Knowledge-based deformable models for medical image analysis

Schmid, Jérôme

UNIVERSITÉ DE GENÈVE

FACULTÉ DES SCIENCES

Département d'informatique — Professeur Thierry Pun

FACULTÉ DES SCIENCES
ÉCONOMIQUES ET SOCIALES

Département de systèmes d'information — Professeur Nadia Magnenat-Thalmann

Knowledge-based Deformable Models for Medical Image Analysis

THÈSE

présentée à la Faculté des sciences de l'Université de Genève
pour obtenir le grade de Docteur ès sciences, mention informatique

par

Jérôme SCHMID
de
Buchs (AG)

Thèse N^{o} 4284

GENÈVE
Atelier d'impression ReproMail

2011

UNIVERSITÉ DE GENÈVE

FACULTÉ DES SCIENCES

Doctorat ès sciences
Mention Informatique

Thèse de *Monsieur Jérôme Frédéric SCHMID*

intitulée :

" Knowledge-based Deformable Models for Medical Image Analysis "

La Faculté des sciences, sur le préavis de Madame N. MAGNENAT THALMANN, professeure honoraire et directrice de thèse (Faculté des sciences économiques et sociales Département de systèmes d'information), de Messieurs Th. PUN, professeur ordinaire et codirecteur de thèse (Département d'informatique), H. DELINGETTE, docteur (Institut National de Recherche en Informatique et en Automatique, Analyse et Simulation d'Images Biomédicales, Sofia-Antipolis, France), N. NAVAB, professeur (Computer Aided Medical Procedures & Augmented Reality, Fakultät für Informatik, Technische Universität Münschen, Deutschland), autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 17 janvier 2011

Thèse - 4284 -

Le Doyen Jean-Marc TRISCONE

N.B. - La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives aux thèses de doctorat à l'Université de Genève".

Acknowledgments

I would like to deeply thank my supervisor Prof. Nadia Magnenat-Thalmann for her permanent support and vision throughout this PhD, during which I had the chance to learn a lot more than scientific knowledge and skills. Many thanks to the jury members: Prof. Thierry Pun (University of Geneva), Dr. Hervé Delingette (INRIA Sophia-Antipolis) and Prof. Nassir Navab (Technische Universität München) for their time and patience in reviewing my manuscript.

I would like also to thank all my former and current colleagues of MIRALab with whom I experienced great moments and shared valuable scientific insights during these past years. I would like to acknowledge first of all the medical group: Caecilia Charbonnier, Lazhari Assassi, Pascal Volino, Jinman Kim, Benjamin Gilles and Gwenael Guillard. I have also really enjoyed the close collaboration with Niels Nijdam and Seunghyun Han from the network and mobility group. Warm thanks to the designers, Marlène Arévalo-Poizat and Nedjma Cadi-Yazli, as well as to Ghislaine Wharton in helping me in management and reporting tasks. Great thanks to the other MIRALab members in keeping a nice working atmosphere: Ugo Bonanni, Bart Kevelham, Mustafa and Zerrin Kasap, Maher Ben Moussa, Sylvain Chagué, Yeonha An, Christiane Luible, Mélanie Montagnol and Etienne Lyard.

This work would not have been possible without the support of the Marie Curie RTN project "3D Anatomical Human" (3DAH) funded by the European Union. In addition to the financial support, the 3DAH project offered me the great opportunity to work with many international early stage and experienced researchers. In particular, special credits go to the 3DAH ESRs José A. Iglesias-Guitián, François Chung, Anders Sandholm, Bailiang Chen, Caroline Öhman and Ioana Ciuciu for their precious collaboration in GPU, segmentation, simulation, MRI acquisition, soft-tissue modeling and knowledge management, respectively. I am also very grateful to all the Scientists in charge and staff of the 3DAH partners who welcomed and assisted me during my different stays.

Many thanks to the University Hospital of Geneva, Dr. Frank Kolo-Crhistophe, Dr. Jacques Ménétrey, Dr. Christoph Becker and Dr. Pierre Hoffmeyer for their clinical support and expertise. Warm thanks to all the anonymous volunteers and dancers from the Ballet of the Grand Théâtre de Genève.

I would like to deeply acknowledge Tobias Heimann, Jinman Kim and Yan Bailly for their scientific expertise, friendship and patience in proof-reading my publications and thesis manuscript. Many thanks also to all my friends of Switzerland, France, Italy, Spain and China for their en-

couragements.

Finally, I would like to thank in my particular my family: my mother, sister and brother for their unconditional love; my aunt, uncle, cousins and grand-mothers for their infallible support. Last but not least, this work would never have been possible without the support, trust and love of my beloved wife Yan: xiexie wode ai.

Jérôme Schmid

Abstract

The treatment of musculoskeletal disorders (MSD) is of paramount importance, as MSDs are chronic pathologies accounting for the largest fraction of temporary and permanent joint disabilities in industrialized societies. Surgical treatments can restore joint mobility but require high precision and time-consuming complex operations, which thus necessitate the use of computer-aided systems (CAS). Accurate modeling of musculoskeletal structures is hence crucial for diagnosis and pre-operative planning and usually relies on Magnetic Resonance Imaging (MRI) for its ability to simultaneously image soft and bony structures composing the articulations. However, MRI images present, in addition to ubiquitous image artifacts, highly heterogeneous bone tissue intensities and articular structures with fuzzy boundaries. Furthermore, clinical MRI images are often acquired with limited field of view (FOV) or low image resolution, seriously compromising its use in automated image processing for CAS.

In this thesis, we focus on the design of efficient image segmentation methods to create subject-specific musculoskeletal models from MRI data. We propose a novel prior knowledge construction for use in discrete deformable models to yield an effective modeling of musculoskeletal structures. The use of prior knowledge is crucial because conventional approaches that do not use knowledge (e.g., thresholding and region growing) or exploit weak priors (e.g., image registration and deformable models) are not optimal for MRI segmentation. To this end, our research innovates in the derivation of new knowledge-based deformable models which preserve the strengths of deformable models (sub-voxel accuracy, computational efficiency) while offering a high level of robustness.

Our methodology covers the key processes of knowledge-based deformable models which are the construction of priors, initialization, and evolution of deformable models. We provide guidelines to build robust statistical shape and appearance priors for structures of interest by addressing the issues of shape correspondence and training database enlargement. The initialization of models in images with varied characteristics (resolution, FOV) is addressed with our proposed registration-based methods that exploit prior knowledge expressed as appearance (atlas) or kinematical constraints. We demonstrate that these methods are robust and accurate for subsequent segmentation based on the deformable model evolution. Our model evolution is based on a multi-resolution scheme that takes advantage of global (point distribution model) and local (Markov random field-based modeling of variations) shape priors to confer to the segmentation robustness and accuracy, respectively. In addition, speed and interactivity of the segmentation are optimized with parallelized implementations using Graphics Processing Units (GPU).

The usability of our algorithm is demonstrated with its application to subject-specific models for physio-anatomical simulations (e.g. soft tissue simulation of hip and knee articulations, neuro-muscular simulation of the lower limbs), as well as the simultaneous registration and segmentation of anatomical structures from multiple images. Furthermore, an interactive segmentation system is developed for concurrent and collaborative environment, in which multiple users simultaneously segment the same image. Our results suggest that this work directly contributes to the challenging objectives of the virtual physiological human research which aims at integrating effective CAS tools into the clinical environment to support e.g. aided diagnosis, surgical planning and navigation, predictive medicine and post-operative treatment.

Résumé

Le traitement des troubles et maladies musculosquelettiques est d'une importance capitale, puisque ces pathologies chroniques représentent la plus grande fraction des handicaps temporaires et permanents dans les sociétés industrialisées. Il existe des traitements chirurgicaux pouvant restaurer la mobilité articulaire mais ils nécessitent une grande précision ainsi que des opérations longues et complexes. Dans ce contexte, l'utilisation de systèmes assistés par ordinateur se révèle être un atout essentiel. La modélisation précise de structures musculosquelettiques est alors cruciale, aussi bien pour le diagnostic que pour la planification préopératoire, et celle-ci repose généralement sur l'imagerie par résonance magnétique (IRM) qui permet d'acquérir à la fois les structures molles et osseuses des articulations. Cependant, le traitement de ces images IRM s'avère difficile car les images sont entachées de nombreux artefacts et les structures osseuses présentent des intensités très inhomogènes ainsi que des contours souvent peu nets dans la région articulaire. Si l'on ajoute le fait que les images IRM cliniques sont souvent acquises avec un champ de vision limité ou possèdent une très faible résolution d'image, le traitement automatique de ces images constitue un véritable défi.

Dans cette thèse, nous nous concentrons sur la conception de méthodes efficaces de segmentation d'image afin de créer à partir de données IRM des modèles musculosquelettiques spécifiques aux patients. Pour atteindre ce but, nous proposons de combiner l'utilisation de modèles déformables discrets avec de la connaissance *a priori*. L'utilisation de cette connaissance a priori est essentielle car les approches conventionnelles n'exploitant pas (par ex. les techniques de seuillage ou de croissance de région) ou peu (par ex. le recalage d'images ou les modèles déformables) de connaissance ne s'avèrent pas optimales pour la segmentation de structure musculosquelettiques à partir d'images IRM. Notre recherche innove dans l'expression de nouveaux modèles déformables basés sur la connaissance a priori qui conservent les points forts des modèles déformables (précision subvoxelique, efficacité en termes de temps de calcul) tout en offrant une grande robustesse.

Notre méthodologie couvre les principaux aspects de ce type de modèles déformables, à savoir l'expression de la connaissance a priori, l'initialisation et l'évolution des modèles. La création de modèles statistiques de forme des structures d'intérêt est notamment étudiée en détails, en abordant les questions de correspondance et d'élargissement de la base de données d'entrainement. L'initialisation des modèles déformables dans des images aux caractéristiques très variables (différences de résolution ou de champ de vision) est réalisée à l'aide de méthodes de recalage qui exploitent des connaissances a priori sur l'apparence et la forme des structures à segmenter.

Nous montrons que ces méthodes d'initialisation sont suffisamment robustes et précises pour être utilisées dans toute segmentation ultérieure basée sur l'évolution des modèles déformables. Notre évolution utilise une approche multi-résolution qui repose sur une modélisation des variations de forme à la fois globale (modèle déformable statistique PDM) et locale (modélisation par champs aléatoires de Markov), ce qui confère robustesse et précision à la segmentation. En outre, la rapidité et l'interactivité de la segmentation sont optimisées par une programmation parallèle basée sur les nouvelles technologies des processeurs graphiques.

L'usage de notre méthodologie est illustré dans le contexte de la simulation physio-anatomique (par ex. la simulation des tissus mous des articulations de la hanche et du genou, ou bien la simulation neuromusculaire des membres inférieurs), ainsi que dans des applications de recalage et segmentation simultanés de structures anatomiques à partir de plusieurs images. De plus, nous présentons aussi un système de segmentation interactive et collaborative qui permet à plusieurs utilisateurs de segmenter ensemble et en même temps une même image.

Notre travail contribue aux objectifs ambitieux de la recherche de l'"humain physiologique virtuel" qui vise à intégrer des outils efficaces informatiques dans le milieu clinique, tels que le diagnostic assisté par ordinateur, la planification et la navigation chirurgicale, la médecine prédictive et le traitement postopératoire.

List of Publications

Most of the material presented in this thesis previously appeared in the following publications.

Peer-reviewed Journals

- J. Schmid, J. Kim, and N. Magnenat-Thalmann, "Robust Statistical Shape Models for MRI Bone Segmentation in Presence of Small Field of View," *Medical Image Analysis*, vol. 15, no. 1, pp. 155–168, 2011.
- J. Schmid, J. Kim, and N. Magnenat-Thalmann, "Extreme leg motion analysis of professional ballet dancers via MRI segmentation of multiple leg postures," *International Journal of Computer Assisted Radiology and Surgery*, vol. 6, no. 1, pp. 47–57, 2011.
- J. Schmid, J. Iglesias Guitián, E. Gobbetti, and N. Magnenat-Thalmann, "A GPU framework for parallel segmentation of volumetric images using discrete deformable models," *The Visual Computer Journal, Special Issue 3DAH*, vol. 27, no. 2, pp. 85–95, 2011.
- F. Chung, J. Schmid, N. Magnenat-Thalmann, and H. Delingette, "Comparison of statistical models performance in case of segmentation using a small amount of training datasets," *The Visual Computer Journal, Special Issue 3DAH*, vol. 27, no. 2, pp. 141–151, 2011.
- S. Han, N. Nijdam, J. Schmid, J. Kim, and N. Magnenat-Thalmann, "Collaborative telemedicine for interactive multiuser segmentation of volumetric medical images," *The Visual Computer Journal, Proc. CGI 2010*, vol. 26, no. 6-8, pp. 639–648, 2010.
- L. Assassi, C. Charbonnier, J. Schmid, P. Volino, and N. Magnenat-Thalmann, "From MRI to Anatomical Simulation of the Hip Joint," *Computer Animation and Virtual Worlds Journal, Special Issue on Physiological Human*, vol. 20, no. 1, pp. 53–66, 2009.

Peer-reviewed Conferences

- J. Schmid, J. Kim, and N. Magnenat-Thalmann, "Coupled Registration-Segmentation: Application to Femur Analysis with Intra-subject Multiple Levels of Details MRI Data," in *Proc. MICCAI*, vol. LNCS 6362, no. 2. Springer-Verlag, 2010, pp. 562–569.
- N. Magnenat-Thalmann, J. Schmid, L. Assassi, and P. Volino, "A Comprehensive Methodology to Visualize Articulations for the Physiological Human," in *Proc. Cyberworlds*. IEEE Computer Society, DOI:10.1109/CW.2010.41, 2010.
- J. Schmid, N. Nijdam, A. Han, J. Kim, and N. Magnenat-Thalmann, "Interactive Segmentation of Volumetric Medical Images for Collaborative Telemedicine," in *Modelling the Physiological Human, Proc. 3DPH*, vol. LNCS 5903. Springer, 2009, pp. 13–24.
- C. Charbonnier, J. Schmid, F. Kolo-Christophe, N. Magnenat-Thalmann, C. Becker, and P. Hoffmeyer, "Virtual Hip Joint: from Computer Graphics to Computer-Assisted Diagnosis," in *Eurographics 2009 - Medical Prize (1st Prize awarded)*. Eurographics Association, 2009, pp. 1–4.

– J. Schmid, A. Sandholm, F. Chung, D. Thalmann, H. Delingette, and N. Magnenat-Thalmann, "Musculoskeletal simulation model generation from MRI datasets and motion capture data," in *Recent Advances in the 3D Physiological Human*. Springer-Verlag, 2009, pp. 3–19.
– J. Schmid and N. Magnenat-Thalmann, "MRI Bone Segmentation using Deformable Models and Shape Priors," in *Proc. MICCAI*, vol. LNCS 5241. Springer-Verlag, 2008, pp. 119–126.
– N. Magnenat-Thalmann, C. Charbonnier, and J. Schmid, "Multimedia Application to the Simulation of Human Musculoskeletal System: A Visual Lower Limb Model from Multimodal Captured Data," in *Proc. MMSP*. IEEE Publisher, 2008, pp. 520–525.

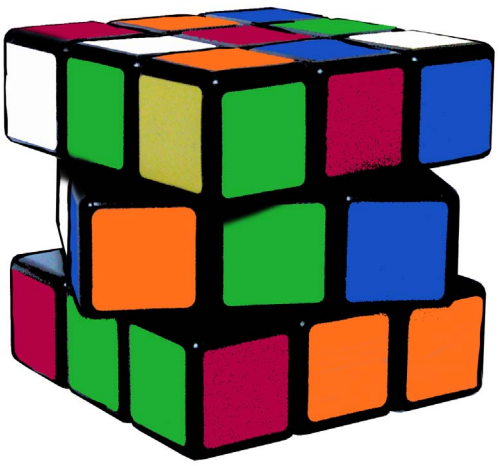
Contents

List of Figures

List of Tables

Chapter 1

Introduction



1.1 Medical Context and Motivations

In the last decade, the demand for computer-assisted support in health and medicine has been constantly increasing. One of the main reasons stems from the *revolution of the digital information* in the clinical environment. Thanks to ground breaking advances in engineering, physics, chemistry and biology, medical equipments have gone through revolutionary transformations. Current equipments are constantly evolving and are able to nowadays produce highly detailed information of the human body. New challenges are introduced by the production of a tremendous amount of information that can only be managed and processed by appropriate computer-based solutions. Another explanation comes from the capability of computer systems to create virtual environments to *model, examine and simulate* medical scenarios. By devising adequate models and simulation procedures, an infinite number of testing possibilities are made available to the scientific and medical community to improve the actual treatment.

In this era of technological revolution, the synergy between Health and IT is really expected to improve the understanding of the human system with related pathologies. Among them, the treatment of musculoskeletal disorders (MSD) is of paramount importance. MSDs are chronic pathologies that can be found at all levels of the neuromusculoskeletal system, from neurological disorders to organ malfunctionalities. They account for the largest fraction of temporary and permanent disabilities and are closely monitored in industrialized societies [RHA+98, WBPR+04, CACC05] since they represent a significant economic burden for western societies [Reg02].

In MSDs, osteoarthritis (OA) is a highly prevalent pathology. OA is characterized by a degeneration of the articular cartilage which yields pain, loss of mobility and stiffness of human articulations. Abnormal bone morphology observed as hip femoroacetabular impingements (i.e. abnormal frictions and collisions between articular structures) is known to be at the origin of early OA [GPB+03, PMD+06], but there is still some idiopathic OA observed in people who routinely perform extreme motions. Motion analysis is hence essential to fully understand OA.

Despite the staggering complexity of the human body, the study of sub-domains such as the musculoskeletal system can be carried out with "simpler" yet realistic models to make the problem tractable and manageable (limited time, computational power, etc.). Furthermore, there is a particular interest in building patient-specific anatomical models to support:

- Medical diagnosis: as previously mentioned, some forms of OA are related to the bone morphology. By reconstructing accurate bone models from radiological images, the computation of morphological measures [PMD+06] can be automated or simplified compared to the time-consuming reading of the images.
- Motion analysis: motion is intimately linked to the musculoskeletal system. The tracking of bone structures in different postures in images [MKP+05] provides essential insights into human motion non-invasively. Moreover, subject-specific bone models are useful to compute scaling factors in neuromusculoskeletal simulation [DAA+07], where a generic model is scaled according to the subject's limb attributes (weight, length, etc.).
- Soft-tissue simulation: the study of a joint involves the simulation of the various structures

in contact such as cartilages, bones and ligaments. Anatomical models are essential since their shape and associated mechanical properties have a direct influence on the simulation outcomes.

MIRALab quickly noticed the huge potentials that lied in the research of anatomical modeling. Since its foundation in 1989 by Prof. Magnenat-Thalmann, MIRALab has expanded and adapted its core researches in computer graphics (e.g., cloth simulation [VT00, VMT05, VMTF09]) and virtual humans towards medical modeling and simulation [KBG+95, GKB+96, MWMTT98, YCGMMT03, MTYCS04, GMMT06, MTCS08]. In particular, soft tissue simulations of the articular structures driven by kinematical information [CLMT08, ACS+09, CAVMT09] were carried out during the projects CHARM [CHA96], Co-Me [COM09] and 3D Anatomical Human (3DAH) [3DA10]. We are particularly involved in this last project which gathers several international partners in a multi-disciplinary framework. Its main objective is the development of realistic functional 3D models of the human musculoskeletal system, with the methodology demonstrated on the lower limbs. In this project, multiple disciplines are integrated such as image processing and acquisition, motion analysis, biomechanical and soft-tissue simulation, interactive visualization and knowledge management.

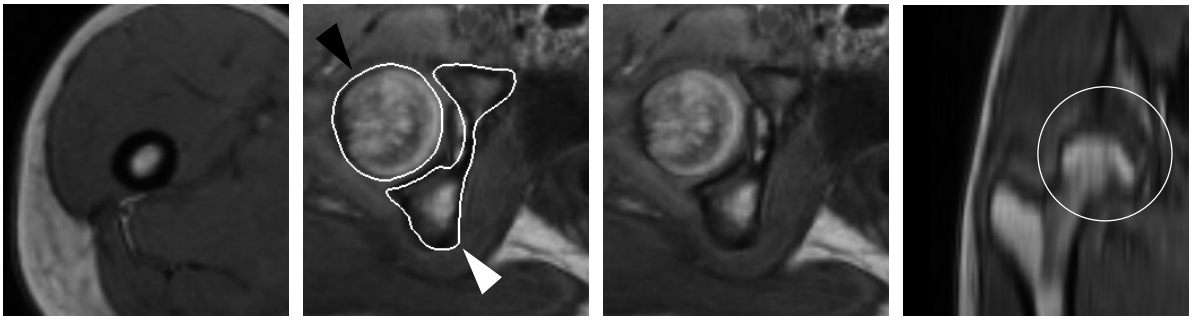
1.2 Research Context and Objectives

Image-based anatomical modeling is one of the most promising way to create subject-specific models [BAGD07, SJL+07]. Anatomical modeling from medical images addresses in particular the problem of image segmentation. Segmentation consists in extracting anatomical structures from medical data acquisitions by dividing an image into non-overlapping and homogeneous regions (each one corresponding to a unique structure of interest, e.g. an organ) and by creating 3D models from the segmented data. A trained operator in reading medical images and with an appropriate anatomical knowledge is able to perform segmentations with high accuracy, despite the presence of image artifacts (e.g., noise, partial volume effect (PVE)), the human body complexity or the large differences persisting between individuals. Conversely, these factors seriously impede computer-based segmentation approaches. Simple approaches based on low-level image processing such as thresholding, edge detection or region growing will very likely fail with clinical images.

In case of the musculoskeletal system, the rich variety of soft and bony structures demands the use of appropriate imaging modalities. Computed Tomography (CT) based on X-ray imagery offers an excellent bone contrast but is improper to highlight soft tissue such as ligaments. Conversely, Magnetic Resonance Imaging (MRI) can simultaneously image hard and soft structures and MRI does not create any harmful ionizing radiation [SC04]. Furthermore, its richness in imaging protocols and capabilities (e.g., muscle fibers direction detection in dtMRI [HSV+05]) as well as its constant technological improvements (e.g., 3T scanners [BSP+07]), make MRI a natural choice for musculoskeletal segmentation.

However, MRI musculoskeletal segmentation introduces a series of challenging issues:

- Numerous musculoskeletal structures such as muscles are in perpetual contact which increases the complexity of the modeling and prevents the formation of clear delineations between similar structures in the images.
- Bone intensity is highly heterogeneous due to differences in cortical and trabecular (cancellous) bone [LFG+98] (Fig. 1.1(a)). Furthermore, subchondral bone, located below cartilage tissue at joint interfaces, tends to present a different structure (e.g., thinner, contrast discrepancies) with respect to the surrounding cortical bone (Fig. 1.1(c)). These factors, combined with any present pathology (e.g., osteoporosis, calcifications) seriously impede segmentation methods [WGE06].
- The segmentation of articular areas is hindered by the strong proximity of bones and the absence of clear boundaries [SVG02] (Fig. 1.1(c)).
- There is a ubiquitous trade-off between image field of view (FOV) and resolution. As a consequence large FOV yields a very low image resolution, which necessarily results in strong partial volume effect (Fig. 1.1(d)).



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

Figure 1.1: Common issues found in MRI musculoskeletal segmentation. (a) In a transverse cross-section, we can clearly see that bone tissue intensity is not homogeneous as cortical bone appears dark while trabecular bone is brighter. (b) At the joint level, bone contours of femur (F) and hip bone (HB) are drawn in white, F and HB being pointed by a black and white arrow, respectively. (c) It can be noticed that bone boundaries are fuzzy and unclear, and that average bone intensities are quite different, since F is much more brighter than HB. Moreover, compared to (a), cortical bone is much more thinner and almost not visible. (d) In the coronal plane, the partial volume effect due to high slice thickness (low image resolution in this direction) is obvious and clearly affects the delineation of the bones, especially at the joint level (circled area).

In this thesis, we aim at addressing these issues efficiently to support both medical and biomechanical studies. In particular, our methodology leads to:

- Improved morphological bone analysis to better detect potential causes of OA.
- Biomechanical simulation of human articulations (knee within 3DAH project [3DA10], hip within Co-Me [COM09] and MIRALab activities [CLMT08, ACS+09, CAVMT09]) based on subject-specific anatomical models to better understand their behavior during extreme and daily activities.
- Motion analysis by bone structure tracking in image sequence to quantify extents and visualize soft tissue deformations.

Our work relies on some of the foundations laid by the work of Gilles [Gil07], who proposed an implementation of simplex meshes [Del99] evolving as dynamic particle systems [GMMT06, GMT10]. Despite some very good results of Gilles *et al.* 's approach, the work presented some

weaknesses. The deformable models proposed by the author were quite sensitive to the initialization which was based on the manual positioning of landmarks in images. In addition to not being automatic, this initialization was in some cases not applicable: varying and especially small image FOV prevented the identification of a sufficient number of landmarks, which ultimately affected the quality or viability of the initialization. Furthermore, poor image resolution and heavy presence of artifacts seriously affected the quality of this deformable model-based segmentation. Whenever possible, the author partially addressed this issue by using simultaneously multiple images of the region to segment, which improved the quality of the image information. However, it was assumed that the images were aligned together, a condition which is not often respected (in practice images can be produced at different moments in time, subjects may move, etc.).

In this thesis, our focus is to improve musculoskeletal modeling by relying on *strong prior knowledge* during the segmentation process. Spanning the last two decades, knowledge-based deformable models combined with statistical shape and appearance priors (i.e. statistical shape models (SSM)), have been established as one of the most advanced and robust segmentation methods [HM09]. Prior-knowledge has the capability to robustify deformable models approaches to better face the image artifacts and variable image characteristics (unpredictable FOV and different image resolution). However, the key success of these knowledge-based approaches stands in the construction and usage of the priors. Our work aspires to fulfill the following research objectives:

- Efficient expression of shape and appearance priors from training data, which might be incomplete or partially erroneous. This can relax the constraint of having a large collection of complete datasets, which is not always guaranteed (especially in 3D [HM09]).
- Design of initialization approaches requiring none or very little user intervention, and which are applicable on images with heterogeneous characteristics.
- Implementation of a robust, accurate and interactive evolution of knowledge-based deformable models to yield patient-specific anatomical models from MRI images. Interactivity is ensured by a fast evolution and the presence of segmentation control mechanisms.

1.3 Contributions

In this thesis, we demonstrate how the limitations of deformable models are efficiently tackled by exploiting prior-knowledge of the shape and appearance of the structures to segment. The resulting statistical shape and appearance models were extensively tested on a large collection of MRI datasets to qualitatively and quantitatively assess their performance in segmenting musculoskeletal structures. Emphasis was put on bone segmentation because a training dataset was more easily constructed for bones. In fact, (semi-)manual segmentation was easier to perform for bony structures than soft tissue, mainly because in many images at our disposal the image quality was not satisfactory enough to accurately segment soft organs.

Nevertheless bone segmentation from MRI was depicted as very challenging in Sec. 1.2 and

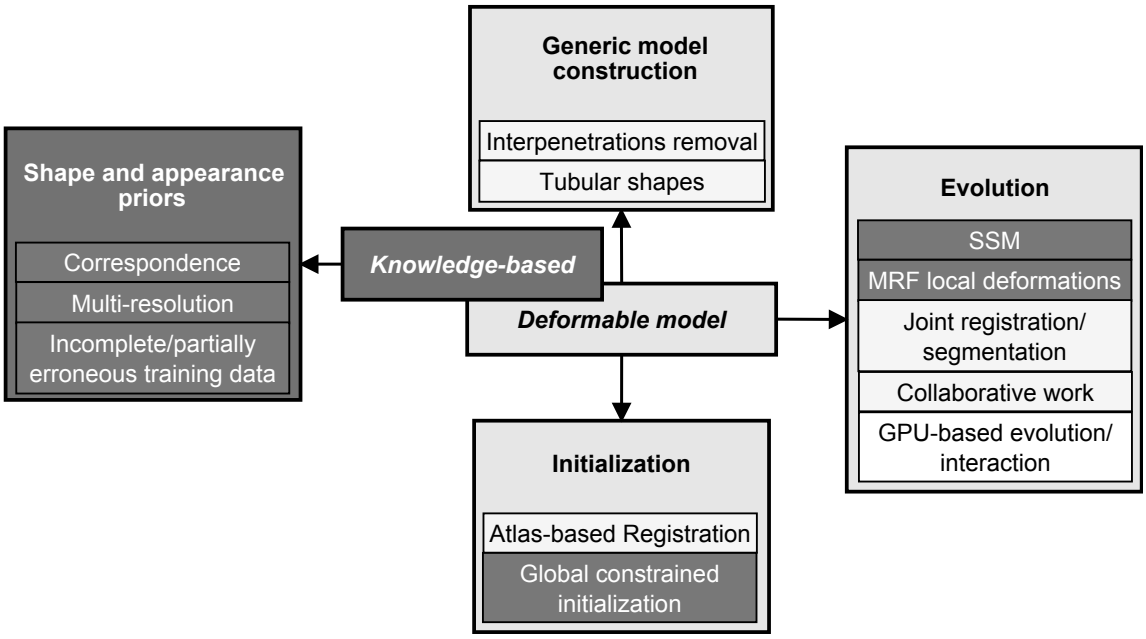


Figure 1.2: Overview of contributions of our knowledge-based deformable model for segmentation. Two main research directions are identified: the first deals with the construction and usage of priors (dark boxes), while the second addresses the key stages of deformable models, which are initialization, evolution and preparation (light boxes).

our work contributes to this very active domain of research. Except from a novel initialization specific to the hip segmentation, our proposed methodology to build and use robust SSMs can be applied without any restriction to the segmentation of other structures. Main contributions are mostly divided into two parts, (i) aspects related to prior-knowledge and (ii) concepts which pertain to deformable models. Particularly, contributions were made in the following topics (see Fig. 1.2):

1. **Generic model construction**
   - Specific approaches to build thin and elongated tubular structures.
   - Inter-penetrations removal among segmented structures.
2. **Knowledge inference from training data**
   - Multi-level SSMs with effective tackling of the point correspondence and usage of different alignment transforms.
   - Robust construction of priors from corrupted data characterized by the absence of information or the presence of non-reliable data.
3. **Deformable model initialization**
   - Exploration of atlas-based registration.
   - Design of an initialization approach for hip joint bone segmentation based on a constrained global optimization procedure.
4. **Deformable model evolution**
   - Effective coupling of SSM with Newtonian evolution of deformable models.
   - Expression of local variations priors based on Markov Random Fields (MRF).

- Support of the simultaneous segmentation of structures from multiple images based on a novel coupled registration-segmentation approach.
- Implementation of discrete deformable models in Graphics Processing Unit (GPU) architecture to improve segmentation interactivity.
- Investigation of concurrent and collaborative segmentation with multiple users over a network.

Part of this work was made in close collaboration with the 3DAH partners, especially people from UCL (MRI acquisition), INRIA (alternative appearance priors) and CRS4 (GPU-based segmentation). The collaborative segmentation was set up with a MIRALab's research team more focused on network and mobility topics, within the EU project InterMedia [INT10]. Segmented data was verified and built in conjunction with the radiological and orthopedic departments of the University hospital of Geneva. Finally, many contributions would not have been possible without the participation of the professional ballet dancers of the Grand Théâtre of Geneva, who gave consent to undergo MRI scanning sessions. Thanks to them, many MRI datasets were acquired with some unique characteristics (e.g., acquisition of split data).

1.4 Manuscript Organization

This manuscript is organized as follows. The first two chapters present related work in medical image segmentation and musculoskeletal modeling. The segmentation literature is notably presented under the perspective of prior-knowledge usage in segmentation approaches. Main advantages and drawbacks of presented methods are reported to better highlight some of the efforts made in this thesis to improve some existing methods. In the presentation of the musculoskeletal modeling, prerequisites in anatomy and medical imagery will be in particular provided. Subsequently, our methodology is thoroughly detailed in Chap. 4. It describes how shape and appearance priors are built from training data, by tackling the issues of generic model and training data construction, correspondence, alignment and statistics inference. The use of this prior-knowledge is then depicted in Chap. 5, where we describe the mechanisms underlying our knowledge-based deformable model in the key phases of initialization, evolution and control. The evaluation of the proposed segmentation approach concludes this Chapter with a series of experiments. We discuss the strengths and limitations of our work in the final Chap. 6, where future perspectives will be given. Finally, appendices are provided to give additional technical details. In particular, Appendix B presents a useful technical background on the domains of statistics and probability theory, which are the core of knowledge-based approaches.

Chapter 2

Prior Knowledge in Medical Image Segmentation

2.1 Introduction

This chapter aims at presenting the literature in the domain of medical image analysis with emphasis on *image segmentation*. Medical image analysis involves the understanding of images for medical purposes. Diagnosis is a typical example in which physicians study medical images to assess patients' health status and to make decisions for possible treatments. The nature of the medical images, i.e. the modalities that generate them, can vary depending on the available resources and the examined anatomical or pathological structures. A discretized version of the acquisition signal is usually required for human understanding, yielding to images with N dimensions. Commonly, 2D (e.g., X-rays, 2D Ultrasound), 3D (e.g., MRI, CT, 3D Ultrasound) and 4D images (e.g., Dynamic MRI, dynamic 3D Ultrasound) are available and routinely used in the clinical environment. For a more detailed description of image modalities please refer to Chap. 3.

Medical image analysis has been very active over the past three decades, and image segmentation is surely one of the most prolific areas of medical image analysis, signal processing and computer vision. Image segmentation is the process to divide an image into homogeneous and non overlapping regions. In case of medical images, these regions usually correspond to anatomical (e.g., an organ) or pathological (e.g., a tumor) entities. This dissertation covers in full details only the segmentation algorithms that are directly relevant. In fact, image segmentation offers a huge variety of methods and the natural trend is to produce specialized reviews depending on, e.g., the segmentation approaches (e.g., registration [MV98, ZF03], deformable models [MT96, JZDJ98], GPU-assisted [HLSB04]), the modality (e.g., MRI segmentation [CVC+95], Ultrasound [NB06]), the structure of interest (e.g., organs such as knee [STZ06], liver [CC07, HvGS+09], or other structures like characters [CL96]), the nature of the image (e.g., color images [CJSW01]) or related actions (e.g., evaluation of segmentation [ZFG08], interaction in segmentation [OS01]).

So far a unique and universal image segmentation approach is not available. Indeed, segmentation approaches tend to be sensitive to the type of modality, the ubiquitous artifacts (e.g., lack of contrast, noise, partial volume effect) and the intra- and inter-subject variability. In this context, the use of *prior knowledge* can considerably improve and assist the segmentation. In this thesis, the accent is hence put in the usage of *knowledge-based* approaches for medical image analysis. By knowledge-based, we refer to approaches that try to solve problems using some prior knowledge. In particular, we will address the use of prior knowledge coupled with discrete deformable models. We will thus rely on review and seminal papers on the usage of prior-knowledge (e.g., statistical shape models [HM09], deformable models [MDA01] coupled with statistical approaches [CRD07], etc.).

The use of knowledge into medical image analysis techniques essentially relies on fundamental principles of statistics and probability theory. As a result, we refer interested users in the technical background on statistics and the probability theory to Appendix B. Based on the more recent classification of Withey and Koles [WK07], we report segmentation approaches in different categories from weak to strong prior knowledge (Fig. 2.1). First, approaches using weak or

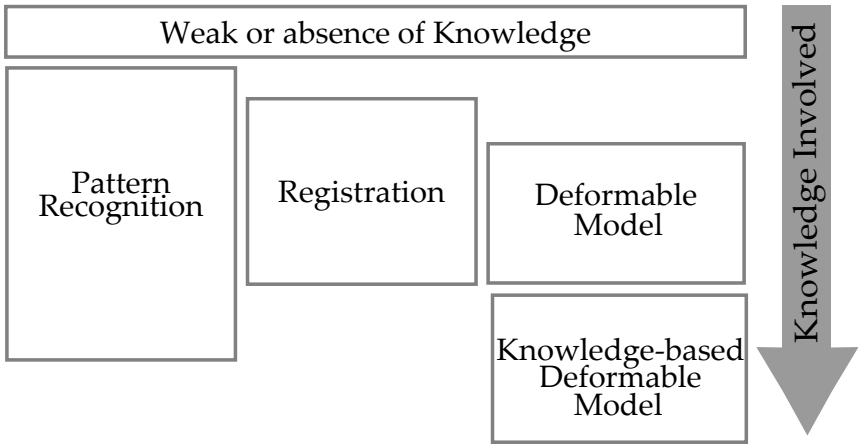


Figure 2.1: Overview of medical image segmentation approaches. Five main groups are identified: approaches using weak or no prior knowledge, registration, pattern recognition, deformable models and knowledge-driven deformable models.

no prior knowledge are briefly described in Sec. 2.2. Subsequently, image registration is presented in Sec. 2.3. Registration is a very prolific area of medical image analysis and only the concepts related to image segmentation and our work will be of interest. Section 2.4 addresses approaches based on deformable models which are in their original form not based on strong prior knowledge. Finally, Sec. 2.5 puts the accent on deformable models based on statistical and probability concepts, which are at the core of this thesis. We purposely ignored the overview of pattern recognition techniques because we would like to maintain the focus on image segmentation. In fact, pattern recognition is a broad field of research from which some approaches are borrowed for segmentation purpose (e.g., classification). We refer interested users in this topic to the comprehensive review of Jain *et al.* [JDM00].

2.2 Weak Prior Knowledge Approaches

Approaches falling into this category use no or very little knowledge of the problem they attempt to solve. These methods are often regarded as low-level techniques and are commonly used as a preprocessing for more advanced approaches because they are simple and easy to parameterize. They belong to what is often referred to as *direct segmentation*. We present them in an order that reflects the degree of involved prior knowledge.

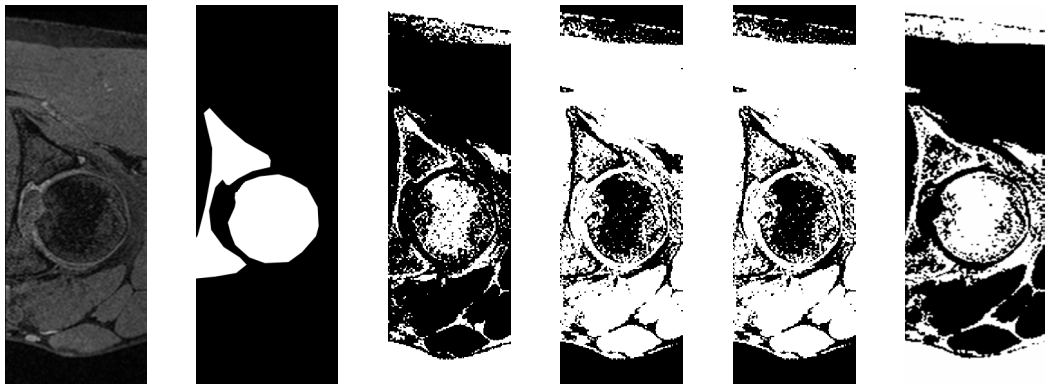
2.2.1 Thresholding and Histogram Analysis

Thresholding is surely one of the most simple but popular concepts to segment an image. Given an image with scalar (e.g., CT voxel/pixel value) or multi-dimensional (e.g., RGB color) values, thresholding approaches define threshold(s) to divide the image information space into various classes. The most common thresholding uses one threshold and groups pixels/voxels into 2 classes whether their intensity is greater or not than the threshold.

Thresholding techniques can be divided into global and local approaches [Wir07]. Global ap-

proaches use the same thresholds over the whole image while local approaches define spatially varying thresholds to particularly tackle the shortcomings of inhomogeneities. In global approaches, choosing adequate thresholds is critical and they are often estimated interactively by visual assessment from an operator [PXP00]. To circumvent this time consuming and subjective task, automatic thresholding techniques were designed, such as the Otsu [Ots79] and Isodata [RC78] methods.

For images presenting little noise and strong contrast between the structures to segment, thresholding is a simple yet effective approach [SSWC88]. However, thresholding techniques also exhibit very poor performances in presence of moderate noise and image inhomogeneities, which are ubiquitous in medical images. This is illustrated in Fig. 2.2, in which MRI bones cannot be segmented by manual or automatic thresholding approaches.



(a) (b) (c) (d) (e) (f)

Figure 2.2: (a) axial MRI slice from which bones need to be segmented. (b) Manually drawn by an expert, (c) manual thresholding, (d) Otsu [Ots79], (e) Isodata [RC78] and (f) connected region growing results. As expected, (naive) thresholding approaches are not appropriate to segment these complex images.

More advanced thresholding approaches are unsupervised, in the sense they do not require any user interaction to determine the thresholds. These approaches rely instead on some basic prior knowledge, such as the number of classes, and are often the result of an analysis on the histogram of the image intensity. As reported in Section B.5.1, the histogram is strongly related to statistical and probabilistic concepts and thus various approaches carried out histogram analysis to yield an automatic selection of the thresholds. For comparisons and detailed formulations of these approaches, we refer interested readers to [SSWC88, Spi93, Gla93].

2.2.2 Edge Processing

Thresholding approaches can greatly benefit from the analysis of local intensity changes, or more exactly the analysis of intensity gradients. Local intensity discontinuities expressed by strong gradients are interpreted as edges and in the context of medical imaging, as boundaries of the structures of interest. Edge-based approaches intent to first estimate these edges and then to provide better estimates of the thresholds. These approaches rely on different stages:

– Detect edges by using e.g., the Marr-Hildreth [MH80] operator based on the detection of 2nd

order derivative zero-crossings obtained by the convolution of the image with a 2nd order derivative Gaussian.
- Discard false positive edges often caused by image noise and artifacts. The Canny edge detector [Can86] is a famous method that attempts to find "optimal" edges by reducing noise and by exploiting hysteresis thresholding, which shares concepts with region growing approaches (Section 2.2.3).
- Follow and assemble edges into continuous boundaries.
- Estimate thresholds from the edge information to be used in subsequent thresholding approaches (e.g., Yanowitz and Bruckstein approach [YB89], local intensity gradient method [Par91])

Although some segmentation studies reported good results in using edge-based detection (e.g., [BHTR90, Ehr90]), these segmentation approaches are unlikely to be efficient in many occasions where edge detection is prone to failures due to the absence of clear boundaries, the presence of strong edges not corresponding to the structures of interest or the highly varying texture of imaged organic tissues.

2.2.3 Region Growing

Region growing approaches attempt to group pixels/voxels with similar characteristics into regions. This is achieved by comparing the pixel/voxel properties (intensity, gradient) against some homogeneity criteria. Region growing is an iterative process, which explores the neighbors of pixels/voxels to add them or not to the region the pixels/voxels potentially belong to. The process is then repeated until no more image element need to be analyzed. Region growing approaches can be classified into two groups according to [Wir07]:

- **seeded vs. non-seeded approaches**: usually region growing starts from selected pixels/voxels denoted as *seeds*. By exploring the neighbors of these seeds, new potential pixels/voxels are added to the regions. When seeds are explicitly specified by an operator, region rowing is termed as *seeded*. Theoretically, one seed per region to segment is sufficient, so the total number of seeds is relatively low as the number of regions to segment in one image is usually not very large. However, multiple seeds are in practice required per region due to image artifacts that interrupt the region growing resulting in incomplete regions. As a result, this yields a time consuming and subjective seed selection task. Non-seeded approaches attempt thus to automatically select appropriate seeds by using e.g., thresholding techniques based on the image intensity as described in Section 2.2.1. Other approaches do not require any seed to start but still require some parameters, like the method of Lin *et al.* [LJT00] which dynamically creates new regions and add image elements to them depending on a distance measure based on the mean intensity of the regions.
- **merging vs. splitting approaches** [Spi93]: the merging algorithm is an ascending strategy that over-segments the image at first by splitting it into many regions (up to the individual pixels/voxels at maximum). Regions are then progressively expanded by including image elements that satisfy the homogeneity criteria, as long as there are pixels/voxels to be processed.

Conversely, the splitting algorithm starts from very few regions (which could be the image itself) and then adopts a descending strategy. Regions are successively divided into subregions if any of the homogeneity criterion is violated, until all subregions are homogeneous. Both approaches tend to be computationally intensive algorithms. Horowitz and Pavlidis [HP76] proposed the Split and Merge algorithm which combines the advantages of both approaches and relies on a quadtree (extended to octree in 3D) data structure.

Region growing combined with thresholding and edge processing is widespread in the segmentation literature, but most of the time it is used as a complement or a start to more advanced segmentation algorithms. Indeed, region growing is accompanied by various pitfalls [Spi93] that include the critical determination of the seeds, the careful design of the homogeneity criteria as well the strategy to explore the image elements. Figure 2.2(f) depicts a seeded region growing algorithm where seeds were placed in bony regions, and for which the homogeneity criterion consisted in keeping voxels whose intensity was within some bounds. It can be observed that some areas of the bones are still not segmented correctly and that a "leaking" occurred in the surrounding tissues (e.g., fat tissue between muscles).

2.3 Registration

2.3.1 Introduction

In this Section, an overview of registration is given. The aim is not to provide a comprehensive review of image registration, for that we refer interested readers to the reviews [MV98, AFP00, ZF03, SMFF07, FM08], but to present the necessary information to grasp the registration concepts used in medical image segmentation. We will see for instance in Sec. 2.5.4 that in order to establish a point correspondence among training shape, registration is a key component.

Registration is the process to find the transformation that maps one entity to another. The form of the registration transform T is hence assumed to be known and its parameters have to be computed. Segmentation of a target image is carried out by propagating a segmentation to the target image by registration: a source entity with segmentation information is registered to a target entity in relation with the target image. Entities can be the images or features extracted from them as discussed in the following.

Registration relies on the following components:

– **Registration Features**: points, lines, models, image information (intensity, gradient), etc.
– **Transform and associated regularization**.
– **Similarity Metric**.
– **Registration strategy**, which includes optimization and interpolation choices.

Figure 2.3 gives an overview of the registration process. In an iterative manner, the similarity between reference and target features is computed. Target features are computed based on the target image and the current transform parameters. Since images are discretized signals, an interpolation is necessary to estimate the target features. Given the computed similarity value

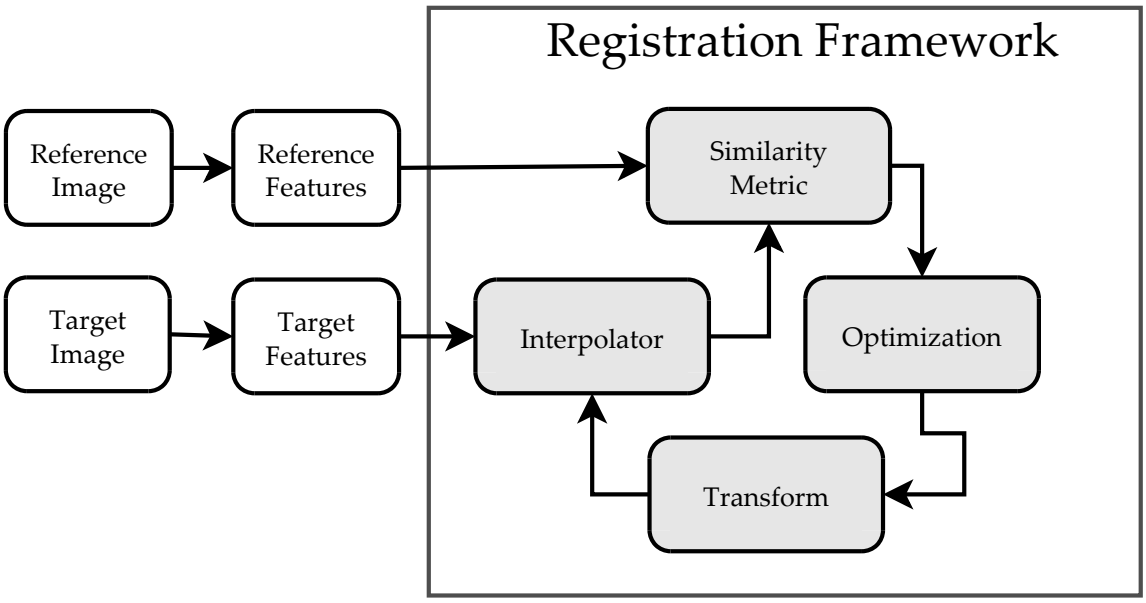


Figure 2.3: Overview of registration framework (derived from [ISNC05])

and the optimization strategies, new values of the transform are computed and the procedure restarts. The optimization process usually will iterate until it considers that the similarity has reached an extremum. Without any loss of generality, let us assume that the optimization always looks for a minimum.

