar, H. K. et al. 2009, 'Image registration techniques: an overview', *International Journal of Signal Processing, Image Processing and Pattern Recognition* **2**(3), 11–28.

Xia, Y., Fripp, J., Chandra, S. S., Schwarz, R., Engstrom, C. and Crozier, S. 2013, 'Automated bone segmentation from large field of view 3d mr images of the hip joint', *Physics in Medicine & Biology* **58**(20), 7375.

Xia, Y., Manjón, J. V., Engstrom, C., Crozier, S., Salvado, O. and Fripp, J. 2017, Automated cartilage segmentation from 3d mr images of hip joint using an ensemble of neural networks, *in* '2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)', IEEE, pp. 1070–1073.

Xu, H., Bazavan, E. G., Zanfir, A., Freeman, W. T., Sukthankar, R. and Sminchisescu, C. 2020, Ghum & ghuml: Generative 3d human shape and articulated pose models, *in* 'Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition', pp. 6184–6193.

Yang, Z., Fripp, J., Chandra, S. S., Neubert, A., Xia, Y., Strudwick, M., Paproki, A., Engstrom, C. and Crozier, S. 2015, 'Automatic bone segmentation and bone-cartilage

interface extraction for the shoulder joint from magnetic resonance images', *Physics in Medicine & Biology* **60**(4), 1441.

Yao, J. and Taylor, R. H. 2001, Construction and simplification of bone density models, *in* 'Medical Imaging 2001: Image Processing', Vol. 4322, pp. 814–824.

Yeo, S., Romero, J., Loper, M., Machann, J. and Black, M. 2018, 'Shape estimation of subcutaneous adipose tissue using an articulated statistical shape model', *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* **6**(1), 51–58.

Yunpeng, L., Wenli, C., Guobin, H., Renfang, W. and Ran, J. 2018, Automatic segmentation of shoulder joint in mri using patch-based and fully convolutional networks, *in* '2018 25th IEEE International Conference on Image Processing (ICIP)', IEEE, pp. 3508–3512.

Zheng, M., Zou, Z., Bartolo, P. j. D. s., Peach, C. and Ren, L. 2017, 'Finite element models of the human shoulder complex: a review of their clinical implications and modelling techniques', *International Journal for Numerical Methods in Biomedical Engineering* **33**(2), e02777.

Zhou, S. K. and Xu, D. 2016, A probabilistic framework for multiple organ segmentation using learning methods and level sets, *in* 'Medical Image Recognition, Segmentation and Parsing', Elsevier, pp. 157–178.

Titre : **Vers un framework pour la modélisation statistique multi-classes de la forme, de l'intensité et de la cinématique dans les images médicales**

Mots clés : Espace latent de forme, pose et intensité, articulation humaine de l'épaule et de la hanche, analyse d'images médicale, reconnaissance de formes.

Résumé : Cette thèse porte sur le développement de modèles basés sur l'apprentissage statistique pour les complexes articulaires humains à partir d'images médicales. La principale contribution est une nouvelle formulation mathématique permettant de modéliser les systèmes musculo-squelettiques à l'aide d'un espace latent unifié qui intègre des caractéristiques de forme, de cinématique et d'intensité à partir d'observations données. Ce nouvel espace fournit un modèle continu capable de générer des instances en tirant parti des corrélations apprises entre caractéristiques. Une méthode d'ajustement est également proposée pour appliquer ce modèle sur de nouvelles données. La méthode complète de modélisation et de prédiction est validée à l'aide de données synthétiques entièrement contrôlées, et les résultats montrent que le modèle encapsule fidèlement les relations morpho-fonctionnelles imposées entre les objets, ainsi que leurs informations structurelles internes. Enfin, le modèle est utilisé pour analyser les articulations de l'épaule et de la hanche à partir d'images CT. L'intérêt clinique est que les corrélations de caractéristiques apprises par le modèle améliorent la prédiction de la forme pré-morbide et la précision de l'estimation du mouvement des articulations, à partir d'images médicales en deux ou trois dimensions.

Title: **Towards a framework for multi class statistical modelling of shape, intensity, and kinematics in medical images.**

Keywords: Shape-pose and intensity latent space, shoulder and hip human joint, medical image analysis, pattern recognition.

Abstract: This thesis focuses on the development of statistical learning-based models for human joints from medical images. The main contribution is a new mathematical formulation to model musculoskeletal systems by developing a unified computational latent space that embeds shape, kinematics and intensity features from given observations. This new space provides a continuous model capable of generating instances leveraging learnt feature correlations. A fitting method to apply models developed using this framework to unseen data is also proposed. The complete modeling and prediction framework is validated using bespoke synthetic data, showing that the framework faithfully encapsulates any prescribed morpho-functional relationships between objects, as well as their internal structural information. Finally, the framework is applied to the analysis of shoulder and hip joints from CT. The clinical interest is that the feature correlations learned by the model improve premorbid shape prediction and joint motion estimation accuracy, from two- and three-dimensional medical images.