Depending on the choices made on the registration components, a large variety of registration methods exist. However, a simple taxonomy of the various methods can be provided as:

- Image-to-Image registration: reference and target images are registered by mostly exploiting iconic features such as intensity, gradients or statistical properties.
- Model-to-Image registration: a reference model represented as a structure with different nature than an image (e.g., a mesh, a pointset, a parametric curve) is registered with a target image.
- Model-to-Model registration: a source and target models, with not necessarily the same nature, are registered mostly based on geometrical features (e.g., position, curvature).

2.3.2 Registration features

iconic features Most common image registration approaches directly exploit the image information, features being often referred to as icons. Image features can be based on voxel information or neighborhoods (e.g., template matching [DGS01], intensity profiles [Del99]), or any possible combinations. Choices are often related to the nature of the image formation process. For instance, Roche *et al.* [RPR+00] jointly used image intensity and gradient when registering CT to US, because US images enhance interfaces between anatomical structures due to the underlying acoustic nature of the signal. It is quite common to pre-process the images to filter out noise (e.g., NL-means [BCM05], anisotropic diffusion [PM90]) or enhance structures (e.g., edges [Can86]). Moreover, registration can be robustify by discarding iconic features in the reference or target images that correspond to noise or non matching structures (e.g. based on region saliency

[Zha08]).

Shape features The most simple feature is the point. Points often referred to as landmarks can be related to a geometrical structure such as a mesh (e.g., Iterative Closest Point (ICP) [BJ85]) or can be extracted from the images via image processing techniques. In that latter case, points can correspond to either manually-positioned fiducials (e.g., radio-opaque markers [NSP+04]) or to artificial ones (e.g., corners [Nob88, BS97, Roh01]). Alternatively, other kind of shape features can be used such as lines which are detected in images as strong edges (e.g., [Can86]) or ridges (e.g., [MVDEV96, MV98]). One of the drawback of using shape features extracted from images is that the registration is clearly dependent on the quality of the feature segmentation (e.g., sensitivity of registration based on L_2 norm with respect to outliers), which may require robust techniques [ZF03].

2.3.3 Transform and Regularization

Transform types Different types of transformations are available and they are commonly categorized based on their ability to capture global or local changes. Local transforms are also denoted as *non-rigid*. Transforms are also characterized by their number of degrees of freedom (DOF), i.e. the number of independent transformation parameters. More DOFs a transform has, more complex changes it can express but more a regularization is required (explained later in this Section). Table 2.1 reports some main 3D transforms. A first group of transforms are the linear transforms, which can be represented by a 4×4 homogeneous matrix. They are more commonly found in intra-patient registration, i.e. when images are coming from the same patient. Affine transform characterized by translation, rotation, scaling and shearing is a first crude approximation of a non-rigid transformation. Similarly, polynomial transforms [WGH+98] can express higher variability by increasing their degree. Various authors proposed piecewise affine transforms [FA96, PBTM06] to have a tradeoff between DOF and non-rigidity. The principle consists in splitting the reference and target images into pieces and to compute the best affine transform to match the pieces. Pieces can be blocks, triangles from meshes triangulations, regions detected by block matching and clustering, etc.

Non-rigid approaches attempt to model more accurately the deformations by increasing the DOF. A first popular approach consists in creating a structured grid of control points in the images, and to create deformations by modifying the position of the control points. An interpolation function is defined to compute the information between the control points. Various functions are possible such as Bezier, Hermite or B-splines, depending on the desired degree of continuity. Rueckert *et al.* [RSH+99] made notably popular the use of free-form deformation (FFD) with B-splines in medical image registration. Alternatively, unstructured point sets can be used to achieve registration of images. Points are first extracted from the images based on image processing. Then, interpolation approaches are once again applied to estimate data at unknown position. Radial basis functions (RBF) are a popular choice and the most common one is the thin-plate splines (TPS) [Boo89]. TPS minimizes the bending energy and has been extensively used in

image registration [RSS+01, Roh01]. One of the drawbacks of using TPS-based registration is the resulting computation burden. As a result, effort was put in optimizing the TPS computation [Flu92, BSB+93].

Type	DOF	Global	Remarks
Rigid	6	Yes	intra-patient, rigid structures
Similarity	7	Yes	Intra-patient, different scales
Affine	12	Yes	coarse approximation of non-rigid transformation
Polynomial	$3\sum_{i=1}^{d} C_{i+2}^{0}$	Yes	d =polynom order, non-rigid approximation
Piece-wise affine	$12N$	No	trade-off DOF/non-rigidity, N =number of pieces
Radial Basis Functions	$3N$	No	N =number of nodes, non-rigid
FFD	$3N_xN_yN_z$	No	N_x, N_y, N_z =grid size, non-rigid
Deformation maps	$3N$	No	N =number of deformation vectors, commonly found in elastic, fluid registration

Table 2.1: Common 3D registration transforms

Regularization Registration is an ill-posed problem. This implies that in many cases a regularization is necessary to yield satisfactory results. Regularization is commonly applied in the optimization stage, where the computation of the next transform parameters follows a strategy based on some regularization constraints. Regularization is commonly found in elastic registration. In elastic registration, the transform is not limited by a parametrization but instead a dense deformation map is computed under some regularization constraints. The deformation map h simply expresses the change from a point $\boldsymbol{x}$ to a point $\boldsymbol{h(x)}$ as the sum with a deformation vector $\boldsymbol{u(x)}$: $\boldsymbol{h(x)} = \boldsymbol{x} + \boldsymbol{u(x)}$. A parametrization is chosen on the deformation vector (e.g., complete orthonormal basis [CRM94], Fourier series representation [Chr99]) and the regularization is done on the deformation map, usually based on some "physical" phenomenon (e.g., elastic [GRB93, Dav97], fluid [Chr94, Chr99]).

2.3.4 Similarity Metrics

Similarity metrics express the quality of matching between reference and target features. They are a key components of the registration as the optimization tries to minimize them. A large variety of metrics exists, depending on the type of shape or iconic features that need to be registered. For example, in case of point-based features, common metrics are L_p minkoski and Hausdorff distances. Some of these distances are more robust against outliers [LE01] than others.

Many image-based metrics were proposed [HHD+99, MV98], based on different assumptions:

- Intensity conservation: sum of square differences (SSD), sum of absolute differences (SAD), cross correlation, normalized cross correlation (NCC)[HHD+99].
- Intensity changes conservation: (normalized) gradients flow [PMV03, HM06].

– Functional relationship between image intensity: (bivariate) correlation ratio (CR) [RPR+00], (normalized) mutual information (MI)[VW97, SHH99].

Mutual information (MI) [VW97] has become particularly popular in multimodal medical image registration [PMV03]. Based on concepts on information theory, MI minimizes the joint density of the gray value distribution by using kernel density estimation techniques (e.g., Parzen Windowing, see Sec. B.5.2.1). To make the MI less sensitive to the overlap of the reference and the transformed target image, the normalized MI (NMI) [SHH99] was proposed. NMI encountered a great success in medical image registration [RFS01, FRSN02, RFS03, OBB+04, VTL+04]

2.3.5 Optimization Strategies

optimization The optimization computes new values of the transform parameters based on the computed similarity and the current registration state. Given a similarity cost function f with d parameters to optimize, a large panel of local or global numerical methods exist to minimize f. Local approaches search for a minimum that is the closest to the initial parameters. A first way is to minimize the cost function by descending along its gradient. Popular gradient-descent approaches [PFTV92] include the classic gradient descent (descent in direction $-\nabla f$), (Quasi-)Newton [Ypm95] and Levemberg-Marquardt [Mor77] (combination of Newton and gradient descent). While these approaches often converge efficiently and quickly to a local minimum, they rely on the computation of gradients which cannot always be computed accurately. As a result, other approaches can be used such as the Downhill simplex [NM65] or the Powell's method [PFTV92].

Since local approaches suffer from attraction to local minima, a popular improvement is to adopt a coarse-to-fine strategy. In case of transforms, this can be achieved by progressively increasing the DOF (i.e. from rigid to non-rigid transform [RSH+99]). In case of images, an image pyramid is built which encompasses various smoothed versions of the original image computed at different resolution levels (e.g., [SL96]). Starting from the lowest resolution, the registration is run various times, where the results from the previous level are used as the initialization of the current iteration. The multi-resolution approach confers faster convergence and robustness against local minima (enlarged capture range, better resistance against image artifacts).

Global methods remove the dependency of a good initialization. Simulated annealing [Kir84] is a probabilistic approach which changes parameters based on an acceptance probability $P(f, T)$, where T is a "temperature" term. When T is high, changes are allowed in the direction of increased energy which avoids local minima. When temperature is low, changes are preferred in a downhill manner, i.e. similarly to a gradient descent. The temperature decrease is determined by the annealing schedule which ensures that a null temperature is reached at the schedule completion. Another type of global approach is based on evolution concepts where various individuals that encode the transform parameters are modified based on mutation, cross-over and selection operators. The individual that "survives" the evolution corresponds to the final solution. Evolution algorithms (EA) notably include genetic algorithms (GA) [Hol75] and differential evolution (DE) [SP95] which proved the be efficient global minimization approaches.

interpolation Interpolation is needed to compute the target features transformed by the current transform estimate. In case of images, this means that image information is interpolated between voxels. Interpolation is used thousand of times in a registration approach strongly affecting the speed of the approach. Moreover, the accuracy of the interpolation influences the smoothness of the metric. As a result, a tradeoff is required between quality and computational complexity of the interpolation. In increasing order of accuracy, typical interpolation techniques are the nearest neighbor, linear and B-splines interpolation. B-splines interpolation uses shifted B-spline kernels, whose order can be chosen based on desired speed or accuracy. Moreover, according to signal sampling theory, the optimal interpolator is the *sinc* function. This function has the inconveniences to have an infinite support and a slow decay which are not really suitable for efficient numerical implementation. Various methods were thus proposed to approximate this *sinc* interpolation [MNPV99, MNV01].

As a last remark, when the target image is transformed into the reference space, the transform that needs to be optimized is defined from the *reference to the target* space. This ensures that no "holes" appear in the resampled target image [ISNC05]. This is something to consider when the transform is not invertible. In fact, the inverse transform may not be always available (e.g., some non-rigid transforms) and as a consequence some features of the target space cannot be easily expressed in the reference space. Note that there exists some non-rigid invertible transforms, notably the so-called diffeomorphisms (see Sec. 2.5.4.4 and 2.5.4.6).

2.4 Deformable Models

Deformable models were originally introduced in the Computer Graphics (CG) context by Terzopoulos *et al.* [TPBF87, TW88], and have been studied and explored in many CG domains [NMK+06] such as in cloth simulation [BW98, VT00, VMT05]. Rapidly, deformable models gave birth to a large variety of image analysis approaches in two- [JZDJ98, MT96], three- [MT96, MDA01] and four-dimensions [MD05].

The organization of this section is as follows. Section 2.4.1 briefly introduces deformable models. Sections 2.4.2 and 2.4.3 present the popular active contours and levelsets deformable models, respectively. Subsequently, the discrete deformable models are discussed in Sec. 2.4.4. Active contours and levelsets are presented for a sake of completeness and to highlight their advantages and drawbacks.

2.4.1 Introduction

Deformable models are entities which undergo geometrical or topological deformations to reach an equilibrium state, which is synonym of segmentation completion. Originally, deformable models possess a weak knowledge of the structures to segment. This knowledge is usually expressed as:

– Model representation, in particular parameterization in case of continuous models (see as follows).

– Boundary/surface smoothness.
– Topology (known genus) and geometrical invariants (e.g., symmetry).

Deformable models can be categorized based on their representation. In [MDA01], a detailed taxonomy of deformable models is provided and Table 2.2 gives some typical examples of deformable models. Some of them are described in details by presenting their respective representation and evolution strategies. Our work is in particular based on discrete deformable models, and they will be discussed in details in Sec. 2.4.4.

The two main representations are continuous and discrete. In the discrete representation, the shape is known at a finite number of points. Discrete models can represent shapes of higher complexity and arbitrary topology. While the discretization of continuous deformable models remains necessary for computational reasons, the continuous representation yields more accurate computations of differential quantities such as normals or curvatures [MDA01].

Continuous representation	
Explicit	***Implicit***
Snakes [KWT88, CKSS97],	Levelsets [OS88, MSV95, CKS97] (Fig. 2.5),
Superquadrics [TM91],	Superquadrics [BCA96],
Spherical harmonics (SPHARM) [BGK95]	Medial [Blu67, Mor95, PEFM98]
Discrete representation	
Meshes	***Particle Systems***
Triangular meshes [SVG02] (Fig. 2.6(b)), Simplex meshes [Del99, MD05] (Fig. 2.6(c))	[ST92, DG96, PKA+05]

Table 2.2: Examples of deformable models representations (based on [MDA01])

Continuous representations commonly rely on some *parameterization* which limits the number of degrees of freedom of the shape variations or imposes the model topology. For instance, spherical harmonics (SPHARM) [BGK95] only describe shapes topologically equivalent to a sphere, and superquadrics [TM91, BCA96] only allow the representation of shapes with an axis of symmetry. Similarly, medial models [Blu67, Mor95, PFJ+03, JPF+02, FPAB04] represent the shape as a series of *medial nodes* with a position, a local frame and a radius, lying on a medial sheet located at the "center" of the shape. Shape surface is reconstructed as the union of spheres whose centers and radii are those of the medial points. This representation is compact (smaller number of medial nodes compared to number of vertices in a discrete mesh) but yields (too) smooth reconstructed shapes. As an example, figure 2.4 illustrates a type of medial model for tubular structures, where the medial sheet is a centerline composed of medial modes. Such a structure has been used in many vessel segmentation approaches [KMA+00, AB02, FPAB04, PJBB05].

In general, parameterization restrictions are adequate for the segmentation of structures with a relatively simple geometry but are not always appropriate for the delineation of complex shapes.

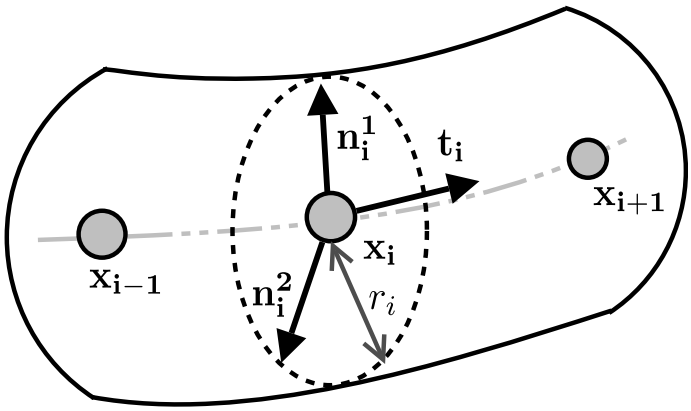


Figure 2.4: Medial deformable model example: a tubular structure is implicitly modeled as a series of medial nodes $\boldsymbol{C_i} = \{F_i, \boldsymbol{x_i}, r_i\}$, where $F_i = \{\boldsymbol{t_i}, \boldsymbol{n_i^1}, \boldsymbol{n_i^2}\}$ is the local frame and r_i is the estimation of the object "radius" at $\boldsymbol{x_i}$.

Nevertheless, this limited knowledge provides some advantages with respect to previous segmentation approaches which quite exclusively relied on image information (direct segmentation and registration). Indeed, in image locations where information is missing or is not reliable, the segmentation can exploit the model shape and is thus more robust to image artifacts. As pointed out by Heimann and Meinzer [HM09], many deformable models could be customized to represent specific shapes but the prior knowledge involved is essentially relying on smoothness or parameterization constraints which are not derived from statistical or probabilistic information. As a result, this overview only briefly covers the literature on such a kind on deformable models. Moreover, we will not discuss about volumetric deformable models (e.g., based on tetrahedral or hexahedral meshes) because they cannot often be found in the context of medical image segmentation, although some works exist (e.g., heart segmentation [PVC+01]). In fact, volumetric deformable models are more frequently used in physical simulations of dense deformable structures. These models rely on various representations such as particle systems [ACS+09], finite element models (FEM) [PVC+01] or finite volume models (FVM) [TSB+05]. They are furthermore very computationally expensive approaches, although various GPU-based implementations [TE05, GSMY+07] were proposed to speed-up the simulation. These simulations seek accuracy and realism, whereas in segmentation the evolution of the underlying deformable models is somehow artificial since it usually does not rely on any real physical process.

2.4.2 Snakes

In their seminal work [KWT88], Kass *et al.* presented their *active contour* deformable models, also known as "snakes". Given a 2D image domain Ω, snakes are parametric curves $\mathcal{C} : [0,1] \to \Omega$ driven by the minimization of an energy function:

$$E(\mathcal{C}) = \frac{\alpha}{2}\int_0^1 \left|\mathcal{C}'(q)\right|^2 dq + \frac{\beta}{2}\int_0^1 \left|\mathcal{C}''(q)\right|^2 dq + \gamma\int_0^1 \left|f(\mathcal{C}(q))\right| dq \tag{2.1}$$

The first two integrals denote the internal energy which enforces model smoothness based on Tikhonov stabilizers of 2nd order [MDA01]. The last integral expresses the external energy which

attracts the contour toward the boundaries of the structures to segment. In its simplest form, snakes are attracted by edges with higher gradient magnitude, i.e. $f(x) = \|\nabla I(x)\|$. Coefficients α, β and γ weight the stretching, bending and image attraction of the snake, respectively. Natural extensions to the 3D case (active surfaces) were proposed [TWK88, CKSS97] based on first and second order derivatives.

Based on the calculus of variation (or variational principle), the snake evolution follows the Euler-Lagrange equation: $\nabla E(\mathcal{C}) = 0$. To solve this equation, an artificial time t variable is introduced and the snake $\mathcal{C}(q)$ is made dynamic [KWT88]:

$$\frac{\partial}{\partial t}\mathcal{C}(t,q) = \alpha\mathcal{C}''(t,q) - \beta\mathcal{C}''''(t,q) + \gamma\nabla f(\mathcal{C}(t,q)) \tag{2.2}$$

This equation can be efficiently solved with finite differences discretization and the iterative solving of the discrete system. This minimization is local and the snake will thus converge to a local minimum.

The main weaknesses of snakes are (i) the necessity to be initialized closely enough to structure's boundaries (local minimization), (ii) their inability to move into concavities of the boundaries (smoothness constraints) and (iii) their fixed topology (parameterization). The first shortcoming was partially addressed by the use of "balloon" forces [Coh91] that pushed the snake to inflate or deflate in the normal direction. The main issue with this force was the need to know a priori whether an inflation or deflation was required, which prevented the initialization of the snake across the structure boundaries. The use of gradient vector flow (GVF) fields [XP97] to model the external energy yielded an increase of the snake capture range and a better delineation of concavities. Davatzikos and Prince [DP95] proposed the use of constraint points to better capture concavities in case of the brain cortex. Finally, the support of topological changes was included in the "T-snakes" of McInerney and Terzopoulos [MT00]. Many other works improved snakes to cope with some of these limitations, among them Nascimento and Marques [NM05] who used an EM-based (See Sec. B.4.4 for details on EM) algorithm to estimate the confidence of the boundary points, therefore robustifying the snake evolution.

2.4.3 Levelsets

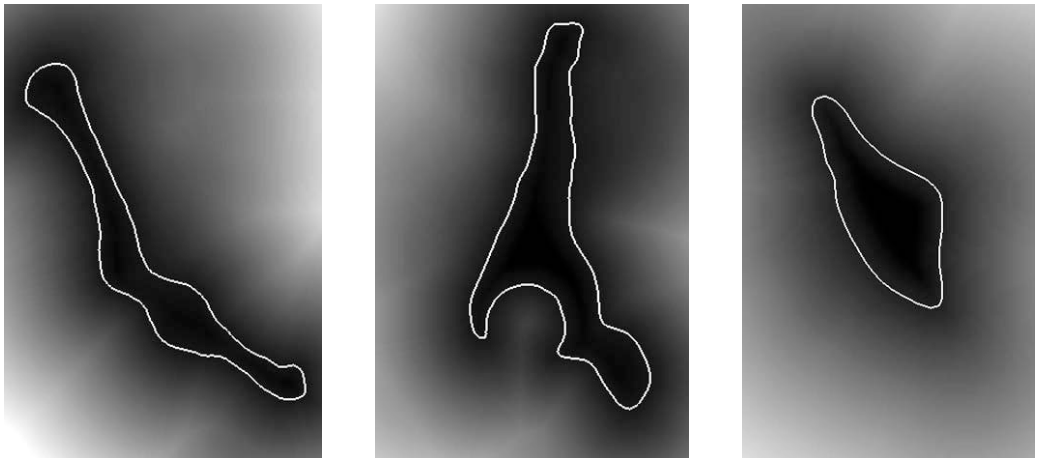
In the Levelsets theory [OS88], the deformable model is embedded in a higher dimensional space. A levelset S is an iso-hypersurface of a function $\Phi : \Omega \to \mathbb{R}$, $\Omega \in \mathbb{R}^d$:

$$S = \{x \in \Omega | \Phi(x) = c\} \tag{2.3}$$

where c is a constant usually equal to zero. When $d = 2$ or $d = 3$ the levelset S is an iso-contour or an iso-surface, respectively. The levelset S is also seen as the boundary ∂R which encloses a region R:

$$\begin{aligned} S &= \partial R \\ R &= \{x \in \Omega | \Phi(x) < c\} \end{aligned} \tag{2.4}$$

This formulation allows thus to define *implicitly* the deformable model with Φ. By evolving function Φ, the deformable model represented by S is hence implicitly modified. The greatest advantage with respect to the snake modeling is that S can naturally undergo topological changes as it can split/ merge into various iso-hypersurfaces during its evolution. Note that S is a set of *closed* hypersurfaces [MDA01]. A common way to define the function Φ is to use the signed Euclidean distance function (SDF) to the deformable model. Given a point $x \in \Omega$, $SDF(x)$ is the signed Euclidean distance between x and its closest point on the deformable model. A negative sign is arbitrarily chosen, e.g. $SDF(x) < 0$ when $x \in R$ given Eq. (2.4). Figure 2.5 illustrates the usage of a SDF to implicitly define a hip bone surface. Malladi *et al.* [MSV95] introduced the



(a) (b) (c)

Figure 2.5: Implicit deformable model representation. A hip bone shape is encoded as an iso-surface of a function Φ which is computed in this example as the signed Euclidean distance (SDF) map to the shape surface. In (a) coronal, (b) sagittal and (c) transverse slices of the SDF map (volumetric image), the hip bone boundary/iso-contour is painted in white.

levelset approach to medical image segmentation. Based on the levelset equation, levelsets can be evolved using, similarly to snakes, a time parameter t:

$$\frac{\partial}{\partial t}\Phi + \vec{V}.\nabla\Phi = 0 \tag{2.5}$$

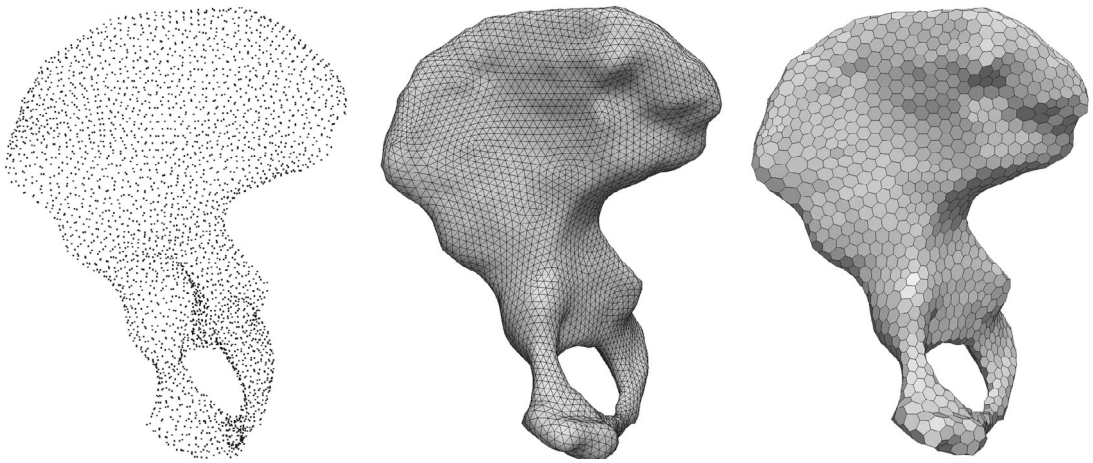
where $\vec{V}$ is an external force field driving the evolution toward the correct image location. This field is computed based on image information such as the image gradient magnitude. This formulation requires hence the need to solve partial differential equations (PDE). Caselles *et al.* [CKS97, CKSS97] combined the snake-based evolution with the levelset formulation into what they called *geodesic active contours*. In addition to the possible topology changes, this synergy brings several advantages with respect to classic snake implementation. First, no re-parameterization is required during the evolution, which is usually necessary with snakes when self-intersections are observed. Second, the function Φ is *independent* from the contour/surface parameterization, which means that the deformation are *only* dependent on the shape of the deformable model and the image.

Levelsets are of course not exempt from inconveniences. In cases where the structure to seg-

ment's topology is known in advance, the ability of the levelset to adopt topological changes can become an issue. It might happen for example that due to image noise or missing information, the levelset creates separate components whereas only one was expected. The other main problem in using levelsets is the computational burden, especially in 3D. Solving the PDEs requires efficient numerical schemes [OS88] which often have tough stability constraints. For instance, in the basic levelset implementation the iso-surface can progress at maximum one pixel/voxel at each iteration. In the special case where the level-set only deform in its normal direction, fast marching method can be applied [Set99]. Other implementations were proposed to tackle more efficiently the general cases (e.g., narrow band [AS95], sparse-field [Whi98], PDE-free [SK05] levelsets), but interactive update rates are often difficult to get with traditional 3D medical images and conventional processors [CLW04]. As a result, a large variety of Graphics Processor Unit (GPU)-based levelset approaches [LCW03, CLW04, LKHW05, Kla07] were designed to speed-up the process while adding the possibility to interactively visualize the deformable model evolution. A more thorough literature review on GPU-based segmentation is presented in Sec. 5.4 where our own implementation of a GPU-based discrete deformable segmentation framework is described.

2.4.4 Discrete Deformable Models

Discrete deformable models are characterized by a series of points lying on the model boundaries. When an explicit connectivity is established between the points, the discrete deformable models are referred to as meshes. Figure 2.6 illustrates three possible representations for a hip bone shape.



(a) (b) (c)

Figure 2.6: Explicit discrete deformable models representation. In this example, a hip bone shape is represented as (a) a point set, (b) a triangular or (c) 2-simplex mesh.

More formally, a discrete model S can be represented in a d-dimensional space as a triplet $S = \{\boldsymbol{X}, \boldsymbol{N}, \mathcal{G}\}$, where:

– $\boldsymbol{X} = \{\boldsymbol{x_1}, \ldots, \boldsymbol{x_M}\}$ is the point set composed of M points $\boldsymbol{x_i} \in \mathbb{R}^d$.

– $\boldsymbol{N} = \{\boldsymbol{n_1}, \ldots, \boldsymbol{n_M}\}$ is the optional normal set, one normal per point.
– $\mathcal{G}$ is the connectivity information which specifies the point indices of edges and the edges of faces. $\mathcal{G}$ is only defined for meshes and is also called graph of the mesh.

Normals can be defined from the point set $\boldsymbol{X}$ only using point-based techniques (e.g. Oriented particles [ST92], Moving Least Squares (MLS) [Lev03], Implicit surface MLS [AA04]) or in conjunction with the connectivity information $\mathcal{G}$ based on neighbors positions (e.g., [MBL+91, Del99]). Exact normals could be computed for discrete models produced from continuous representation, but since deformable models are constantly evolving an update of the continuous representation is necessary. Thus, these models would fall into the continuous category.

2.4.4.1 Particle systems

Particle systems only exploit a point set and are often coined meshless approaches. Particle systems consider that each particle $\boldsymbol{p_i}$ has a state composed of a mass m_i, an acceleration $\boldsymbol{a_i}$, a velocity $\boldsymbol{v_i}$ and a position $\boldsymbol{x_i}$. Additional parameters can be added to their state depending on the application. Particle systems are commonly used in computer graphics to simulate highly deformable bodies [DG96] and physical phenomena such as fluids [FM96, PTB+03], solid fractures [PKA+05] or skin deformations [WTMT95]. Particles are subject to internal and external forces. Internal forces denote the interaction between particles of a same deformable model. These forces are in practice computed based on local descriptors expressed from the particle and the state of its neighbors. External forces indicate external actions that affect the particle system evolution (e.g., gravity, collisions between different deformable models). Particle systems evolution is dictated by the Newtonian law of motion:

$$\begin{array}{rcl} \frac{\partial}{\partial t}\boldsymbol{x} & = & \boldsymbol{v} \\ m\frac{\partial}{\partial t}\boldsymbol{v} & = & \boldsymbol{f}(\boldsymbol{v}, x) \end{array} \tag{2.6}$$

where the force vector $\boldsymbol{f}(v, x)$ represents the internal and external forces that depend on the particle position x and velocity v (damping effect). In case of massless particles with identity damping matrix, the dynamics are expressed as a first order Lagrangian evolution [TM91, MDA01] which is equivalent to a gradient descent strategy [PFTV92]. The differential equations for each particle are first time-discretized with some time-step dt [VMT05]:

$$\begin{array}{l} \boldsymbol{Q_{t+dt}} - \boldsymbol{Q_t} = \boldsymbol{Q'_{t+\alpha dt}}dt \\ \boldsymbol{Q_t} = (\boldsymbol{x_t}, \boldsymbol{v_t})^\top, \boldsymbol{Q'_t} = (\boldsymbol{v_t}, m^{-1}\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{v}))^\top \end{array} \tag{2.7}$$

where α is the "implicity" coefficient as explained in the following. This system of equations (2.7) is numerically solved by using an integration scheme. Various integration schemes are available [HES03, VMT05, NMK+06] for which a tradeoff between accuracy, stability and computational cost exists. When an integration scheme yields a bounded solution, it is coined stable [HES03]. From simulations with analytical dynamic models (e.g., spring-mass model for accuracy and Dahlquist's test equation for stability [HES03]), accuracy and stability quantitative and qualitative indicators are usually derived [HES03, VMT05]. The accuracy is dependent on the stability of the system. Indeed, some approaches, like the explicit Euler integration (α=0

in Eq. (2.7)), are *conditionally* stable, meaning that if the time-step exceeds a specific value, the numerical integration will diverge, resulting in non-plausible results. Explicit Euler is a cheap method in terms of computations that can be easily extended into improved explicit schemes (e.g., Runge-Kutta [PFTV92], Verlet [Ver67] methods). Implicit methods are more appropriate since the stability is significantly improved and they allow larger time-steps, thus yielding faster convergence. For example, the implicit Euler [BW98] approach (α=1) is *unconditionally* stable but demands more computational and implementation resources. Similarly, the implicit midpoint integration scheme [VMT05] (α=0.5) is a very efficient approach, but compared to the pure implicit Euler, it requires damping to be stable. Moreover, implicit approaches can be as accurate as explicit counterparts, but they might suffer from a loss of accuracy due to numerical damping. Backward differential formulas [HE02] can be used to improve implicit methods' accuracy by exploiting the particles' state history. While this increases the accuracy, the drawback is that sudden changes (e.g., a collision response) may induce instabilities. A small comparison of the different presented methods is reported in Table 2.3.

Method	**Accuracy**		**Stability**	**Work**
	Large dt	Small dt		
Explicit (forward) Euler	−−	++	−−	−−
Runge-Kutta	−−	++	−	−
Verlet	−	++	−	−
Implicit Euler	++	+	++	+
Implicit midpoint	+	++	+	+
Backward differential formulas	++	++	+	++

Table 2.3: Numerical integration methods comparison. Accuracy is computed at equilibrium with large and small time-steps (dt). Work defines computational cost.

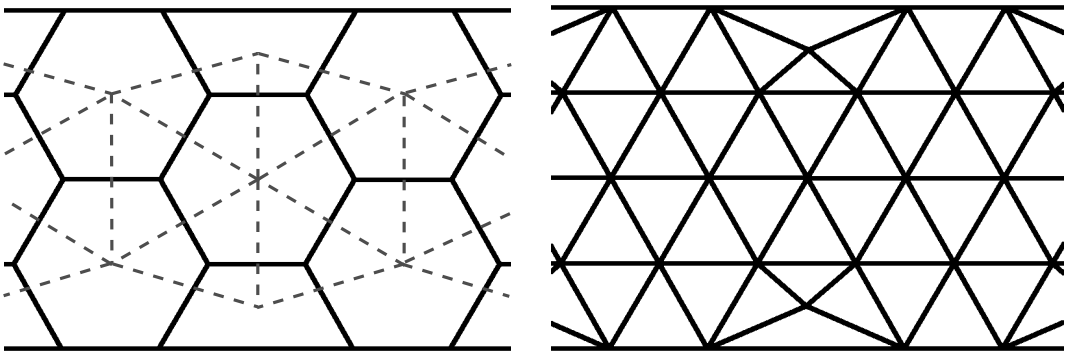
Compared to some very expensive approaches, such as finite element models, particle systems are very attractive and have been successfully applied in simulation and object reconstruction. However to date there is no segmentation approaches that fully exploited the particle system approach. The tendency is to use deformable models with parameterization or connectivity information and apply the mechanics of particle systems to them as explained in the following.

2.4.4.2 Meshes

Meshes are popular in many computer science fields such as medical image analysis, computer graphics and simulation. Triangular meshes (Fig. 2.6(b)) are two dimensional manifolds characterized by triangular faces. They are the most common type of surface meshes. In the computer graphics area, triangular meshes are natively supported by graphic hardware. In the segmentation domain, they are often preferred to other representations of deformable models due to their ability to represent shapes of arbitrary topology and complexity.

Miller *et al.* [MBL+91] proposed their geometrical deformable models that evolved under external (simple image force based on thresholding, balloon-like expansion [Coh91]) and internal (topology preservation based on local curvature estimation) constraints. Local curvature was

estimated from the neighbors of a vertex. The evolution was carried out with an algorithm that moved vertices in the direction of steepest descent along the cost surface computed from the external and internal constraints. Lachaud and Montanvert [LM99] designed a coarse-to-fine segmentation approach using triangular meshes able to undergo topological changes via Eulerian topological transformations of creation, deletion or inversion. In their work, models evolution followed Newtonian dynamics and was numerically solved with the Runge-Kutta integration scheme as described in the previous Section. Similarly, Snel *et al.* [SVG02] used multi-resolution deformable triangular meshes to segment carpal bones in MRI images based on Lagrangian mechanics. Triangular meshes were also commonly used with statistical shape models [LSHD04, HMMW07, SKH+08, KLZH09, FGXL10], this topic being further discussed in Sec. 2.5.5. Triangular surfaces were also implicitly expressed in [PB07], where the surface was embedded in the Delaunay tetrahedralization of interface points. Similarly to the levelset approach (Sec. 2.4.3), the surface was manipulated through a higher dimensional deformable model and topological changes (splitting and merging) were supported.



(a) (b)

Figure 2.7: 2-Simplex mesh example. (a) thick lines represent the edges of a 2-simplex mesh where each vertex has exactly three neighbors. The dotted line is the triangular mesh dual of the 2-simplex mesh, which is obtained by connecting the centers of the 2-simplex faces. (b) An alternative way to tessellate the 2-simplex mesh is produced by connecting all vertices to the center of the corresponding 2-simplex face. This tessellation has the advantage to keep the points of the 2-simplex mesh.

Another interesting family of meshes are those based on H. Delingette's *simplex meshes* [Del99]. k-simplex meshes are discrete models with a constant vertex connectivity of $k+1$ neighbors. Surfaces are modeled with 2-simplex meshes as depicted in Fig. 2.6(c) and 2.7(a). 2-Simplex meshes are topologically dual to triangulations: by connecting the centers of three adjacent faces, a triangle is formed as shown in Fig. 2.7(a). Other tessellation schemes are possible. Gilles [Gil07] proposed a preservative scheme which keeps simplex mesh vertices in the tessellated mesh (Fig. 2.7(b)). The fixed vertex connectivity of simplex meshes brings advantages in computing local descriptors, such as curvature and normals. The most interesting ones are the *simplex parameters* [Del99, Gil07] which are 3 independent coefficients invariant under similarity transform that able to describe the local configuration of each vertex. More exactly, a vertex is uniquely defined with respect to its three neighbors with the simplex parameters. These parameters yield the creation of "shape memory" constraints that can be enforced during the deformable model evolution. This

provides a stronger knowledge compared to the usual smoothness or topological constraints seen in other deformable models.

Various authors successfully exploited the simplex mesh representation in segmentation applications. Gilles *et al.* [GMMT06, GMT10] segmented musculoskeletal structures from MRI by using a Newtonian evolution of the meshes. In their work, external forces were computed based on the maximization of *intensity profiles* similarity [Del99] where vertices were attracted to locations at which intensity profiles were the most similar to pre-defined ones given the NCC measure (Sec. 2.3.2). Similarly, Costa *et al.* [CDNA07] performed bladder segmentations from CT scans by using 2-simplex meshes driven by intensity profiles and Lagrangian mechanics. In [MD05], 2-simplex meshes were efficiently extended into 4D deformable models to segment a time series of cardiac images by introducing time-dependent constraints. A benchmark of deformable models segmentation was done in [Toh02] where the simplex mesh representation was chosen to model discrete meshes. Simplex meshes were also combined with region-based snakes by Tejos *et al.* [TICB09] who applied their method to MRI knee cartilage segmentation.

2.4.5 Discussion

We presented various deformable model types by introducing the most popular approaches. This overview discussed the properties of deformable models in image segmentation which have the following attributes.

Strengths:

- robustness: the weak knowledge on the shape bring robustness against image artifacts.
- subpixel accuracy: unlike voxel labeling approaches, deformable models have a subpixel accuracy.
- shape modeling: deformable models can model complex shapes with the adequate representation. Moreover, the shape reconstruction is carried out during the segmentation, avoiding the need of post-processing such as the Marching Cubes method [LC87].
- computation speed: except for some levelset approaches, deformable models often offer satisfactory computation times.

Weaknesses:

- initialization (capture range): most deformable model approaches require close initialization to converge to the correct location.
- poor shape variations control: various deformable models cannot capture fine details only if internal (smoothing and topology) constraints are relaxed. This requires ad-hoc tuning and there is the risk that the model will loose its robustness against image artifacts.

We refer interested readers to the reviews [JZDJ98, MT00, MDA01, NMK+06] for more details on deformable models.

2.5 Knowledge-based Deformable Models

In knowledge-based approaches deformable models are enhanced with prior knowledge derived from statistical and probabilistic analysis. In particular, statistical deformable models will be discussed by detailing stages to construct and use the statistical knowledge. The organization of this section is from many aspects inspired by the recent review of Heimann and Meinzer [HM09]. First, the selection of features to define the knowledge is described in Sec. 2.5.1. Subsequently, Sec. 2.5.2 explains how features are aligned after the correspondence has been established as described in Sec. 2.5.4. After that, Sec. 2.5.3 details how knowledge from the features is exploited to build underlying statistical and probabilistic models. To conclude, the different strategies to exploit these knowledge-based deformable models will be discussed in Sec. 2.5.5.

2.5.1 Features Selection

When knowledge-based models are built, the first choice to make is the selection of the information to be used. We denote this information as "features" and these are partly similar to those defined in Sec. 2.3 covering registration approaches. We can identify three main groups of features:

- shape-based features: points, curves, medial sheets/points, etc.
- appearance-based features: image information (pixel/voxel intensities, gradients, etc.).
- transformation-based features: relative pose between models, non-rigid deformations fields, etc.

In this thesis, we are mostly interested in the first and second groups, i.e. in the creation of *statistical shape and appearance models* (SSM and SAM), as detailed in the next Sections.

2.5.1.1 Shape-based Features

Features are extracted from the deformable models. In case of discrete models (Sec. 2.4.4), points are commonly selected as features. In the SSM's literature, points are referred to as *landmarks*. Only a subset of the point set can be selected as landmarks. For meshes, additional information derived from the connectivity can be included to allow the reconstruction of the surface from the landmarks, which is necessary in some correspondence (Sec. 2.5.4) and evolution algorithms (Sec. 2.5.5.2). Cootes *et al.* [CHTH93] were pioneers in exploiting the use of landmarks in SSMs in proposing the so-called *Point Distribution Models* (PDM). Landmarks were also used to study shape variations across population individual to establish anthropomorphic, anatomical or pathological differences [PLNS04, SXESG07].

Medial parameters [Blu67, PEFM98, PFJ+03] can also be used as features. In [HAGM04], various parameters (e.g., distance between medial nodes, radius) were jointly used to build hierarchical statistical shape models for segmentation of brain structures. Fletcher *et al.* [FLJ03] proposed a rigorous mathematical model to build SSM based on M-rep (3D medial models) and illustrated it with shape variations of the kidney. Similarly, continuous M-reps were exploited in [YZG05] to build a statistical shape model.

Another type of shape features could be defined. Instead of using features of the shapes, let us consider the space in which they are embedded. This concept is related to implicit deformable models and especially levelsets, which were presented in Sec. 2.4.3. In case of levelset, this means for instance that each pixel/voxel of the SDF map is used as a feature as in [LGF00, Cre06].

2.5.1.2 Appearance-based Features

The idea here is to extract appearance information relative to the shape. Appearance is built upon image information and can be very diversified: image intensity, gradient, texture, moments, etc. Region features cover the interior of the shape. The spatial sampling of these features usually depends on the image from which the features are extracted. Region features are very common in statistical shape models for e.g. face recognition [TP91, BDBS02, BMS02, YDB02, NL05], often referred to as "eigenfaces".

A major problem faced with region features is their use with 3D data. In addition to the problem of the curse of the dimensionality (Appendix Sec. B.5.3), the memory consumption as well as the computation effort are dramatically increased. Associated segmentation algorithms are thus affected by these constraints and local features are instead considered. Local features are typically defined around the boundaries of the shape, and are thus coined boundary features. The most common boundary feature is obtained by regularly sampling the image intensity/gradient along the normal direction of the shape as shown in Fig. 2.8. The collected values vectors are called intensity (gradient) profiles (IP), and they have been extensively used in knowledge-based deformable models [CHTH93, Coo01, BMVS02, HMMW07]. Compared to region features, boundary features yield more efficient algorithms in terms of speed and memory footprint. However, since only local information is used, the deformable model becomes more sensitive to the initialization as many image locations, not on the shape boundaries, might present similar local intensity neighborhoods.

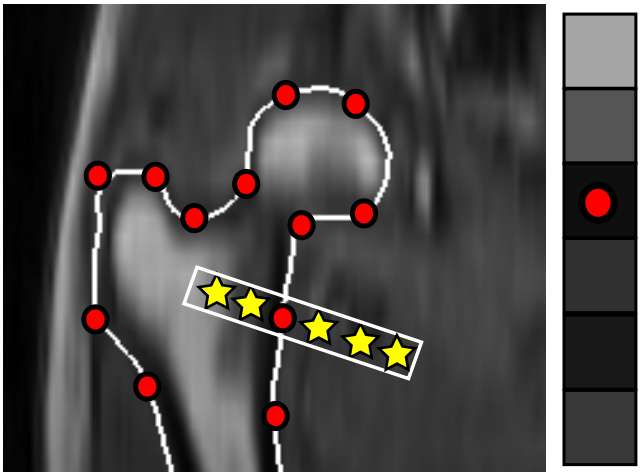


Figure 2.8: Boundary appearance feature. For each shape landmark position (circle) an intensity profile is extracted by sampling intensity values along the normal direction of the mesh at regularly spaced positions (stars). The intensity of the extracted profile is enlarged in the bar (right).

2.5.1.3 Transformation-based Features

Transformation parameters can be also used as features. Such features are not found in classic registration, since registration considers transformation parameters as variables to be estimated. Some authors considered the parameters of linear transforms as indicators of variability. In [BCP+06], local rigid transformations between consecutive vertebrae were used as features to build a *statistical deformation model* (SDM) of the spine. Since rotation matrices do not form a linear vector space, the Fréchet mean was used (see later on the discussion on linearity of statistical models construction in Sec. 2.5.3.4). Recently, rigid transformation parameters of the hip joint were embedded into SSMs by Kainmüller *et al.* [KLZH09] and Yokota *et al.* [YOT+09].

Non-rigid deformation parameters (B-Splines control points, dense deformation vectors) can also be selected as features. Rueckert *et al.* [RFS01, RFS03] particularly proposed to exploit the deformation field extracted from their FFD-based approach to build SDMs. Their approach was successfully exploited and extended in various works [KL04, FS06]. Further details on SDM are given in Sec. 2.5.4.4.

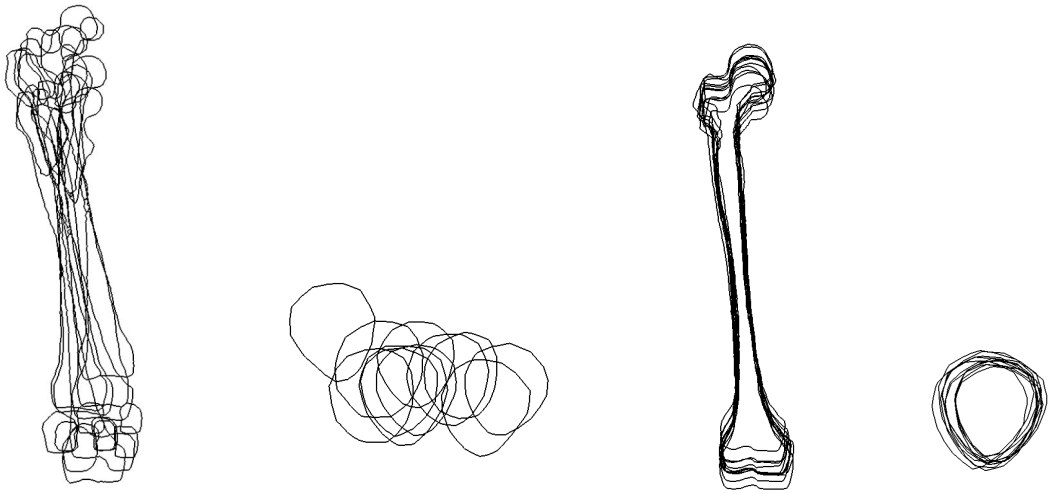
2.5.2 Alignment

Alignment attempts to remove the effects of pose changes that are not related to the intrinsic changes of features. The notion of shape is invariant under similarity transform, thus training shapes are commonly aligned with a similarity transform. In case of linear transforms (rigid, affine, similarity), the most popular approach is the generalized Procrustes approach [Gow75] (GPA), which iteratively aligns shapes by registering their corresponding point sets. The unknown transform $T^\star$ aligning two shapes $\boldsymbol{X}$ and $\boldsymbol{Y}$, is determined by a least square minimization:

$$T^\star = \underset{T}{\operatorname{argmin}} \sum_i \|\boldsymbol{y_i} - T(\boldsymbol{x_i})\|^2 \tag{2.8}$$

This least square problem can be efficiently solved by using Horn's quaternion [Hor87] and Umeyama's Singular Value Decomposition [Ume91] approaches to estimate rigid, affine and similarity transforms. Fig. 2.9 illustrates GPA with rigid alignment applied on femur shapes. GPA has the nice characteristic to not depend on any reference shape as all shapes are iteratively aligned with the mean $\bar{\boldsymbol{X}}$ which is constantly updated. Furthermore, GPA usually quickly converges after a few iterations.

However, since GPA is based on the L_2 norm, it is not robust against outlier points, as reported by Larsen and Eriksson [LE01] who proposed the use of alternative norms such a the L_1 or L_∞ norms. Furthermore, to preserve the linearity of aligned shapes during the procedure, Cootes *et al.* suggested to project the aligned shapes $\boldsymbol{X_i}$ into tangent space by applying the scaling $1/(\boldsymbol{X}.\bar{\boldsymbol{X}})$. While GPA is commonly used with shapes, similar approaches can be devised for appearance features. In [CET98b, Ste04], region features were processed with a variant of GPA to tackle changes in scale and offset among intensity samples. Projection to the tangent space was also performed.



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

Figure 2.9: Rigid alignment example. Ten not aligned femur shapes are represented in (a) coronal and (b) transversal views (here zoom factors differ between both views). Based on a generalized Procrustes (GPA) analysis with rigid transform, the shapes are aligned as illustrated in (c) coronal and (d) transversal views.

2.5.3 Knowledge-based Model Construction

Features that are in correspondence and aligned can be used as *training* feature sets from which statistical or probabilistic models of the features *variation* across the training sets can be derived. A subtle difference between both types of models can be made as follows [Koi06]. In probabilistic models, model parameters are usually assumed to be independent to ease model description and calculations, whereas in statistical models this independence assumption is rejected by considering that parameters might be inter-correlated (see Sec. B.3.4 for definitions and relationships between independence and correlation). As a result, statistical deformable models tend to create more compact representation of the feature variation. This has computational advantages as fewer parameters need to be controlled to impose valid variations to the features. In this Section, statistical models based on the principal component analysis (PCA) will be discussed. Subsequently, shape models and appearance models will be described. Note that in this Section, it is assumed that features are in correspondence. Details on correspondence methods are afterward given in Sec. 2.5.4 because they share many concepts of knowledge-based model construction, which needs to be presented first.

PCA, or Karhunen-Loève transform, is probably the most well-known multivariate data analysis. It was first introduced by K. Pearson in 1901 [Pea01] to find the best fitting of lines to a point cloud. Many research areas such as psychology, ecology, economics, biology, bioinformatics to mention a few, have been interested in PCA. The main PCA formulations are presented in Sec. 2.5.3.1 as well as its main extensions and variations in Sec. 2.5.3.2 and 2.5.3.3, respectively.

2.5.3.1 PCA Formulations

Let us consider each feature set as a vector $\boldsymbol{X_i} = (\boldsymbol{x_{1i}}, \ldots, \boldsymbol{x_{Mi}})^\top$ composed of M features of dimension d. A feature set is related to one *training sample*, hereon referred to as an *observation* to make a link with the statistics and probability theory. For instance, a training sample can be all the points of a discrete mesh ($d = 3$). Given N training samples, the mean $\bar{\boldsymbol{X}} = (\bar{x}_1, \ldots, \bar{x}_{dM})^\top$ is computed:

$$\bar{\boldsymbol{X}} = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{X_i} \tag{2.9}$$

The $dM \times N$ centered data matrix X is constructed by subtracting the mean to each sample:

$$X = \left(\boldsymbol{X_1} - \bar{\boldsymbol{X}}, \ldots, \boldsymbol{X_N} - \bar{\boldsymbol{X}} \right) \tag{2.10}$$

In other words, $X_{ij} = x_{ij} - \bar{x}_i$. The corresponding $dM \times dM$ covariance matrix Σ_X is:

$$\Sigma_X = \frac{1}{N-1} X X^\top \tag{2.11}$$

This covariance estimator is unbiased (Sec. B.4.1.2 in Appendix B for more details) and, when data follows a normal distribution, it corresponds to the Maximum-Likelihood (ML) estimator (Sec. B.4.3). The covariance matrix is symmetric positive definite, that is Σ_X is symmetric and $\forall \boldsymbol{x} \neq 0,\ \boldsymbol{x}^\top \Sigma_X \boldsymbol{x} > 0$. Theoretically, the matrix is invertible, but in practice it might happen that the estimator becomes unstable when $N << dM$. That is Σ_X will no longer be invertible by using numerical methods. A solution to this problem is to use regularization approaches [LLPS05, SS05, Bou06]. These approaches replace the covariance matrix elements with new ones ensuring that the covariance matrix is invertible. Now that we have defined the essential quantities, let us present the various PCA formulations. In the following, we set $m = dM$.

Change of Basis Formulation The aim of PCA is to find an orthonormal matrix P where $Y = PX$ such that Σ_Y is diagonal [Shl03]. It can be proven that:

$$\begin{aligned} &\Sigma_Y = diag(\lambda_1, \ldots, \lambda_m). \\ &\lambda_i \text{ are the eigenvalues of } \Sigma_X. \\ &\text{The rows of } P \text{ are the eigenvectors } \Phi = \{\boldsymbol{\phi_1}, \ldots, \boldsymbol{\phi_m}\} \text{ of } \Sigma_X. \end{aligned} \tag{2.12}$$

This change of basis means that any observation $\boldsymbol{X_i} \in \mathbb{R}^m$ can be defined in a new basis $\{\boldsymbol{\phi_i}\}$ as a new vector $Y_i \in \mathbb{R}^m$ (and vice-versa):

$$\boldsymbol{X_i} = \bar{\boldsymbol{X}} + \Phi \boldsymbol{Y_i} \tag{2.13}$$

$$\boldsymbol{Y_i} = \Phi^\top (\boldsymbol{X_i} - \bar{\boldsymbol{X}}) \tag{2.14}$$

In practice, any observation is a linear combination of the eigenvectors $\boldsymbol{\phi_i}$, which are called *principal axes* or principal directions. Since the covariance matrix Σ_Y is diagonal, vectors $\boldsymbol{Y_i}$ are *inter-uncorrelated*. The components of $\boldsymbol{Y_i}$ are the *principal components*. PCA selects the principal directions in the $m-$dimensional space along which the variance in X is maximized.

The effective computation of PCA is hence performed by an eigenvalue decomposition (EVD) of Σ_X, which is commonly carried out with the Cyclic Jacobi's method [GVL96]. Eigenvectors are sorted according to the eigenvalues (always positive since Σ_X is symmetric definite positive) in the decreasing order, so that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m$. If $m >> N$, it is better to perform the eigenvalue decomposition of the transposed covariance matrix $\Sigma_X^\top$. In fact, if $\boldsymbol{u_i}$ and l_i are respectively the eigenvectors and eigenvalues of $\Sigma_X^\top$, $X\boldsymbol{u_i}$ and l_i will be the N first eigenvectors and eigenvalues of Σ_X, respectively. To get higher numerical stability, it is also preferable to replace the EVD with the Singular Value Decomposition (SVD) performed on the centered data matrix X [Shl03].

It is possible to look for a change of basis that will achieve a *dimensionality reduction*. In fact, by keeping the first k eigenvectors of the covariance matrix Σ_X, an affine *subspace* $\mathbb{E} \subset \mathbb{R}^m$ of dimension k is defined. A projection operator $P_\mathbb{E}$ onto the subspace is then specified [Bou06]:

$$\begin{aligned} P_\mathbb{E} &: \mathbb{R}^m \longrightarrow \mathbb{R}^m \\ P_\mathbb{E}(\boldsymbol{X}) &= \Phi_k \Phi_k^\top (\boldsymbol{X} - \bar{\boldsymbol{X}}) + \bar{\boldsymbol{X}} \end{aligned} \tag{2.15}$$

where Φ_k is the $m \times k$ matrix equal to Φ without the last $m - k$ columns. Based on this operator, $\hat{\boldsymbol{X}} = P_\mathbb{E}(\boldsymbol{X})$ is the *reconstructed* observation based on the PCA statistic. When $k = m$, $\boldsymbol{X} = \hat{\boldsymbol{X}}$ as expected.

Dimensionality reduction is beneficial (i) to reduce the complexity of the model and (ii) to retain the more "meaningful" information by filtering out possible noise in the data. Complexity reduction is necessary to devise efficient algorithm in terms of speed and memory consumption, while noise filtering yields reconstructed values that better reflect the real (noise-free) underlying model. The difficulty stands in choosing an appropriate value for the number k of retained modes. Many approaches based on heuristic and statistics were proposed [Cat66, EK82, Jac93, Fer95, VLQ99]. Recently, Cangelosi and Goriely [CG07] tested many of them with synthetic and cDNA microarray datasets, and proposed a new one based on Shannon Entropy. Their conclusions were data dependent and they did not give a definitive answer on which method should be selected. They aptly quoted the words of Ferré [Fer95] and Jolliffe [Jol02], who said that a best approach does not exist and that simple approaches can be as efficient as more complex ones in some situations. In image segmentation context, it is moreover difficult to predict the impact of the value k of retained components in a segmentation algorithm because other factors come into play (image quality, subject anatomical and pathological characteristics, etc.). A simple and very popular approach to determine k uses the ratio "total variance explained by the k principal components/total variance expressed by all principal components":

$$p = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{m} \lambda_i} \tag{2.16}$$

By choosing p, commonly in the interval $[0.9, 0.99]$, the corresponding value of k is computed. With k chosen and the principal axes matrix Φ_k computed, the PCA defines a new statistic which provides a *generative model* to create new samples $\boldsymbol{X}$:

$$\forall \boldsymbol{Y} \in \mathbb{R}^k,\ \boldsymbol{X} = \bar{\boldsymbol{X}} + \Phi_k \boldsymbol{Y} \tag{2.17}$$

Figure 2.10 depicts various shapes generated by using the generative model (2.17) based on a point distribution model (features are points). Shape variations is obtained by varying the first shape coefficient y_1 in the interval $[-3\sqrt{\lambda_1}, +3\sqrt{\lambda_1}]$. This interval is derived from constraints applied on the parameter vector $\boldsymbol{Y}$ to ensure the correctness of the generated samples as explained in the following probabilistic formulation.

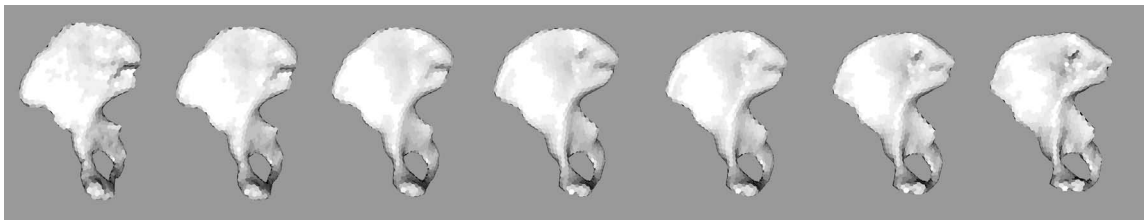


Figure 2.10: Generative model of PCA illustrated with hip bone shape variations. In this example, a point distribution model of a hip bone shape has been built, and various shape instances are generated by varying the first shape coefficient between $-3\sqrt{\lambda_1}$ (first shape on the left) and $+3\sqrt{\lambda_1}$ (last shape on the right). Middle shape represents the mean shape of the PCA model.

Least Squares Fitting Formulation In multivariate data analysis, a $m \times N$ model matrix F is fitted to the centered data matrix X under certain constraints. It is proven that PCA achieves the optimal fitting of the data X in a *least squares* sense [Kie97, AOC07]. This fitting is thus equivalent to finding F such that the following function e is minimized:

$$e(F|X) = \underset{F}{\operatorname{argmin}} \sum_{i=1}^{N} \sum_{j=1}^{m} (F_{ji} - X_{ji})^2 \tag{2.18}$$

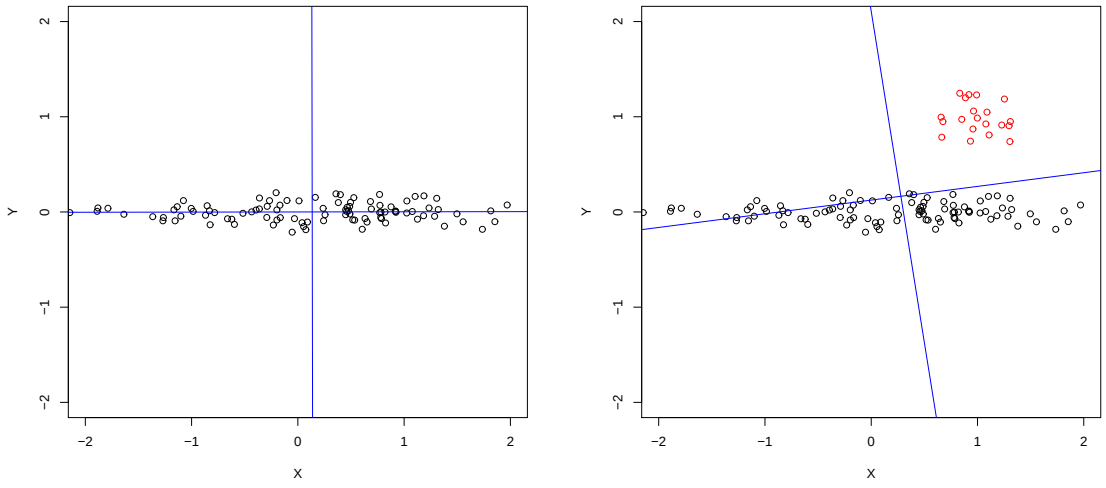
Function e is called the reconstruction error and it is greater than 0 when the dimensionality reduction is applied. In that case, the constraints on F are easily enforced by writing $F = AB^\top$, where A and B are $M \times k$ and $m \times k$ matrices, respectively [Kie97]. Given Eq. (2.15), natural choices are $A = \Phi_k$ and $B^\top = \Phi^\top(X - \bar{X})$. In Fig. 2.11, the least squares fitting is illustrated with points drawn from a Normal distribution. The L_2 norm is known to be sensitive to outliers as exemplified in Fig. 2.11(b). In Sec. 2.5.3.2 we will discuss about robust extensions of PCA that identify and tackle outliers.

Probabilistic formulation Tipping and Bishop [TB99] proposed a probabilistic formulation of PCA, coined PPCA. They considered a *linear generative model* which assumes that the samples $\boldsymbol{X_i}$ are observations of a Random Variable $\boldsymbol{X}$ generated by

$$\boldsymbol{X} = A\boldsymbol{S} + \boldsymbol{V} \tag{2.19}$$

where $\boldsymbol{S} \in \mathbb{R}^k$ $(k \leq m)$ follows a unit isotropic Gaussian distribution with zero mean and A is the $m \times k$ matrix of factor loadings. Noise $\boldsymbol{V}$ is also Gaussian ($\boldsymbol{V} \sim \mathcal{N}(\boldsymbol{0}, \Sigma)$, with $\Sigma = diag(\sigma_1^2, \ldots, \sigma_m^2)$) and is assumed to be independent from $\boldsymbol{X}$. Given this formulation, $\boldsymbol{X}$ also follows a normal distribution: $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{0}, C)$, with $C = \Sigma + AA^\top$.

The distribution of $\boldsymbol{S}$ is defined in a space of lower dimension, and it models interdependencies between the observed variables *independently of the noise*. Choi [Cho04] stressed that this was a



(a) (b)

Figure 2.11: PCA least squares fitting. (a) 100 points drawn by a 2D Normal distribution ($\sim \mathcal{N}((c,\mu)$, $c = (0,0)$, $\mu = (1,0.01)$) are considered as observations and analyzed with a PCA. Black lines represent the principal axes (PA) passing through the mean point, the first PA being the almost horizontal one. This figure illustrates how the first PA captures most of the variance and how PCA tries to fit at best the data in a least squares sense. (b) 20 additional points (upper right point cluster in red) are added to the observations of (a) (in black). These new observations are not drawn from the same distribution and could be considered as outliers. This figure illustrates how PCA is sensitive to outliers since the PAs underwent a rotation to satisfy at best the new observations.

main difference with respect to classic PCA, which identically treats the inter-variable dependencies and the noise. In case where the noise is isotropic, i.e. $\sigma_1^2 = \ldots = \sigma_m^2 = \sigma^2$, it can be shown [TB99] that there exists an ML estimator A_{ML} of the matrix A which *spans* the principal subspace of dimension k. More specifically, the columns of A_{ML} are linearly independent vectors that are *rotated* and *scaled* eigenvectors of the centered data matrix X (2.10) (which should not be mixed with the RV $\boldsymbol{X}$).

In practice this implies that A is generally not the matrix of principal directions Φ_k (2.15), as well as the corresponding $\boldsymbol{S_i}$ will not represent the principal components. A post-processing is necessary to recover the principal axes and eigenvalues. Ahn *et al.* [AOC07] and Choi [Cho07] proposed alternative approaches that calculate the principal axes without any rotational ambiguity. Both algorithms are partially based on the minimization of the integrated-square-error (in contrast with the square-error of PCA as described in Sec. 2.5.3.1). Choi's wake-sleep PCA (WS-PCA) relies on an Helmoltz machine that couples generative and recognition models, while Ahn's EM-ePCA approach is based on the EM-algorithm of Roweis [Row98]. The EM-algorithm is derived from the limiting case of the PPCA when $\sigma \to 0$. It was shown [TB99, Row98] that in this special case PPCA becomes equivalent to PCA. In particular, the likelihood of the RV $\boldsymbol{X}$ is uniquely dominated by the squared Euclidean distance between X and its reconstruction $A\boldsymbol{S}$ [Cho04]. This observation supports the least square fitting formulation of PCA. The steps of the

EM-PCA algorithm are as follows [Row98]:

$$\text{E-Step: } S = (A^\top A)^{-1} A^\top X \tag{2.20}$$

$$\text{M-Step: } A^{new} = XS^\top (SS^\top)^{-1} \tag{2.21}$$

where S is the $k \times N$ matrix $S = (\boldsymbol{S_1}, \ldots, \boldsymbol{S_N})$ and X the centered data matrix (2.10). Albeit EM is a local minimization approach, Tipping and Bishop [TB99] proved that such a algorithm converges to a *global* minimum. In terms of computations and memory consumption, EM-PCA is a more attractive algorithm compared to PCA approaches which explicitly compute the sample covariance matrix Σ_X (2.11).

Final interesting property of considering the PCA under a probabilistic angle is that constraints can be applied on the principal components to enforce the "validity" of reconstructed samples. This is to ensure that new samples are drawn from the statistic inferred from the observations. The constraints are applied on the parameters vector $\boldsymbol{Y}$ in Eq. (2.17). Two main different types of hard constraints [CHTH93] can be applied. The first type of constraints is the *clamping* constraint, which clamps the components y_j to the interval $[-3\sqrt{\lambda_j}, 3\sqrt{\lambda_j}]$. This relation derives from the assumption that each component follows a normal distribution $y_j \sim \mathcal{N}(0, \lambda_j)$. The second kind of constraints assumes that vector Y is the realization of a RV distributed according to a multivariate Gaussian. As a result, the sum of squared components follows a χ^2_k distribution. The *scaling* constraint scales the y_j so that $\sum(y_j^2/\lambda_j) \leq M$, where M is computed from the χ^2_K distribution. These constraints will be particularly essential in the usage of statistical deformable models.

2.5.3.2 PCA Extensions

PCA has been extended into many variants in order to improve its computational behavior (speed, memory consumption, sequential processing) as well as to support new features.

Iterative PCA In its original form, PCA requires the computation of the covariance matrix and the eigenvalue decomposition. EVD with the Cyclic Jacobi's method [GVL96] has a complexity around $O(m^3 + m^2 N)$ [SP07b]. SVD directly operates on the centered data matrix and has a lower complexity of $O(4m^2 N + 8mN^2 + 9n^2)$. When m becomes large, both EVD and SVD are hence very expensive approaches. Iterative approaches that compute the principal axes (PA) iteratively are interesting alternatives since they can be stopped after the calculation of the first k necessary PAs. The probabilistic PCA algorithms [Row98, TB99] fulfill these requirements, but they require post-processing since the obtained subspace only spans the principal subspace (Sec. 2.5.3.1). In [SP07b], a fast PCA approach based on the fixed-point algorithm of Hyvarinen and Oja [HO97] was proposed. This approach ensures the PAs' orthonormality but still requires an initial computation of the covariance matrix Σ_X, which has a complexity of $O(m^2 N)$.

The power method [GVL96] and the nonlinear iterative partial least squares (NIPALS)-PCA [WEG87] are other iterative approaches. In its original form, the power method is only designed to compute the first PA so is somehow limited [1]. The NIPALS-PCA extracts sequentially the

1. Power method can be extended to compute the other PAs through deflation, becoming similar to the NIPALS-PCA

first k PAs while preserving the orthogonality. However in case of large samples, NIPALS-PCA suffers from rounding errors that ultimately degrade the orthogonality of the PAs, as reported by Andrecut [And09]. The author hence proposed to enforce orthogonality in the algorithm by using the robust Gram-Schmidt orthogonalization process. The new GS-PCA algorithm being naturally parallelizable, a GPU implementation based on NVidia's Compute Unified Device Architecture (CUDA) architecture was moreover proposed by the author.

incremental PCA Aforementioned approaches are batch methods. This means that when new observations are made available, these methods need to be executed from scratch with the old and new data as input. In addition of being very time consuming, this raises the issue of storage because the old original *high-dimensional* observations must be kept. This also affects on-line computing as a remote workstation performing the calculation needs to get a copy of the observations first. As a consequence, incremental (or sequential) PCA methods were proposed. Hall *et al.* [HMM02] described arithmetic rules for addition and subtraction to be applied on eigenspaces both for the EVD and SVD algorithms, which inspired the incremental PCA of Skočaj and Leonardis [SL03]. In [LXMJ03], an incremental approach to accept new observations was described but it required the experimental setting of a weight parameter to "blend" old with new data. Similarly, Choi [Cho04] extended the EM-PCA into a sequential approach, with the extra feature of producing rectified subspaces (i.e. non negative principal axes and modes). Costache *et al.* [CCP09] proposed an algorithm based on [HMM02, LXMJ03] which was able to enlarge a PCA with new observations as well to fuse multiple PCAs into a new one.

robust PCA As previously mentioned in Sec. 2.5.3.1, PCA is quite sensitive to *erroneous* features. Erroneous features are usually the result of (i) *missing* or (ii) *perturbed* data. Missing data is for instance quite common in computer vision, where due to occlusions, some image information may be absent. Perturbed data can be the result of noise in the process which creates or estimates the features. Since PCA needs to have feature vectors of same size, a naive approach simply replaces the missing features with arbitrary values, which in practice yields erroneous principal axes and components. Another solution consists in discarding the observations with some erroneous features from the training dataset. While this might be acceptable when many training samples are available, there is often in practice a lack of sufficient training data, especially for 3D segmentation [HM09]. Hence, the rarity of the training data demands alternative PCA approaches. Robust PCA methods attempt to estimate the principal directions and components from data with erroneous features.

Supervised approaches know which features are erroneous and what are the confidences for each features. This is simply translated as a weight w_{ij}. The lower the weight, the less the jth feature of the ith observation is reliable. Many approaches aim thus to minimize the *weighted* reconstruction error h:

$$h(F|X) = \sum_{i=1}^{N} \sum_{j=1}^{m} w_{ji}^2 (F_{ji} - X_{ji})^2 \tag{2.22}$$

approach

where $F = \Phi_k Y$, Φ_k being the matrix of k principal axes, and Y the matrix of principal components. Skočaj *et al.* [SLB07] presented two supervised robust approaches. The first one, wEM-PCA, extends the PCA's EM iterative algorithm (2.21) to support weights. As a result, this algorithm yields non orthogonal components which requires a post-processing. The second one (mdEM-PCA) is again inspired by the EM approach, but during the M-step the principal axes are estimated by computing a classic PCA on corrected data, in which missing information is replaced by reconstructed values. Since classic PCA is used, principal axes are orthogonal and eigenvalues are computed. In [Kie97], classic PCA was also used in a robust PCA algorithm (wPCA), but with a different approach than the EM-PCA. The described method exploits the minimization of a majorizing function to compute iteratively an estimate of the corrected centered data matrix X which is then used in a classic PCA step to get the principal axes and components. This majorizing function is simpler to compute than h (2.22), and is thus easier to minimize. Grabiel and Zamir [GZ79] also proposed a method to directly weight the SVD to achieve robust PCA.

Unsupervised approaches try first to identify erroneous features and then propose mechanisms to handle them. Main approaches are based on robust estimators that are applied e.g. on the covariance matrix Σ_X [CH00] or on the reconstruction error (2.22) [DlTB01, LXMJ03]. Based on their missing data algorithm, Skočaj *et al.* [SLB07] also proposed an unsupervised robust algorithm which was tested in a computer vision scenario. More information on robust approaches can be found in [Che02].

2.5.3.3 PCA Variations

The presented PCA extensions modified the classic approach to improve it and always aimed at computing the (noise-free) principal subspace. There exists some other approaches that intentionally change the PCA behavior to yield new properties. These methods result in new set of directions and components that are no longer strictly the "principal" ones.

Sparse PCA In general, PCA computes global modes that simultaneously affect all the features. In case of shape models, this means that the variation of one mode will influence all the model landmarks. The process of computing modes that dictate the behavior of a limited and preferably locally clustered number of features is called *sparse* PCA (SPCA) [HM09]. Sparsity enhances human understanding, reduces computational cost and yields a better generalization in learning methods [MWA06]. Sparse PCA can be seen as the maximization of the quadratic form $\boldsymbol{x}^\top \Sigma_X \boldsymbol{x}$, where $\boldsymbol{x}$ is a sparse vector having a maximum of k non null elements. Sparse PCA will thus produce principal axes that differ from the original ones, for which strict [SSL06b, JNRS10] or relaxed [ZHT06, SH08] orthogonality can be enforced. In [SSL06a], sparse PCA was for example presented as more appropriate to understand the variability of some anatomical structures such as the lungs, heart or brain structures.

Independent Component Analysis (ICA) [Hyv99] also yields sparsity by producing non orthogonal axes and *independent* principal components. ICA is useful when observations are not drawn

by a Gaussian law but requires new approaches to select and order the principal axes. No consensus has been reached on stating which of PCA or ICA is better for segmentation or recognition purposes. In [ÜFS+03], ICA-based appearance models were superior to PCA-based priors in segmenting heart structures for cardiac MR. Baek *et al.* [BDBS02] however observed better performance for PCA in a human face recognition scenario, although Bartlett *et al.* [BMS02] concluded the opposite on the same FERET database. Probabilistic ICA also exists and has been exploited in the segmentation context [dRKD00].

Hierarchical PCA In *hierarchical* PCA approaches, local changes are modeled by organizing features data into sub-groups in a vertical (multi-scaling) or horizontal (division) manner. The multi-scale approach decomposes the features using appropriate schemes such as the wavelet transform [DXD03]. This yields a coarse-to-fine PCA modeling. Features space partitioning is also commonly adopted, like the Sub-pattern PCA of Tan and Chen [TC05] or the collective PCA of Kargupta and Huang [KHSJ01]. Clustering approaches were moreover proposed to address the unsupervised selection of sub-groups. Heap *et al.* [HH97] notably devised a clustering method that yielded SSMs with a better Specificity, i.e. a better expression of valid instances (see Sec. 2.5.4.7 for details on this measure).

Size-constrained PCA As discussed in the sparse PCA paragraph, the variation along one mode usually affects many variables simultaneously. In shape-based PCA, it is often observed that changes in size are correlated with changes in shape. This property is named *allometry*. There is however often an interest in morphological shape analysis in partitioning components of variation related with size from components associated with shape [Sun89], resulting in *isometry*. With that perspective, size-constrained PCA [Som89] attempts to find the first axis summarizing variation in isometric size alone. This approach was later on improved by Cadima and Jolliffe [CJ96].

2.5.3.4 PCA Discussions

In this Sec. on PCA, we have seen various ways to interpret, improve or modify PCA. Main PCA formulations and extensions are summarized in Table 2.4. The presented approaches do not constitute an exhaustive list but gives a clear picture of the richness of PCA from which segmentation could profit. Wrapping up, PCA relies on the following assumptions [Shl03]:

- *Linearity*: the change of basis formulation highlights the data linearity assumption of PCA, which might be violated with real data. To tackle this, the kernel PCA approach [SSM98] were proposed and efficiently coupled with deformable models [CKS03, DRT06]. Kernel PCA transforms (by means of kernels) features into a linear space in which conventional PCA can be carried out.
- Mean and variance are *sufficient statistics* (Sec. B.3.3): this implies that observations should be exponentially distributed, otherwise the generative model of PCA (2.17) might create invalid instances. PCA can be thus generally extended to other distributions of the exponential family. Collins *et al.* [CDS02] proposed other distributions of the exponential family that are more

adapted to non Gaussian observations, such as the Poison and Bernoulli distributions (Table B.1) for integer- and binary-valued data, respectively.

– Large variances are *meaningful*: PCA sorts the principal directions depending on the value of eigenvalues, and thus the data variances. By adopting the dimensionality reduction, modes associated with low variances are implictly considered as less interesting information.
– Principal directions are *orthogonal*: this allows the use of linear algebra techniques to solve PCA and also offers an efficient way to project data in the PCA space. It is thus advised to favor methods yielding orthogonal directions or to apply post-processing to recover the orthogonality (e.g., for EM-based PCAs).

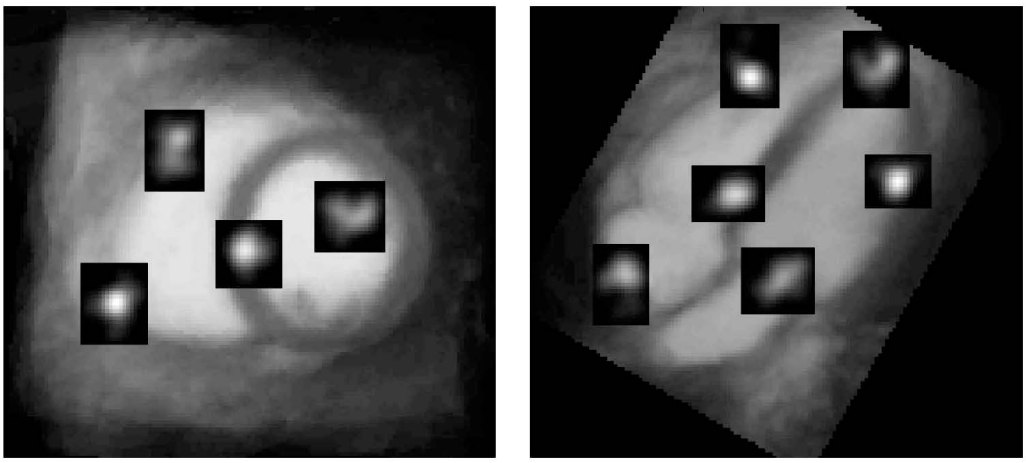
Name	**Ortho.**	**Batch**	**Covariance**$^{(-1)}$	**Noise**	**Robust**	**References**
PPCA	No	Yes	No	Yes	No	[TB99]
EM-PCA	No	Yes	No[a]	No	No	[Row98]
EM-ePCA	Yes	Yes	No	No[b]	No	[AOC07]
WS-PCA	Yes	Yes	No	No	No	[Cho07]
Fast PCA	Yes	Yes	Yes	No	No	[SP07b]
NIPALS-PCA	Yes[c]	Yes	No	No	No	[WEG87]
GS-PCA	Yes	Yes	No	No	No	[And09]
Seq. EM-PCA[f]	No	No	No	No	No	[Cho04]
Mul. Inc. PCA[f]	Yes	No	Yes[d]	No	No	[CCP09]
Inc. PCA[f]	Yes	No	Yes	No	No	[SL03]
Inc. Rob. PCA[f]	Yes	No	Yes[e]	No	Yes	[LXMJ03]
wEM-PCA	No	Yes	No	No	Yes	[SLB07]
wPCA	Yes	Yes	Yes	No	Yes	[Kie97]
mdEM-PCA	Yes	Yes	Yes	No	Yes	[SLB07]
RPCA	Yes	Yes	No	Yes	Yes	[DlTB01]
Rob. Cov. PCA[f]	Yes	Yes	Yes	No	Yes	[CH00]

[a] Many EM-based methods, do not *explicitly* compute Σ_X, but rather calculate some of its values
[b] The authors also present a version which estimates the noise variances
[c] Possible loss of orthogonality due to rounding errors
[d] Classic PCA is initially performed at the various sites
[e] Classic PCA is carried out in some stages
[f] seq.: sequential, mul.:multiple, inc.: incremental, rob.: robust, cov.: covariance

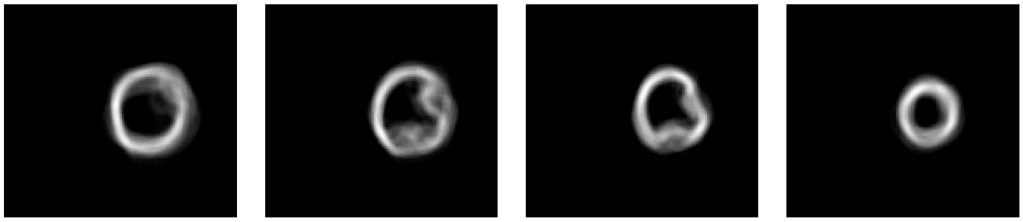
Table 2.4: Different PCA formulations and extensions. Names are presented in the text. **Ortho.**: ensures orthogonality of the computed principal axes; **Batch**: is a batch method, i.e. it needs to recompute information from scratch when new data is added; **Covariance**$^{(-1)}$: might require the computation of the (inverse) covariance matrix Σ_X; **Noise**: can estimate noise in data based on a noise model; **Robust**: support detection/correction of outliers.

2.5.3.5 Shape Models

In previous sections, we briefly presented point distribution models (PDM) [CHTH93]. PDMs are simply models based on landmark features, usually associated with a triangularization, and from which a shape variations model is inferred by PCA. These models are very popular in medical image analysis. Additional details on how to establish features correspondence and use PDM are given in Sec. 2.5.5 and Sec. 2.5.4, respectively.



(a)



(b)

Figure 2.12: Probabilistic shape variations modeling. (a) In case of heart ventricles and atrial surface points variability modeling, non-parametric PDFs of the surface points variations are estimated by KDE based on Parzen windowing [Par62] (See Appendix Sec. B.5.2). In each small box, the non-parametric PDF of the surface point in the center of the box is shown. (b) Heart left ventricle surface variations modeling by probabilistic atlas. From left to right, a slice of the probabilistic atlas is depicted from the valve level to the apical level. Bright intensity indicates high probability. Images courtesy of Juha Koikkalainen [Koi06].

Other approaches were proposed to model shape variations. A first approach consists in estimating Probability Density Functions (PDF, See Appendix B.3.2) of shape variations across a dataset. PDFs can be first assumed to be known and the Gaussian model is commonly chosen to express variations. For instance, in [SD96] a multivariate Gaussian distribution describes the variability of the surface modeling parameters based on parametric Fourier descriptors. However, many biological shape changes cannot be limited to Gaussian processes. As a result, instead of assuming the form of the PDFs, non parametric estimation was chosen to approximate the real PDFs. For example, Koikkalainen [Koi06] expressed the surface points variability of the heart by using a kernel density estimation (KDE) based on Parzen windowing [Par62] (See Appendix Sec. B.5.2), as illustrated in Fig. 2.12(a). Similarly, instead of computing a PDF per point, the author built

a probabilistic atlas of the shape variations (Fig. 2.12(b)). A probabilistic atlas is constructed by (i) aligning the shapes to remove variations not related to the shape, (ii) converting shapes to images by a grid discretization process and (iii) averaging the images. Alternatively, instead of using shape surfaces, the signed distance maps of shapes [SCD+03] can be used to build the atlas.

Other shape models were proposed. Kervrann and Heitz [KH98] proposed to model global shape variation by a PDM model and local variation by a Gauss Markov process. They applied their method to video object tracking. In [Pen90, PS91], shape variations were derived from a modal decomposition based on FEM. A base shape was chosen, such as an ellipsoid, and adequate modes of vibration were built to capture variability in the training shapes dataset. They proposed in particular a closed-form solution [PS91] to model complex 3D shapes.

2.5.3.6 Appearance Models

As presented in Sec. 2.5.1.2, two main kinds of features are available: region and boundary features. One of the most popular appearance models assumes that features follow a Gaussian distribution. Mean $\boldsymbol{\mu}$ and covariance matrix Σ_X are thus estimated from the training features X. The goodness of fit of a feature $\boldsymbol{u}$ to this appearance model is then simply achieved by calculating the Mahalanobis distance d_M: $d_M(\boldsymbol{u}) = (\boldsymbol{u}-\boldsymbol{\mu})^\top \Sigma_X^{-1}(\boldsymbol{u}-\boldsymbol{\mu})$. By building a PCA from the features, the Mahalanobis distance can be instead expressed as a function of the eigenvalues and the eigenvectors matrix [CHTH93]. Furthermore, by keeping the first k modes, a more robust appearance model to noise is devised. This appearance model is part of the Active Shape Models (ASM) of Cootes *et al.* [CT04]. Cootes *et al.* [CHTH93] and Behiels *et al.* [BMVS02] used intensity and (normalized) gradient boundary profiles to build PCA models, and exploited the Mahalanobis distance. These two studies drew different conclusions with respect to the best choice of profile (intensity vs. normalized gradient), suggesting that choices strongly depended on the type of image modalities (video vs. X-rays) and structures (face vs. bone).

When region features are used, the most popular appearance model is Active Appearance Models (AAM) [CET98b, CT04]. The AAM appearance model also uses PCA modeling to express the modes of variations of the shape interior appearance, usually denoted as texture. To decouple the appearance from the shape and to be able to apply PCA, textures need first to be transformed into a reference shape (commonly the mean shape of the shape PCA model) by using a warping approach. Various warping strategies exist, such as mesh-based and TPS [CT04, Rog01] or piecewise affine [Ste02, Ste04] warpings. Linear [CT04] and non-linear [BML+02] texture intensity normalizations are usually required to cope with strong intensity variations among training images. Also, alternative intensity features (e.g., edges, corners, and gradients [SCT03]) less sensitive to these intensity changes can be included into the appearance model. Similarly, multiple views of a structure can be considered in the AAM construction [OLU+03] to improve the appearance model robustness. Region-based models with AAM were successfully exploited in the literature [OLU+03, MLvdG+00, LvdGM+02, ÜFS+03, SFK+08], but AAMs have their shortcomings which may demand appropriate evolution strategies (e.g., Robust AAM [BBLS05]).

Similarly to shape models, non Gaussianity of region features can be addressed by building probabilistic atlases of appearance [Koi06, SCD+03]. The choice of the reference image can affect the results as bias is introduced [RFS03]. Various authors proposed to carefully select it based on distance and similarity criteria [PBHM05], or to remove it by achieving the simultaneous registration of all training images [JDJG04]. More details on atlas-based construction can be found in Sec. 2.5.4.4.

In case of boundary features, non-parametric approaches are available to build non-normally distributed appearance models. In particular, the k-Nearest Neighbor (k-NN) classifier [Cov68] was used in a segmentation context [dBGVN03, BGNV03, HMMW07]. To achieve this, sets of boundary $\mathcal{S}$ and non-boundary $\urcorner\mathcal{S}$ training profiles are computed during the training phase to define the two classes *border* and $\urcorner$*border*, respectively. During the segmentation, a profile u of dimension d belongs to a boundary if its probability $P(border|u)$ is over a threshold. The k closest "neighbor" profiles in $\urcorner\mathcal{S} \cup \mathcal{S}$ to u are computed based on a profile distance, typically the Euclidean distance in $d-$dimension. The moderate k-NN probability [AK02] is then expressed as $P(border|u) = (K+1)/(k+2)$, where K is the number of neighbor profiles that belong to S. Since all training profiles need to be constantly explored to compute the k nearest neighbors, efficient (approximate) k-NN computation methods were proposed [War96, AMN+98, AASK08]. The use of k-NN classifier is theoretically applicable with any type of appearance features. For instance, Van Ginneken *et al.* [VGFS+02] replaced intensity profiles by 2D square grids centered at the landmark points. As reported by Heimann and Meinzer [HM09], any classifier returning a probability can be in general used and combined for segmentation purpose, as exemplified in [LZJ04, LI05, NSC+05]. Fusion of multiple classifiers is a topic which has been extensively studied in pattern recognition. We refer interested readers to [LS97, ACK01, KA03].

To improve appearance models and to minimize the effects of the curse of the dimensionality, clustering approaches can be adopted. In clustering approaches, appearance features are clustered into classes from which statistics are inferred, via e.g. PCA. This yields various advantages such as the increase of training samples number for each landmark (good for statistical inference), the reduction of memory consumption and computational cost since during segmentation extracted profiles need only to be compared against the clusters. Brejl and Sonka [BS00] were the first to propose the clustering of appearance feature vectors for appearance models. They used the fuzzy C-means algorithm [BE+84] to estimate K clusters from all landmarks in all training images. Similarly, Stough *et al.* [SBPC07b, SBPC07a] also used the fuzzy C-means to cluster appearance features based on intensity quantile histograms. In [PEW05], features computed as the dot product between surface normals and image gradients were clustered and used in heart segmentation from CT cardiac images. In [HMMW07], the accuracy of the k-NN classifier was improved by using the k-Means clustering algorithm [HW79]. Chung and Delingette [CD09] later on proposed an EM-based algorithm, initialized with fuzzy C-means, to achieve clustering and classification of intensity profiles. In particular, their approach proposed an automatic selection of the best number K of clusters. Moreover, a spatial smoothing was applied because appearance-based approaches did not consider neighborhood information, which might lead to a

non-spatially smooth distribution of the intensity profiles into the different clusters. The process is illustrated in Fig. 2.13.

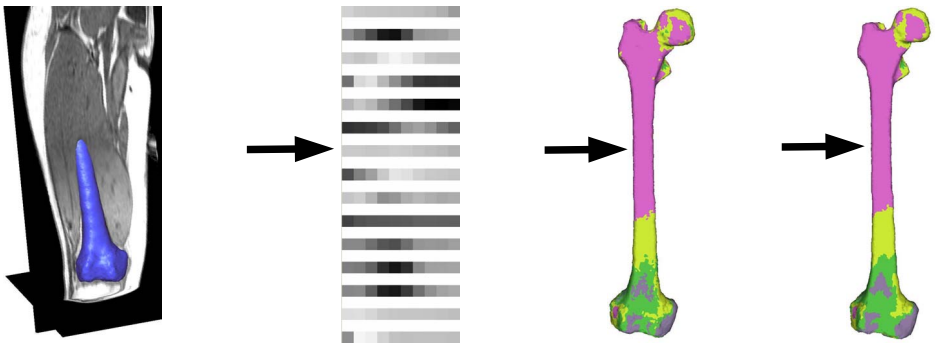


Figure 2.13: Intensity profiles clustering. (a) From a MRI segmented femur mesh, (b) intensity profiles are extracted and an (b) EM-based algorithm is used to achieve profiles clustering by optimally choosing the best number of clusters. (c) Finally a last smoothing processing is applied to account for neighborhood information. Images courtesy of François Chung.

Similarly, Ho and Gerig [HG04] considered a multi-scale spatial clustering to hierarchically gather profiles based on neighborhood information. Finally, various hybrid appearance models can be found in the literature. For instance, Yushkevich *et al.* [YZG05] exploited the continuous medial representation to combine appearance ASM- and AAM-like appearance models.

2.5.4 Correspondence

In the previous Sections it was assumed that the features were aligned and in correspondence. Since some correspondence approaches are based on concepts of PCA, it appears as more natural to detail now the various correspondence approaches. We present various correspondence approaches that are suitable for appearance correspondence (e.g., image-based correspondence Sec. 2.5.4.4) with accent on shape correspondence method based on landmarks. Moreover, when shape correspondence is achieved, appearance correspondence might be directly available.

Approaches that exploit the correspondence of features across training samples require accurate and meaningful correspondence. Otherwise the resulting model will capture the errors introduced by the (absence of the) correspondence procedure. So far, in the literature no consensus has been reached on the choice of the best correspondence approach. Usually, the application context drives users' choices. Human annotation is often regarded as a gold standard (like manual segmentation) to define correspondence but it is unpopular. This is mainly due to the time-consuming and tedious aspect of the approach (especially in 3D) [HM09] and the ubiquitous inter-users variability which affects the reproducibility of the results. Most correspondence methods are based on the registration of the features across the various shapes. Depending on the features' nature, different correspondence strategies are more appropriate than others. Based on the recent review of Heimann and Meinzer [HM09], each correspondence category will be presented.

2.5.4.1 Approaches Overview

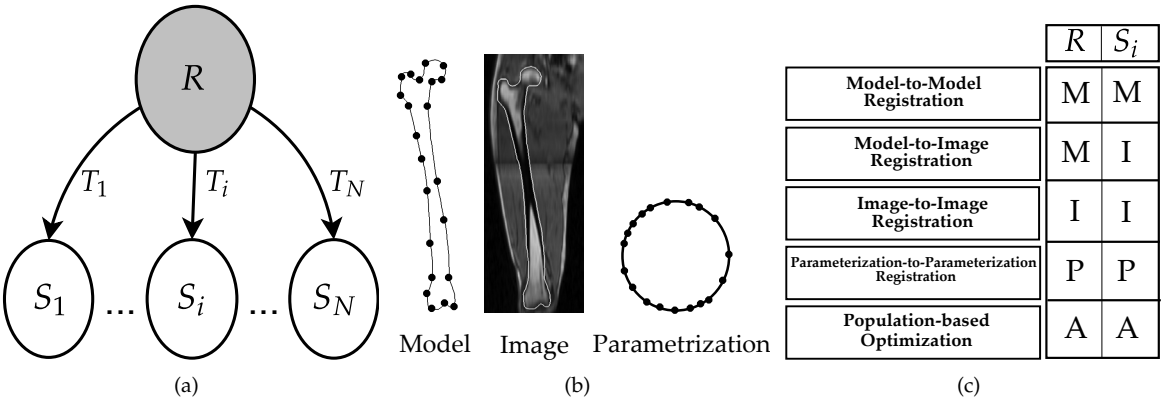


Figure 2.14: Correspondence approaches [HM09]. (a) In general correspondence scheme, a reference R is registered with N targets S_i, whose representation depends on the correspondence approach. (b) General types of reference/target entities: **M**odels, **I**mages and **P**arametrizations. In this example, a bone shape is modeled as either a contour with landmarks, or an image from which the shape was reconstructed (white overlay) or a parameterization e.g. the expression of the contour into a circle (parameterization is not real, here only illustrative). (c) Five correspondence groups, where nature for reference R and targets Si is specified. The population-based optimization can use all types of representations (A) and iterates until a certain criterion of correspondence optimality is met.

Figure 2.14(a) presents the overall principle of a correspondence method. A correspondence approach mostly relies on the registration of a reference entity R to N targets S_i extracted from the training dataset (see Sec. 2.3). The reference having features, the registration naturally propagates them to the targets, establishing the correspondence. Reference and targets can be divided into models, images and parameterizations, as exemplified in Fig. 2.14(b). A parameterization expresses the shape in an alternative way which is supposed to be more appropriate for the correspondence method. Depending on the nature of the targets, the source and the registration transform T_i, various correspondence methods are thus naturally defined (Fig. 2.14(c)). Generally, registration is made on a pair-wise manner, i.e. the reference R is registered with a target S_i independently of the registration with another target S_j, where $j \neq i$. This might introduce some bias with respect to the chosen reference. Population-based correspondence attempts to jointly and simultaneously find the correspondence across all shapes based on some optimization approaches. The optimization tries to find an "optimal" correspondence according to chosen criteria.

2.5.4.2 Model-to-Model Registration

In this category, the source is a model that needs to be registered with the other target models. Typically models adopt the discrete representation of Sec. 2.4.4 and features of interest are landmarks. Correspondence methods are naturally derived from the rich variety of point-/surface-based registration approaches [AFP00]. When transforms T_i are linear, the iterative closest point (ICP) of Besl and Jain [BJ85] is often chosen as registration approach. In [HMK+02], the skeletons of knee bones were used as models. Skeletons [Mor95] are extracted from their corresponding images and follow the principles of medial models (Sec. 2.4.1). The reference

skeleton is thus registered with the target training skeletons. Hufnagel *et al.* [HPE+07] proposed an EM formulation of the ICP algorithm and demonstrated it with the automatic correspondence of kidney and putamen data. One of the criticisms of these approaches is the bias induced by the selection of the reference shape. In [VBA+04], this problem was partially solved by running a second ICP with the mean shape as a reference, similarly to a Procrustes alignment. Brett and Taylor [BT00a] proposed a better approach based on a symmetric version of the ICP algorithm and the construction of binary trees. In a bottom-up fashion, The tree starts with the leaves corresponding to the training shapes and each upper level is built by merging each pair of training shapes. The merging uses the symmetric ICP, and the root is the result of all merged shapes. In a top-down fashion, the root is decorated with landmarks and these are propagated to the leaves.

ICP being restricted to linear transforms (rigid, affine, similarity, etc.), research was done to incorporate non rigid changes in the correspondence process to better capture shape variability. Fleute *et al.* [FLJ99] successfully registered a template mesh to a set of training shapes (represented as point clouds digitized from real femurs) by using a Multi-resolution approach based on the Octree Splines of Szeliski and Lavallée [SL96]. Gold *et al.* [GRL+98] used a combination of deterministic annealing and a proposed "softassign" algorithm to non-rigidly register pointsets of unequal sizes while computing point correspondence. In [STA98], skull shapes were non rigidly registered by using as features the crest lines of the shapes and an extension of the ICP algorithm to non-rigid transforms. Wang *et al.* [WPS00, WPS03] also exploited alternative features in a two stage algorithm: a sparse set of corresponding points were matched based on local geometrical features of the brain surface (first stage) which provided an initialization to a non rigid registration procedure (second stage) exploiting geodesic paths (shortest distance between two points over a surface). Similarly, Lorenz and Krahnstöver [LK00], also used a first initialization stage based on homologous points that were non-rigidly registered by the TPS transform (Sec. 2.3.3), followed by a projection of a reference triangular mesh on the registered shapes. Despite a relaxation scheme that corrected potential flipping triangles caused by the projection, users needed to manually define the initialization points which limits the use of the approach for large training set. More recently, Fu *et al.* [FGXL10] proposed an hybrid approach which combined ICP, to initialize the alignment of training shapes, with an active surface segmentation to refine the deformation of a reference mesh.

2.5.4.3 Model-to-Image Registration

In medical image analysis, it is very common to get training models from corresponding medical images by using a reconstruction process such as the Marching Cubes [LC87]. The idea of model-to-image correspondence is hence to register a reference model to the corresponding images of the training models. The registration approach is thus similar to a deformable model-based segmentation in which a template deforms to match a structure of interest in an image. Images can be either original (gray-valued) images, or the segmented binary images or pre-processed images to e.g. highlight edges. Kaus *et al.* [KPL+03] deformed a reference mesh by (i) estimating the similarity pose between the mesh and the image (initialization) and by (ii) locally displac-

ing the mesh vertices (mesh "reconfiguration"). Both stages were based on the minimization of energies which accounted for the image data (gradients with same directions as mesh normals) and mesh regularity (curvature preservation). This approach was later on improved in [ZT05], where the bias induced by the choice of the reference mesh was addressed. Based on the deformable models approach of Joshi *et al.* [JMG97] for brain structures, Styner *et al.* [SXESG07] registered a reference model of the hippocampus composed of 26 landmarks to the training images. In [CDNA07], SSMs were built by fitting a deformable template to CT bladder and prostate training images. Recently, Dam *et al.* [DFP08] devised an iterative process which progressively updated an SSM, refining thus the correspondence, based on medial shapes. The main idea was to use the mean of the current SSM as the reference model to be fitted to the training images, yielding new shapes to define a new SSM. As pointed out in [HM09], as long as the reference model do not exhibit self-foldings, a homeomorphic mapping between the training shapes is guaranteed.

2.5.4.4 Image-to-Image Registration

Here, the idea is to only rely on the training images rather than on extracted models. When a reference image with a corresponding segmentation is available, we dealt with atlas-based registration. The atlas can be composed of a scalars-valued (i.e. "raw" image such as CT, MRI grey-valued image) and corresponding segmentation information. The segmentation information can be a labeled image with an associated model such as a triangular mesh, and features for correspondence can hence be defined (e.g., points densely sampled on the labeled image borders, vertices of the mesh, etc.). The transform obtained by the registration of the atlas image (raw or labeled) with a training image can be thus applied on the features to propagate them on the training images. In [FRSN02, OBB+04], the multi-resolution scheme of Rueckert *et al.* [RFS01] based on FFD and B-Splines was applied to establish the correspondence between labeled images. The traditional normalized mutual information (NMI) similarity measure was in particular replaced with an alternative metric based on label consistency and the K statistics to better cope with labeled images. Instead of labeled images, Vrtovec *et al.* [VTL+04] used gray-valued training images into a hierarchical elastic registration with the NMI metric to build PDM of the spine.

An alternative to altas-based registration is to directly register the raw images and to not use segmented training images. This approach implies that it is no longer possible to specify features on a reference image with corresponding segmentation. As a result, the idea is to use instead the deformation field between images as features (so called SDMs, see Sec. 2.5.1.3). Seminal works on SDMs include the approaches of Rueckert *et al.* [RFS01, RFS03] who non-rigidly registered a reference image based on the NMI similarity measure. In [FS06], similarity was expressed by means of the Thirion's Demons algorithm [Thi98]. A possible inconvenience with non-rigid registration methods is that they might not yield invertible transforms. Christensen and Johnson [JC02, CJ01] proposed thus to simultaneously estimate the forward and reverse non-rigid transforms, modeled in their works by thin plate splines. Moreover, since no segmentation information is available, these registration approaches tend to require images in which the structures

of interest (SOI) are clearly visible (high contrast, large coverage). In fact, the registration will "blindly" register *all* parts of the image without giving a special importance to the SOI. This might result in non perfectly registered SOI and thus in non optimal correspondence. A last remark in using SDMs concerns the risk of some non-linear registration algorithms yielding deformations still embedded with a linear component. As pointed out by Zikic *et al.* [ZHG+08], this persistence of linearity decreases the quality of the resulting SDM as it will capture information not related to the shape, the shape being invariant under similarity transform. The authors proposed thus a post-processing stage to recover an optimal non-linear field based on the concept of minimal displacements.

2.5.4.5 parameterization-to-parameterization Registration

Parameterization is a bijective mapping between the shape and an appropriate base domain [HM09]. The idea of using parameterizations is to carry out the correspondence in the base domain. The choice of the parameterization often constrains the topological genus of the shapes. For instance, closed shapes topologically equivalent to a sphere (genus 0) have been intensively used in parameterization-based correspondence methods. In [KSG98], shapes are expanded into a series of spherical harmonics (SPHARM [BGK95]) and the main axis of the first order shape descriptor (an ellipse) is used to align the parameterized shapes. Similarly, Thompson and Toga [TT97] extracted brain sulcii and mapped them to a sphere before applying a warping algorithm to establish the unknown correspondence. Harmonic maps [GVSS00] were also used in [BT00b, PSS01] to create shape parameterizations and establish correspondences. Conformal maps, which ensure a bijective mapping, are equivalent to harmonic maps in case of genus 0 [GWC+03].

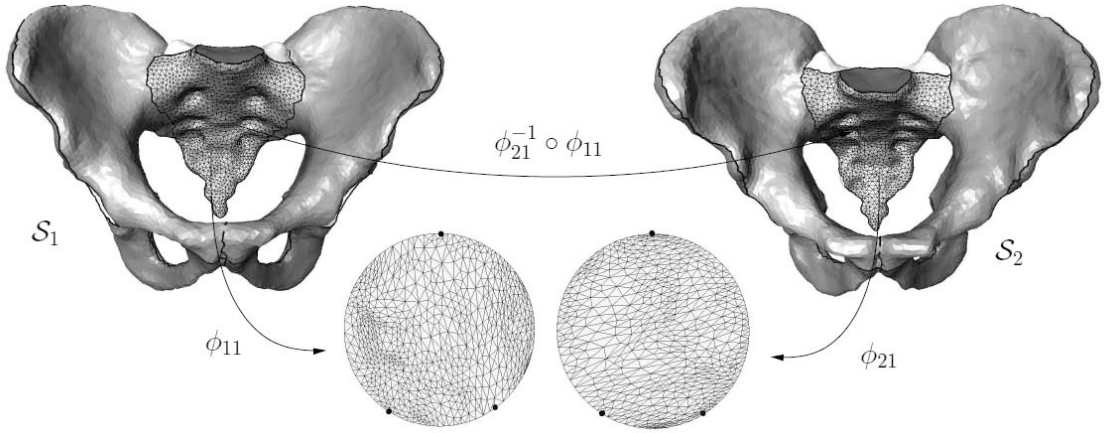


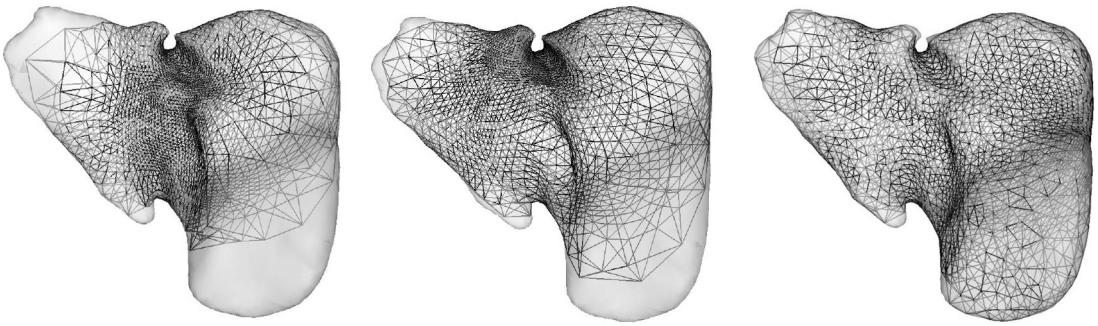
Figure 2.15: Parameterization-to-parameterization registration by patches: Shapes S_1 and S_2 are put in correspondence by (i) dividing the shapes into patches, (ii) parameterizing the patches as discs and (iii) establishing correspondence between the discs. Image from [LSHD04] used by permission.

Lamecker *et al.* [LLS02, LSHD04] got rid off topological constraints by first dividing the shapes semi-manually into patches homeomorphic to discs, and then by computing the correspondence

(Fig. 2.15). Their approach was successfully exploited to build liver [LLS02] and bone [LSHD04, SKH+08] SSMs.

2.5.4.6 Population-based Optimization

Population-based optimization is characterized by (i) the removal of pair-wise registrations and (ii) the use of optimization approaches that update the shapes representations or parameterizations based on some optimality criteria. Kotcheff and Taylor [KT98] presented the first group-wise correspondence approach in which they preferred correspondences resulting in compact SSM, i.e. with small shape variation space. The objective function to minimize was based on the logarithm of the determinant of the covariance matrix Σ_X, which quantifies compactness as $ln|\Sigma_X| = \sum \lambda_i$ = total variance. It was thus naturally called DetCov approach. Despite its good results in 2D, the approach uses an expensive genetic algorithm which has the disadvantages to be slow and to not always converge. Later on, Davies *et al.* [DTC+02] proposed a more rigorous objective function based on the MDL principle (See Appendix Sec. B.4.5), which endeavors to balance the complexity and the compactness of the statistical model. Based on the SPHARM parameterization [BGK95], the approach re-parametrizes the shapes until a (local) minimum of the MDL-based objective function is reached. This seminal work gave birth to many variants and improvements (simplified objective function [Tho03, HWWM05], consideration of curvature [TO03, KE06] and Cartesian differential invariants [WQL05], fastest convergence by gradient descent optimization [HMMW07]), and was applied in various segmentation tasks such as in bone segmentation [FBCO07, FCWO07, ZSG10]. An example of result from Heimann *et al.* 's correspondence method [Hei08, HWM07] is depicted in Fig. 2.16.



(a) (b) (c)

Figure 2.16: Shape correspondence based on MDL and conformal mapping parameterization [Hei08, HWM07]. From liver training shapes (transparent mesh) landmarks (and corresponding triangulation) are automatically created from shape parameterizations based on (a) conformal representation [GWC+03] (2D map) and (b) diffusion mapping. Conformal and diffusion maps express in a 2D representation the desired vertex/landmark density over the shape (e.g., few landmarks in areas of low curvature). Despite the improvement of diffusion to yield more uniform distribution of vertices over the surface, the shape representation remains deficient. By using the SPHARM parameterization to parameterize shapes into spheres, spheres are divided into 2 half-spheres which are then mapped to 2D maps on which the diffusion is applied. By considering all shapes contemporaneously and their diffusion maps, average maps can be created and new triangulations can be mapped back to all the training shapes establishing the correspondence. The resulting training shapes after MDL-based optimization exhibit (c) uniform landmarks distribution. Images from [Hei08] and [HWM07] used by permission.

One of the shortcomings of Davies *et al.* 's [DTC+02] approach and its variants, is the restriction to have shapes that are already topologically equivalent to a sphere. Various alternative approaches were thus proposed to address this point. In [CFS+07], a non-parametric particle system independent of the shapes topology exploited the notion of shape entropy to compute an optimal correspondence. Syrkina *et al.* [SGS07] replaced the SPHARM parameterization with a similar approach to [LLS02], in which shapes were automatically divided into patches mapped to discs, and used the simplified MDL-based objective function [Tho03, HWWM05]. Another type of approaches not imposing topological constraints are the landmarks sliding correspondence methods [WKR04, DMW+07, DJM+09]. These approaches update the position of landmarks over the shapes by locally sliding them until a cost function is minimized. The cost function is based on the thin plate bending energy and on a series of constraints imposed on the landmarks sliding. The minimization is formulated as a quadratic programming problem that can be solved efficiently [GW03].

Correspondence based on group-wise image registration is also becoming popular. In [TWT+04], Tsai *et al.* exploited the optimization of an alignment cost functional which simultaneously aligned labeled images by minimizing the difference between any pair of images in the training collection. Based on the diffeomorphisms theory [MY01], various approaches were proposed [JDJG04, CMT+04, HHL07, CTBT08]. In addition to support group-wise correspondence, these methods produce invertible non-rigid transformations. Eventually, other hybrid techniques exist. For instance, Luo and Hancock [LH03] designed an objective function based on the cross-entropy between probability distributions for alignment and assignment errors. This approach was only presented in 2D case.

2.5.4.7 Evaluation

As mentioned in the introduction, no consensus has been reached in choosing the best correspondence approach. This is mainly due to the fact that the true correspondence for biological shapes is usually unknown. As a result, various evaluation approaches were proposed in order to objectively assess the quality of a correspondence approach (and hence the corresponding statistical deformable models). The following review focuses on landmark-based statistical models (PDMs). Davies *et al.* [Dav02] proposed three measures function of the number k of retained principal modes (See Sec. dimensionality reduction in Sec. 2.5.3.1) to compare correspondence methods [MDW08]:

- *Compactness* $C(k)$: measures the amount of variance by $C(k) = \sum_{1 \leq j \leq k} \lambda_j$.
- *Generality* $G(k)$: measures the ability of the SSM to represent shapes not present in the training set by using k modes. This is usually computed by using a leave-one-out (LOO) cross-validation strategy which computes the distance of the omitted shape to the closest match of the SSM built on the remaining training shapes.
- *Specificity* $S(k)$: measures the capacity of the SSM generative model to yield only valid shapes by using k modes. This measure is estimated by generating random parameters values $\boldsymbol{Y} = \{y_1, \ldots, y_k\}$ in the generative model (2.17) according to a zero mean Uniform distribution (i.e.

$y_i \sim \mathcal{N}(0, \lambda_i)$) and by computing the distance of the generated shape to the closest training shape.

Given two SSMs built from the same training set and feature type, and with a maximum K of principal modes, SSM 1 is considered better than SSM 2 according to measure $X \in \{C, G, S\}$ if [MDW08]:

$$\begin{aligned} &\forall k \in \{1, 2, \ldots, K\}, X_1(k) \leq X_2(k), \text{ and} \\ &\exists k \in \{1, 2, \ldots, K\} / X_1(k) < X_2(k) \end{aligned} \tag{2.23}$$

By varying k, curves for Compactness, Generality and Specificity can be plotted and correspondence methods can be visually compared. In 1998, Neumann and Lorenz [NL98] already compared 2D SSMs with various representations (PDMs, wavelets, etc.) by using a reconstruction measure similar to the Davies's Specificity S. In [SRN+03], Styner *et al.* used these measures to rank SPHARM parameterization, MDL and DetCov approaches with a training set of brain structures. The conclusion was that MDL and DetCov outperformed SPHARM. Landmark sliding correspondence approaches [WKR04, DMW+07] were also compared to MDL and SPHARM in [DMW+07, DJM+09] with hippocampus data.

Davies's measures were nevertheless criticized by various authors which proposed alternatives. Karlsson and Ericsson [KE06] reported the issue of intersecting curves which yield ambiguity between benchmarked approaches (violation of Eq. (2.23)). They proposed to use a ground truth-based correspondence measure from synthetic or real datasets with manually located landmarks for which correspondence was known. Various MDL-based approaches were investigated. In [SXESG07], Image-to-Image deformable registration, MDL and SPHARM were compared by using a local shape analysis based on non-parametric permutation tests [PLNS04]. The Compactness measure was criticized because the MDL objective function optimizes shape compactness, which might create a bias in favor of MDL. Munsell *et al.* [MDW07, MDW08] proposed to generate a large collection of artificial shapes from a ground truth SSM and to apply various correspondence approaches. Corresponding PDMs were then built and compared between them to evaluate the accuracy of the shape correspondence. One of the issues of this approach is that in many real cases, a ground truth PDM is not available. The authors proposed to address this issue by modifying the framework to exploit instead a large collection of real aligned shapes. Although this modification removed the need of a ground truth, difficulties were raised in finding a sufficient number of shapes and in aligning them without prior point-correspondence.

Last but not least, the quality of SSMs and their correspondence approach can also be evaluated in a segmentation context by assessing their efficiency. In [VADF+06], van Assen *et al.* tested various correspondence/landmarks distributions to build SSMs of the heart ventricle. Subsequently, these SSMs were tested on cardiac MRI images and produced segmentation results were compared with manual ones.

2.5.5 Statistical Deformable Models

Let us present now more in details, statistical deformable models based on landmarks and appearance models. We already stressed that implicit shape representations are intentionally briefly covered in this thesis. Implicit models requiring special approaches, these will not be detailed in this chapter. In the following, we assume that the SSMs have been built from a set of shapes with aligned landmarks in correspondence. A statistical shape model based on the PCA statistics has been derived and the SSM is ready to be used for medical image analysis. The various steps for using SSMs will be now described.

2.5.5.1 Initialization

For statistical shape models derived from deformable models special attention must be paid to their initialization. The first way to initialize models is through manual intervention [KSG99, LSHD04, PFJ+03]. In case of SSM, the mean shape can be manually positioned in the image (e.g., [LSHD04]) or special cues (e.g., landmarks [HJ08]) can be specified by an operator to drive the initialization.

The major inconvenience of manually-based initialization is the loss of automation which is essential when large datasets need to be processed. Segmentation approaches with low prior knowledge can be used to automate the initialization. In [CDNA07], bladder segmentation of CT images is initialized by first performing a locally affine registration on bone structures extracted by thresholding, followed by a region growing algorithm. Similarly, Fripp *et al.* [FCWO07] and Koikkalainen *et al.* [KPL+04] carried out atlas-based affine registrations to respectively initialize SSMs of knee and heart structures in MRI. Generalized Hough transform was also applied to automatically fit the mean shape, as reported in [SKH+08, EPS+08, KLZH08, KLZH09].

Aforementioned approaches did not fully exploit the generative aspect of the SSM. In fact, based on the generative aspect of SSM, it is possible to estimate initial alignment (pose) and shape parameters by optimization techniques. For instance, an edges extraction algorithm was exploited in [YOT+09], where a cost function expressed the goodness of fit between a shape generated by an SSM and detected edges. Similarly, Ding *et al.* [DLH07] attempted to initialize a 3D SSM of femur with 2D X-rays by using extracted edges as image cues and by optimizing projection and shape parameters. In [KH98], 2D structures to be tracked in video images were initialized by formulating the optimization process as a Maximum *a posterior* (MAP) problem (See Appendix Sec. B.4.2), solved by a global optimization based on simulated annealing.

Another type of popular automatic initializations are those based on the concepts of evolutionary algorithms (EA), and notably its popular class of genetic algorithms (GA) [Hol75]. The idea is to represent the pose+shape parameters (genes) as an individual (chromosome) of a population. Generation after generation, these individuals undergo several transformations through operators of selection, mutation and crossover based on the concepts of evolution. Generally, EA algorithms are randomly initialized and are global approaches since they can avoid being stuck in local minima. Successful initialization of SSMs based on EA principles were reported

in [HT92, HMMW07]. In a similar fashion, particle filtering was used by de Bruijne and Nielsen [dBN04] to achieve vertebrae segmentation from spine radiographic images. Compared to GA, particle filtering adopts a variant of the mutation scheme and does not have a cross-over operator [HM09].

2.5.5.2 Evolution

In this Section, we will review various evolution (or search) strategies used for knowledge-based deformable models and in particular PDMs. Special emphasis will be made on the Active Shape (ASM) and Active Appearance (AAM) model searches which are often used for image segmentation.

Active Shape Model (ASM) Originally presented by Cootes *et al.* [CHTH93], ASMs are particularly popular in medical image segmentation. Based on PDM, the ASM search assumes that evolving models $\boldsymbol{X}$ are shapes expressed by the PDM and transformed by an alignment transform T:

$$\boldsymbol{X} = T(\bar{\boldsymbol{X}} + \Phi_k \boldsymbol{Y}) \tag{2.24}$$

The alignment transform type must be identical to the one used in the alignment procedure (Sec. 2.5.2). During the evolution, along the normal $\boldsymbol{n_i}$ of each mesh point $\boldsymbol{x_i}$, a similarity score E is computed at various positions $\{\boldsymbol{z_i^0}, \ldots, \boldsymbol{z_i^m}\}$ that are regularly sampled based on the chosen appearance model. Among these m values, the "optimal" position $\boldsymbol{z_i^*}$ with respect to the appearance model is expressed as the position with highest score. The next step consists in finding a shape $\boldsymbol{X^*}$ generated with Eq. (2.24) that satisfies at best the shape $\boldsymbol{Z^*} = \{\boldsymbol{z_1^*}, \ldots, \boldsymbol{z_M^*}\}$. To achieve this, Cootes *et al.* proposed an iterative process composed of the following steps:

(S1) Initialize the shape parameters vector $\boldsymbol{Y^0}$ $(i = 0)$ to 0.

(S2) Generate a shape instance $\boldsymbol{X} = \overline{\boldsymbol{X}} + \Phi.\boldsymbol{Y^i}$ (Eq. (2.17)).

(S3) Find the transform T that best fits $\boldsymbol{X}$ with $\boldsymbol{Z^*}$.

(S4) Use its inverse to project $\boldsymbol{Z^*}$ in the common frame, $\hat{\boldsymbol{Z}} = T^{-1}\boldsymbol{Z^*}$.

(S5) Update the shape parameters to match $\hat{\boldsymbol{Z}}$: $\boldsymbol{Y^{i+1}} = \Phi^T(\hat{\boldsymbol{Z}} - \overline{\boldsymbol{X}})$ (Eq. (2.15)).

(S6) Constraint the shape parameters vector $\boldsymbol{Y^{i+1}}$ (Sec. 2.5.3.1).

(S7) if $RMS(\boldsymbol{Y^{i+1}} - \boldsymbol{Y^i}) > \epsilon$ return to step (S2).

(S8) End of optimization, $\boldsymbol{X^*} = \hat{\boldsymbol{Z}}$.

RMS stands for root mean squares error. The iterative process stops when the variation of shape parameters between two successive iterations are within a significance level ϵ, which is empirically derived.

One of the criticisms of ASM is that optimal image-driven positions $\boldsymbol{z_i^*}$ are computed without any prior-knowledge of the shapes. As a result, if many model points become wrongly attracted by the image artifacts, the shape model will try to satisfy these points and will inevitably yield incorrect segmentation results. A first solution to tackle this issue was to use a multi-resolution approach on the image [CT94, HQS06]. The principle is to evolve the deformable model on

different image resolutions from coarse to fine, improving accuracy and speed of convergence. Alternatively, erroneous positions can be detected and corrected. Niu *et al.* [NSC+05] identified outliers based on a confidence metric and dynamically updated the underlying PDM model to reflect the reliable point positions. Duta and Sonka [DS97] compared the influence of a new point position on the shape parameters against an average influence to express point reliability. If it was too large, the point was flagged as erroneous and its position was estimated from its reliable neighbors. Instead of detecting and correcting erroneous points, their influence can be reduced by the use of weights, as exemplified in [RG02], in which the use of M-estimators vs. Random sampling (RANSAC [FB81]) was compared.

In order to improve the SSM's ability to delineate fine details, several authors proposed to relax the global effect of PDMs by devising *local* models of variations. By combining global and local models, the SSM preserves its specificity while improving its generality to capture fine details. A popular approach to model local variations is based on Markov Random Fields (MRF). MRFs consider the local relationship that exist between neighbors in a formulation suitable for Bayesian inference. From a computational viewpoint, these MRF-based approaches are more attractive than methods that attempt to *simultaneously* and globally optimize the positions of *all* points with respect to the appearance model. Nevertheless, compared to the traditional ASM search, these MRF-based models are more computationally expensive and are more commonly found in 2D segmentation [KH98, LWTX06, HLM07, TBA+09].

Active Appearance Model (AAM) The AAM search simultaneously exploits region and shape features by building a combined PCA model. Let us consider the following generative models for shape $\boldsymbol{X_s}$ and texture $\boldsymbol{X_t}$ (all textures being warped to the mean shape):

$$\begin{aligned} \boldsymbol{X_s} &= \overline{\boldsymbol{X_s}} + \Phi_s \boldsymbol{Y_s} \\ \boldsymbol{X_t} &= \overline{\boldsymbol{X_t}} + \Phi_t \boldsymbol{Y_t} \end{aligned} \tag{2.25}$$

The combined PCA model recovers the correlation between the shape and the texture by performing a PCA on the principal components of the two models [CET98b, Ste04]:

$$\boldsymbol{Y} = \begin{bmatrix} W_s \boldsymbol{Y_s} \\ \boldsymbol{Y_t} \end{bmatrix} \tag{2.26}$$

$$\boldsymbol{Y} = \Phi_c \boldsymbol{C} \tag{2.27}$$

where, $\boldsymbol{Y}$ is a vector combining the two principal components and which is expressed with a third PCA model (2.27). The matrix W_s weights the principal components of the shape model with respect to those of the appearance model. It can be simply approximated as $r^2 I$, where I is the identity matrix and r^2 denotes the ratio of the total intensity variation to the total shape variation [Coo01]. Any shape-appearance instance can be thus generated with the combined

PCA model:

$$\begin{aligned} \boldsymbol{X_s} &= \overline{\boldsymbol{X_s}} + \Phi_s W_s^{-1} \Phi_{c,s} \boldsymbol{C} \\ \boldsymbol{X_t} &= \overline{\boldsymbol{X_t}} + \Phi_t \Phi_{c,t} \boldsymbol{C} \\ \Phi_c &= \begin{bmatrix} \Phi_{c,s} \\ \Phi_{c,t} \end{bmatrix} \end{aligned} \tag{2.28}$$

Dimensionality reduction can also be applied by truncating Φ_s, Φ_t and Φ_c. Any instance of a AAM is eventually produced by using Eq. (2.28) and a pose transform T as in ASM's Eq. (2.24). Let us consider a vector $\boldsymbol{P}$ which contains the AAM model and pose parameters. We name $\boldsymbol{G_i}$ the texture generated by warping the texture of the current shape to the mean shape and $\boldsymbol{G_m}$ the texture generated with Eq. (2.28). The residuals are thus given by $\boldsymbol{r(P)} = \boldsymbol{G_i} - \boldsymbol{G_m}$. The main assumption of AAM is to consider a constant linear relationship between residuals and variations $\boldsymbol{dP}$ of model parameters:

$$\boldsymbol{dP} = -R\boldsymbol{r(P)} \tag{2.29}$$

where R is referred to as the prediction matrix. During the segmentation, a new deformation of the model is thus computed by adding the prediction $\boldsymbol{dP}$ to the current state $\boldsymbol{P}$ based on the residual $\boldsymbol{r(P)}$. The estimation of the prediction matrix is crucial in the AAM search, and various approaches to compute it were hence proposed. Cootes *et al.* [CET98b] initially employed a multivariate regression approach by randomly generating a large number of perturbations on the training images. Although the regression could be efficiently computed by principal component regression, a simpler approach based on a first order Taylor approximation was preferred [Coo01, Ste04]. In addition to it being faster, it was also claimed to be more reliable. In [DRL+06], canonical correlation analysis was instead used to compute the prediction matrix. The computed matrix by this approach was depicted as more accurate than the Taylor expansion, which resulted in a faster convergence of the AAM search. AAM search has been successfully used in various segmentation approaches, e.g. [OLU+03, MLvdG+00, LvdGM+02, ÜFS+03, SFK+08], and similarly to ASM, many variants were proposed. For instance, Beichel *et al.* [BBLS05] described ways to robustify the AAM search by detecting and rejecting outliers based on an unsupervised clustering of the residuals. Similarly, Gross *et al.* [GMB06] presented a robust search based on the gradient descent algorithm, coined robust normalization algorithm, which was able to efficiently track faces in presence of occlusions. More details with comparative studies on AAM can be found in [CET98a, CKN02, HM09].

Other works In addition to ASM, AAM and ASM-AAM hybrid approaches (e.g., [KMB+03, YZG05]), various alternatives that exploit shape and appearance priors are found in the literature. Many approaches define the segmentation problem as a maximum *a posteriori* (MAP) problem formulated in a Bayesian framework. Leventon *et al.* [LGF00] used a MAP estimation of the pose and shape parameters to derive the optimal model configuration from image information and shape priors built from a PCA on SDF maps. Ding *et al.* [DLH07] used the same approach but added the use of dynamic directional gradient flow to guide the model evolution. In [DGZ07], a 3D SSM of the proximal femur was first coarsely fitted on an X-rays image. Then, a bone contour with landmarks was created from the 3D projected model. A local refinement of the

landmarks position was finally carried out by a MAP estimation, which accounted for image information (gradient) and terms penalizing shape deformations that were too far from the initial contour. Seise *et al.* [SMRW06] also studied bone segmentation from X-rays and proposed a MAP estimation in which shape and appearance models were expressed as PDM and Gaussian mixtures, respectively.

Heimann *et al.* [HMMW07] explored the use of SSM and graph-cut techniques to achieve CT liver segmentation based on Li *et al.* 's method [LMW+05]. This method can be seen as a good 3D alternative of the 2D profile search optimization presented by Behiels *et al.* [BMVS02]. The principle consists in building a directed graph based on the current mesh points and its candidate point positions (i.e. positions regularly sampled along direction of intensity profiles). Graph weights are then computed based on smoothing constraints and the image similarity at each candidate positions. By performing a graph-cut, optimal point positions of the model at the next iteration are produced. This approach achieves a *global* optimization based on the appearance model, whereas traditional ASM search optimizes each landmark displacement independently from each other.

Other examples include the approach of Pizer *et al.* [PFJ+03] who first globally fitted an SSM built on medial models (m-reps). Then, local refinements were done by varying the positions of the medial nodes. In [HAGM04], medial models were also exploited in the construction of hierarchical SSM. However, the evolution was not similar since the authors devised an *organisms*-based approach. This kind of methods derives from concepts of artificial life in which entities are provided with sensory modules, behavioral routines, and decision-making strategies [MH06].

2.5.5.3 User Control

Despite the great performances of knowledge-based segmentation approaches, segmentation remains error-prone, especially in situations where significant changes with respect to the training data are present (strong pathological structures altering appearance and shapes, new imaging protocols, etc.). In that context, the use of user intervention to guide the segmentation is of paramount importance [OS01]. However, the required user-intervention should be reasonable and intuitive. User constraint points have been already successfully combined with free deformable models, such as the classic snakes [KWT88] or the live-wire approach [BM97].

The use of constraint points with statistical deformable models is however more complex as user constraints might indeed diverge from the shape expressed by the shape models. The challenge stands in giving flexibility to the shape variations to account for user interaction while keeping the specificity of the model against image artifacts. Cootes and Taylor [CT01] proposed to include user priors on point positions in a probabilistic framework to enhance the AAM search. In [VGDBLV03], an iterative procedure was used to constraint shapes to have boundaries lying on user control points. The constrained problem was solved with a dynamic programming approach. While this approach returned good results, in some cases the global aspect of PCA (see Sec. 2.5.3.3 on sparsity) lead to the displacement of all landmarks in order to satisfy user

interaction, pushing away the model from the correct boundaries. Neumann and Lorenz [NL98] minimized an error energy which considered PDM and user constraint points. In [EKW09], user interaction was achieved by the placement of control points that pulled the mesh landmarks in a Gaussian neighborhood. A regularization of the deformations was achieved based on a shape energy which prevented the evolving mesh to deviate too much from the local configuration of a reference shape. This approach ensured that all shapes created with this process shared a satisfactory point correspondence.

2.5.6 Training Dataset Size

2.5.6.1 The Problem

We have seen that knowledge-based deformable models rely on training data from which knowledge is inferred. We already stressed that a small number of observations causes many issues throughout the construction of priors: computation/eigenvalue decomposition of the (inverse) covariance matrix for PCA (Sec. 2.5.3.1) or PDF estimations (Sec. 2.5.3.5 and 2.5.3.6), poor Generality of the statistical deformable models (Sec. 2.5.4). In the segmentation context, poor priors affect the capability of the segmentation algorithm to adapt itself to new data (generality) and results in "over-constrained" behavior. As reported in [HM09], various authors [LSHD04, HMMW07] observed a strong correlation between segmentation accuracy and number of training samples.

The creation of (high-dimensional) priors is intrinsically doomed by the so-called curse of dimensionality (see Appendix Sec. B.5.3), where the required number of samples grows exponentially with the dimension. In simple words, a significant amount of data is necessary to produce good priors, especially when the observations are high-dimensional features. However, as reported in Sec. B.5.3, the number of required training samples does not necessarily always scale exponentially with respect to the number of dimensions. For instance, a biological structure is a $N-$dimensional entity when its shape is modeled by N landmarks. But, its shape might vary a little among individuals, and as a result it is very likely that a satisfactory number to express shape variations will be significantly smaller than N. Still, biological variations can be important and in practice a satisfactory number of 3D (processed) datasets is rarely available in a medical environment [HM09] due to the difficulty to gathering and processing medical acquisitions (e.g, time consuming (manual) segmentation hampered by limited clinical resources).

2.5.6.2 Enlargement Approaches

When not sufficient training data is available, adapted approaches are necessary in order to improve, at best, the quality of the statistical models. A first solution consists in using special estimation approaches. For instance, Everson and Roberts [ER00] proposed a Bayesian inference-based approach, sharing concepts with PPCA, to estimate the eigenvalues spectrum in presence of noise and limited number of observations.

Another idea is based on improving the generality aspect by artificial means. The most popular

approach is to modify the covariance matrix by generating synthetic shapes variations. In [CT95], FEMs were used to synthesize shape variations that were jointly included into the covariance matrix. Similarly, Wang and Staib [WS00] merged the original covariance matrix with a synthetic one. In [BH95], local variations were created by adding noise to the covariance matrix, which corresponded to the addition of isotropic Gaussian noise to the training shapes.

Another approach is to directly create new artificial shapes from the existent training shapes. In [LAL+05, TKLL06] various approaches of enlargement were tested to build SSMs of the heart. The best approach was obtained by non-rigidly perturbing the shape landmarks in randomly located neighborhoods represented by spheres. Alternatively, some authors proposed to reduce the dimension of the training shapes by using hierarchical representations. Davatzikos *et al.* [DXD03] used for example a wavelet transform to decompose the shapes into sub-bands on which principal modes were computed. The same idea was exploited in [ZT05], were shapes were divided into small regions before a PCA was applied on each of them. To ensure shape continuity between the regions, training shapes were projected as curves on a hyperspace in which continuity was established by using a curve alignment scheme.

2.6 Discussion

In this state-of-the art, we have seen how prior-knowledge can significantly improve segmentation approaches, notably deformable models. In fact, while deformable models offer some advantages (sub-pixel accuracy, occlusion robustness, complex shape modeling, etc.) they suffer from a poor control of the shape variations (e.g., ad-hoc smoothing controlled by parameters). The consequence is their incapacity to segment highly curved details and an increased sensitivity to poor initialization. However, the use of prior-knowledge in shape and appearance helps in correcting these limitations.

The counterpart is of course the effort required to build efficient statistical deformable models. Correspondence must be established, models have to be aligned and priors need to be computed. All these stages have their weaknesses and requirements. In this thesis, we propose new solutions to devise robust and efficient knowledge-based deformable models by addressing the key stages which are model construction, initialization, evolution and control.

We purposely did not detail the construction and usage of knowledge-based implicit deformable models. The principal reason is that special approaches are necessary. Indeed, one of the main aspects to care with implicit models is that PCA is theoretically not applicable with features not belonging to an Euclidean vector space, which is a linear space. For example, Signed Euclidean distance function maps (SDF maps), previously presented for levelset-based implicit deformable models (Sec. 2.4.3), do not form a linear space. By creating linear combinations of SDF maps, chances are that invalid shapes will be implicitly defined. Nevertheless, Leventon *et al.* [LGF00, Lev00] reported successful results with implicit SSM based on SDF maps and inspired many extensions [RP02, TYW+03]. Later on, Pohl *et al.* [PFS+06] proposed a solution to this problem based on the embedding of the SDF maps into the Logarithm Odds space. Similarly, implicit

representations based on medial models (Sec. 2.4.1) require special care as the medial nodes components do not belong to an Euclidean vector space but to the Lie group [FLJ03]. More information on shape priors for implicit models are found in [CTT+02, CRD07].

Chapter 3

Human Musculoskeletal Modeling

3.1 Introduction

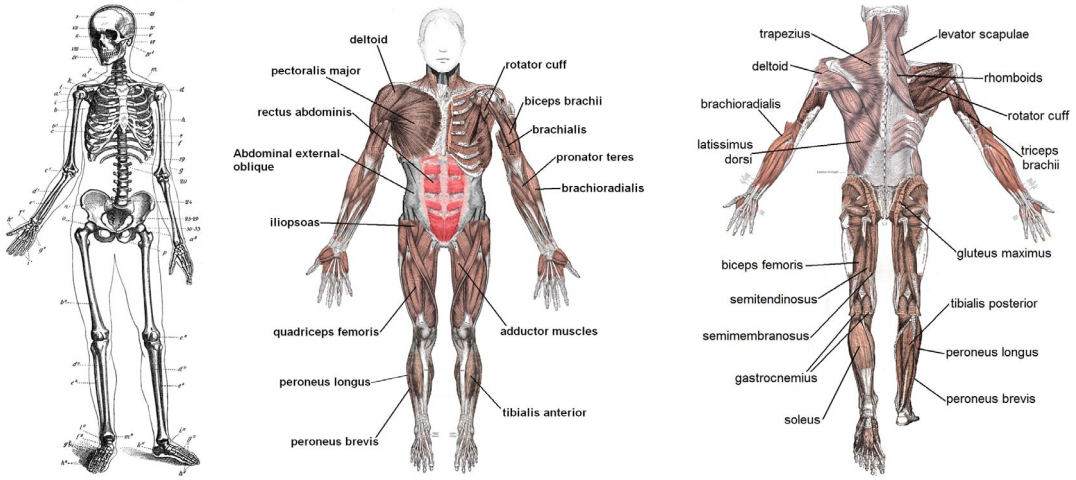
In this Chapter, we will present the human musculoskeletal system. In Sec. 3.2, an anatomical description will is by presenting the main type of structures of the human musculoskeletal system. Then, Sec. 3.3 will report most common acquisition modalities to image the musculoskeletal structures. These information are provided to better understand the challenges faced by the different computer-based studies involved in modeling the musculoskeletal system. These studies are detailed in Sec. 3.4, where special attention is given to image-based modeling since it is closely related to the topic of segmentation previously detailed in Chap. 2.

3.2 Anatomy

Medical information have been gathered from a collection of sources such as Wikipedia (`http://www.wikipedia.org`), Henri Gray's anatomy book (old [Gra18] and new [S+05] editions), as well as other relevant literature e.g. [Cae10].

3.2.1 Overview

The human musculoskeletal system is mainly composed of the *skeleton* and *muscular* systems, and it provides locomotion, stability and protection. The skeleton (Fig. 3.1(a)) is composed of bones on which skeletal muscles (Fig. 3.1(b) and 3.1(c)) are attached via tendons. Bony segments are connected with each other at joints, which provide different degrees of freedom depending on their structure. Joints are stabilized either by fibrous tissue at the articular interface or by ligaments. Ligaments are flexible and passive structures that play a major role in the joint stability. Indeed, ligaments transfer some loads and quickly increase their rigidity in presence of abrupt motion changes, thus reducing the risk of articular injuries. Cartilages are another essential structures of the joints, which are also vital in the early development of many bones. Cartilages prevent bone-to-bone contacts and act as shock absorbers by allowing fluid to be squeezed out, which ensures good lubrication of the joint. Tendons attach muscles to bones by strong insertion into bone tissue, and are flexible like ligaments. Their main mechanical characteristic is to be inextensible to efficiently transmit muscle pulling forces to the bones. Cartilages, ligaments, tendons and bones belong to the category of connective tissue. Conversely, muscles are composed of very particular cells that form the muscular tissue. Human muscles are mostly composed of skeletal muscles which react to conscious control to produce locomotion and enforce posture. By changing the size of their muscle cells, muscles can contract (i.e. muscles pull) and are typically paired together in sets that work in conjunction with one another. In the following, we describe more in details each of these structures.



(a) Skeleton system (b) Muscular system: anterior view (c) Muscular system: posterior view

Figure 3.1: The human musculoskeletal system is mainly composed of the skeleton and muscular systems. Images hosted at Wikipedia: (a) Hubert Lüdwig, 1891; (b) and (c) collage of various Gray's Anatomy [Gra18] muscle pictures by Mikael Häggström.

3.2.2 Cartilage

During the early fetal life, cartilage tissue mostly composes the human skeleton and is progressively replaced by bone tissue. At young and adult ages, cartilage is still present at joint interface where potential contact between bones exists (hyaline cartilage) or as a supportive tissue (elastic cartilage, e.g. in ears, or fibrocartilage, e.g. as the hip's labrum). Cartilage plays a major role in synovial joints. Synovial joints are a particular type of joints that link articulated bones and are mainly classified based on their shape (plane, hinge, pivot, bicondylar, spheroidal, etc.). Cartilage cells called chondrocytes produce Type II collagen which represents up to 50% of the dry weight of cartilage. Collagen mostly confers to the tissue flexibility and resistance. As a result, cartilage is simultaneously stiff and inflexible, which allows the transmission and repartition of loads, acting as a shock absorber in the articulation. Coupled with synovial fluid, the cartilage moreover decreases friction within the joint. Synovial fluid is secreted by the inner layer of the articular capsule and it also contributes to the maintenance of the living cells in articular cartilage. In fact, cartilage tissue is avascular and chondrocytes need to be supplied by a diffusion process which is helped by the pumping of synovial fluid into the cartilage under compression. Being a structure constantly stressed and without blood vessels, cartilage exhibits hence a slow regeneration process.

Cartilage completely covers the bony areas in contact to prevent painful and mechanically inefficient bone-to-bone contacts. Other cartilaginous structures might be also present to improve the joint efficiency. This is the case e.g. of the fibrocartilaginous labrum rim that is attached to the margin of acetabular cartilage of the hip (See Sec. 3.2.6 and Fig. 3.3(c)). Its role is to deepen the acetabular cup and to reduce the friction between articular cartilages. As a consequence, mechanical fatigue is decreased and the joint becomes more stable and more congruent [FBGI00],

where the joint congruency quantifies the degree of lateral joint subluxation. Articular cartilage thickness is related to the joint congruence, as thiner cartilages are usually found in congruent joints (e.g. ankle) [SS99]. Cartilage thickness is also altered by some pathologies and especially osteoarthritis (OA) which progressively degenerates cartilage tissue until it is completely broken down. As such, cartilage thickness evolution is often monitored with computer-assisted techniques [LEHE97, KGG+03] to detect early signs or progression of articular diseases.

3.2.3 Bone

Bones are usually composed of two parts, compact and trabecular (also known as cancellous or spongy) bone tissues. Compact bone is usually present in the external shell of mature bones (cortical bone) and it is much more denser than cancellous bone. Indeed, cancellous bone is usually found at the bone organ interior and presents many cavities containing the marrow. Similarly to cartilage, bone is also made of collagen which provides some elasticity to the tissue. This special architecture confers to the bone an excellent resistance with a low weight (the skeleton represents around 14% of the total body weight of an adult). In fact, a first role of bone is obviously to protect the different internal soft organs and to support the weight of the different structures. Moreover, bones contribute to locomotion by transfering muscle forces and moments via the system of joints and tendons. Bone can also transmit vibrations to support the hearing system.

Based on the shape and proportions of compact to cancellous bone, there are different categories of bones:

- *Long* bones such as the femur (Fig. 3.2) and most bones of the limbs. Their main characteristic is to have a shaft named the diaphysis that is much greater in length than width.
- *Short* bones such as those of the wrist or ankle, roughly shaped as cubes. *Sesamoid* bones are short bones that present the characteristic to be embedded in tendons, like the knee patella. They influence the distance or angle between the tendon and its joint thus possibly increasing the magnitude of the joint moment, which enhances the muscle leverage.
- *Flat* bones such as the skull or the hip bone (Fig. 3.2) are usually thin and curved.
- *Irregular* bones are the bones which do not fit into the above categories, such as the spine and some parts of the pelvis (e.g. coccyx).

In addition to their mechanical functions, bones provide 99% of the metabolic calcium and phosphate. Furthermore, bone marrow plays a primordial role in the metabolic system. The marrow is composed of red and yellow types. Red marrow is reponsible for the production of blood cells (e.g., white and red cells) while the yellow marrow essentially stocks fat cells. In case of high demand for blood cells due to e.g. a heavy blood loss, yellow marrow can be converted to red marrow.

Bone growth follows either the intramembranous ossification or endochondral ossification processes. Intramembranous ossification is responsible for the formation of flat bones, while endochondral ossification produces long bones. The process of endochondral ossification, which transforms cartilage into bone and produces the different tissues such as the marrow, explains

for example why at bone extremities, the compact bone tissue is much more thinner with respect to the one observed in the diaphysis of long bones. Bone located below the articular cartilage, named subchondral bone, tends thus to present a thin cortice. It is important to be aware of these bone density and thickness variations resulting from the ossification processes. In fact, they explain some strong bone intensity inhomogeneities observed with different imaging techniques (See Sec. 3.3). This knowledge can thus help us in devising more adequate segmentation techniques.

3.2.4 Ligament

Most ligaments are articular ligaments, i.e. they attach the bones together. Articular ligaments can be either extra-capsular or intra-capsular, i.e. located outside or inside the synovial capsule, respectively. Extra-capsular ligaments stabilize the joints, such as the iliofemoral ligament of the hip (Fig. 3.3(d)) which is the strongest ligament in the human body. Intra-capsular ligaments also improve joint stability and also contribute to having a larger range of motion (e.g., the cruciate ligaments of the knee). Joint stability is essential as instabilities can eventually lead to wear of the cartilage and thus to OA.

Ligaments mainly provide mechanical support to the joint. Like tendons, ligaments transmit longitudinal loads (they are almost only tensile) and are wrapped in a layer of loose connective tissue to allow sliding with little friction. However compared to tendons, ligaments are loaded heavily only occasionally, typically to avoid joint damage (dislocation, etc.). Ligaments are thus elastic to allow progressive deformations but only for a certain amount to avoid injuries. Additionally, they are characterized by hysteresis, or in other words, they are viscoelastic materials. Furthermore, ligaments have high water content and are thus almost incompressible under pressure.

Various factors affect the mechanical behavior of ligaments. Ligaments are for instance altered with age. Developing ligaments are more compliant and less resistant than the ones of adults [WOG+86]. Further, the older the person gets, the less the tissue deforms before rupture and the more its stiffness decreases [WHA+91]. Gender may also influence the mechanical properties of the ligaments, where significant difference is only observed for the structural properties (force, displacement) [TMM78] rather than the material properties (stress, strain, etc.) [BLS+07].

3.2.5 Muscle and Tendon

There are three types of muscle in the human body:

- *smooth* muscles, which are present in the walls of hollow organs, vessels and respiratory pathways. Their contraction is unconscious and they are thus referred to as involuntary muscles. In some cases, their fasciculi are associated with those of skeletal muscles (e.g., urinary bladder).
- *skeletal* muscles, which are attached to bones via tendons. Skeletal muscles are made of bundles of parallel fibers that can adopt different configurations, yielding a large variety of skeletal muscle types based on their form and fascicular architecture (quadrilateral, strap, tricipital,

triangular, bipennate, etc.).

– *cardiac* muscles, which are uniquely found, as their name indicates, in the wall of the heart. This type of muscle is peculiar because it presents a tissue architecture similar to skeletal muscle but differs from being an involuntary muscle. Furthermore, fiber bundles are less regular than in skeletal muscles.

Muscle cells contain a contractile protein called myofibrils which contracts under nervous stimuli. Muscle cells possess a cytoskeleton whose elements can thus lengthen or shorten. In case of skeletal muscles, generated forces are transmitted to the surrounding connective tissue (myofascia) to eventually yield movement via the bone-tendon coupling. Tendons are attached to the periosteum of bones (i.e. the outter membrane of bone surface) and are made up of collagen fibers which give them strength and elasticity. Tendons present various shapes and sizes depending on their role and location. Bone-tendon attachment locations are correlated to bone shape to ensure mechanical efficiency and present a little inter-individual variability [KvdH04].

There are approximately 639 skeletal muscles (varying number because medical sources group muscles differently) in the human body and they account for the largest fraction of the human muscles. Several muscles participate in the production of movements although some movements might be very simple. This apparent muscle redundancy highlights the puzzling capability of the musculoskeletal system to deliver high performance with a complex muscles network. The usage of different muscles might be explained by the need to achieve a balanced reaction at joints so that stability depends less on passive structures such as ligaments. Similarly, the limited range of operation of some muscles may require the combination of several muscles to deliver appropriate forces over the full range of movements. In this context, advanced numerical simulations [TAD03, SP07a] are a great tool to investigate different models of the neuromuscular system.

3.2.6 Hip Joint

Hip joint structures are covered in depth since many of our modeling approaches have been specifically designed for this articulation. The hip joint is mainly composed of:

– Two bones: the *femur* and the *hip bone* (Fig. 3.2). The hip bone is divided into three bony areas named ilium, ischium, and pubis (top of Fig. 3.2(b)). The femoral head (bottom of Fig. 3.2(b)) is inserted into the acetabulum to form a "ball-and-socket" synovial joint.

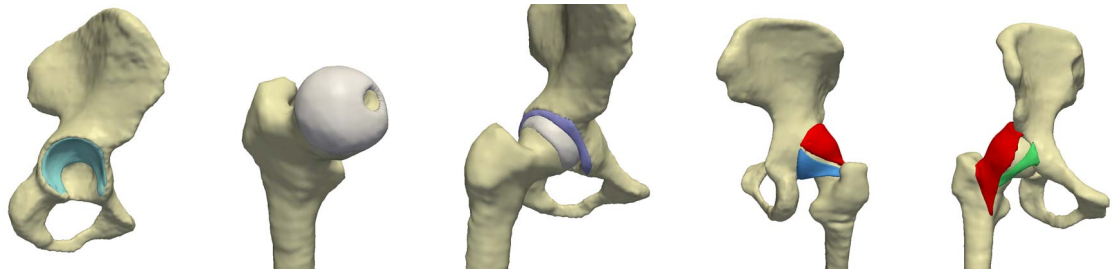
– Two articular cartilages and one fibrocartilage (Figs. 3.3(a)-(c)): the *femoral* and *acetabular cartilages*, and the *labrum*. Femoral and acetabular cartilages cover the femoral head and the acetabulum, respectively. The labrum is a supportive structure that improves the joint efficiency and stability.

– Four ligaments (Figs. 3.3(d) and (e)): the *iliofemoral*, *ischiofemoral*, *pubofemoral* and *round* ligaments which attach the femur to the ilium, ischium, pubis and acetabular notch (top of Fig. 3.2(b)), respectively. The first three ligaments are strong extra-capsular ligaments that stabilize the hip joint. The round ligament (formally referred to as ligamemtum teres femoris) is an intra-capsular ligament that connects the hip bone's acetabular notch and the femur's fovea

capitis (bottom of Fig. 3.2(b)) to mainly prevent serious subluxation from occurring.
The hip joint is similar to the shoulder joint for the lower limbs and allow multiplanar movements. The hip however has to bear around 2/3 of the body weight and requires thus a stronger stability ensured with a deeper joint articulation and an efficient network of muscles and ligaments.

(a) (b)

Figure 3.2: Bones of the hip joint. (a) from left to right: anterior, lateral, posterior and medial views of the right hip joint bones. (b) top: hip bone divided into three bony regions, pubis (green), ischium (blue) and ilium (red). Acetabular notch (AN) is located below the acetabular (region delineated with white dots). bottom: proximal femur with greater (green) and lesser (blue) trochanters, and femoral head (red) with fovea capitis (FC).



(a) Acetabular cartilage | (b) Femoral cartilage | (c) Labrum | (d) Ischiofemoral and iliofemoral ligaments | (e) Iliofemoral and pubofemoral ligaments

Figure 3.3: Cartilages and ligaments of the hip joint. (a) acetabular cartilage, (b) femoral cartilage, (c) labrum, (d)-(e) ischiofemoral (blue), iliofemoral (red) and pubofemoral (green) ligaments.

Various musculoskeletal pathologies and associated disorders affect the hip. Among them, osteoarthritis (OA) is a very common disease, which has several known causes. Abnormal joint morphology is one of the most common causes, which includes hip dysplasia [RSGP06], hip dislocation and femoroacetabular impingements (FAI) [GPB+03, PMD+06]. Others factors such

as repetitive micro-trauma and some characteristic movements (e.g., torsion, twisting, hyper-abduction) have been also associated with the development of hip OA [LS06]. The relationship between OA and motion is dependent on various factors, such as the subject characteristics (e.g., gender, age, weight) and performed activities (frequency, intensity) [LB00]. In some cases, moderate activities might not accelerate the development of OA and may even be beneficial [GEG+04].

Conversely, activities with extreme motions have been often associated with an increasing risk for the development of OA. As an example, a study from the University of Bern [KSWM09] corroborates the idea that athletes participating in intense sports possess a higher rate of hip OA and an earlier onset of this disease compared to the general population. Using karate athletes as subjects, a strong correlation between the early OA and FAI was established. The study between OA and extreme motion is also subject to intense research at MIRALab [ACS+09, CAVMT09] and part of this thesis particularly aims at supporting this research direction.

3.3 Imaging Modalities

The most common imaging modalities for musculoskeletal imaging and modeling are X-rays imagery, Computed Tomography (CT), Magnetic Resonance Imaging (MRI) and ultrasound (US). We report in this Section sthe primary characteristics, as well as strengths and limitations of each modality. It is especially important to understand the capabilities and constraints of these modalities with respect to our goal of modeling the musculoskeletal structures.

3.3.1 X-rays, Fluoroscopy and Computed Tomography

X-rays imagery quantifies the absorption of X-rays from the various tissues under X-rays exposure. It is thus an invasive modality as ionizing radiation is observed and it is known to yield to cancers in case of over- or repetitive exposures. Still, it remains a widely utilized modality [BMVS02, DGZ07, ZSG10] for its ability to be used intra-operatively and its relative simplicity and low-cost. Fluoroscopy is a dynamic modality based on X-radiography which produces X-rays movies. Both X-rays and fluoroscopy images suffer from their inherent 2D projections which seriously limits the spatial interpretation of structures as well as the accurate tracking of motion in dynamic images. Furthermore, these modalities are not always the most appropriate when it comes to gauging small bone changes due to musculoskeletal disorders. In these cases, MRI (Sec. 3.3.2) or ultrasound imagery (Sec. 3.3.3)[BKS+99] are better alternatives.

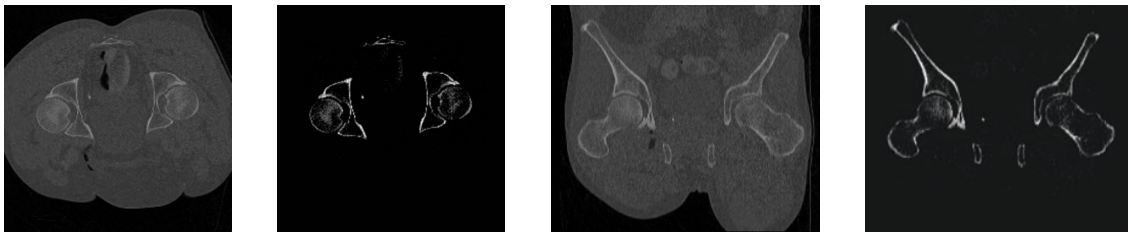
The 3D static counterpart of X-rays images is Computed Tomography (CT), which is often preferred for the reconstruction of skeletal structures [KEK03, KK08], depending on the structures to be imaged (e.g., X-rays remains the common choice for trauma imaging of the hand and wrist [Wil05]). CT can deliver images with high resolution (below mm) and offers imaging in the transverse plane and multiplanar reconstructions. Moreover, in CT, cortical bone can be clearly differentiated from trabecular tissue which yields a better detection and characterization of complex fractures [Fel00]. CT intensity of different tissues can be predicted with a reasonable

accuracy by referring to their typical values in the Hounsfield scale. The common Hounsfield Units (HU) of different substances and tissues are reported in Table 3.1.

Substance/Tissue	**HU**
Air	≈ 1000
Fat	≈ 120
Water	0
Muscle	40
Bone	> 400

Table 3.1: Typical Hounsfield units of various substances and tissues (source: http://www.wikipedia.org)

Bone HU are typically very high compared to those of other structures. This characteristic eases the design of bone segmentation approaches based on techniques such as thresholding, region growing and deformable models (See Sec. 3.4.2). Figure 3.4 illustrates how a simple thresholding gives a good delineation of the bone contours. Finally, different research directions jointly combine some X-ray variants to derive new acquisition techniques, like the fusion of CT with fluoroscopy to assess in vivo motion [YWT+07].



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

Figure 3.4: Computed Tomography (CT). (a) transversal CT slice of the hip articulations, which is easily thresholded as shown in (b). (c)-(d) same but in the coronal plane.

3.3.2 Magnetic Resonance Imaging

In MRI, hydrogen atoms, which are essentially found in water molecules within the tissues, are aligned under the influence of a powerful external magnetic field. Radio frequency (RF) fields are then used to perturb the alignment of the atoms which emit a resonance signal detected by the scanner. The collection and processing of these signals yield the creation of volumetric images. Such magnetic-based imaging does not create any harmful ionizing radiation [SSG+03, SC04], which makes MRI adapted to conduct studies on healthy volunteers.

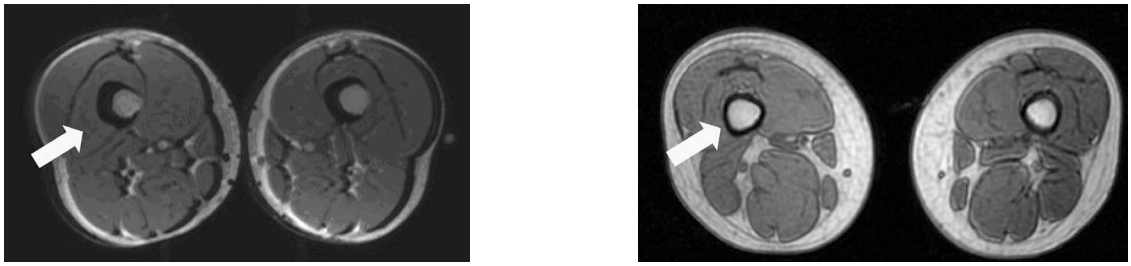
The diversity of factors accounted by the MRI signal (e.g., proton density, T1 and T2 relaxation times, fluid flow, etc.) grants MRI with unique versatility. For instance, a fluid flow/ tissue relationship is exploited in diffusion tensor MRI (dtMRI) to extract muscle fiber direction, providing an essential biomechanical parameter [HSV+05]. Similarly, dtMRI can be used to image

bone marrow disorders [Stä05]. Unlike CT and X-rays imagery, MRI is furthermore able to simultaneously capture and depict bony and soft structures (e.g., cartilages and ligaments). This characteristic of MRI has played an important role in the study of numerous human joints, such as the hip [KEK03, ZSN+04, PKB06, GMT10], knee [KBG+98, KHC05, LFG+98, KEK03], shoulder [BR05, NLBA07], elbow [CS05] and wrist [SVG02, STCK03]. Figure 3.5 illustrates how different high resolution images of the articulations can be obtained by varying the magnetic field strength (e.g., 1.5 T or 3 T) or the MRI acquisition sequences.

An increase of the magnetic field strength significantly improves the signal-to-noise ratio and hence the image quality. In fact, the increase in the signal-to-noise ratio allows to image smaller field of views with equivalent quality in the same acquisition time [Sch05]. As a consequence, MRI musculoskeletal imaging of joints clearly benefits from stronger magnetic fields, as depicted in Figs. 3.5(b) and 3.5(c). These figures also show that by varying the acquisition protocol, the imaged intensities can drastically change between acquisitions as e.g. the bone tissue appears darker than the muscle tissue with the MERGE protocol 3.5(b) whereas the opposite happens with the CUBE sequence 3.5(b). This great flexibility of MRI compared to CT is very useful for clinical diagnosis but seriously complicates the task of image segmentation for structure delineation. Indeed, some kind of prior knowledge on the appearance is necessary to be able to delineate structures with such inhomogeneous intensities (See Fig. 1.1 in Chap. 1). Further, the presence of ubiquitous imagery artifacts (e.g. those resulting from the chemical shift effect as shown in Fig. 3.6(a)) demand fine tuning of acquisition sequences and the design of more robust modeling approaches.

(a) (b) (c)

Figure 3.5: MRI acquisitions of human articulation. (a) slice of an MRI volume (slice resolution: $0.39 \times 0.39mm$) of the ankle acquired with a 1.5T High resolution FLAIR sagittal sequence. Knee acquired at high resolution on a 3T scanner with (b) MERGE (multiple echo recombined gradient echo) and (c) FSE-Cube sequences. These different sequences acquired with a more powerful magnet allow better imaging of the thin and soft structures, such as knee fibrocartilaginous menisci and ligaments. Images courtesy of Bailiang Chen and Prof. Andrew Todd-Pokropek from UCL, and Prof. Wady Gedroyc, Mr. Warren Casperz and Ms. Lauren Sundblom from St. Mary's Hospital.



(a) (b)

Figure 3.6: Example of MRI acquisition artifacts. In (a) artifacts due to the chemical shift effect are visible: bone (white arrow) does not present uniform cortical and trabecular bone layers, some kind of "directional" artifacts are indeed produced. As depicted in (b), this problem is clearly resolved with a more adapted acquisition protocols. Images courtesy of Bailiang Chen and Prof. Andrew Todd-Pokropek from UCL, and Prof. Wady Gedroyc, Mr. Warren Casperz and Ms. Lauren Sundblom from St. Mary's Hospital.

3.3.3 Ultrasound

Ultrasound (US) emits sound waves into the medium and computes the amplitude and the time required for the waves to bounce back. Based on an estimation of the sound speed into the corresponding medium, distance from the emitter can be estimated. By casting multiple beams, a 2D image of the internal organs can be reconstructed. The acquisition process is extremely fast, thus real-time and dynamic images of the human body can be produced. Similarly, by rotating the emitter in 3D space or by using arrays of transducers, (dynamic) 3D US is nowadays possible and commonly used in fetal echography. US can also profit from a Doppler mode to track flow of fluids in time by using a color scale which indicates the fluid direction and speed.

US is low-cost, non-invasive and portable which makes it very attractive as an imaging device. Unfortunately, US suffers from low image quality due to speckle noise (scattering elements with dimension below the wave length), geometric distortion due to incorrect assumptions on wave velocity and most importantly shadowing effects due to the incapacity of sound waves to traverse some media. This is particularly the case of bone tissue whose interior cannot be imaged with US and which occludes surrounding structures. US is also handicapped by the quick decay of the emitted waves which prevents the clear imaging of deep structures and results in small image field of view. Still, US is more and more used in diagnostic of musculoskeletal soft tissue to detect e.g. muscle, tendon and nerves disorders [BM05]. Figure 3.7 shows an example of a 3D US of the leg. Image quality is undoubtedly poor compared to MRI or CT, and only a trained eye can detect the interfaces of some structures such as the tibia or the calf muscle.

3.3.4 Image resolution and field of view

In the clinical environment, acquisition time is often a critical factor and many acquisition protocols strive to reduce it, often at the cost of (i) reduced field of view (FOV) [MI05] or (ii) low image resolution. The FOV is typically kept small in arthrography [EW04] in order to obtain the best resolution and signal-to-noise ratio around the joint (e.g., Fig 3.5). A low image resolution is often necessary to cover large FOV in order to fully image the bones of the joints, as reported

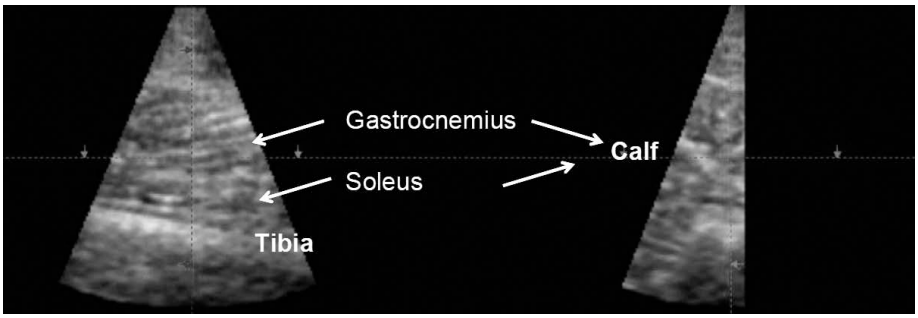


Figure 3.7: 3D ultrasound imaging of the leg. This picture shows different planes of the 3D US volume. The tibia and calf muscle (composed of the gastrocnemius and soleus muscles) can only be distinguished by a trained operator due to the poor image quality and limited structures visibility. Image courtesy of Dr. Jing Deng from UCL.

in some biomechanical studies [MKP+05]. As a consequence, a large slice thickness is produced, which yields ubiquitous partial volume effect (PVE).

A solution to simultaneously acquire a large FOV, while having fine details for small structures such as the cartilages is to combine together different acquisitions with different FOVs and image resolutions. This allows an overall reduction of the acquisition time but also introduces other issues related to the registration of the different acquisitions. Modern scanners can record the table localization information which can be used for the registration. However, other problems remain such as the movement of the patient between acquisitions and the many imagery artifacts which hinder the registration.

While most acquisition protocols are designed for best clinical practice, they are not necessarily well suited for purposes of segmentation. Even with the ability to register multiple images, segmentation is a complex task. In fact, we will later on demonstrate that image resolution and FOV have effects in all the stages of knowledge-based deformable models: creation of priors, initialization and evolution of deformable models. This will demand the design of new solutions to address these difficulties.

3.4 Related Work on Modeling

In this Section, related work on musculoskeletal modeling will be presented. Instead of classifying the existing methods based on the modeled structures (e.g., muscles, bones, etc.), we present them based on their modeling approaches to show the strong ties with the previous Chapter. We mainly distinguish two categories:

– **Computer graphics (CG)-based methods**: approaches that use a collection of graphic tools to build musculoskeletal models.
– **Image-based methods**: approaches that rely on medical images to carry out the modeling.

The main distinction between both categories is the support of *subject-specific* modeling. In the first category, anatomical information (anatomical books, sketches, etc.) is exploited to ensure that a minimum of fidelity is achieved in building the musculoskeletal models. However, the produced models are somehow "generic" anatomical models. That means that they express general

anatomical traits shared by all individuals. These models are then used in various ways, such as teaching of medical student or artist, 3d models for computer graphics (games, animation), etc.

Conversely, image-based modeling has the constraint to stick as close as possible to the anatomy of the subject in the images. Image-based modeling has been described as the most promising way to yield subject-specific musculoskeletal models [BAGD07]. It is nowadays the only way to capture human anatomy on living subjects in a less invasive way. Image-based modeling is intimately linked with medical image segmentation presented in previous Chap. 2. In fact, the quality of the models strongly depends on the segmentation results. Eventually, in order to fulfill the simulation (e.g., mesh element quality and quantity) or modeling (e.g., model smoothness) requirements, additional post-processing steps are usually required.

3.4.1 CG-based Modeling

The increasing demand for better human realism strongly contributed to the development of CG-based anatomical modeling. Fostered by the revolution in CG hardware, anatomical modeling and simulation was addressed by various research groups. Scheepers *et al.* [SPCM97] modeled human skeleton and musculature based on simple geometric entities such as ellipsoids. Anatomical concepts were considered, such as the bone-muscle attachments via tendons. Furthermore, simple muscle simulation (e.g., isometric muscle contraction, volume preservation) was also presented. Nedel and Thalmann [NT98] proposed the use of *action lines* to model the muscles. Action lines describe muscles as as set of lines which account for the muscle fiber directions and bone-muscle attachment areas (with "attachment" and "via" points). Action lines are still nowadays found in advanced neuromuscular simulation [TAD03, DAA+07, SP07a] because they significantly simplify the simulation which remains very time consuming. In [AT01], a similar direction was chosen by modeling the upper body musculature with a two-layers abstraction. The first level described muscles as action lines to ease muscle interactive modeling and simulation, while the second level expressed the surface of the muscles. Shao and Ng-Thow-Hing [SNTH03] proposed to better model human articulations to assist CG designers. By limiting the motion of joints based on anatomical considerations, designers have better capabilities to produce realistic movements.

(a) (b) (c)

Figure 3.8: CG-based musculoskeletal modeling. (a) lower limb volumetric templates [FMT+04]; (b), (c) anatomical modeling and simulation of the upper limbs [TSB+05]. Images used by permission.

Some approaches made use of medical images to design muscle templates with better realism. For instance, the human visible dataset (VHD) [Ack95] was used in many occasions to model musculoskeletal structures of the upper [GKB+96, MT99, TSB+05] or lower [NTHF02, FMT+04] limbs, as depicted in Fig. 3.8. Ng-Thow-Hing and Fiume [NTHF02] described the fitting of B-spline solid models based on extracted contours in each slice of the segmented VHD. The Free-form deformation approach was used in [FMT+04, OMSA09] to deform volumetric templates (Fig. 3.8(a)) built from the VHD to match the muscle and bone geometries of another subject. In [BD05], generic templates were additionally decorated with biomechanical properties such as the fiber directions. The direct construction of volumetric models is of interest for realistic simulations such as finite element models (FEM). By assigning mechanical and neuromuscular properties to these models, complex modeling and simulations can be indeed produced. In particular, Teran *et al.* [TSB+05] (Fig. 3.8(b) and 3.8(c)) proposed an advanced simulation based on Finite Volume Elements which was able to robustly handle extreme conditions such as tetrahedra collapse.

In summary, CG-based modeling is providing powerful *interactive* tools to model and simulate the musculoskeletal system. Although some approaches achieve complex biomechanical simulations (e.g. [TSB+05]) based on medical images, these approaches are not really appropriate for subject specific modeling. They tend to require significant user intervention, which make them very time consuming and non applicable on large subject database. Moreover some approaches rely on over-simplifications, which are not satisfactory for medical or clinical use. Nevertheless, some of these approaches could be combined with the following image-based approaches to e.g. simulate realistic models for CG purposes.

3.4.2 Image-based Modeling

In the following, image-based approaches are reported based on their segmentation categories.

3.4.2.1 Direct Segmentation

As discussed in the previous Chapter, direct segmentation approaches are usually easy to understand and to use, but in order to be successful, specific conditions and requirements are necessary. These include operator intervention and correction, as well as adapted imaging protocols. These factors often limit, in practice, the efficiency of direct segmentation in real conditions. Still, several works can be found in the musculoskeletal modeling literature which made use of direct segmentation, often as an initialization process to more complex approaches (e.g., use of SSMs with thresholding in [FBCO07]).

Specific acquisition protocols [BFS+07, FBCO07, DAWW07] are commonly adopted to ensure the well-behavior of classification [FOP+05, FBCO07] and thresholding [LRN+08] techniques. Otherwise, post-corrective techniques are usually used such as in [KVP+10] where after a thresholding procedure, an adapted method was applied to separate the fused bony structures in pathological glenohumeral joints. Edge-based techniques can be found in MRI cartilage [LEHE97] or CT bone

[YAGE05] segmentations, especially in the case where clear and highly contrasted boundaries are visible. Region-growing is also very popular [LZZ+00, HM02, KEK03, ZSN+04, NLBA07] when structure boundaries are less reliable. Operators of mathematical morphology (dilation, closing, opening, etc.) were often used in conjunction with region growing to filter and correct the results (e.g., holes closing, false positives removal, etc.).

3.4.2.2 Registration-based Segmentation

First type of registration-based approaches is the fitting of geometrical primitives with image data. The relatively simple nature of the primitives only allows the modeling of simple anatomical structures or can serve for the computation of some biomechanical parameters. For instance, the femoral head is often approximated by sphere fitting from which the hip joint center is simply estimated as the center of the fitted sphere. This approach was adopted by Overhoff *et al.* [OEvJ99] to segment the femoral head in US images. Ehrhardt *et al.* [EHW+00] also adopted sphere fitting techniques but this time on segmented models of the femur and hip bone (Fig. 3.9(c)). The segmentation resulted from an initial affine registration with an atlas and was followed by a free-form registration based on Thirion's demons algorithm [Thi98], as depicted in Fig. 3.9(a) and 3.9(b). Similarly, atlas-based affine registration was used in [FCWO07] to initialize a more efficient segmentation algorithm based on ASM.

(a) (b) (c)

Figure 3.9: Registration-based musculoskeletal modeling. (a) a CT pelvic bones atlas (corresponding segmented models in white contours) is registered with a test subject image. Result is shown in (b). Registration of geometrical primitives can also be used to recover some biomechanical parameters like in (c): a sphere is fitted to the femoral head or acetabulum to estimate the hip joint center. Reproduced from [EHW+00], © 2000, with permission from IOS Press.

Pettersson *et al.* [PKB06] exploited non-rigid registration of CT images to segment the hip joint bones. The registration was based on the morphons algorithm [KA05] which relies on a local descriptor referred to as quadrature phase. In [SJL+07], muscle geometry was extracted from MRI images by using a non-rigid atlas registration based on the mutual information. The same approach was exploited to compute subject specific muscle action lines [SLS+09]. In [GP08], a non-rigid model-to-image registration was devised to segment musculoskeletal structures of the lower limbs. Non-rigid deformations were expressed by blending rigid transforms of overlapping mesh regions defined as model vertex clusters.

3.4.2.3 Deformable Model-based Segmentation

As described in Chap. 2, direct segmentation and registration approaches are usually unable to preserve model smoothness and capture fine details, respectively. In this context, deformable models (See Sec. 2.4) are more appropriate. Bone modeling had greatly benefited from deformable models, especially in CT where usually highly contrasted boundaries are visible. As a result, a deformable model closely initialized to the structure contours often produced good results. Active contours are very popular and are often used in an interactive manner, where contours can be locally modified to correct possible errors. Various works [SM99, KGG+03] used a slice-by-slice segmentation approach where contours were propagated between slices. Despite the claim of effectiveness of this approach, a loss of coherence between slices was often observed which yielded unsmooth models as depicted in Fig. 3.10. Smoothing approaches are of course available to correct these results, but special care must be taken in ensuring that the smoothing preserves good segmentation accuracy.

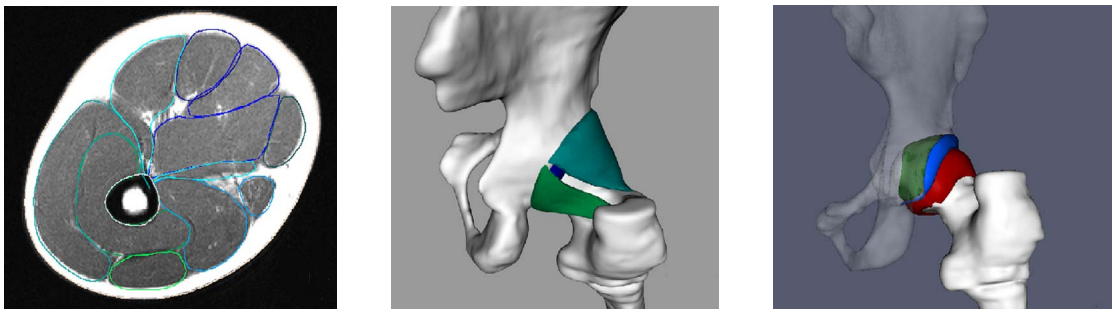
Figure 3.10: Active contour-based femur segmentation [SM99]: by segmenting the structure slice by slice (here proximal femur), there is a high probability to produce non-smooth models. © 1999 IEEE.

Bone segmentation remains difficult in articular areas due to the close proximity of the bones. Evolving contours get attracted by the wrong contours which produce incorrect segmentation results and inter-penetrating contours in case of multiple bones segmentation. Active contours in their original formulation are unable to tackle these issues because (i) they use a simple gradient magnitude criterion and (ii) no interaction between evolving contours is considered. Ballerini and Bocchi [BB03] extended the active contours approach by using a more appropriate image energy that was able to discriminate edges more intelligently. Furthermore, instead of using the variational calculus to iteratively minimize the snake energy, they made use of genetic algorithms. These had the advantage to be global optimization approaches, which were more robust in presence of local minima. Finally, a binding was created between adjacent snakes based on elastic forces. Their approach was successfully illustrated with hand bones segmentation from X-rays images, although the extension to 3D of their approach might not be straightforward and as efficient. In [STCK03], attention was also paid in devising specific deformable models for skeletal structures segmentation. The authors proposed an approach which coupled active contours and region growing, in which contours inter-penetrations were resolved with a region competition algorithm.

Geodesic active contours and implicit deformable models based on the levelset evolution were

used to segment the knee [LFG+98, KHC05], spine [KK08] and skull [RBH+00] bones. As previously mentioned in Sec. 2.4.3, levelsets have the capacity to undergo topological changes which offer more flexible ways to initialize the deformable models with respect to snakes. Nevertheless, levelsets are still facing the issues of strong bones proximity as exemplified in [KK08], where some spine vertebrae were wrongly fused together. Computational burden is also an issue, and fast levelsets approaches [KHC05, KK08, UPB08] are often necessary. Finally, Lorigo *et al.* [LFG+98] also reported the incapacity of some implicit deformable models to distinguish cortical from trabecular bone.

Discrete deformable models have a more flexible representation. As a consequence, various techniques can be combined to better address the aforementioned issues. Gilles *et al.* [GMMT06, GMT10] proposed to consider 2-simplex meshes [Del99] as a particle system. Each vertex was a lumped mass particle with a dynamic state (position, velocity and acceleration), which was subject to internal and external forces. Internal forces were based on some weak prior information that relied on assumptions about surface regularity (smoothness, curvature) and shape. The weak shape prior used the simplex parameters which encode the local configuration of a vertex with respect to its neighbors [Del99]. External forces considered image information (e.g., intensity profiles and gradients) and non-penetration constraints. Non-penetration was ensured thanks to collision detection and response techniques between the different models. A Newtonian evolution was solved by a stable implicit integration scheme, which confered robustness and the possibility to use large timesteps. This approach was successfully used to segment bones, cartilages, ligaments and muscles (Fig. 3.11) from MRI images. More details will be given in Chap. 5 as part of our work is based on Gilles et al.'s approach.



(a) (b) (c)

Figure 3.11: Musculoskeletal modeling with discrete deformable models [GMT10]. (a) Axial MRI slice with segmented muscles overlay. 3D models of segmented bones with (b) ligaments and (c) cartilages of the hip joint.

In Snel *et al.* [SVG02], wrist bones segmentation from MRI was carried out with discrete triangular meshes that follow Lagrangian mechanics. Images forces were computed based on a 1-D convolution of the image intensity with a second-order normalized Gauss filter. The scale of the Gaussian kernel was evolved in a multi-scale fashion to filter out noise (large scale) while detecting edges accurately (small scale). Regulation of meshes was based on curvature forces which preserved the model smoothness. An inflation force (similar to the balloon force in the

snakes model [Coh91]) was also available. Coupled with a remeshing scheme, this inflation strategy allowed to place "seeds" (small tetrahedra) in the image and let them inflate until the bone boundaries were reached. This gave a significant flexibility to the initialization process. However the downside of this approach was that a first binarization of the image was necessary to guide the inflation process. Depending on the quality and availability of the binary image, the segmentation might not produce satisfactory results.

3.4.2.4 Knowledge-based Deformable models Segmentation

Knowledge-based deformable models being very efficient and robust approaches, they have quickly become a great success in musculoskeletal modeling. Bone modeling has particularly benefited from them since bone variations among individuals are less important compared to those of soft organs, which allow the creation of satisfactory training datasets.

Musculoskeletal modeling from digitized data can be also considered as image-based modeling. Digitized data are commonly obtained by reconstructing a structure surface by means of optical capture technology. Typically, bone surfaces are partially reconstructed in intra-operative setups where only part of the bones is visible in the operative field. SSMs expressed as PDMs are used in an optimization process which varies the shape parameters until the generated instance fits the digitized data optimally [FLJ99, RST+07, ZSG10].

Generally speaking, SSMs are appropriate to segment structures which are partially visible. This is valid for X-rays imagery which is a widely utilized modality for its ability to be used intra-operatively. Since X-rays images suffer from their inherent 2D projections SSMs shape priors are very useful. The idea is similar to digitized data, i.e. shape and pose parameters are optimized to have an instance which best matches the imaged structure in the X-rays image. This can be seen as a 2D (X-rays image) to 3D (instance of the SSM) registration process. Different strategies are thus possible based on the various registration components (Sec. 2.3). Dong *et al.* [DGZ07] registered a 3D femur SSM to a 2D contour extracted by Bayesian inference. Similarly, Lamecker *et al.* [LWH06] extracted pelvis contours from X-rays images (simulated from CT in their experiments) with a Canny filte [Can86], and matched them with optimized instances of a pelvis SSM. Cresson *et al.* [CCB+09] segmented pathological spines from X-rays images by using 2D registration features extracted from the images. Signed distance maps built from the extracted 2D contours were also used in [KNO+09] to express the similarity criterion. 3D to 3D registration is also possible, Heinze *et al.* [HMK+02] used an ICP-driven registration to optimize shape and pose parameters in segmenting knee bones from CT images.

In [SCT07, HJ08, SFK+08], the use of contour or features extraction was avoided by registering the X-rays image with a simulated one, called Digitally Reconstructed Radiograph (DRR). The simulated image was built from a statistical model similar to AAM which encompassed shape and appearance obtained from CT data. Given a point source and calibration parameters, the DRR creation process mimics the principle of X-rays image formation. In [DLvB+10], DRRs were not created because only an SSM was available. However, a projective image to be registered

with the X-rays was created from the SSM following the same principle of ray casting. One of the difficulties of using 2D-3D registration approaches with X-rays images is that in some cases the calibration parameters (intrinsic and extrinsic parameters) of the X-rays acquisition device are unknown. As a result, additional parameters might need to be optimized, which increases the search space dimension, the computation time and the risks of being stuck in local minima.

Alternatively, X-rays and other types of modalities can be segmented with evolving statistical deformable models. To segment femurs in X-rays, Behiels *et al.* [BMVS02] used active shape models. In [DLH07], statistical implicit deformable models (i.e., levelsets) yielded the extraction of hip bones and femurs from X-rays. These approaches were strongly inspired by the seminal work of Leventon *et al.* [LGF00] which illustrated the efficient embedding of levelsets into SSM to segment femurs and vertebrae from MRI and CT, respectively. ASM and AAM evolutions were coupled in [KMB+03] to efficiently segment carpal bones in X-rays images. Knee structures were successfully segmented from MRI images based on ASM [FCWO07, FBCO07], although specific acquisition protocols were necessary. Pelvic bones were also successfully segmented from CT images by Lamecker *et al.* [LSHD04], who combined an SSM of the pelvis with an intensity profiles-based evolution, as shown in Fig. 3.12.

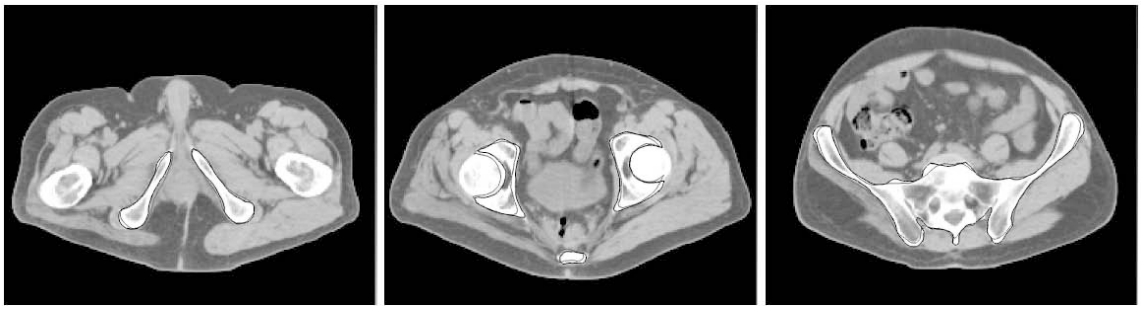
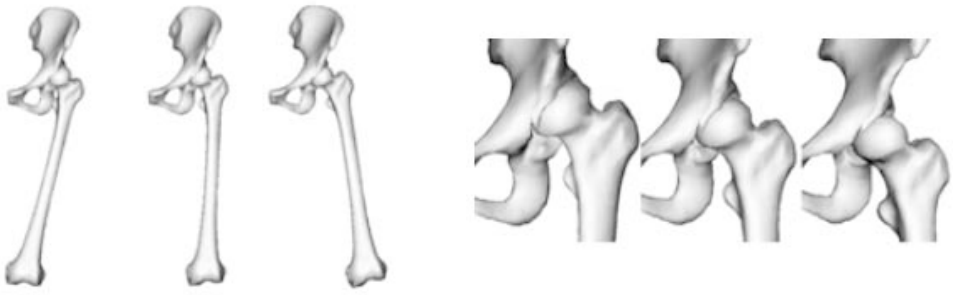


Figure 3.12: CT pelvis segmentation based on SSM [LSHD04]. Gray contours indicate the final segmentation results (black ones should not be considered here). Images used by permission.

Seim *et al.* [SKH+08] extended this approach by adding in particular an automatic initialization relying on the generalized Hough transform. In [KLZH09] and [YOT+09] hip joint kinematics were moreover included into SSMs (Fig. 3.13(a)). Yokota *et al.* [YOT+09] additionally considered pathological conditions of the hip joint (osteoarthrosis caused by hip displasia) where bones appeared to be fused in the images. To do this, a separate SSM of pathological articular areas was built and combined with the pose-shape statistical model (Fig. 3.13(b)). In both works, such enriched SSMs yielded significant improvements in hip bones segmentation, in particular in the difficult articular area, where bones are in close proximity. Seise *et al.* [SMRW06] also tackled the issue of knee bones proximity in X-ray by devising shape and appearance models to be used in a Bayesian framework.



(a) (b)

Figure 3.13: SSM of the hip considering shape, pose and pathology [YOT+09]. (a) First mode of variations of SSM which clearly expresses relative pose between the hip bone and the femur. (b) Fourth mode mainly represents pathological deformations of osteoarthrosis caused by hip dysplasia. Images used by permission.

3.5 Discussion

This Chapter has underpinned the great complexity of the human musculoskeletal system. Different soft and bony structures interact together to ensure the stability, protection and locomotion of the human body. This complex machinery is studied in vivo with acquisition modalities that have their advantages and drawbacks. In particular, MRI is the chosen modality in our work due to its capacity to simultaneously image soft and bony structures in an non-invasive manner.

We explained that computer-based musculoskeletal modeling is addressed in different ways depending on the desired requirements. If complex models need to be created without the constraint of being subject-specific, CG-based modeling is a good solution. In fact, generic models can be carefully built with interactive tools and subsequently resized or altered to create different models. Coupled with some simulation capabilities, these models improve the realism of virtual humans. Similarly, these models can be used for teaching purpose to a targeted audience (e.g., biomechanics or medical students) as long as these models provide the desired degree of realism and complexity.

When medical diagnosis or assistance is in stake, subject-specific modeling is fundamental. Image-based modeling approaches should hence be considered and chosen depending on the segmentation context. For instance, highly contrasted images might be correctly processed with direct segmentation approaches. However as previously explained in Chap. 2, many real problems require the use of more robust approaches, such as those based on prior-knowledge. These approaches are very powerful but have their pitfalls such as the correspondence and training size problems. These factors strongly motivate our research in devising efficient and robust knowledge-based deformable models for musculoskeletal modeling.

As a summary, Table 3.2 reports the cited image-based modeling approaches with used acquisition modalities and structures. The following abbreviations are used to describe the main techniques involved in each approach:

– **DS**: direct segmentation (Sec. 2.2), where:

- *class.*: classification
- *edge*: edges-based
- *fitting*: simple primitive fitting, like sphere
- *grow.*: region growing
- *morph.*: mathematical morphology
- *prot.*: special acquisition protocol
- *thresh.*: thresholding

– **Reg.**: registration (Sec. 2.3)
– **DM**: deformable model (Sec. 2.4), where:
- *AC*: active contour

– **KDM**: knowledge-based deformable model (Sec. 2.5), where:
- *PDM*: point distribution model, might be used or not with the active shape model search (ASM)
- *SDF*: SSM built with signed distance functions
- *AAM*: SSM built on appearance and using the active appearance model
- *DF*: SSM built with a non-rigid deformation field

The term *digitized* refers to digitized point clouds or partial surfaces of structures acquired with an optical technology.

Table 3.2: Image-based musculoskeletal modeling approaches.

Reference	Modality	Structure	Techniques
[LEHE97]	MRI	Cartilage: knee	DS: edge
[HM02]	CT, MRI	Bone: spine	DS: thresh., morph., grow.
[KEK03]	CT, MRI	Bone: hip, skull	DS: grow., morph.
[ZSN+04]	MRI	Bone: femoral head	DS: grow., morph., fitting, class.
[YAGE05]	CT	Bone	DS: edge
[FOP+05]	MRI	Cartilage: knee	DS: class.
[BFS+07]	MRI	Bone: knee	DS: class., prot.
[DAWW07]	MRI	Bone	DS: thresh., prot.
[FBCO07]	MRI	Bone, cartilage: knee	DS: SVM class., prot., KDM: PDM
[NLBA07]	MRI	Bone: shoulder	DS: edge, grow.
[LRN+08]	CT	Bone	DS: thresh., graph cut
[KVP+10]	CT	Bone: shoulder	DS: thresh., special region growing
[OEvJ99]	US	Bone: femoral head	Reg.: sphere fitting
[EHW+00]	CT	Bone: hip	Reg.: atlas affine, demons
[PKB06]	CT	Bone: hip	Reg.: atlas non-rigid
[FCWO07]	MRI	Bone: knee	Reg.: atlas affine, demons; KDM: PDM

Table 3.2: Image-based musculoskeletal modeling approaches (continued)

Reference	Modality	Structure	Techniques
[GP08]	MRI	Bone, muscle, ligament, cartilage: hip	Reg.: shape matching
[SLS+09]	MRI	Muscle: action lines points	Reg.: atlas non-rigid
[MLZ+10]	CT	Bone: thoracic vertebra	Reg.: shape matching
[GXDX10]	X-rays	bone: femur	Reg.: shape matching; DM: AC
[SM99]	CT	Bone: proximal femur	DM: AC
[KGG+03]	Synthetic MRI	Bone, cartilage: knee	DM: AC
[BB03]	X-rays	Bone: hand	DM: genetic AC
[STCK03]	CT	Bone: wrist	DM: AC; DS: grow.
[LFG+98]	MRI	Bone: knee	DM: geodesic AC
[RBH+00]	MRI	Bone: skull	DM: levelset
[KHC05]	MRI, CT	Bone: knee	DM: levelset
[KK08]	CT	Bone: spine, pelvis	DM: levelset
[UPB08]	X-rays, MRI	Bone: hand, knee	DM: geodesic AC
[SVG02]	MRI, CT	Bone: wrist	DM: discrete
[GMMT06, GMT10]	MRI	Bone, cartilage, ligament, muscle: hip	DM: discrete
[FLJ99]	Digitized data	Bone: femur	KDM: PDM
[LGF00]	MRI, CT	Bone: femur, vertebrae	KDM: SDF
[HMK+02]	CT	Bone: knee	KDM: PDM
[BMVS02]	X-rays	Bone: femur	KDM: PDM
[KMB+03]	X-rays	Bone: hand	KDM: PDM + AAM
[LSHD04]	CT	Bone: pelvis	KDM: PDM
[LWH06]	X-rays	Bone: rib cage	KDM: PDM
[SMRW06]	X-rays	Bone: knee	KDM: PDM, appearance models
[DGZ07]	X-rays	Bone: femur	KDM: PDM, appearance models
[RST+07]	Digitized data	Bone: femur	KDM: PDM
[DLH07]	X-rays	Bone: hip	KDM: SDF
[SCT07]	X-rays	Bone: pelvis	Reg.; KDM: PDM
[FCWO07]	MRI	Bone: knee	DS: prot.; KDM: PDM
[HJ08]	X-rays	Bone: femur	Reg.; KDM: PDM, AAM
[SFK+08]	X-rays	Bone: femur	KDM: PDM
[SKH+08]	CT	Bone: pelvis	KDM: PDM
[KLZH09]	CT	Bone: hip	KDM: SSM
[YOT+09]	CT	Bone: femur	DS: edge, thresh.; KDM: PDM
[KNO+09]	X-rays	Bone: femur	KDM: PDM
[CCB+09]	X-rays	Bone: spine	KDM: PDM

Table 3.2: Image-based musculoskeletal modeling approaches (continued)

Reference	Modality	Structure	Techniques
[ZSG10]	X-rays	Bone: femur	KDM: PDM
[DLvB+10]	X-rays	Bone: rib cage	KDM: PDM
[KMG+10]	US	Bone: spine	KDM: DF

Chapter 4

Construction of Shape and Appearance Priors

4.1 Introduction

In this Chapter, we will describe the process of building appearance and shape priors. As discussed in Sec. 2.5 of Chap. 2, several steps are required to express these priors. First the nature of the training features must be specified and computed for each training samples. In our case, segmented meshes with corresponding images define a training sample for a given subject. Shape and appearance features can then be derived and used to express corresponding priors, formulated as statistical models. Alignment and correspondence procedures are usually required to build accurate models.

This Chapter is organized as follows: the mesh representation along with the approaches to create the training data from the images are detailed in Sec. 4.2. Subsequently, Sec. 4.3 and 4.4 describe the construction of shape and appearance priors, respectively. Then, an evaluation Sec. 4.5 depicts the behavior and efficiency of the proposed methodology followed by a discussion Section.

4.2 Generic Model Creation

This Section describes the creation of what we refer to as generic models. A generic model expresses the main anatomical features shared by individuals and provides a weak prior of the future structure to be segmented. The construction of a generic model is similar to that of some CG-based musculoskeletal modeling approaches that rely on medical images (Sec. 3.4.1). Medical images are exploited to delineate the structures of interest before post-processing techniques are applied to improve the generic models. Quality criteria are defined according to the future use of these models. In our case, we aim at creating models that are smooth, quasi-regular and with no inter-penetration of the structures.

4.2.1 Simplex Meshes

In this thesis, shapes are modeled as discrete deformable models by using a 2-simplex mesh representation. Simplex meshes were first proposed by H. Delingette [Del94a, Del94b, Del99] and were later on extended by J. Montagnat [MD98, MD05] and B. Gilles [Gil07, GMT10]. In this Section, we focus on our contributions around the use of 2-simplex meshes, while details on simplex meshes are given in Sec. A.3 of Appendix A.

4.2.1.1 Introduction

A 2-simplex mesh presents the characteristic of having vertices with exactly three neighboring distinct vertices. Meshes are composed of faces which have an arbitrary number of vertices and as a result these are not necessarily planar as depicted in Figure 4.1. 2-simplex meshes, which will be simply referred to as simplex meshes, can represent any surface with arbitrary topology. Furthermore, they are topologically dual with triangulations. This means that it is possible to convert a simplex mesh to a triangular mesh and vice versa based on geometrical schemes. The

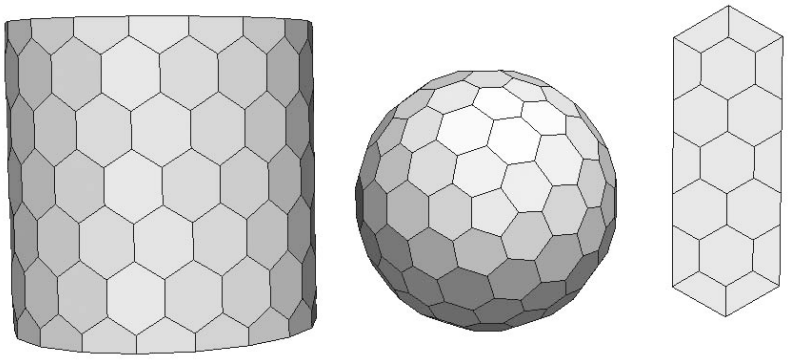


Figure 4.1: Examples of 2-simplex meshes.

conversion is not ensured to preserve the exact geometry but the topology is kept. To convert a simplex to mesh to triangular mesh we chose the preservative scheme of Gilles *et al.* [Gil07] (See Fig. A.2 in appendix).

4.2.1.2 Simplex Parameters

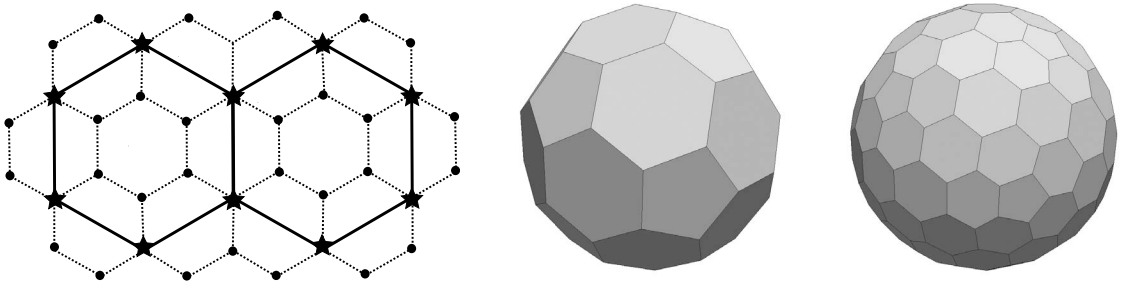
Three independent parameters are sufficient to express in a unique manner the position of a vertex P with respect to its neighboring vertices P_i. These parameters coined *simplex parameters* are invariant to similarity transform [Del99]. In [Gil07], B. Gilles reported that the uniqueness of the simplex parameters was only respected with some mesh configurations, and proposed hence alternative simplex parameters (Sec. A.3.3), which we used in this work.

4.2.1.3 Operators for Mesh Transformation

In [Del94a, Del99], H. Delingette described four basic independent operators to transform a k-simplex mesh (Sec. A.3.5.1). Some of these operators do not preserve the mesh topology by performing e.g. operations of mesh merging or splitting. Macro-operators can be defined from these basic operators. In [Gil07, GMT10] a variety of these operators was proposed to yield an optimization procedure that automatically alters simplex mesh geometry to satisfy quality constraints. For instance, quasi-regularity of the faces can be enforced by aiming at a desired average edge length. This allows the creation of meshes with a specific resolution and quality as described in Sec. 4.2.2.

4.2.1.4 Multi-resolution Scheme

Various resolutions of a shape can be generated with a multi-resolution scheme. We adopted the conservative scheme presented in [Gil07] which is illustrated in Fig. 4.2. Given a mesh $\boldsymbol{X^k}$ at resolution k with N^k vertices x_j^k, this scheme preserves the vertices of the lower resolutions: $x_j^{k+1} = x_j^k, \forall j \in [1, N^k]$. When shape points are contiguously stored in memory as a 1-D array, point indices are consistent across resolution levels. In the following, we may use the notation $y = x_{|p}$, which means that y is built by taking the first p values of x. With such a notation, $\boldsymbol{X^k} = \boldsymbol{X^{k+1}}_{|N^k}$.



(a) (b) (c)

Figure 4.2: Selected simplex mesh multi-resolution scheme [Gil07]. (a). A 2-simplex mesh at resolution k ($\star$ vertices and thick edges) yields a new 2-simplex mesh at resolution $k+1$ ($\bullet$ vertices and dotted edges). With this scheme, the resolution $k+1$ contains the vertices of the lower resolution k. (b)-(c) Example of mesh resolution increase with this scheme.

4.2.2 Interactive Model Construction

In this Section, we assume that a model can deform to perform different actions (e.g., smoothing in Sec. 4.2.2.2, topological re-meshing in Sec. 4.2.2.3, etc.), the main principles of the model evolution being given in the next Chapter.

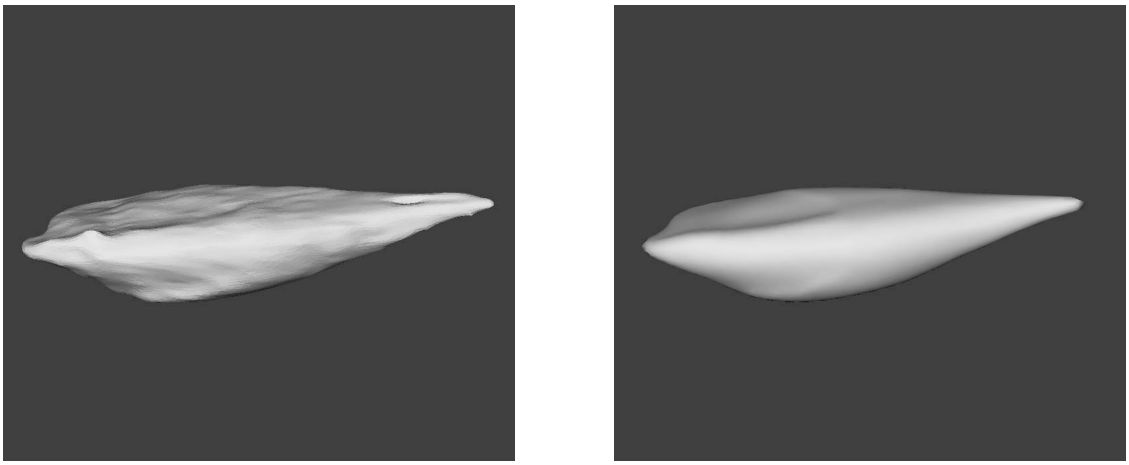
4.2.2.1 Supervised Segmentation

The creation of generic models is based on MRI acquisitions which image the structures to model. The first step consists of a (semi-)manual segmentation of the structures in the images: an operator manually delineates the structure contours in each slice, or she/he does it only in some slices and subsequently uses an interpolation technique to approximate the remaining contours in the non-segmented slices. For our work, manual segmentations were carried out with the free application ITK-SNAP[1]. The Marching Cubes algorithm [LC87] is subsequently applied on the labeled image to produce a triangular mesh. Eventually, we used the standard conversion scheme of Delingette [Del99] to produce a simplex mesh from the triangular mesh (See Sec. A.3.2).

4.2.2.2 Smoothing

Generally, the generated mesh surface is not smooth, as depicted in Fig. 4.3(a). This is mostly attributed to the dataset resolution (especially in case of poor inter-slice distance) and the subjective labeling in manual segmentation. Ideally, the resolution should be as high as possible to provide detailed images and to avoid big jumps between consecutive MRI slices. This is not always feasible (e.g., device limitations, acquisition time restrictions) as previously explained in Sec. 3.3.4. Furthermore, a lower resolution might be preferred to reduce the memory footprint and ease the interactive segmentation. Finally most anatomical structures are smooth, therefore the objective is to find a solution that makes the generated mesh surface appear as such.

1. http://www.itksnap.org



(a) (b)

Figure 4.3: Mesh smoothing example. (a) A model of soleus muscle is reconstructed from MRI by the Marching Cubes [LC87] algorithm of the manually segmented labels. The model presents some bumps and pits not inherent to the muscle anatomy. (b) These flaws are corrected with a conservative smoothing approach.

To tackle this issue, we apply the smoothing approach presented in [Gil07]. For a given point P, the smoothing computes a target point $\tilde{P}$ based on a weighted Laplacian smoothing coupled with an additional term accounting for the average local elevation of the neighboring vertices P_i:

$$\tilde{P} = \frac{\sum_{i=1}^{3} S_i P_i}{\sum_{i=1}^{3} S_i} + h_n \boldsymbol{n} \tag{4.1}$$

where $\boldsymbol{n}$ denotes the normal at P and S_i is the area "covered" by a vertex P_i, which is defined as the sum of the areas of all triangles sharing this point P_i according to the chosen triangular tessellation (See Fig. A.2(b) in Annex). The average local elevation $h_n = \sum_i \|P_i - P_{\perp i}\| /3$ (Fig. A.4) reduces the shrinking effect recurrent with Laplacian smoothing. Figure 4.3(b) shows an example of smoothed mesh.

4.2.2.3 Re-meshing Optimization

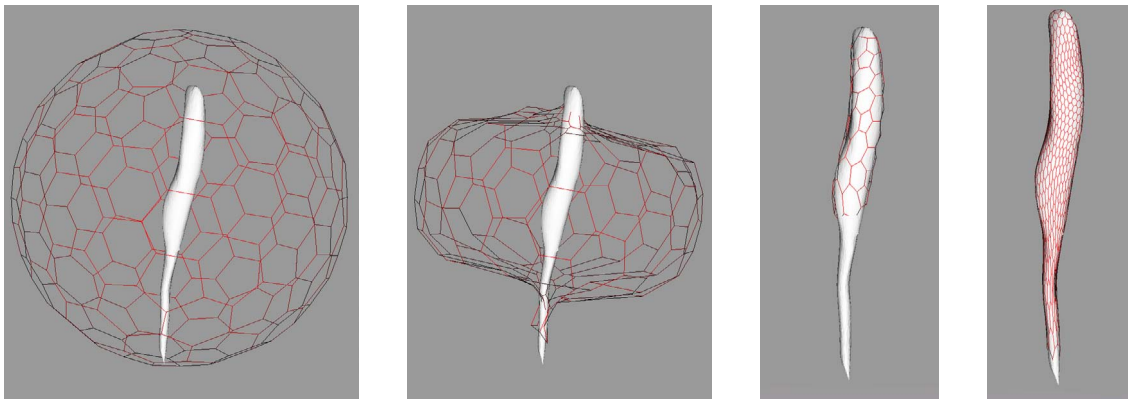
Once the model is smoothed, we must take care of its geometric (uniformity of vertex repartition) and topological (uniformity of edge number among faces) qualities [GMT10]. In fact, the smoothed model may present highly irregular faces and a bad vertex repartition over its surface (e.g., localized vertex clusters of high density despite a low curvature). This is often due to the Marching Cubes approach which tends to produce too much elements of unequal quality. We would like to have meshes with quasi-regular faces to ensure a better stability of the segmentation and a better expression of shape priors (See Sec. 4.3.1.2). In [Gil07], the author described an approach which consisted of the following main steps:

1. Initialization of a simple and low resolution primitive to match the smoothed model.
2. Optimized topological and geometric re-remeshing of the primitive.
3. Computation of different mesh resolutions from the optimized mesh.

In the first step, a simple primitive (usually a sphere) is centered on the center of gravity of the mesh and sized according to the longest principal axis (Fig. 4.4(a)). Then, a deformable model procedure based on an ICP-based force deforms the primitive so that is best matches the model. In the second step, the resulting fitted primitive undergoes an optimization procedure which alters its geometric and topological configurations according to defined quality and control criteria (e.g., desired average edge length). The result is a mesh which fits the initial model and presents a lower number of faces with a better quality. In the last step, a high resolution version of the optimized primitive is computed and it is once again accurately fitted to the initial model. Finally, different lower resolutions are produced from this high fitted resolution model.

4.2.2.4 Tubular and Elongated Structure Modeling

While the optimized re-meshing usually produces meshes with desired quality, some type of meshes cannot be correctly processed. These are e.g. meshes which are tubular-like structures with thin extremities such as the tibialis anterior muscle depicted in Fig. 4.4(a), where the deformation of a sphere primitive faces some difficulties. Indeed, the ICP-based deformation process cannot correctly capture some parts as exemplified in Fig. 4.4(c). This is due to (i) the low resolution of the initial sphere, which is purposely not set too high in order to not penalize the re-meshing process, and (ii) the presence of thin tubular extremities which do not create sufficient "attraction features" for the ICP-based matching. The consequence is that the optimized re-meshing will only be applied in some parts of the model (e.g., upper part of tibialis anterior muscle in Fig. 4.4(c)). During the last stage, the higher resolution version of the optimized model is able to recover the missing parts by referring to its finer resolution, but at the cost of a strong deformation of the mesh (Fig. 4.4(d)). Eventually, the resulting mesh is of poor quality.

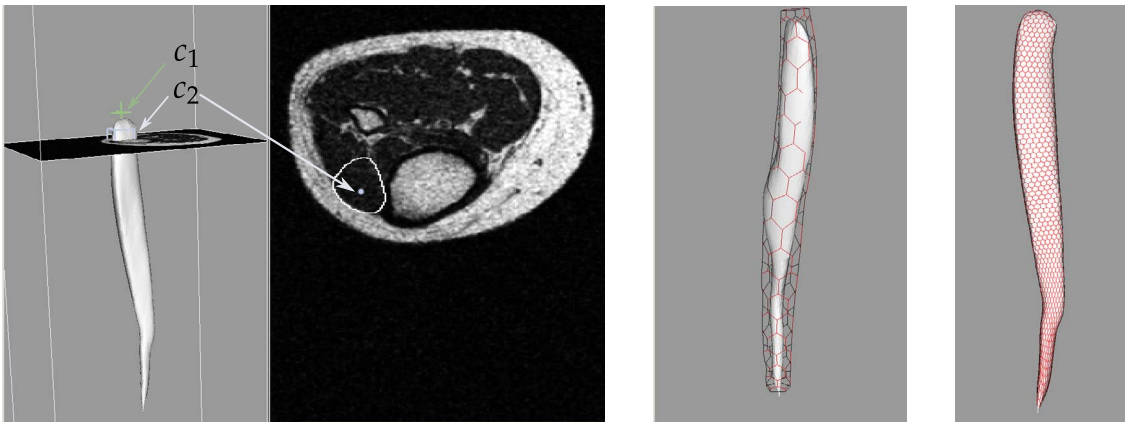


(a) (b) (c) (d)

Figure 4.4: Issues with elongated models. (a) A sphere primitive (wireframe) is initialized and will be deformed to match an initial model of a tibialis anterior muscle, which is characterized by an elongated shape. (b) The primitive evolves to match the model based on an ICP approach. (c) The optimized re-meshing faces difficulties in keeping the primitive fitted with the lower part of the initial model. (d) The higher resolution of the primitive manages to fit tightly the whole initial model but ultimately presents faces of poor quality.

To address this issue, we propose to use another type of initial primitive that fits more closely

to the initial model. To achieve this, the initial primitive is modeled as a tubular structure which is simply characterized by a centerline. A centerline is a succession of points c_i with an associated radius r_i. In the literature of vessel segmentation, various vessel tracing approaches [AB02, PJBB05] create, on the fly, the centerline by tracking the ridge of the vessels based on the values of the eigenvalues of the Hessian intensity matrix. When applicable, these methods can be used to automatically build the centerline of the structures of interest. In the more general case, the operator can quickly place a few centerline points c_i as shown in Fig. 4.5(a) and define for each of them an approximate radius. From the centerline, a tubular mesh is hence created (Fig. 4.5(b)). The centerline definition does not need to be highly accurate as the fitting process is subsequently applied. We found that if the initial tubular mesh is appropriately created, the re-meshing optimization process is no longer necessary. In fact, after the initial tubular mesh has been fitted, we simply increase the resolution and apply a final fitting process to get a final mesh of good quality, as illustrated in Fig. 4.5(c).



(a) (b) (c)

Figure 4.5: Fitting with tubular primitives. (a) An initial tubular primitive is coarsely initialized by first (automatically) defining centerline points c_i on different image slices. (b) Then, a simplex tubular mesh is created from the centerline. (c) After fitting and resolution increase, the final model accurately matches the initial model and presents faces of good quality.

The creation of a tubular structure from a centerline composed of N points c_i is done in an iterative manner. For each pair of consecutive centerline points c_i and c_{i+1}, a simplex oriented truncated cone (TOCone) is built. All the TOCones are then joined together to build the final tubular structure. The construction of a TOCone is performed as follows (Fig. 4.6(a)). First, a vector n_i is computed for each point c_i as $n_i = (c_{i+1} - c_i)/\|c_{i+1} - c_i\|$ (for the last point c_N, we set $n_N = n_{N-1}$). This vector n_i is the normal of the disc f_i at one of the TOCone extremities. Then, an arbitrary point s_i is chosen on the contour of f_i. The vector $\boldsymbol{c_{i+1}s_i}$ is subsequently projected on the disc f_{i+1} to compute the point s_{i+1}. For each disc f_i, the point s_i is used as a start point to uniformly sample at M positions p_i^j the contour (i.e., $p_1^j = s_i$) with an orientation kept consistent for all the discs. The sampling characteristics (e.g., number of samples) are chosen based on a desired average edge length. With the segments $[p_i^j, p_{i+1}^j]$ available, the standard procedure to build a simplex cylinder can be applied as shown in Fig. 4.6(b).

We noticed that the estimated point s_{i+1} was critical in getting a simplex TOCone with faces of good quality. A "twisting" effect was indeed often observed, especially with very oblique TOCones (i.e. a large angle between the normals n_i and n_{i+1}). To address this issue, we optimized the position of S_{i+1} so that the sum of the segment distances $\sum_{j=1}^{M} \left\| p_i^j, p_{i+1}^j \right\|$ was minimal. An example of corrected result is visible in Fig. 4.6(c).

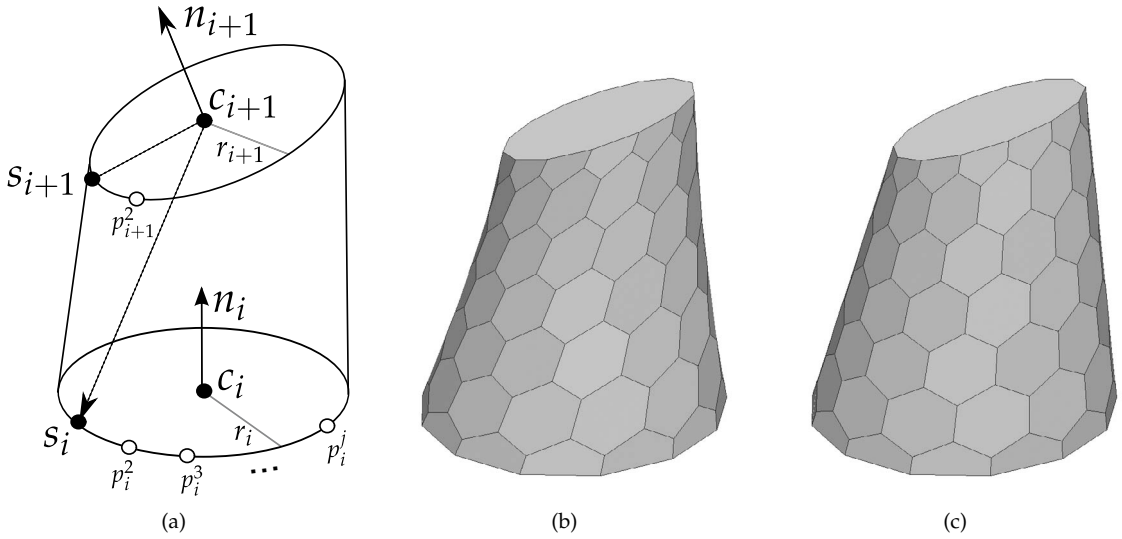


Figure 4.6: Simplex truncated oblique cone creation (TOCone). (a) A TOCone is characterized by two discs with center c_i and c_{i+1}, normals $\boldsymbol{n_i}$ and $\boldsymbol{n_{i+1}}$, and radii r_i and r_{i+1}, respectively. The contours of the discs are regularly sampled at positions p_i^j and p_{i+1}^j from initial points s_i and s_{i+1}. (b) Given the segments $[p_i^j, p_{i+1}^j]$, the simplex TOCone is easily created but may present some "twisting" effect which affects the simplex face quality. (c) A corrective process is applied to improve the model quality.

4.2.2.5 Attachment Modeling

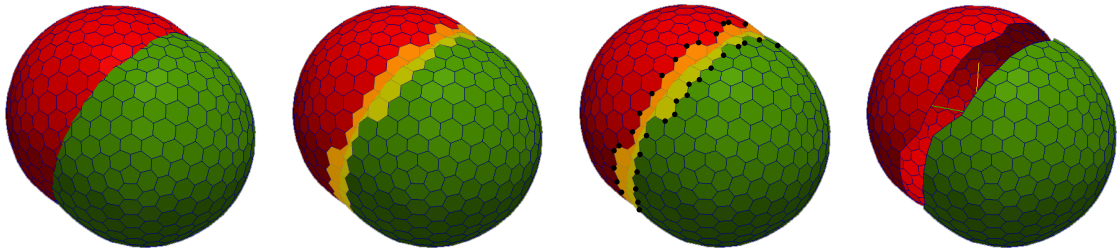
Anatomically, an attachment is defined as the linking region between two structures. For instance, bone/tendon attachments are characterized by the penetration of the tendon tissue into the periosteum of bones. The tendon attachments are thus defined as the tendon tissue sharing a common region with the bone tissue. Comparatively, it means that the meshes modeling these structures should be stuck together which is not the case *a priori*. Indeed, as accurate as the segmentation may be, generated meshes are not guaranteed to be attached together, especially after the smoothing procedure. We thus re-used the method described in [Gil07] to manually place splines curves to define attachment regions. This provides a simple but efficient way to model attachment regions. Meshes are then deformed until they are joined together.

4.2.3 Model merging

Sometimes it might be desirable to merge two intersecting models into a new combined model. This confers a greater flexibility to the construction process as the two parts of the new model can be created independently. In order to perform the merging operation, different steps are necessary:

1. *Cutting contours identification*: given the continuous intersection curve defined between the two models, discrete contours are defined on each model.
2. *Model cutting*: given the cutting contours, a new instance of each model is created such that the faces present in the interior of the other respective model have been removed.
3. *Model merging*: cut models are merged into a single model.

In the first step, intersected faces are computed for each mesh. Given a mesh $\mathcal{M}_1$, intersected faces correspond to those which have vertices inside and outside the other mesh $\mathcal{M}_2$, i.e. the faces of $\mathcal{M}_1$ that are crossed by the intersection curve. This is illustrated in Fig. 4.7(b) and 4.7(c) where the intersected faces of two intersecting spheres are highlighted with a bright yellow color. From these intersected faces, specific vertices are chosen to define the cutting contours. Given a mesh $\mathcal{M}_1$, these vertices are defined as the vertices that (i) belong to both intersected and non-intersected faces of $\mathcal{M}_1$ and (ii) have their three neighbors not inside $\mathcal{M}_2$ (See black discs in Fig. 4.7(c)).



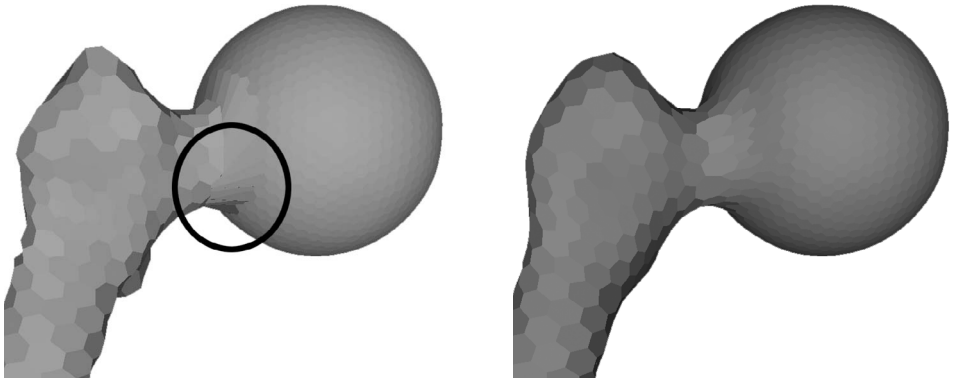
(a) (b) (c) (d)

Figure 4.7: Cutting of intersecting meshes. (a) Two intersecting spheres. (b) Intersected faces of both models are highlighted with a bright yellow color. (c) Black discs indicate the selected vertices of the intersected faces that are used to build the cutting contours. (d) Resulting spheres cut along their respective cutting contours.

In the second step, new meshes are created by cutting along the contours, i.e. by removing the faces inside the cutting contours. This is achieved by using the simplex mesh meta-operator T_3^2 (Sec. A.3.5.1) [Del94a]. An example of result is depicted in Fig. 4.7(d).

In the third step, the simplex mesh meta-operator T_4^2 (Sec. A.3.5.1) is used to merge the two models with the procedure described in [Del94a]. This merging procedure attempts to distribute uniform vertices at the junction of the models to preserve the regularity of the mesh. However, it might happen that the resulting merging area presents largely distorted faces which affect the model smoothness and quality. This is illustrated in Fig. 4.8(a) where a sphere and a proximal femur, cut with the two previous steps, have been merged. Some faces of the merging area (circled) appear as over-stretched. This issue can be addressed by using a corrective post-processing that will produce more regular faces. In Fig. 4.8(b), a smoothing procedure is adopted, resulting in more regular faces in the merging area. However, in this case the smoothing is applied to the whole model which has the negative effect to fade out some anatomical features such as the trochanters. To address this issue, vertices not belonging to the merging area are blocked during the smoothing process .

When there are multiple intersection curves (e.g., a torus intersecting a sphere), the whole process is simply applied on each intersection.



(a) (b)

Figure 4.8: Merging of cut models. (a) A sphere and a truncated proximal femur models are merged with the procedure based on the meta-operator T_4^2 [Del94a]. As shown in the black circle, some faces are too elongated. (b) A corrective process based on smoothing improves the merging area.

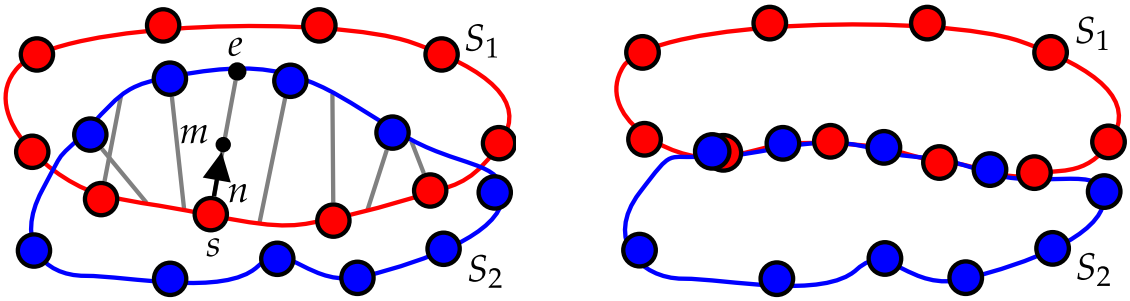
4.2.4 Inter-penetration Removal

The various construction procedures previously described can create non realistic inter-penetrations between meshes. In order to remove them, we propose to apply a geometric post-processing. Let us consider the surfaces S_1 and S_2 in Fig. 4.9(a) that are inter-penetrating each other. The aim of the procedure is to move the colliding points of each surface (i.e. points located in the interior of another mesh), to reach a non inter-penetrating state. Given a point $s \in S_1$ with inner normal n, the line passing through s and directed by n intersects S_2 in e. This point e must be inside S_1 to be valid, i.e. the procedure is not applicable when S_1 is completely at the interior of S_2. The new corrected position m of the point s is then chosen as:

$$\begin{cases} m & = & \underset{x}{\operatorname{argmin}}\, f(x),\ x \in [s,e] \\ f(x) & = & (l_1 * d_1(x) - l_2 * d_2(x))^2 \end{cases} \tag{4.2}$$

where d_i depicts the Euclidean signed distance from x to the surface S_i and l_i is a rigidity parameter. For instance, a configuration with $l_1 \ll l_2$ leads to final points closer to surface S_2. This can be useful to resolve penetrations between a very stiff (e.g., a bone) and soft (e.g., a muscle) material. Alternatively, the stiffness can also be seen as the confidence given to the position of a surface: if S_2 is more likely to be correct compared to S_1, we set $l_1 \ll l_2$ so that S_2 will deform less.

While this technique was inspired from [KMH+04] it presents several differences. First, the point e is not defined as the projection of s on S_2. Indeed, by using such a projection, the point m usually does not correspond to the new position of s in the non colliding state. By using instead the normal direction n, the point m will provide a better approximation. As a result, this technique avoids the usage of a re-sampling scheme as described in [KMH+04], since



(a) (b)

Figure 4.9: Inter-penetration removal. (a) Surfaces S_1 and S_2 are inter-penetrating each other. (b) Result after the inter-penetration removal process.

points just need to be moved to their new positions (Fig. 4.9(b)). Second, with our approach a perfect contact between the surfaces cannot be ensured due to the absence of re-meshing but the results are usually satisfactory when mesh resolution is not too low. However, the future use of these models for segmentation purpose does not require such a perfect contact. Physically-based simulations using these meshes may require an accurate contact, but in any case an additional meshing procedure would have been necessary to convert meshes to volumetric models. Lastly, implicit surfaces [AA04] are not used in our approach since the hypothesis of vicinity to surfaces may be invalid, i.e. large inter-penetrations might be present. In fact, in [KMH+04], the inter-penetration removal is used in a collision response context where inter-penetrations tend to be less important as the collision detection is applied with small time steps.

Usually, the inter-penetration removal process must be repeated a few times until surface changes are below a small threshold which is empirically chosen. To speed up the process, precomputed signed distance maps can be used as well as the efficient golden section search technique [PFTV92] to minimize f.

4.3 Shape Priors

4.3.1 Statistical Shape Models

4.3.1.1 Construction of The Training Shapes

Given the multi-resolution shape description of Sec. 4.2.1.4, a shape $\boldsymbol{X}^k$ at resolution k is modeled as a 2-simplex mesh with N^k vertices x_i^k and outward normals $\boldsymbol{n}_i^k$. As previously presented in Sec. 2.5, the creation of an SSM requires a training set of M shapes $\boldsymbol{X}_j^k$ from which statistics are inferred. Vertices of a training shape $\boldsymbol{X}_j^k$ at level k are written as x_{ij}^k, $i \in [1, N^k]$. To build these shapes, training images were segmented by progressively deforming a generic mesh to match the structure boundaries in the images. These generic meshes were built based on our methodology and were not used as test datasets in order to provide fair evaluations.

Supervised by a radiologist from the University hospital of Geneva, we used the deformable

models approach of [GMT10] to automatically deform the template mesh and interactively corrected erroneous results by using constraint points [Gil07, SNH+09]. This approach relaxed the need for completely manual segmentation while retaining equal accuracy. Since we based our work on this deformable model approach, details will be given in the corresponding Chapter 5.

4.3.1.2 Shape Correspondence and Alignment

Correspondence Building an SSM requires shapes that are aligned to a common frame, which implies that the shapes must be in correspondence. Our method used to construct the training shapes is classified as a "mesh-to-volume" registration correspondence method [HM09] (Sec. 2.5.4). In fact, the advantage of using the same generic model is that all training shapes are identical in topology and have the same vertex indices. However, since such a method is the result of a supervised technique that does not necessarily seek point correspondence optimally, we adopted an additional correspondence method. Various works using SSMs to model musculoskeletal structures (e.g., [RST+07, ZSG10]) are based on Davies *et al.* [DTC+02] Minimum Description Length (MDL) formulation, which aims to balance the compactness of the model with respect to its complexity. However, these approaches suffer from a fixed topology of genus 0 (i.e. shape topologically equivalent to a sphere), and do not apply e.g. to a hip bone shape.

In [SGS07] this constraint was removed by partitioning the shape into different topologically compliant patches. In this thesis, we preferred the approach of [DMW+07] as it did not require any specific shape topology and was appropriate for the purpose of refining a previous correspondence. This method has thus the advantage to not apply any re-meshing process, which avoids the need to update the topology of the various mesh resolutions (Sec. 4.2.1.4) with their possible associated attachments (Sec. 4.2.2.5) since vertex indices are preserved. This approach assumes that the point correspondence is known and consists in sliding landmarks on the surface of the shapes so that a shape-correspondence error based on the Thin Plate Spline (TPS) energy is minimized. Only vertex positions are thus altered. The TPS energy is computed between landmarks $\boldsymbol{U_L}$ of an arbitrarily chosen template shape (selected from the training shapes) and the landmarks $\boldsymbol{V_L}$ of the shapes that still need to be optimized by sliding.

In a multi-resolution context, we usually seek for the correspondence refinement of the shapes with the highest resolution since these meshes model the most faithfully the anatomical structures. From shapes with the highest resolution, the correspondence refinement for shapes of lower resolution is computed by a simple update of vertex positions from the higher resolution shapes. Landmarks can form a subset of the shape vertices of the highest resolution. In our setup, we selected the points x_{ij}^1 of the lowest resolution of each shape $\boldsymbol{X_j^1}$ as landmarks as they were regularly distributed and located near key anatomical locations (Step (1) Fig. 4.10). This justifies the need to have quasi-regular meshes, condition which has been fulfilled in the generic model creation (Sec. 4.2). Further, the use of the lowest resolution decreases the computation time of the correspondence optimization. However, the mesh at the lowest resolution may approximate the real shape too coarsely and thus the finest resolution was used for a landmark surface re-projection after sliding each landmark in order to closely match the training shapes.

When the procedure was complete, the TPS transformation was used to generate all the other resolution levels from the optimized coarsest resolution, followed by a last re-projection on the original higher resolution surface (Step (3) Fig. 4.10). A more technical description with some implementation details of this correspondence approach is given in Sec. A.4 of Appendix A and Fig. 4.10 summarizes the different steps.

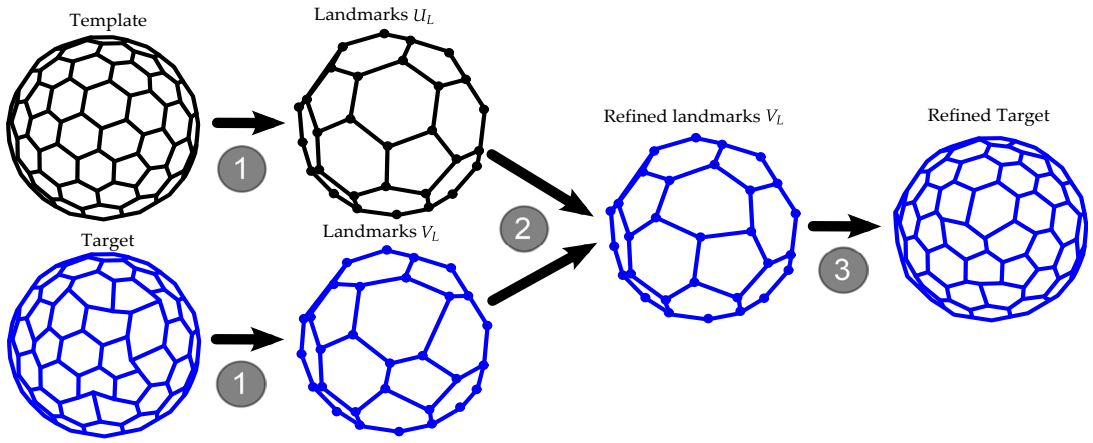


Figure 4.10: Shape correspondence scheme. The correspondence between the template and the target shapes is refined based on the correspondence by landmarks sliding [DMW+07]. (1) Template and target landmarks U_L and V_L are extracted as the vertices of the lowest resolution. (2) The applied correspondence method slides the landmarks of V_L over the surface of the target shape. (3) Based on the TPS transform and the resolution change scheme, the final refined target shape is computed.

Alignment Corresponding shapes were then aligned using the generalized Procrustes approach [Gow75], which iteratively aligned shapes by registering their corresponding point sets (Sec. 2.5.2). We used different types of alignment transforms and computed them based on Horn's [Hor87] quaternion and Umeyama's [Ume91] Singular Value Decomposition approaches. These approaches give close-form solutions of the alignment problem, which can be quickly and efficiently computed. In this thesis, we denote an SSM built with a rigid, affine or similarity alignment transforms, as a "rigid", "affine" or "similarity" SSM, respectively. In the next Section, we show how different behaviors can be conferred to the SSM by choosing different alignment schemes.

4.3.1.3 Multi-resolution Statistical Shape Models

Generative Model Given M aligned training shapes $\boldsymbol{X}_j^k$ at a resolution level k, a point distribution model (PDM) is built based on the principal component analysis (PCA) of the points of all shapes. An arbitrary shape $\boldsymbol{Y}^k$ is approximated from the computed statistics by using the PCA generative model (2.17) presented in Sec. 2.5.3.1:

$$\boldsymbol{Y}^k \approx T^k(\overline{\boldsymbol{X}^k} + \Phi^k \boldsymbol{b}^k) \tag{4.3}$$

where $\boldsymbol{b}^k$ is the vector of shape parameters, T^k denotes the alignment transform, $\overline{\boldsymbol{X}^k}$ is the mean shape and Φ^k is the matrix of principal directions. Dimensionality reduction is applied by

selecting only K^k ($K^k \ll N^k$) principal directions. In this thesis, we commonly used 99% of the total variance to select K^k based on Eq. (2.16).

In Sec. 2.5.3.1, we reported the use of clamping and scaling constraints to restrict the variation of the shape parameters. In practice we observed little differences between both constraint schemes as depicted in Fig. 4.11 where the two types of constraint are applied to an instance of a hip bone generated by a corresponding SSM. However as expected, the scaling constraint is more restrictive as it considers all shape parameters simultaneously. To compute the scaling threshold from the $\chi^2_{K^k}$ distribution we used the DCDFLIB [2] library. During a segmentation, we can choose to apply either one type of constraints or both of them in a sequential manner.

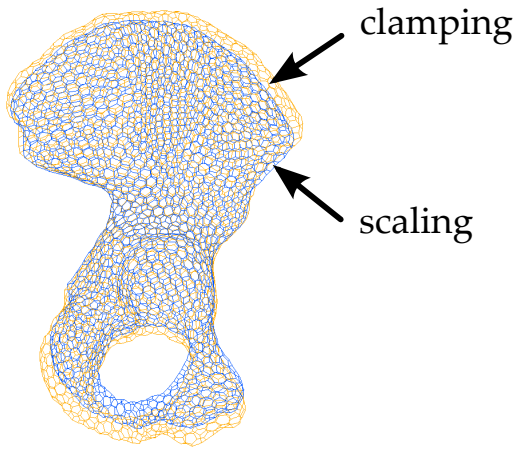


Figure 4.11: Shape constraints illustration. A hip bone instance is generated from a similarity SSM by varying the shape parameter of the first mode, which mostly affects the size of the shape. The scaling constraint (in blue) results in a slightly smaller shape compared to the clamping constraint (in orange).

Regularization Model Given an SSM and an arbitrary shape $\boldsymbol{Y}$, the iterative procedure based on ASM search [CHTH93] (Sec. 2.5.5.2) is used to estimate the unknown parameters $\boldsymbol{b}$ and T that generate a shape from the SSM that is the closest as possible to $\boldsymbol{Y}$ (in a least square sense). To prevent the generation of invalid shapes, clamping and scaling constraints are applied. The replacement of a shape by its closest shape counterpart after finding the corresponding T and $\boldsymbol{b}$ is hereafter referred to as *shape regularization*.

A Remark about Eigenvalues and Eigenvectors It is straightforward to prove that if a shape $Y^k = \overline{\boldsymbol{X^k}} + \Phi^k \boldsymbol{b^k}$, any corresponding shape Y^l at a lower resolution level $l < k$ is expressed by retaining the first N_l elements of the principal components ϕ^k_i at level k:

$$Y^l = \overline{\boldsymbol{X^l}} + \Phi^k_{|N^l} \boldsymbol{b^k} \tag{4.4}$$

In practice, this means that a unique SSM built at level k could be used to generate shapes at levels $l < k$. However, it must be stressed that the vectors $\phi^k_{|N^l}$ are in general not equal to the principal components computed at level l (i.e., $\phi^l \neq \phi^k_{|N^l}$), as e.g. the orthonormality is not necessarily preserved for the vectors $\phi^k_{|N^l}$. Furthermore, eigenvalues are also not necessarily

2. http://people.sc.fsu.edu/~jburkardt/cpp_src/dcdflib/dcdflib.html

conserved across levels. It is only possible to bound the eigenvalues by using the inclusion principle:

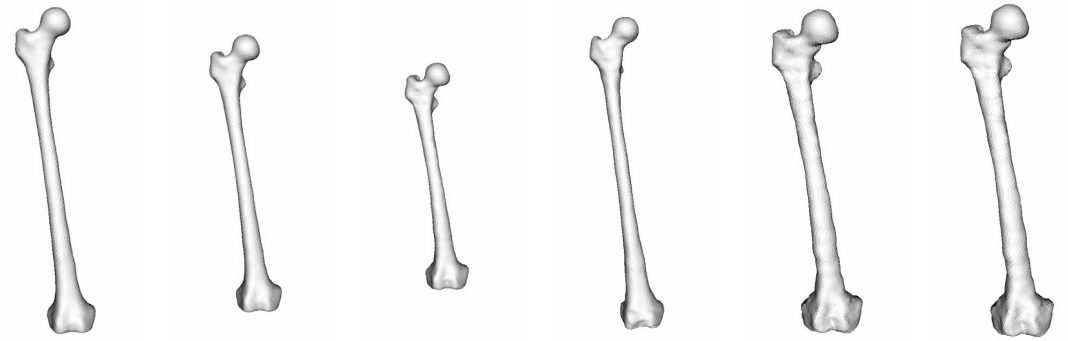
$$\forall l \leq k, \forall i \in [1, N^l], \lambda_i^k \geq \lambda_i^l \geq \lambda_{i+N^k-N^l}^k \tag{4.5}$$

This last remark derives from the observation that given our multi-resolution scheme, a covariance matrix Σ_X^{k+1} is composed of nested covariance matrices at lower levels:

$$\Sigma_X^{k+1} = \left(\begin{array}{c|c} \Sigma_X^k & u_X^k \\ \hline u_X^{k\top} & v_X^k \end{array} \right) \tag{4.6}$$

More exactly, the covariance matrices Σ_X^j, $\forall j \leq k$ are leading principal submatrices of Σ_X^{k+1}. Given these remarks on the eigenvalues and eigenvectors, we preferred to compute for each level an SSM in order to preserve its specificity and to give the possibility to use a different align transform at each level.

SSM Locality Indeed, at each mesh resolution level k, a different alignment transform can be used to build the SSM. The alignment transform affects what we call the "*locality*" of the SSM. The locality is an intuitive notion: rigid SSMs capture global shape changes while affine or similarity SSMs better express local variations as depicted in Fig. 4.12. A multi-resolution SSM adopts different mesh and locality levels based on multi-resolution alignment schemes, such as the rigid-affine-affine (RAA) and similarity for all levels (SSS) schemes. RAA means that the SSM at resolution 1 is rigid, at resolution 2 affine, and so on. The appropriate selection and impact of the multi-resolution alignment will be studied in the segmentation context in next Chapter.



(a) $-3.0\sqrt{\lambda_0}$ (b) mean shape (c) $3.0\sqrt{\lambda_0}$ (d) $-3.0\sqrt{\lambda_0}$ (e) mean shape (f) $3.0\sqrt{\lambda_0}$

Figure 4.12: Variation modes of femur shape with different alignment schemes. (a)-(c): first mode of variation of a rigid SSM. This mode clearly captures the length of the femur. (d)-(f): first mode of variation of a similarity SSM. This mode mainly expresses the thickness of the femur.

4.3.2 Robust Statistical Shape Models

4.3.2.1 SSMs for Images with Varying FOV

We have presented in Sec. 3.3.4 the direct relationship between image field of view (FOV) and resolution. It is quite common to have images with different FOV in order to image the structures

of interest with a desired image resolution. In these images, structures of interest are often only partially visible in the image, especially long bones such as the femurs which cover large body areas.

There are two kinds of segmentation results from images with small FOVs (Fig. 4.13):

– *incomplete* shapes, i.e. shapes that only represent the visible parts of the bone in the image.
– *corrupted* shapes, i.e. complete shapes with some points that are not reliable or not part of the image region to be segmented. These unreliable points are hereafter referred to as *erroneous* points.

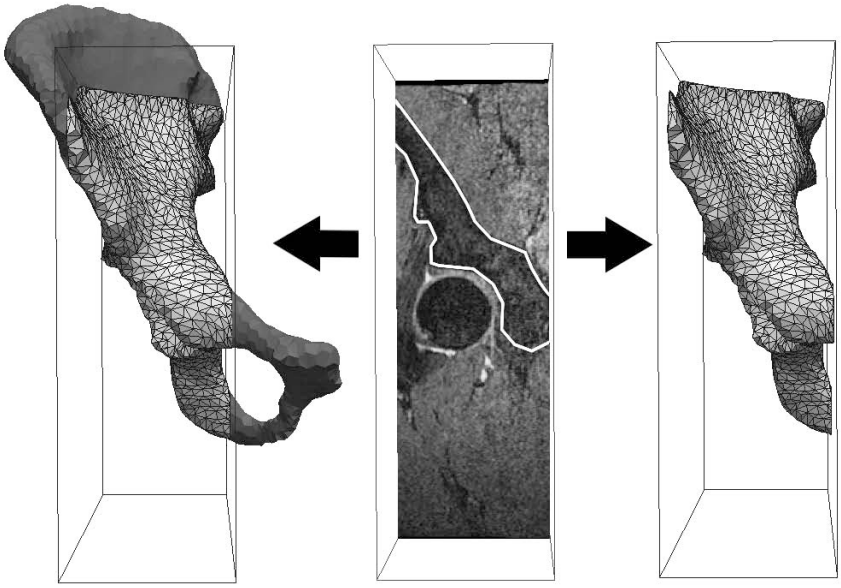


Figure 4.13: Possible shapes resulting from segmentation on images with small FOVs. Center: MRI image with a small FOV and with hip bone segmentation overlaid (white); Left: *corrupted* shape, where only shape vertices in the image FOV (bright wire-framed part) are reliable; Right: *incomplete* shape, where the shape is cropped to only represent reliable parts.

We also previously mentioned in Sec. 2.5.6 the fact that SSMs need to be built with a certain amount of shapes to yield satisfactory results. However, the collection of data in 3D is in practice a problematic and tedious task [HM09], usually due to the lack of sufficient segmented images. Thus it is advantageous to be able to exploit segmentation results from a wide range of images, including those with small FOVs, in order to build and improve the SSMs.

Hence, when we deal with images in which the structures of interest are partially visible, there are two things to consider:

1. How to build an adapted SSM to be used to segment these images.
2. How to re-exploit the segmented results (complete, incomplete or corrupted shapes) from these images.

A common approach to tackle the SSM construction for these images is by creating and using incomplete shapes depending on the image FOV [DLH07, DGZ07]. As reported by Kainmüller *et al.* [KLZH09], this approach is time consuming and often prevents automation, because unseen images which cover a larger part than those of the training shapes cannot be segmented with

the same SSM. It could be argued that it would be sufficient to build one SSM with complete shapes but we just mentioned that there are often too few (complete) training samples and it is hence a real "waste" not to reuse the segmentation results from images with different FOV. Our solution to address these two aspects is to build robust statistical shape models from complete *and* corrupted data as explained in the next Section.

4.3.2.2 Robust Models based on Corrupted and Complete Shapes

Incomplete shapes cannot be directly mixed together because they do not necessarily present identical topology and same number of points. The idea instead is to always use complete shapes with deformable models to segment images with different FOV. However, the resulting shapes are usually corrupted with some erroneous points. Hence, a major challenge is combining these corrupted shapes with complete ones into the same SSM while preserving the SSM's Generality and Specificity, conventional construction approaches being very sensitive to erroneous points (e.g., PCA sensitivity to outliers depicted in Fig. 2.11). We proposed to address this problem as follows.

Robust Alignment Alignment is necessary when applying the generalized Procrustes during SSM construction (Sec. 4.3.1.2). In the case of shapes with erroneous points, a standard alignment is affected since the least square minimization is not robust to outliers. A common way to find the optimal alignment transform $T^\star$ between a current estimate of the mean training shape $\overline{\boldsymbol{X}} = \{\bar{x}_1, \ldots, \bar{x}_N\}$ and a training shape $\boldsymbol{X_j}$ is to adopt a weighted formulation:

$$T^\star = \underset{T}{\operatorname{argmin}} \sum_i w_{ij}^2 \left\| \bar{x}_i - T(x_{ij}) \right\|^2 \tag{4.7}$$

where each training shape $\boldsymbol{X_j}$ is associated with a weight vector $\boldsymbol{w_j} = \{w_{1j}, \ldots, w_{Nj}\}$ that indicates the degree of reliability of each shape point.

As a side remark, this weighted alignment can also be used to regulate a shape $\boldsymbol{X_j^l}$ at a resolution level l with an SSM built from shapes of a higher resolution level k, $l < k$. Indeed, the shape regularization (Sec. 4.3.1.3) uses an alignment step which can also be weighted for our needs. The regularization is done here by creating first an artificial shape $\widetilde{\boldsymbol{X}_j^k}$ with N^k points, where $\widetilde{\boldsymbol{X}_j^k}_{|N^l} = \boldsymbol{X_j^l}$ and a dummy value is set for the remaining points $\widetilde{x_{ij}^k}$, $\forall i \in]N^l, N^k]$. Then, a weighting vector $\widetilde{\boldsymbol{w}_j}$ is built by zeroing the weights $\widetilde{w_{ij}}$, $\forall j \in]N^l, n^k]$ to be used in the weighted alignment. From the resulting regulated shape $\widetilde{\boldsymbol{X}_j^k}$, the vertices of the regulated shape at resolution l are simply updated from the first N^l vertices $\widetilde{x_{ij}^k}$.

Robust PCA The robust PCA is then computed as follows. Let us consider the training set S as the union of non-corrupted $S^\bullet$ and corrupted S° shape sets, where $S^\bullet$ might be empty. Before any robust PCA approach, the training shapes are aligned with a weighted version of the generalized Procrustes alignment. This is achieved by replacing the classic mean and alignment calculations by their respective weighted versions (i.e., $\bar{x}_i = \sum_j w_{ij} x_{ij} / \sum_j w_{ij}$ for mean, and the abovementioned weighted alignment (4.7)). The missing data approach (mdEM) of [SLB07] is then utilized as the

robust PCA method. This method extends the iterative expectation-maximization (EM) algorithm for PCA [Row98, TB99] by allowing it to support binary weights. In the E step, the shape parameters of the training data are estimated based on the current estimation of the principal directions. In the M step, principal directions are estimated by computing a classic PCA on corrected data in which missing information is replaced by reconstructed values. This process thus estimates the "non-corrupted" value of the corrupted points based on reconstructed values. This robust approach was chosen based on the comparison of various robust PCA approaches that will be reported in Sec. 4.5.4.

During both alignment and PCA, a weight w_{ij} is set to 0 if point i is an erroneous point in shape j, and to 1 otherwise. We did not need to use any technique to detect and correct erroneous points as in [RG02] or [BBLS05], since in our case these points were known in advance. In fact, we only flag points out the image FOV as erroneous.

It must be stressed that this robust PCA method usually converges to a local minimum. Tipping and Bishop [TB99] showed that in the case of the EM algorithm the only stable local solution spanned the principal subspace, but this property is not necessarily valid for a weighted EM algorithm. The initialization of the EM algorithm is thus very important. A standard PCA could be used to begin with, but for $|S^\bullet| \geq 0$, we proposed the initialization obtained from a PCA as follows:

1. Build a PCA $P^\bullet$ from $S^\bullet$.
2. Replace corrupted points of shapes in S° with reconstructed points from $P^\bullet$.
3. Rebuild a final PCA from $S^\bullet$ and corrected S°.

4.3.3 Local Shape Variation Modeling by Markov Random Field (MRF)

In this Section, we investigate an alternative approach to model additional local deformations that cannot always be expressed and captured with statistical shape models built from the shape vertices, i.e. with PDM.

Let us consider $\boldsymbol{X} = \{x_1, \ldots, x_M\}$ and $\boldsymbol{Y} = \{y_1, \ldots, y_M\}$, respectively of current and *true* point positions of a shape. Positions $\boldsymbol{X}$ are commonly the result of a procedure which exploited the modeling based on PDM. The objective is now to model the last discrepancy between current and real model. We define $\boldsymbol{Y} = \boldsymbol{X} + \boldsymbol{\delta}$ where $\boldsymbol{\delta} = \{\delta_1, \ldots, \delta_M\}$ represent local deformations. By adapting the idea depicted in [KH98], the local deformation distribution is modeled by a first-order Gauss-Markov random process (See Sec. B.6 for details on MRF):

$$P(\boldsymbol{\delta}) = \frac{1}{Z_m} \exp -\frac{1}{2} \sum_{i=1}^{M} \left[\frac{1}{\eta^2} \sum_{j \in \mathcal{N}(i)} \left\| \delta_i - \delta_j \right\|^2 + \frac{1}{\sigma^2} \left\| \delta_i \right\|^2 \right] \tag{4.8}$$

where $\mathcal{N}(i)$ denotes the index neighborhood of point i and Z_m designates the partition function which is a constant. In case of 2-simplex meshes, the neighborhood size is fixed to 3. Parameters η^2 and σ^2 control the smoothness and the magnitude of the deformations, respectively. The term $\boldsymbol{X}$ is considered as deterministic, $\boldsymbol{Y}$ hence follows a first-order Gauss-Markov process as well. It

is an unknown random process that will be inferred from observed data. The idea behind this model is that local variations are more likely to be expressed based on local quantities, in this case the neighborhood configuration. In Sec. 5.3.3.4 we will illustrate how we can translate this prior into forces to be applied in the segmentation context.

4.4 Appearance Priors

4.4.1 Appearance Feature Construction

Similarly to [GMT10, HMMW07, BMVS02], we chose to exploit local appearance features (Sec. 2.5.1.2) expressed as intensity profiles (IP). From a collection of M segmented shapes $\boldsymbol{X}_j$ with corresponding images I_j, we compute at each vertex x_{ij} an IP $\boldsymbol{q}_{ij}$ with m and p intensity values sampled by tri-linear interpolation in the interior and exterior, respectively (Fig. 4.14). Each sample is spaced by a step s. This gives an IP of length $d = m + p + 1$. In this context of local appearance construction, we clearly understand the importance of good correspondence as non-corresponding vertex positions will produce non-matching IP and hence bad appearance priors.

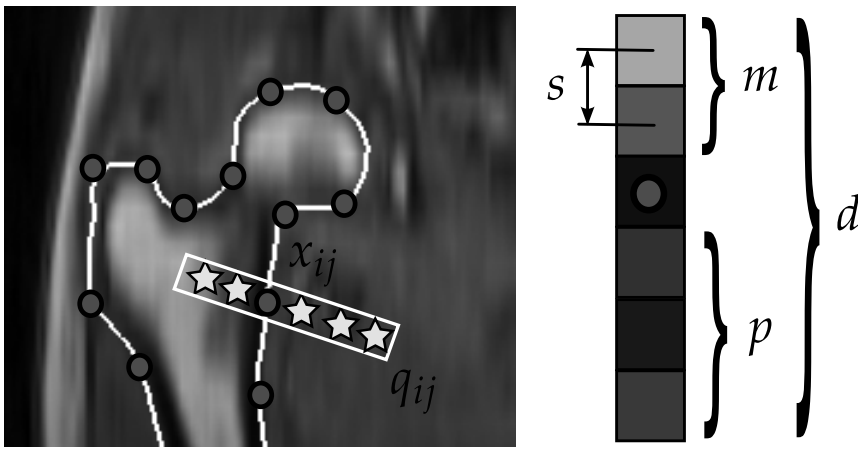


Figure 4.14: Local appearance feature construction. An IP is extracted at each vertex along the normal direction. A series of m and p values are regularly sampled with a step s in the interior and exterior of the mesh, respectively.

4.4.2 Appearance Prior Computation

For a given vertex x_{ij}, the simplest statistic is to take the average $\bar{\boldsymbol{q}}_i = \sum_j q_{ij}/M$. This prior is expected to be very poor for training images with varying intensity. As a result, it has been suggested to also consider the 2nd order statistic by computing the covariance matrix Σ_{qi} of the M IPs $\boldsymbol{q}_{ij}$. The ML estimator (Sec. B.4.1) is commonly used to compute the covariance matrix:

$$\Sigma_{qi} = \frac{1}{M} \sum_j (\boldsymbol{g}_{ij} - \bar{\boldsymbol{q}}_i)^\top (\boldsymbol{g}_{ij} - \bar{\boldsymbol{q}}_i) \tag{4.9}$$

When the profile size $d \ll M$, this estimator is unstable which means that the covariance matrix may no longer be symmetric definite positive. This causes problems when the inverse of the

covariance matrix needs to be computed as the matrix is singular. This condition often happens as it is quite common to have large IPs to robustly capture the local appearance. To tackle this issue, we implemented the shrinkage approach of Schafer *et al.* [SS05] which corrects the covariance matrix so that it is always invertible.

4.4.3 Appearance Prior Usage

Given computed profiles and associated statistics, various strategies exist to exploit these data, which were reviewed in Sec. 2.5.3.6 and Sec. 2.5.5.2. These strategies mostly attempt to find in the neighborhood of the shape IPs which are the most similar to those drawn by the statistics in order to deform the mesh to reach these IPs. For instance, in MRI bone segmentation context, Gilles *et al.* [GMT10] showed that a simple reference IP coupled with the Normalized Cross Correlation (NCC) [HHD+99] similarity measure produces excellent result. In this context, a reference IP can be considered as an average IP. Such a poor IP statistic can be efficient, which is attributed to the fact that the NCC measure is robust to linear intensity changes, and thus, compensates the weakness of the priors.

We decided thus to investigate the use of second order statistics by modeling the IP distribution as a multivariate Gaussian distribution (MGD) and by exploiting the Mahalanobis distance. Given the statistics $(\bar{\boldsymbol{q}}_i, \Sigma_{qi})$ and a test IP $\boldsymbol{u}$, the Mahalanobis distance $d_M(\boldsymbol{u})$ is:

$$d_M(\boldsymbol{u}) = (\boldsymbol{u} - \bar{\boldsymbol{q}}_i)^\top \Sigma_{qi}^{-1} (\boldsymbol{u} - \bar{\boldsymbol{q}}_i) \tag{4.10}$$

Given that the inverse of the covariance matrix is needed to compute this distance, this justifies the use of regularization techniques of the covariance matrix. Cootes *et al.* [CHTH93] proposed an alternative approach to compute the Mahalanobis distance based on PCA performed on the IPs. Given a PCA based on N profiles q_{ij} of a vertex x_{ij} expressed by the principal matrix Φ_i, the m eigenvalues λ_{il} and the mean vector $\bar{\boldsymbol{q}}_i$, the Mahalanobis distance is rewritten as [CHTH93]:

$$d_M(\boldsymbol{u}) = \sum_{l=1}^{m} \frac{b_{il}^2}{\lambda_{il}} \tag{4.11}$$

where $\boldsymbol{b}_i = (b_{i1}, \ldots, b_{im})^\top$ is the model parameter vector of the best fit $\hat{\boldsymbol{u}}$ of $\boldsymbol{u}$ given the PCA model: $\hat{\boldsymbol{u}} = \bar{\boldsymbol{q}}_i + \Phi_i \boldsymbol{b}_i$. A dimensionality reduction can be applied by selecting only the first k modes. In that case, the Mahalanobis distance is approximated as:

$$\begin{array}{rcl} d_M(\boldsymbol{u}) & \approx & \sum_{l=1}^{k} \frac{b_{il}^2}{\lambda_{il}} + 2\frac{R_i^2(\boldsymbol{u})}{\lambda_{ik}} \\ R_i^2(\boldsymbol{u}) & = & (\boldsymbol{u} - \hat{\boldsymbol{u}})^\top (\boldsymbol{u} - \hat{\boldsymbol{u}}) \end{array} \tag{4.12}$$

where $R_i^2(\boldsymbol{u})$ is the squared residual error between the profile $\boldsymbol{u}$ and its best fit $\hat{\boldsymbol{u}}$ computed from the truncated PCA. The advantages of using the PCA-based Mahalanobis distance are:

- The covariance matrix does not need to be regulated, the PCA approach being perfectly capable to handle the condition $d \ll M$.
- By applying the dimensionality reduction, the memory footprint is reduced and some noise in the data is removed.

We decided hence to use the PCA-based Mahalanobis distance by retaining 95% of the total variance.

4.5 Evaluation

In this Section, we will evaluate and illustrate the methodology around the creation of generic models and priors. Regarding the construction of priors, we will only assess the correspondence quality of the shape priors. The performance of local shape variations and appearance priors will be assessed in a segmentation context (See next Chapter). In particular, the choice of the correct appearance priors with associated strategies is dependent on the image type and structures to segment compared to shape priors.

4.5.1 Evaluation Metrics

4.5.1.1 Mesh-to-mesh Comparison

When a mesh-to-mesh comparison is necessary, e.g. when a segmentation result is compared with a ground truth, we need to use relevant comparison metrics. Given two meshes $\boldsymbol{X}$ and $\boldsymbol{Y}$, we used the following metrics proposed by Heimann et al. [Hei08, HvGS+09]:

- *Average asymmetric surface distance* $AASD(\boldsymbol{X}, \boldsymbol{Y})$ computed in mm: $\boldsymbol{X}$ is chosen as the reference mesh, and the distance is expressed as the average of all the point-to-surface distances $d_{\mathcal{S}}(x_i, \boldsymbol{Y})$ between points x_i of $\boldsymbol{X}$ and surface of mesh $\boldsymbol{Y}$. The distance point-to-surface $d_{\mathcal{S}}(x, \boldsymbol{Y})$ denotes the Euclidean distance between point x and the surface of mesh $\boldsymbol{Y}$ which is simply computed by projection of x on the surface of $\boldsymbol{Y}$. As its name indicates, this error is asymmetric since a reference mesh is selected.
- *Average symmetric surface distance (ASSD)* $ASSD(\boldsymbol{X}, \boldsymbol{Y})$ computed in mm: it removes the asymmetry by taking the average of both asymmetric distances:

$$ASSD(\boldsymbol{X}, \boldsymbol{Y}) = 0.5(AASD(\boldsymbol{X}, \boldsymbol{Y}) + AASD(\boldsymbol{Y}, \boldsymbol{X}))$$

  This error is generally chosen to compare two meshes.
- *Average symmetric root mean square symmetric surface distance* $ASRSD(\boldsymbol{X}, \boldsymbol{Y})$ computed in mm: instead of averaging the point-to-surface distances, the root square of the average of all squared point-to-surface distances is instead considered. Compared to the ASSD, this ASRSD additionally measures the error variations over the shape by including the variance in its calculation.
- *Maximum Symmetric Surface Distance* $MSD(\boldsymbol{X}, \boldsymbol{Y})$ also known as the Hausdorff distance [HKR93] expressed in mm: it expresses the maximum point-to-surface distance among all the possible ones. It returns hence the worst distance. It can be used to identify large distances localized in small regions when a small ASSD and a large MSD values are simultaneously obtained.

It is possible to perform voxel-based instead of mesh-based computations by transforming the meshes into labeled images (i.e. rasterization) and by replacing the point-to-surface distance by a voxel-to-voxel distance [HvGS+09]. We used the mesh rasterization approach described in [Gil07] which combines a 3D implementation of the Bresenham line algorithm [Bre65] with

a region growing filling to quickly perform the rasterization. In case of image segmentation, the spacing (i.e. the voxel dimensions) of the labeled image is usually chosen as the same as the image on which the segmentation algorithm was applied. The advantage of using labeled images is that a distance computation is expressed as a closest neighbor retrieval which can be quickly approximated with an ϵ precision by using the approximate nearest neighbor technique [AMN+98]. In case of labeled images, another similarity measure can be computed:

– *Volumetric overlap index* $VOE(\boldsymbol{X}, \boldsymbol{Y})$: it expresses the overlap between the respective labeled images I_X and I_Y as $1 - |I_X \cap I_Y|/|I_X \cup I_Y|$. This volumetric overlap is related to the well-know Dice coefficient D [Dic45] by $VOE = 1 - D/(2 - D)$.

For any of these distances, a zero distance is synonym of identical meshes.

4.5.1.2 Statistical Significance Test

In many situations we want to compare two populations of meshes with respect to the ground truths (GTs). We can thus compute for each individual of a population the mesh-to-mesh distances with respect to these GTs and then compare them between the two populations. This can be useful to assess e.g. which population is the closest to the GT, or in the case of segmentation, which segmentation approach is the most accurate. This comparison is properly done from a statistical perspective by means of hypothesis testing procedures.

We decided to follow the guidelines described by Ruxton [Rux06] which described the appropriate computation of p-values to establish the statistical significance of the results with respect to a chosen significance level α, which was set to 5% in all our experiments. The p-values were computed with two-tailed unequal variance t-tests on raw or ranked data depending on the normality of the data samples. We used the R statistical package for these computations [R D10].

4.5.1.3 Correspondence Metrics

To compare two correspondence methods, we used the correspondence metrics of Davies et al. [Dav02] which are Compactness, Generality and Specificity. These metrics were presented in detail in Sec. 2.5.4.7, and they can be only applied on statistical shape models with same shape topology and number of points. In our case, we want to compare statistical shape models built with and without landmark sliding refinement. Since our shapes present a quasi-uniform vertices distribution, we avoid the pitfall of non-uniform landmark distribution with correspondence measures described by Heimann [Hei08]. Except for the Compactness measure, these metrics require a mesh-to-mesh comparison metric. We chose the ASSD distance.

4.5.2 Generic Model Construction

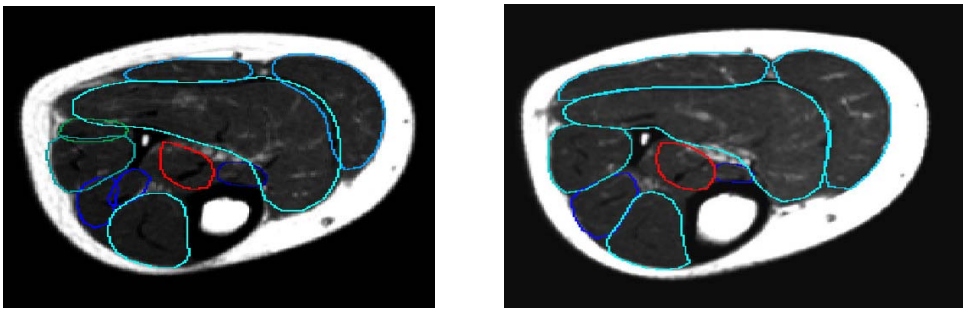
4.5.2.1 Lower Limb Modeling

To illustrate the anatomical modeling, we applied our methodology on datasets produced from a female and healthy subject (age: 24, height: 1m68, weight: 58kg) acquisition. The protocol is an Axial 2D T1 Turbo Spin Echo, TR/TE = 578/18ms, FOV/FA = 40cm/90°, matrix/resolution =

512x512/0.78x0.78mm. To acquire a full lower limb, two different sessions took place and during each of these sessions, three acquisitions were performed. Each acquisition was based on the same protocol, but with varying slice thickness. First session consisted of the hip (thickness: 2mm), thigh (10mm) and knee (4mm); second session acquired the knee (2mm), leg (10mm) and foot (4mm). All acquisitions of a session were merged into a unique volume, and the two session volumes were registered together according to an appropriate knee overlap. The two resulting datasets were created with the following dimensions (resolution): $202 \times 398 \times 595$ ($0.78 \times 0.78 \times 2.46$ mm^3) and $151 \times 213 \times 582$ ($0.68 \times 0.68 \times 2.5$ mm^3) for the first and second datasets, respectively. All acquisitions were performed at the University Hospital of Geneva on a 1.5T MRI device (Philips Medical Systems). The generic modeling has been done in collaboration INRIA who assisted us in the creation of the models before optimization, i.e. in the tasks of manual segmentation and model smoothing. The institutional medical-ethical committee approved the study and the subject gave written informed consent.

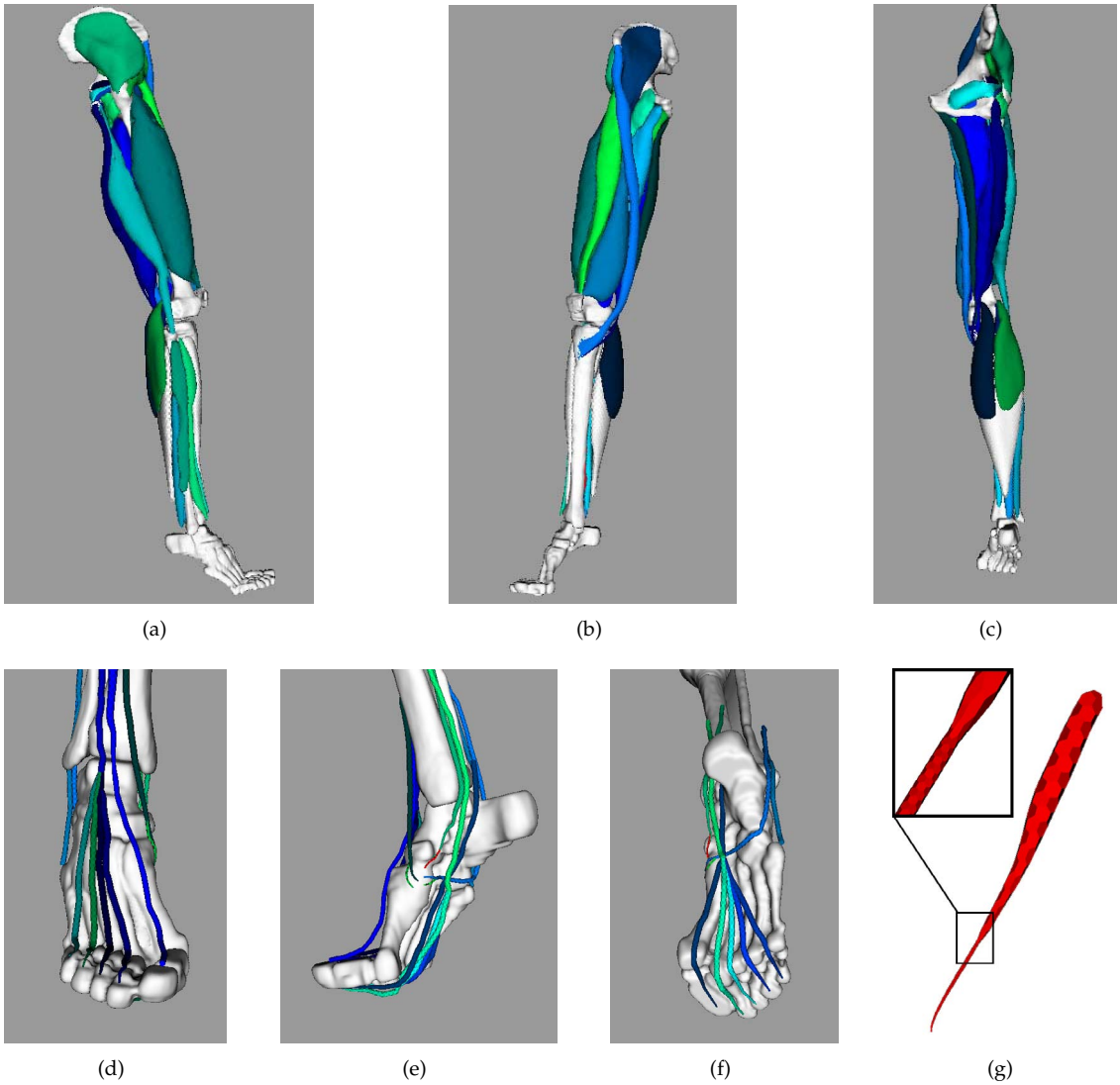
From the first dataset that covers the iliac crests to the knees, the thigh was segmented with Gilles' interactive deformable model approach [GMT10]. In this process, bones, muscles, ligaments and cartilages were reconstructed. Our modeling approach was instead applied on the second dataset which covered the area from the knee to the foot which could not be automatically segmented with the previous approach. Similarly, we reconstructed soft and bony structures. We decided to use both approaches to yield a complex lower limb dataset composed of 33 muscles and 6 bones.

A manual segmentation was first conducted with ITK-SNAP to produce a labeled volumetric image (Sec. 4.2.2.1). Then, we reconstructed triangle meshes from the different labels and produced 2-simplex meshes based on the simplex mesh conversion approach. Subsequently, meshes were smoothed (Sec. 4.2.2.2) followed by re-meshing optimization (Sec. 4.2.2.3) to get models of good quality and of various resolutions. Bones were initially segmented as they were easier to distinguish in the images, then soft tissues were segmented. The lower limb is composed of many muscles in close contact. As a result, the corresponding models were inter-penetrating to each other, particularly after the smoothing procedure (Fig. 4.15(a)). This was corrected with our approach (Sec. 4.2.4) as illustrated in Fig 4.15(b). During this process, we chose rigidity parameters that prevented muscles from diffusing into the bone models. The resulting generic models are visible in Figs. 4.16(a), (b) and (c). Tendons were segmented by running first a tubular tracing method [AB02, PJBB05]. This approach produced different centerlines which were in some occasions incomplete due to the poor image signal-to-noise ratio apparent in some slices of the images. As a consequence, we manually corrected them by interactively placing missing centerline points as explained in Sec. 4.2.2.4. These steps ultimately yielded a very good modeling of the complex tendon network of the foot which is depicted in Fig. 4.16(d)–(f). Tendons were then merged to the corresponding muscles by using the fusion process presented in Sec. 4.2.3. For example, Fig. 4.16(g) illustrates the merging of the tibialis anterior muscle and its tendon. Eventually, muscles and tendons were attached to corresponding bones by using the described procedure in Sec. 4.2.2.5. Attachments were identified on the images whenever possible or estimated from relevant anatomical literature.



(a) (b)

Figure 4.15: Inter-penetration removal of muscle models. (a) Before processing, some muscles models clearly inter-penetrate to each other. (b) Results after correction.



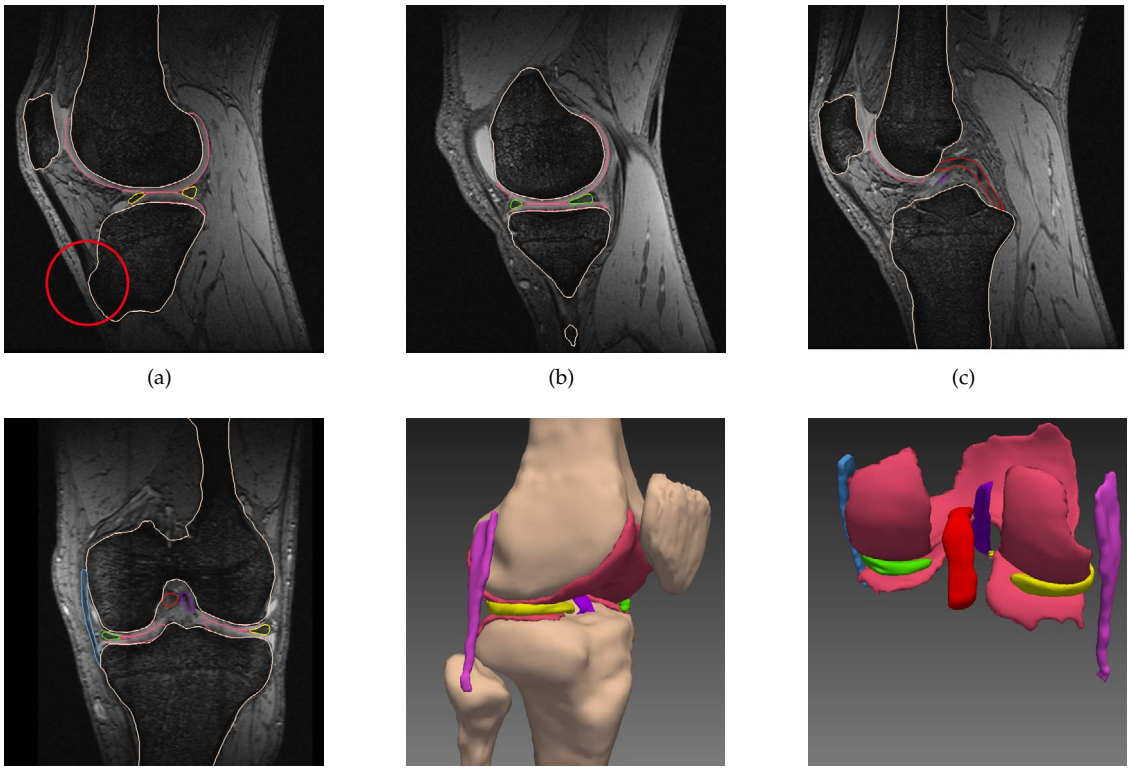
(a) (b) (c)

(d) (e) (f) (g)

Figure 4.16: Lower limb generic model construction. (a)-(c) 3D models of reconstructed lower limb with bones and attached muscles. (d)-(f) Foot tendon models. (g) Tibialis anterior muscle model merged with its tendon, where the closeup view shows the merging area.

4.5.2.2 Knee Modeling

A high resolution MRI image of the knee was sagittaly acquired with dimensions of $512 \times 512 \times 143$ and resolution $0.39 \times 0.39 \times 1$ mm^3. This volumetric image offered better contrast of soft articular structures compared to the previous full lower limb images. By applying the generic model construction techniques, we produced 3D models of the articular soft structures (Fig. 4.17(f)) such as ligaments and cartilages, including meniscii, and bones (Fig. 4.17(e)). Bones were segmented with our automatic segmentation technique depicted in the next Chapter while soft tissues were reconstructed with the semi-manual process. After the smoothing and re-meshing optimization, inter-penetration removal (Sec. 4.2.4) was successfully applied to get smooth models in contact as illustrated in the labeled knee Figs. 4.17(b), (c) and (d) where articular cartilages and meniscii are in contact. A Clinical Senior Lecturer of the academic unit



(a) (b) (c)

(d) (e) (f)

Figure 4.17: Knee generic model construction. (a) This sagittal slice shows an example of tibia segmentation qualified as poor in its anterosuperior aspect (circled area). (b) and (c) sagittal slices in which cartilages and ligaments are notably visible, respectively. (d) Labeled frontal slice of the knee. Notice how structures are in close contact. (e) 3D meshes of the generic knee model. (f) articular soft tissues of the generic knee model. Images rendered with MITK software.

of child health from University of Sheffield, thoroughly analyzed each of the 512 sagittal slices. In each slice, a score was given to each structure using the evaluation scale "poor", "fair" and "good". Furthermore, a series of comments were reported which provided precious insights to the given scores. Table 4.1 reports these scores for each type of structures as percentages. Given a structure X and an evaluation mark S, the percentage quantifies the count of slices, among those

in which X is visible, which were attributed the mark S. Scores are pretty satisfactory as more than 70% of the structure segmentation was denoted in average as good. In particular, cartilages were very well delineated (95.7% of good mark). When both good and fair segmentations were considered, the average percentage for all structures was around 84%.

The percentage of poorly segmented structures was relatively low (e.g., cartilages), except for ligaments. Errors of ligament reconstruction were mostly observed in the attachment areas which were sometimes difficult to identify in the images despite the adapted imaging protocol. In the majority of cases, the lowest scores were attributed to small defaults of the modeled structures as depicted in Fig. 4.17(a). In this Figure, only the anterosuperior of the tibia was poorly segmented. Although the remaining of the tibia delineation was very satisfactory, the segmentation in this slice was reported as poor. As a result, the scoring procedure was quite strict which may explain some low values. Nevertheless, the radiologist expressed an overall positive assessment of the reconstructed structures, which looked very realistic and representative of healthy generic models.

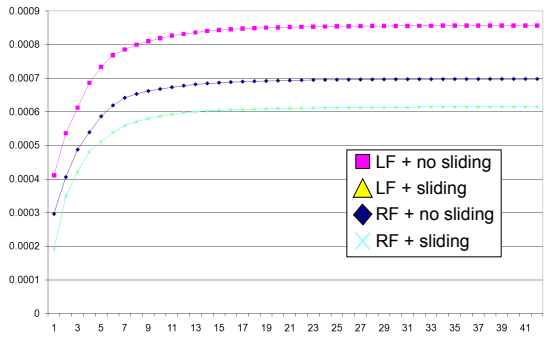
	Bone	Cartilage	Ligament
Poor	11.5%	3.7%	34.7%
Fair	25.7%	0.6%	11.8%
Good	62.6%	95.7%	53.5%

Table 4.1: Radiological evaluation of knee generic model construction. Percentage is related to the average number of slices in which the segmentation was labeled as poor, fair or good depending on the type of structure.

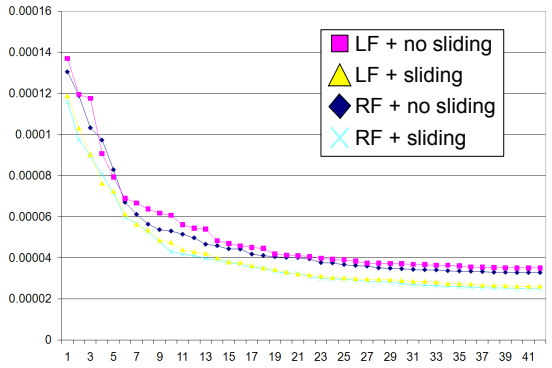
4.5.3 Shape Correspondence

We applied the correspondence method by landmark sliding described in Sec. 4.3.1.2 on a training collection of 43 non-pathological samples, each of which was composed of left and right femurs and hip bones. These samples were created with the training shape construction method which deforms a generic template under the supervision of a user. We recall that this procedure can be seen as a form of mesh-to-volume correspondence method. The aim is to see whether the correspondence can be improved by landmark sliding.

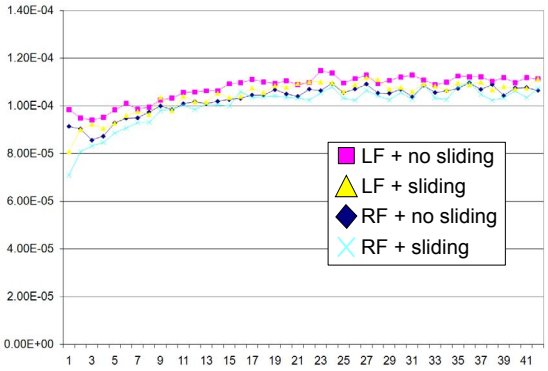
Three resolutions were created for each bone. The lowest resolution of the femur (hip bone) was composed of 514 (814) vertices, whereas the highest resolution has 8K (13K) vertices. Subsequently, shapes were aligned and similarity SSMs were built by retaining 99% of the total variance. One hundred random samples were used to compute the Specificity measure. Figures 4.18(a), 4.18(b) and 4.18(c) report Specificity, Generality and Compactness measures for the right and left femur SSMs, respectively. Note that the ASSD (vertical axis in Figures) is measured after similarity transform, which means that the value does no longer correspond to a physical value since an isotropic scaling was applied during the alignment. The scaling value was approximatively 1E-4 which does not really matter as we are interested in comparing correspondence with and witout landmark sliding for each number of retained components (horizontal axis in Figs. 4.18). These figures show that both left and right femur SSMs reported smaller metrics with the



(a) Compactness of femur SSM



(b) Generality of femur SSM



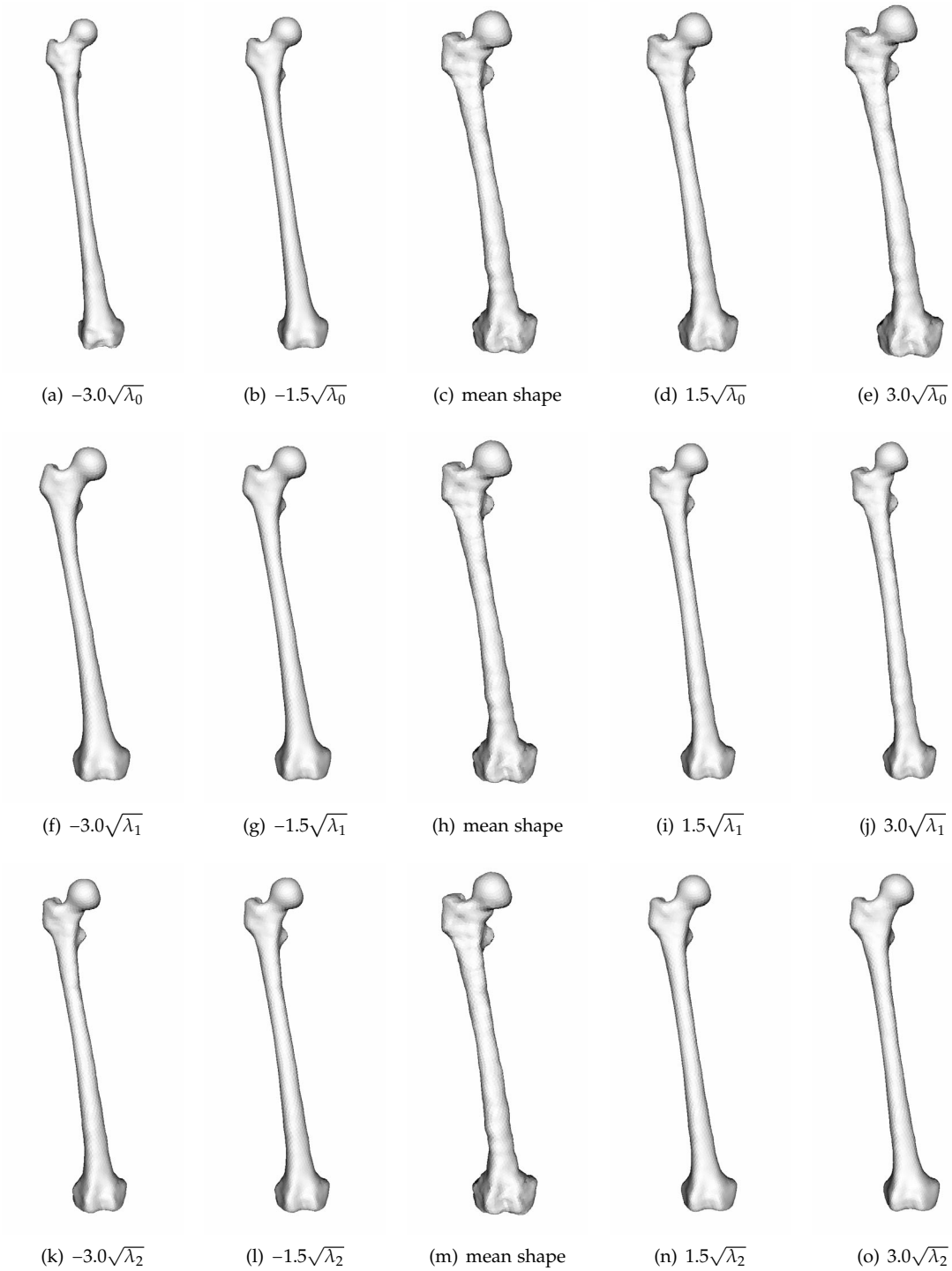
(c) Specificity of femur SSM

Figure 4.18: Correspondence metrics for SSM built with and without the correspondence by landmark sliding, referred to as "sliding" and "no sliding" methods, respectively. RF and LF denote right and left femur, respectively.

use of the correspondence by sliding. A few curve crossings between correspondence with and without sliding are visible in the Specificity (Fig. 4.18(c)). This can be explained by the fact that the number of random samples (100) may not be sufficient to get smooth curves.

In Section 2.5.4.7, we reported various works which expressed reservations about these measures being always appropriate indicators of correspondence quality. We decided thus to observe the modes of variations of the generated SSMs with and without the correspondence by landmark sliding. In Fig. 4.19, we considered the SSMs of right femurs refined with sliding and varied the shape parameters of the first three main principal modes of the SSMs. Each row corresponds to one mode. The generated shapes appear as natural and the modes affect some variations that can be identified quite well. In fact, first mode captures femur thickness, second mode seems to control the angle between the shaft and femoral neck while the third mode encompasses some kind of twisting of the femur around its shaft. The same analysis was performed with the right femur SSM built without applying the landmark sliding (Fig. 4.20). In contrast, the modes of the non-refined correspondence method are less easy to interpret. The first mode does not model the usual scaling change which seems instead to be more distinguishable in the second mode. Third mode also captures some scale variations but also yield shapes less probable (e.g., not very

smooth).

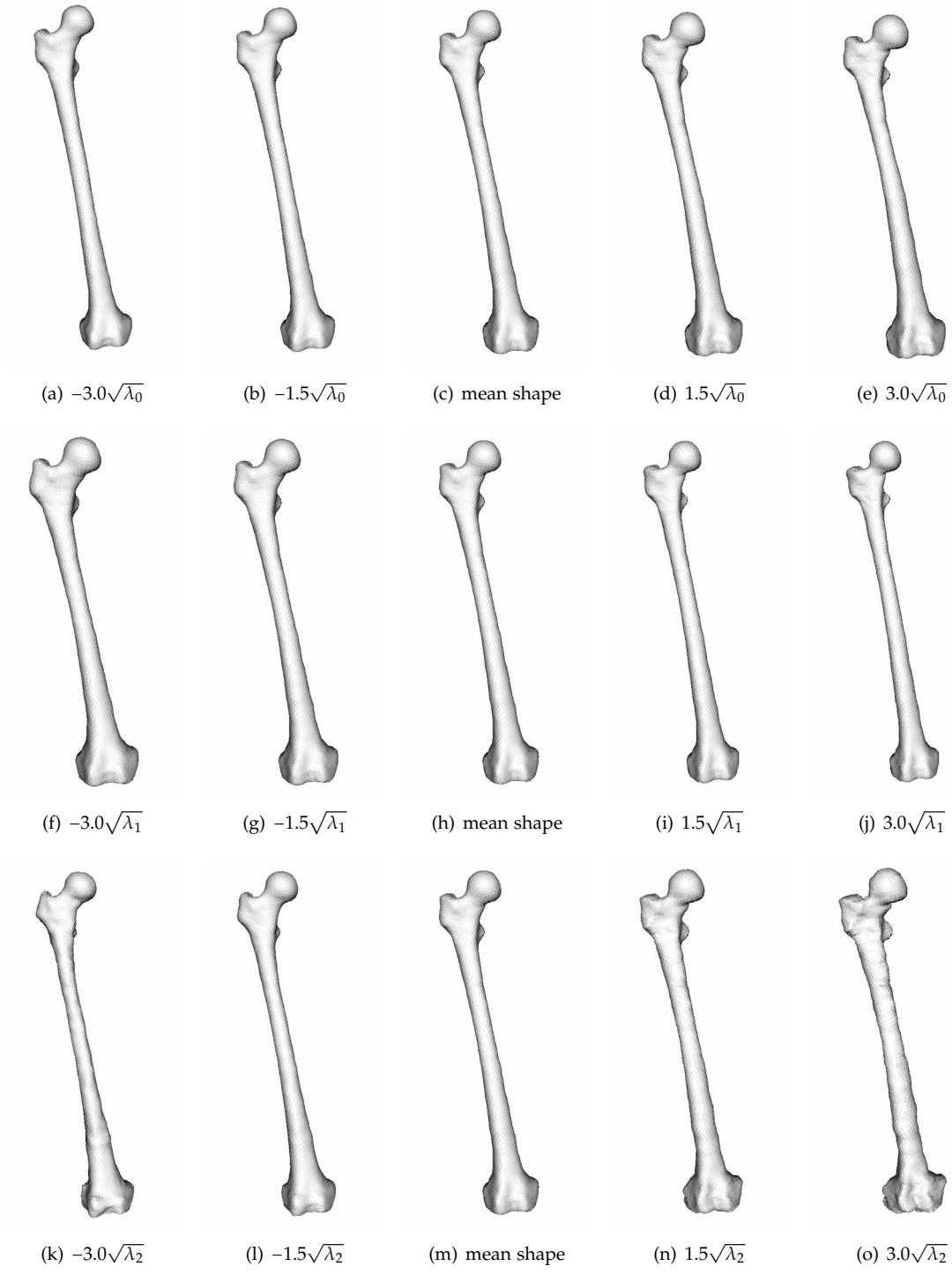


(a) $-3.0\sqrt{\lambda_0}$ (b) $-1.5\sqrt{\lambda_0}$ (c) mean shape (d) $1.5\sqrt{\lambda_0}$ (e) $3.0\sqrt{\lambda_0}$

(f) $-3.0\sqrt{\lambda_1}$ (g) $-1.5\sqrt{\lambda_1}$ (h) mean shape (i) $1.5\sqrt{\lambda_1}$ (j) $3.0\sqrt{\lambda_1}$

(k) $-3.0\sqrt{\lambda_2}$ (l) $-1.5\sqrt{\lambda_2}$ (m) mean shape (n) $1.5\sqrt{\lambda_2}$ (o) $3.0\sqrt{\lambda_2}$

Figure 4.19: Variation modes of right femur SSM with refined correspondence by landmark sliding.



(a) $-3.0\sqrt{\lambda_0}$ (b) $-1.5\sqrt{\lambda_0}$ (c) mean shape (d) $1.5\sqrt{\lambda_0}$ (e) $3.0\sqrt{\lambda_0}$

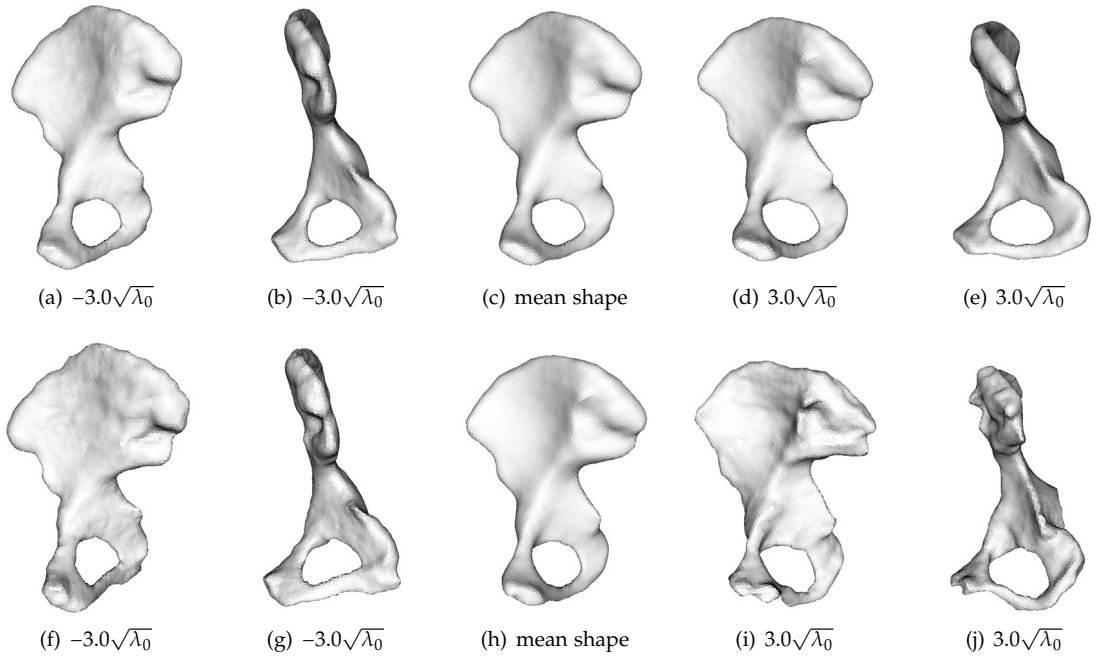
(f) $-3.0\sqrt{\lambda_1}$ (g) $-1.5\sqrt{\lambda_1}$ (h) mean shape (i) $1.5\sqrt{\lambda_1}$ (j) $3.0\sqrt{\lambda_1}$

(k) $-3.0\sqrt{\lambda_2}$ (l) $-1.5\sqrt{\lambda_2}$ (m) mean shape (n) $1.5\sqrt{\lambda_2}$ (o) $3.0\sqrt{\lambda_2}$

Figure 4.20: Variation modes of right femur SSM without refined correspondence by landmark sliding.

Regarding the hip bones, the results of Generality and Specificity were similar to those of the femur. In fact, these measures were lower after correspondence by landmark sliding. The Compactness however did not improve with the correspondence refinement. It is difficult to give a clear explanation to this observation as the shape but also the initial correspondence can be responsible. Nevertheless, the use of the correspondence refinement yielded smoother and more

natural instances as depicted in Fig. 4.21. Without refinement with landmark sliding, extreme instances with $-3.0\sqrt{\lambda_0}$ (Fig. 4.21(f)) and $3.0\sqrt{\lambda_0}$ (Fig. 4.21(i)) presented indeed some deformations that did not appear as genuine and that were not present in the training shapes. Thus, although the Compactness did not improve the generated instances seemed more typical and thus legitimate than those obtained without correspondence refinement.



(a) $-3.0\sqrt{\lambda_0}$ (b) $-3.0\sqrt{\lambda_0}$ (c) mean shape (d) $3.0\sqrt{\lambda_0}$ (e) $3.0\sqrt{\lambda_0}$

(f) $-3.0\sqrt{\lambda_0}$ (g) $-3.0\sqrt{\lambda_0}$ (h) mean shape (i) $3.0\sqrt{\lambda_0}$ (j) $3.0\sqrt{\lambda_0}$

Figure 4.21: First mode of variations of right hip bone SSM. First and second rows show the results of SSMs refined with and without correspondence by landmark sliding, respectively. Subfigures (b), (e), (g) and (j) depict rotated versions of the instances in subfigures (a), (d), (f) and (i), respectively. Variations of the first mode of the SSM without refinement lead to instances that are less natural (i.e. "edgy" aspect).

The correspondence refinement approach was computationally efficient using a standard PC (single core 3.4 GHZ Intel Pentium 4 with 2 GB RAM). Forty-five femurs consisting of 514 landmarks took in average 30 mins. Similarly, hip bone shapes with 814 landmarks required approximatively 90 mins. It must be stressed that the correspondence of all shape resolutions was refined during this process. As a final remark, no special observations were made with respect to the side of the bone, i.e. right or left. Indeed, results were pretty similar regardless of the side, which was expected as healthy bones are mostly symmetric and the construction method of the training shapes is not particularly biased toward a side.

4.5.4 Robust Statistical Shape Model

A series of synthetic experiments was conducted to verify the efficiency of the robust SSM presented in Section 4.3.2.2. Twelve aligned shape pairs of left and right femurs and hip bones formed a test dataset, while 30 training sets were synthetically corrupted with a perturbation scheme based on vertex translation was applied. A random vertex c was selected on some training shapes, and all the vertices present in the ball of radius R centered on c were perturbed by

translation. Two different types of translation vectors were investigated:

– "$\boldsymbol{d}$-perturbation" (Fig. 4.22a): points were translated by a constant noise vector $\boldsymbol{d}$ of magnitude η,
– "$\boldsymbol{n}$-perturbation" (Fig. 4.22b): vertices were translated by a noise vector $\boldsymbol{n}$ perpendicular to the shape with a random magnitude (uniform distribution $\mathcal{U}(0,\eta)$)

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

Figure 4.22: Examples of a) $\boldsymbol{d}$-perturbation and b) $\boldsymbol{n}$-perturbation on a right hip bone; Example of closest shape with a SSM built on c) a PCA with non perturbed shapes (reference result), d) a classic and e) a robust PCA (mdEM) with $\boldsymbol{d}$-perturbed shapes.

Perturbed shapes were subsequently aligned given our weighted alignment procedure described in Sec. 4.3.2.2. For each bone types (femur or hip bone), sides (left or right), and perturbation types, 10 trials were performed (the perturbation location changed each time). In each trial, a standard and robust SSMs were built based on the perturbed shapes. All the test shapes were then compared to their closest match of the standard or robust SSM. A "gold-standard" error was computed from the SSM built without corrupted shapes.

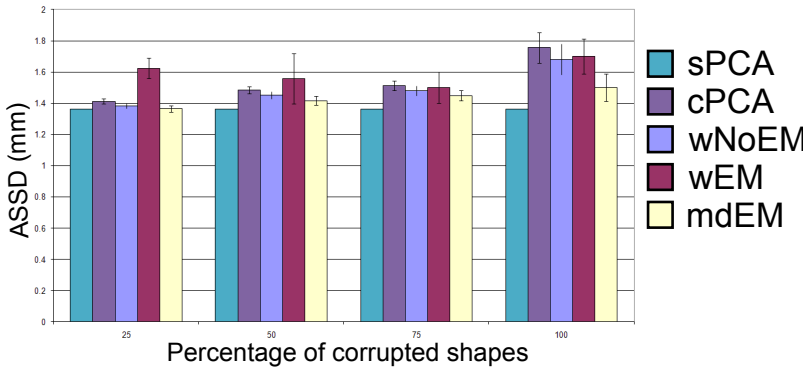
In Sec. 2.5.3.2 we presented various robust PCA approaches aiming at minimizing the weighted least square error h (2.22). In our experiments, weights w_{ij} were known and they were set to 1 for non-corrupted vertices and to 0 otherwise. We chose to investigate three different robust PCA approaches:

– *wEM* approach: method of Skočaj *et al.* [SLB07], which extends the EM-PCA algorithm [Row98, TB99] to support general weights (i.e. not necessarily restricted to 0 and 1 values).
– *mdEM* approach: second method of Skočaj *et al.* [SLB07], which is a specialization of the previous approach for missing data, i.e. data with weights exclusively equal to 0 or 1.
– *wNoEM* approach: method of Kiers [Kie97], which exploits the minimization of a majorizing function. This approach is based on a standard PCA and it is not formulated as an EM problem.

Results are reported in Table 4.2 where $R = 80$ mm, $\eta = 30$ mm and $\vec{d} = (1,0,0)^\top$. With these empirically chosen testing parameters, the corrupted shapes presented strong perturbations (Fig. 4.22(a) and 4.22(b)) in large regions. Results with standard PCA (sPCA) applied on non-corrupted shapes are reported as reference results, while sPCA with corrupted data is coined cPCA and is used to demonstrate the strong sensitivity of PCA to the corrupted data. Four degrees of perturbations were used. For instance, a percentage of 25% meant that only 25% of the shapes underwent the perturbation process. The reported results confirmed the sensitivity of classic PCA to corrupted points (Fig. 4.22(d)), especially when the number of corrupted

% corrup.	PCA	cPCA	wNoEM	wEM	mdEM
25	1.36	1.41±0.02	1.38±0.02	1.62±0.07	1.37±0.02
50	1.36	1.49±0.02	1.45±0.02	1.56±0.16	1.42±0.03
75	1.36	1.51±0.03	1.48±0.02	1.50±0.10	1.45±0.03
100	1.36	1.76±0.10	1.68±0.08	1.70±0.11	1.50±0.04

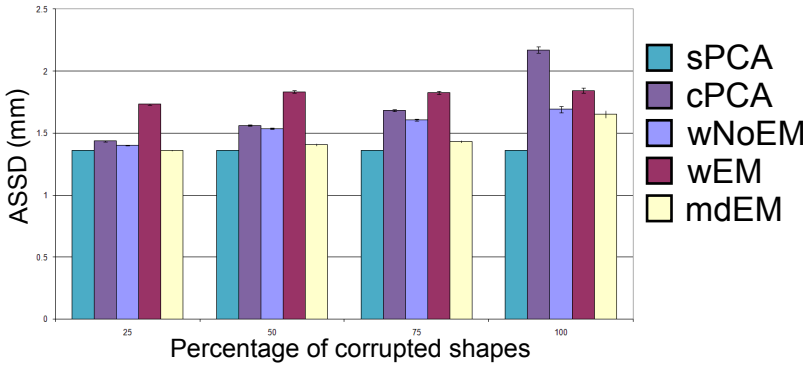
(a) ***d***-perturbation



(b) ***d***-perturbation

% corrup.	PCA	cPCA	wNoEM	wEM	mdEM
25	1.36	1.41±0.02	1.40±0.01	1.73±0.01	1.36±0.02
50	1.36	1.49±0.02	1.54±0.02	1.83±0.01	1.41±0.03
75	1.36	1.51±0.03	1.60±0.02	1.83±0.01	1.43±0.04
100	1.36	1.76±0.10	1.69±0.04	1.84±0.02	1.65±0.05

(c) ***n***-perturbation



(d) ***n***-perturbation

Table 4.2: Comparison of standard PCA vs. various robust PCA implementations (wEM and mEM from [SLB07], and wNoEM from [Kie97]) in the presence of ***d***- and ***n***-perturbations. Standard PCA without corrupted shapes is denoted as sPCA, while cPCA refers to the standard PCA with corrupted shapes. Error metric is ASSD in mm.

shapes was high (cPCA had the highest ASSD with 100% of corrupted shapes). Among all robust approaches, the wEM was the worst regardless of the perturbation type. It only outperformed the cPCA when all shapes were perturbed. mdEM was the approach which was consistently more reliable (Fig. 4.22(e)) with respect to the other two robust implementations, although the difference was not statistically significant for low percentages of corrupted shapes. Nevertheless, the error difference between the mdEM and the sPCA results remained below 10% of the optimal

value regardless of the percentage of corrupted shapes.

The ***n***-perturbations appeared to be more problematic for the cPCA and the robust PCA variants compared to the directional perturbation. It is difficult to clearly explain this discrepancy since perturbations are randomly located over the shapes. Nevertheless, the mdEM approach still produced satisfactory results as the error difference remained below 20% of that of sPCA. Finally, no significant differences were observed depending on the bone types or sides.

4.6 Discussion

4.6.1 Generic Model Construction

The creation of generic models is greatly facilitated by the proposed techniques such as tubular modeling (Sec. 4.2.2.4) and automatic inter-penetration removal (Sec. 4.2.4). However, the issue with the initial segmentation of the images remains. Although semi-automatic segmentation tools are available, such as the interpolation of segmentation results between slices or the use of constraint points, the processing of 500-1000 slices is a tedious task. In fact, a lot of time and effort are spent in checking first the results in a global manner (i.e. analysis of the global shape of the segmented structure) and subsequently in a local fashion by doing a slice-by-slice verification.

The accuracy of the segmentation of course depends of the context. We have seen that the evaluation of our modeling by an expert radiologist was satisfactory. Despite some reconstruction errors observed in the images, 3D models were considered as very good from a modeling viewpoint. The accuracy in building generic models is generally not as critical as in the segmentation of medical images for clinical diagnosis. In the clinical context, results presenting too large deviations from the ground truth can indeed mislead the diagnosis and put patients at risk. Similarly, the creation of appearance priors requires an accurate modeling because the appearance around the border of the models is captured. A poor modeling will inevitably yield non-representative priors which will be useless in a segmentation method.

Instead, what really matters in the creation of generic models is the quality of the models. Models must be smooth, must not inter-penetrate each other and must offer a good trade-off between mesh quality and number of vertices. A mesh quality indicator is the quasi-regularity of the mesh faces and quasi-uniformity of the vertex repartition. Our construction method fulfills these conditions. They are essential to create good shape priors (correspondence) and to have generic models that well behave during a deformable model evolution. Our ultimate goal is the use of deformable generic models coupled with shape priors to segment clinical data.

4.6.2 Shape Correspondence

Based on the experimental results reported in Sec. 4.5.3, we can consider that the correspondence method by landmark sliding is beneficial with respect to the initial mesh-to-volume correspondence method. This was confirmed by the performance metrics and analysis of the modes of

variations. Although the SSMs built with hip bone shapes appeared as less compact after correspondence by landmark sliding, the Specificity and the Generality measures improved. Furthermore, generated instances of hip bones with a refined correspondence appeared as more natural and smoother.

Our metrics presented some curve crossings which could lead to ambiguity [KE06]. In the literature, alternative performance metrics were hence proposed [MDW07, MDW08]. Among them, the testing of the SSM in the segmentation context was considered [VADF+06]. This is a good choice as ultimately we are interested in the performance of our knowledge-based segmentation approach. This investigation will be conducted in the evaluation Section of next Chapter.

The correspondence by landmark sliding approach was in particular appreciated due to its capability to refine a training set of shapes with arbitrary topology. In fact, compared to MDL-based approaches, no constraints are made on the shape genus. Moreover, the resulting refined shapes remain totally unchanged from a topological viewpoint (e.g., same number of points), which is particularly convenient. By using the lowest shape resolution, which offered a good compromise between shape quality and number of vertices, we were able to compute the correspondence across all resolution levels by using projection, TPS transform and resolution change.

The speed of this correspondence method is also an asset. We were able to process 45 right/left femur and hip bone training shapes in approximatively 4 hours in a single core 3.4 Ghz Pentium 4 with 2 GB of RAM. This computation time is considered to be very satisfactory for a shape correspondence approach and it was possible thanks to the use of the lower resolution, which produced a low number of landmarks, and thanks to the efficient solving of the quadratic problem (Appendix Sec. A.4). If we compared it to the fast MDL-based correspondence approach of Heimann [Hei08], we reach similar, if not lower computation times. In fact, Heimann's implementation needed between 30 to 50 hours (depending on the use or not of an optimization process) to process 45 shapes with 2562 landmarks. We ran a similar correspondence procedure on our 43 training shapes with the same number of landmarks in 27.7 hours approximatively. Since the correspondence scales linearly with the number of training shapes, we can extrapolate a total processing time of 29 hours for 45 shapes.

4.6.3 Robust Statistical Shape Model

The synthetic results of Sec. 4.5.4 stressed the need to use robust SSM construction methods. The standard PCA was indeed too sensitive to erroneous vertices. The EM-based robust approach (mdEM) which replaced missing data with reconstructed values was the most effective. Given the strong artificial perturbations that were applied, the robust approach mdEM is expected to behave efficiently with real data in which the amplitude of perturbations tend to be smaller. Real perturbations have good chances to be modeled as a combination of $\boldsymbol{d}-$ and $\boldsymbol{n}-$ perturbations. In fact, $\boldsymbol{d}-$perturbations model the effect that could be observed if vertices outside the image extent were kept fixed while the deformable model was evolving inside the image FOV. This would result in a "shift" effect visible with the $\boldsymbol{d}-$perturbation (Fig. 4.22(a)). The $\boldsymbol{n}-$ perturbation

attempts to mimic erroneous deformations, due to misleading image artifacts or poor appearance priors, which will occur along the vertex normal direction.

The robust PCA error was generally close to the reference PCA error which would suggest that an SSM built from perturbed shapes with a robust PCA could be used for segmentation purpose. However, this experiment illustrated only the good generality aspect of the robust SSM. It is unsure indeed whether the mixing of complete S° and corrupted $S^{\bullet}$ shapes into a robust SSM will preserve its specificity and efficacy in the segmentation context. As a result, this will be assessed in the evaluation Section of next Chapter. In the following of this thesis, the mdEM approach is chosen as the official robust PCA.

Chapter 5

Knowledge-based Deformable Model

5.1 Introduction

In this Chapter, we will present our deformable model algorithm supported by the prior knowledge expressed in the previous Chap. 4. We will address all the important steps presented in our literature review of statistical deformable models (Sec. 2.5.5), which are *initialization* (Sec. 5.2), *evolution* (Sec. 5.3) and *interactive control* (Sec. 5.4). After the methodology, a large series of experiments will be conducted in the evaluation Sec. 5.5 followed by a discussion on the performance and limitations of our knowledge-based deformable model in Sec. 5.6.

While most of the presented techniques are independent of the nature of the images or the anatomical structures, we will describe our algorithm specifically optimized for the segmentation of hip joint bones in MRI images as the main focus of this thesis.

5.2 Initialization

The quality of deformable model-based segmentation is often dictated by the reliability of the initialization. A good initialization should be as automatic as possible and accurate enough so that the deformable model converges to the right result. Usually, an initialization method is crafted based on some prior knowledge of the segmentation task: the image modality and its characteristics (resolution, field of view (FOV)), the nature and pose of the structures of interest, etc. In particular, the FOV has a significant impact on the initialization and it will be carefully analyzed in the evaluation experiments.

We explored and categorized different initialization approaches based on the size of the image FOV with respect to the structures to segment. Our main contribution in this Sec. is the design of an efficient initialization approach for images of the hip joint with small FOV or with unconventional subject limb postures.

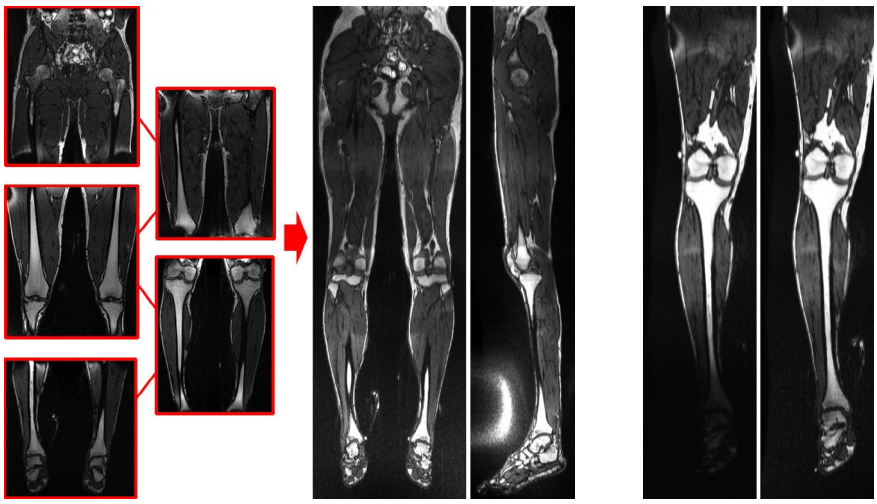
5.2.1 Image Preprocessing

As previously explained in Sec. 3.3.4, an MRI acquisition of large body parts, such as the whole lower limb, cannot be performed in a single scan. Consecutive scans, known as slabs, are necessary to cover its entire length (left of Fig. 5.1(a)). The registration of these slabs is required, in order to generate a uniquely assembled MRI image (right of Fig. 5.1(a)). The registration is computed based on a sufficient overlap among the slabs which is used to compute the alignment between them. When the alignment information was not provided by the scanner, we used a rigid registration based on the manual placement of landmarks.

To cope with differences in intensity distribution that appear between registered slabs, an intensity normalization is necessary. This issue is often due to different MRI intensity ranges between the slabs. Also, the presence of strong artifacts at the image boundaries is likely to affect the intensity distribution, as well as the ubiquitous bias field effect. Our corrective procedure consists in putting artifact intensity values into the background, so that when intensities are normalized

between slabs, these artifacts do not bias the correction. The normalization is performed using an intensity scaling factor. For that, we calculate for each slab the histogram of the main structures of interest (i.e. bones, muscles and fat) and compute the factor so that histograms are similar (Fig. 5.1(b)). Finally, a bias correction approach [SBSG02] is applied on the full image to correct intensity inhomogeneities. When alignment is difficult and prone to errors, our coupled registration-segmentation (Sec. 5.3.4) can be applied as described in the corresponding experiments (Sec. 5.5.5).

The last step consists in smoothing the image to filter for noise. We used an anisotropic diffusion smoothing [PM90] for its good tradeoff between computation time and efficiency (i.e. noise removal with preservation of features such as edges).



(a) (b)

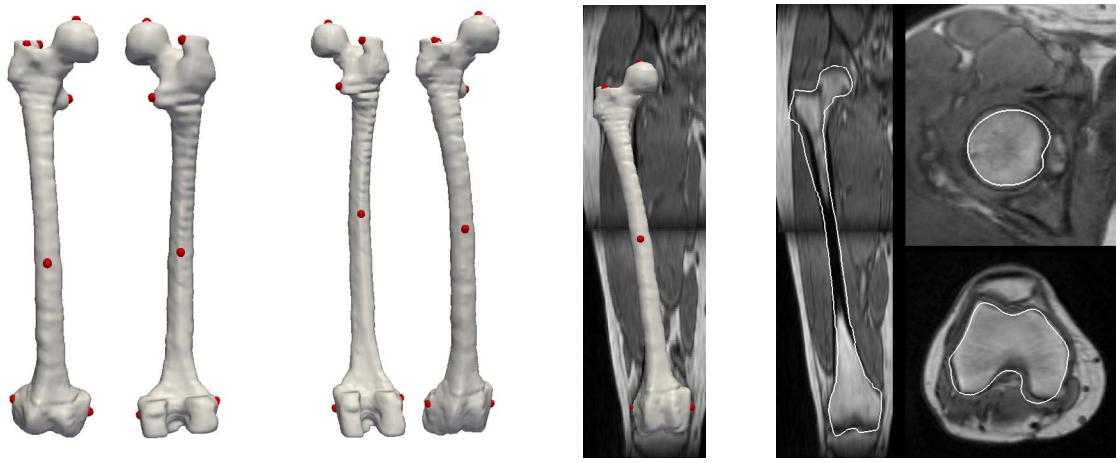
Figure 5.1: Image preprocessing. (a) Different acquisitions, denoted as slabs and covering the full limb, are registered based on a sufficient overlap. (b) Intensity normalization is applied so that structures of interest share the same intensity distribution between the slabs (here the normalization is applied on the left subfigure).

5.2.2 Model Initialization in Image with Large FOV

5.2.2.1 Landmark-based Initialization

A simple yet efficient approach to initialize a model into an image is to use an interpolation technique based on landmarks. We used the approach proposed in [Gil07] which warps a generic model based on source landmarks defined on the model (5.2(a)) and target landmarks selected on the image (Fig. 5.2(c)). The warping exploits an interpolation based on the Thin Plate Spline (TPS) transform. Source landmarks are located at key anatomical locations as shown in Fig. 5.2(a) in order to ensure that they are well preserved in the interpolated model (Fig. 5.2(b)). The placement accuracy is not very demanding and initialized model of sufficient quality can be often obtained as depicted in Fig. 5.2(d). Nevertheless, landmarks should be chosen in such a way that they appear as features easy to identify in the images. These can be located in areas of high

curvature (e.g., extremities of trochanters) or at measurable positions (e.g., landmark located at the anterior middle of the femur shaft).



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

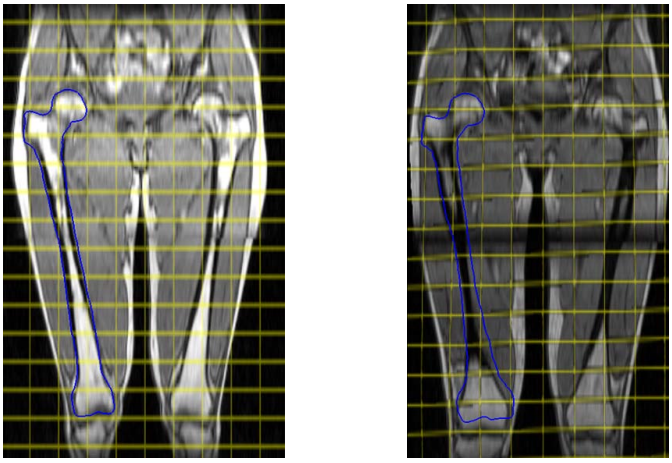
Figure 5.2: Landmark-based initialization. (a) Generic right femur model with corresponding landmarks placed at key anatomical locations such as the greater and lesser trochanters, the top of the femoral head, and the medial and lateral condyles. (b) Resulting interpolated model with TPS transform, where the destination landmarks were placed into the image as shown in (c). Here, the quality of the initialization can be assessed with the mesh overlays visible in the subfigure (d).

5.2.2.2 Atlas-based Registration

Instead on relying on a reference model to be interpolated, we can exploit a reference image. In such initialization, the idea is to achieve an atlas-based registration by estimating the unknown transform T between the reference atlas and the subject image. The resulting T is then used to transform the atlas' segmented meshes into the space of the subject image, achieving the initialization.

It is essential to use a non-rigid transform T to account for inter-subject variability. In this thesis, we used the volumetric registration toolset ElastiX [KSP07]. Initially, an affine alignment was performed followed by a non-rigid B-Splines transformation to result in a registered subject image to the atlas. The affine alignment also needed an initialization, which was simply estimated by aligning the image centers. It was hence assumed that the subject had a similar position (body, lower limbs) in both images.

Both the affine and B-Splines transformations were optimized with standard gradient descent optimizer and a multi-resolution strategy with 3 resolution levels was used to avoid the local minima, where 100 iterations were applied at each level. For the B-Splines transform, a grid of $16 \times 16 \times 16$ voxels was used. These parameters were derived from experimental trials which assessed the impact of the parameters with respect to the final segmentation results (Sec. 5.5.4). Figure 5.3 depicts this process, where a right femur shape is initialized in the subject image (Fig. 5.3(b)) by performing a non-rigid registration with the atlas image (Fig. 5.3(a)). The resulting deformation is illustrated by a deformed grid.



(a) Atlas image (b) Subject image

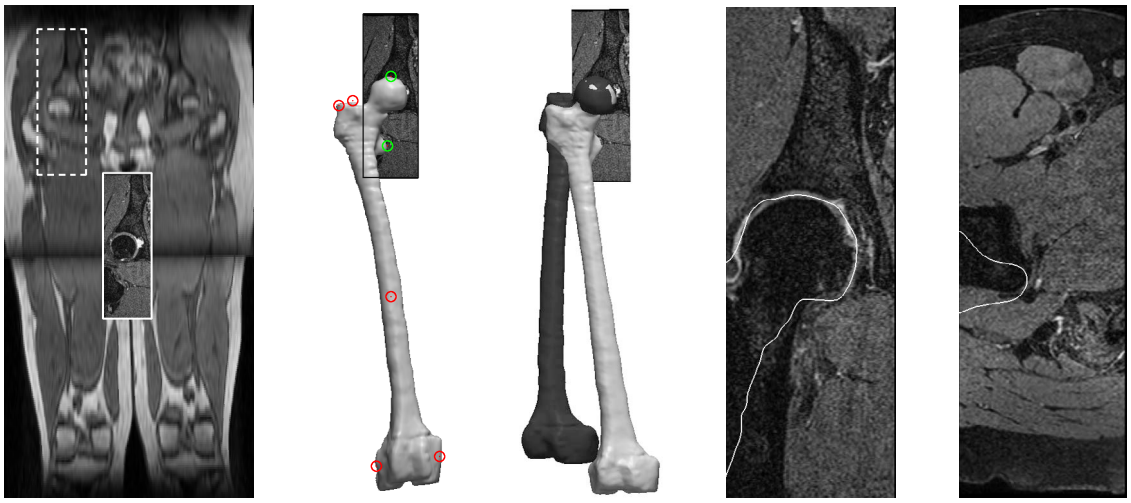
Figure 5.3: Example of initialization by atlas-based registration. A non-rigid registration between (a) the atlas and (b) the subject MRI is performed. The blue overlay represents (a) the atlas and (b) transformed right femur with the computed transformation T. The yellow grid illustrates the non-rigid deformation.

5.2.3 Model Initialization in Image with small FOV

5.2.3.1 The Need of Prior Knowledge and Regularization

Most initialization approaches, and in particular those that were just presented, often rely on a good visibility of the structure to be segmented. As a result, they usually cannot be applied to images with a small FOV. Studholme *et al.* [SHH99] studied for instance the influence of FOV in the case of image registration. If the fixed and moving images present very different FOVs, the registration is likely to fail, especially when simple heuristics are used such as the alignment of the images center (Fig. 5.4(a)). In order to use it with small FOV, we observed that the registration needs images to be aligned relatively close, thus the automatic alignment capability of the registration vanishes since an operator has to provide a good initial alignment. Similarly, landmark-based initialization also faces difficulties. Due to a small image FOV, it may be indeed impossible to locate a sufficient number of landmarks in the image (Fig. 5.4(b)). This can prevent the interpolation from giving satisfactory initializations (Fig. 5.4(c), (d) and (e)). A solution to improve the interpolation would consist in having more markers in the FOV. However, in this example the sphericity of the femoral head makes quite difficult the definition and location of additional markers. Eventually, this would demand to define different landmark set based on the different FOVs. This seriously prevents automation and ease of use as the FOV is not a clinical standard.

Similarly, approaches found in CT segmentation literature may not be applicable. Such approaches include the manual placement [LSHD04] of an initial model which is difficult without adequate visual cues and is also time consuming. Some automatic initialization approaches have higher chances to fail with reduced FOVs, such as the initialization of Seim *et al.* [SKH+08] based on the general Hough transform, which might not be as efficient when the template (in their



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

Figure 5.4: Initialization issues with small FOV. (a) Atlas-based registration is likely to fail when the atlas and the subject images have largely different FOVs, espcially when the initialization of the registration uses the image centers. Here, the dotted rectangle depicts the expected final location of the registered subject image. (b) By using images with small FOV, most landmarks used in a landmark-based initialization [Gil07] are out of the image FOV (red vs. green landmarks). (c) The right femur resulting from the TPS interpolation based on the landmarks present only in the image FOV (dark model) significantly differs from the final segmentation result (bright model). White mesh overlay in (d) and (e) illustrate the poor quality of the initialization.

case the pelvis) is strongly incomplete in the images (as other partially visible structures might produce a higher vote).

Therefore, it is clear that in order to initialize deformable models in images with small FOV, prior knowledge and regularization are necessary. The idea is to devise an initialization which borrows principles from model-to-image registration in which the shape and the pose of the model need to be estimated. By using an SSM, the shape is easily controlled by the shape parameters $\boldsymbol{b}$ of its generative model. The nature of the transform parameters depends on the alignment type of the SSM, i.e. rigid, similarity, affine, etc. Given a generic shape and an associated SSM, optimal shape $\boldsymbol{b}^*$ and transform T^* parameters are hence simultaneously optimized so that the resulting shape is the best fit for the structure depicted in the image. The degree of fitting is computed from a cost function dependent on the shape and transform parameters.

The minimization of the cost function is usually done with global optimization approaches to avoid local minima. For instance, Kervrann and Heitz [KH98] minimized a cost function by simulated annealing in order to localize a structure of interest in the first image of a video sequence. de Bruijne and Nielsen [dBN04] made use of particle filtering to initially fit an SSM in radiographic images. Heimann *et al.* [HMMW07] proposed one of the first approaches of global fitting applied on 3D images in order to initialize a deformable model of the liver. In 3D, the problem is much more challenging as the dimensionality of the space grows and the number of local minima is increased. It is thus necessary to regulate the registration process by constraining the pose parameters. This regularization is really dependent on the context and on the available

prior information.

In our context of bone segmentation of the hip joint, we devised a specific regularization of model-to-image registration based on a constrained global minimization technique. Such approach was inspired by the kinematic behavior of the hip joint. The main steps of this initialization are reported in Fig. 5.5 and detailed in the following Sections.

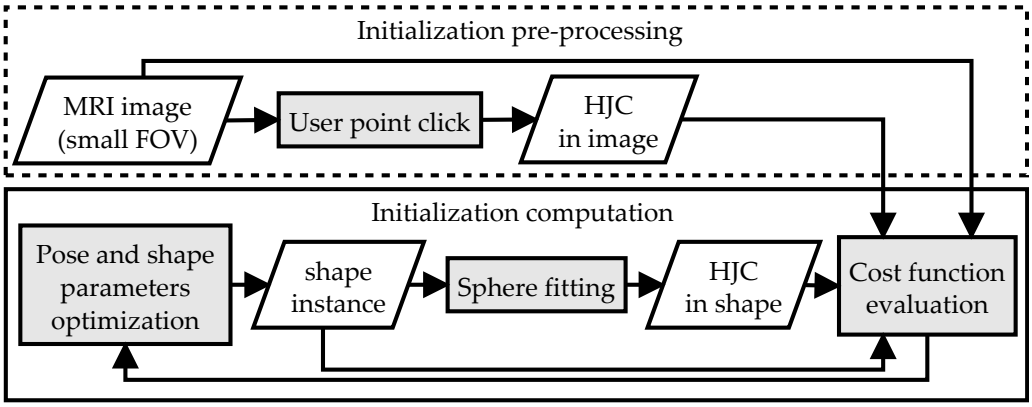


Figure 5.5: Constrained global initialization workflow.

5.2.3.2 Hip Joint Center Estimation

In order to operate, the initialization requires the estimation of the hip joint center. The femoral head is usually largely visible in clinical hip MRI allowing for correct clinical diagnosis of pathologies. Therefore, the hip joint center (HJC), denoted as $\boldsymbol{O_{image}}$ is easily approximated using a simple point and click procedure on the image to select the femoral head center with an error of $\delta = [\delta_x, \delta_y, \delta_z]$ mm^3. During the optimization process, instances of the femur or hip bone are produced by varying their shape parameters. From these instances, the HJC is automatically estimated by spherical least-square fitting [EHW+00]. In this case, the HJC $\boldsymbol{O_{shape}}$ is taken to be the center of the sphere (Fig. 5.6(b)) that approximates the best predefined points corresponding to the femoral head (FH) and the acetabulum (AB) for femur and hip bone instances, respectively (Fig. 5.6(a)). These points are defined as a set of indices that need to be computed only once in one arbitrarily chosen instance of the corresponding SSM (e.g., the mean shape) because there is a point correspondence across generated instances.

5.2.3.3 Cost Function Definition

The cost function C_f expresses the quality of fitting of a current instance generated by transform T and shape parameters $\boldsymbol{b}$. The cost function is computed as the sum of two image energies:

$$C_f = (1-\beta)E^{\text{ip}} + \beta E^{\text{g}} \tag{5.1}$$

Image energies are only computed at points present in the image, i.e. inside the FOV. The first energy $E^{\text{ip}}(y_i)$ computes the similarity between intensity profiles (IPs) p_i of the current shape positions y_i and reference profiles q_i^{ref} as previously presented in Sec. 4.4. The reference

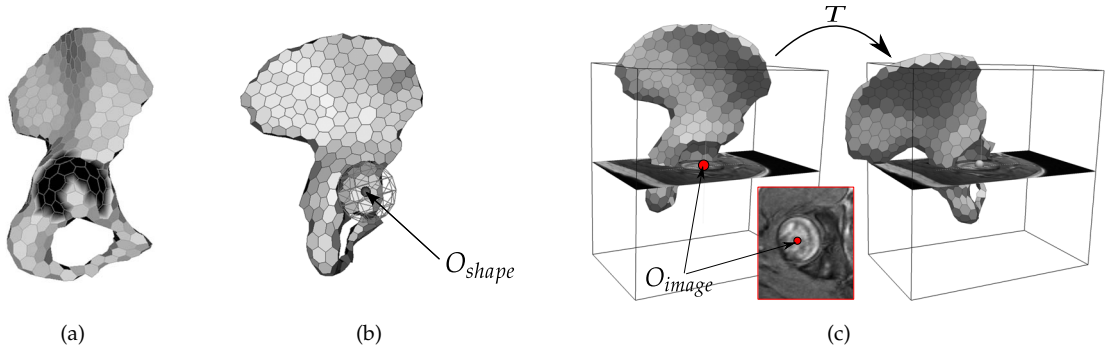


Figure 5.6: Constrained initialization cost function evaluation. Given a hip bone shape instance that is defined by the current shape parameters as in (a) together with the predefined vertex indices of its acetabular region (dark region), the hip joint center (HJC) $\boldsymbol{O_{shape}}$ is estimated from the shape as the center of a sphere fit to these predefined vertices as shown in (b). In (c), a user locates the HJC $\boldsymbol{O_{image}}$ within the image. The two estimated HJC positions $\boldsymbol{O_{shape}}$ and $\boldsymbol{O_{image}}$ are then made coincident which yields the placement of the shape in the image. Subsequently in d), the current transform T, whose rotation center is $\boldsymbol{O_{image}}$, is applied on the shape. The cost function is finally evaluated with the resulting shape.

profiles can be computed from training data by choosing, for example an average IP for each point. Similarly to [GMT10], the NCC is chosen as the similarity measure between extracted and reference profiles because of its robustness to intensity variations. The NCC varies in the interval [-1, 1], where an NCC value of 1 corresponding to identical profiles p_i and q_i:

$$NCC(p,q) = \frac{\sum_j (p_i^j - \bar{p})(q_i^j - \bar{q})}{\sqrt{\sum_j (p_i^j - \bar{p})^2 (q_i^j - \bar{q})^2}} \tag{5.2}$$

The energy E^{ip} to minimize is hence computed from the negated NCC, i.e. $E^{\text{ip}}(y_i) = -NCC(p_i, q_i^{\text{ref}})$. The second image energy E^{g} considers the image gradient direction:

$$E^{\text{g}}(y_i) = \epsilon * \boldsymbol{n_i} \nabla_\gamma I(y_i) \tag{5.3}$$

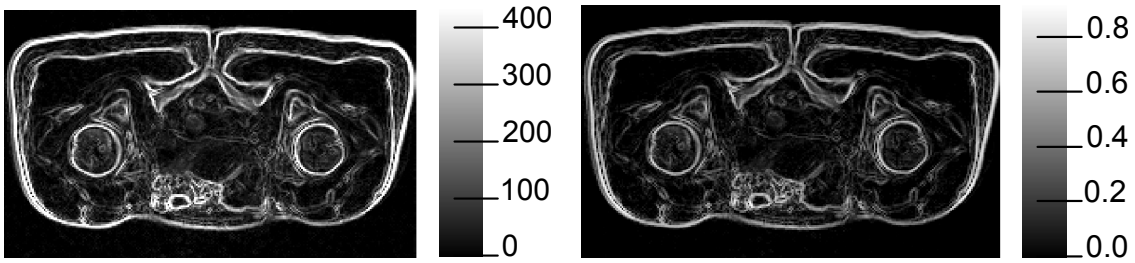
$$\nabla_\gamma I(y) = \frac{\nabla I(y)}{\sqrt{\nabla I(y)^\top \nabla I(y) + \gamma^2}} \tag{5.4}$$

where $\nabla_\gamma I(y_i)$ corresponds to the normalized image gradient at y_i and ϵ equals 1 when the expected gradient points inward with respect to the models surface (-1 when gradient points outward). Gradient normalization is necessary to ensure the image energy is bound within the interval [-1, 1], so that the two image energies can be correctly summed and the weight parameter β can vary in the interval [0, 1]. The normalization is hence of paramount importance. We used the regularized normalized gradient $\nabla_\gamma I$ of Haber and Modersitzki [HM06], in which the γ parameter makes the normalized gradient more robust to smaller gradient values and better differentiable in areas with constant intensity. Moreover, this regularized normalization discards low gradients that are in reality noise in the data. Without it, most of the image would have a regularized amplitude of 1. To automatically select an adequate γ value, we exploited the approach of Ascher *et al.* [AHH06] who used the average gradient over the image domain Ω

regulated by an estimated noise level η:

$$\gamma = \frac{\eta}{|\Omega|} \int_{\Omega} |\nabla I(x)| dx. \tag{5.5}$$

We approximated the noise level η by using the automatic pseudo-residuals technique [GSJS86, CYP+08]. Figure 5.7(a) illustrates the gradient magnitude image of an MRI slice for which the magnitude goes from 0 to > 400. In comparison, Fig. 5.7(b) depicts a regularized normalized gradient magnitude of the same MRI slice. As expected, the gradient magnitude remains in the interval [0, 1] and the normalization correctly preserves the gradients.



(a) Non-normalized gradient magnitude (b) Regularized and normalized gradient magnitude

Figure 5.7: Comparison of non-normalized and regularized normalized gradient on an MRI slice.

5.2.3.4 Cost Function Evaluation

The initialization procedure attempts to find the best pair of rigid transform and shape parameters $(T^*, \boldsymbol{b}^*)$ that minimize a cost function $C_f(T, \boldsymbol{b})$. A rigid SSM is chosen since global variations will be conferred to the changing shape and realistic shape will naturally be instantiated. In fact, a similarity SSM and transform may produce aberrant shapes (e.g., an extremely small shape) if the transform parameters are not correctly controlled.

The key aspect of our constrained initialization method is the way the cost function C_f is evaluated. Given search parameters $(T, \boldsymbol{b})$ and the HJC $\boldsymbol{O_{image}}$ located on the image, the following steps are performed (Fig. 5.6):

1. Generate the shape $\boldsymbol{y}$ from $\boldsymbol{b}$ and estimate the HJC $\boldsymbol{O_{shape}}$.
2. Translate the shape by the vector $\boldsymbol{O_{shape}} - \boldsymbol{O_{image}}$.
3. From T, use the rotation components to rotate $\boldsymbol{y}$ with respect to $\boldsymbol{O_{image}}$.
4. From T, use the translation parameters $\boldsymbol{t}$ to translate the rotated $\boldsymbol{y}$.
5. Evaluate C_f with the resulting shape $\boldsymbol{y}'$.

Given $T = (R, \boldsymbol{t})$, steps 2 to 4 are equivalent to applying the rigid transform with translation vector $\boldsymbol{t} + \boldsymbol{O_{image}} - R.\boldsymbol{O_{shape}}$ and rotation R to $\boldsymbol{y}$. In step 4, we allow small translations $\boldsymbol{t}$ in the rigid transform to account for any HJC placement error. This confers greater flexibility when placing the HJC in the image.

The approach described is thus inspired by the hip joint kinematics (ball-and-socket joint), where the femur and hip bone rotate mainly around the HJC. By limiting the rigid transform T essen-

tially to a rotation centered on the HJC the search space is reduced, thereby increasing the chances of convergence. For better efficiency, a low resolution shape is used for the initialization. Higher resolutions of a specific shape are obtained from the initialized shape by progressively increasing the shape resolution. Each shape is regulated with an affine SSM to filter out artifacts caused by the resolution increase.

5.2.3.5 Cost Function Minimization

The differential evolution (DE) algorithm [SP95] was chosen to minimize the cost function. This algorithm combines the concepts of evolutionary algorithms [Hol75] and a self-adaptive mechanism. In DE, a population member u_i^G represents the search parameters T and $\boldsymbol{b}$, which are evolved from the generation G to the next by applying the mutation, crossover and selection operators. The mutation operator will generate new individuals v_i^G based on the linear combination of various population members u_i^G and a chosen mutation strategy. Storn and Price [SP95] proposed 5 different strategies which are all characterized by the use of a weighted differential variation $F(u_k^G - u_l^G)$, where the indices k and l are randomly chosen among the member indices of the current population. F is a parameter which scales the differential variation. The crossover operator will choose whose values from the mutant individual v_i^G will be preserved in the individual u_i^{G+1} of the next generation $G+1$. This random selection is controlled by a crossover factor C. Eventually, the selection operator accepts this new individual u_i^{G+1} for the next generation only if its cost function is strictly smaller than the original individual u_i^G, i.e. $C_f(u_i^{G+1}) < C_f(u_i^G)$. For this reason, DE is considered as a greedy algorithm.

DE is a global stochastic optimization approach which does not require any initialization. Indeed, parameters of the individuals in the first generation are simply uniformly computed within the variation intervals. Furthermore, DE is appreciated for its simplicity and by the fact that few parameters need to be defined, namely F, C, and the numbers of individuals and generations. Moreover, the variations of the individual parameters, in our case the shape and transform parameters, are easily controlled by using constraints. When parameters independently vary in respective intervals, out-of-bounds parameters can be clamped to the interval extrema or replaced by valid random values [Lam02]. When some parameters must satisfy a constraint, the violation approach can be used [BM06]. For instance, if we choose to encode the rotation parameters of the transform T as a unit quaternion q, we may use for any generated individual the constraint $\|q\| \leq 1$.

In our specific situation, we chose to use Euler angles and a translation vector to encode T. In practice, we bound the rotation angles to the full rotation space $[0,\ 2\pi]$, the translation is limited to the HJC error interval $[-\delta,\ +\delta]$ (Section 5.2.3.2) and shape parameters are clamped according to the "clamping" constraints (Sec. 4.3.1.3) to preserve the Specificity of the SSM. To improve the robustness of the optimization, the angle parameters can be restricted to a smaller interval when some prior knowledge is available (e.g., subject leg is in neutral position).

5.3 Evolution

5.3.1 Dynamic Evolution

From the initialized models, a deformable model is used to refine the segmentation results. The core of our deformable model evolution is based on the approach of Gilles [Gil07, GMT10]. This approach considers each mesh vertex as a lumped mass particle with parameters of position, velocity and acceleration. The evolution is a process which iterates over the following steps:

- **Force computation**: internal and external forces are computed based on particle state and external information (e.g., image).
- **Numerical integration**: forces are applied to these particles and the resulting discrete Newtonian differential equations are solved by an Euler implicit numerical scheme.
- **Mesh update**: mesh vertices are updated and mesh attributes such as normals or simplex parameters are computed.

During the evolution, the various resolutions of the meshes are used in a multi-resolution approach, which increases the mesh resolution after a predefined number of iterations. The number of iterations for each level was chosen based on experiments that determined how many iterations were necessary on average to reduce the vertex displacement of a mesh between two iterations below a small threshold (e.g., 0.5 mm). This yields what we refer to as the "iteration schedule" which provides the number of iterations per resolution level.

Furthermore, models are simultaneously evolved (e.g., femur and hip joint together) and interpenetrations among the models are prevented with collision detection and response techniques based on k-discrete oriented polytopes [MKE03] and distance maps for deformable and static interfaces, respectively. The collision response can be translated as force, position or speed alterations. A damping force inversely proportional to the particle velocity is also considered to avoid potential oscillations.

In this thesis, any force $\boldsymbol{f_i}$ is expressed as a Hookean spring force, which is aimed at moving a point x_i towards a target point $\tilde{x}_i$, i.e. $\boldsymbol{f_i} = \alpha(\tilde{x}_i - x_i)$. Here, the force weight α acts as a stiffness parameter. In order to be used with the implicit Euler scheme, these forces need to be derived with respect to the particle position. In [Gil07, GMT10] additional details on the model evolution are given, such as approximation methods to compute force derivatives for isotropic or anisotropic forces. Generally, internal forces (Sec. 5.3.3) are isotropic, whereas external forces (Sec. 5.3.2) are anisotropic. In the following, we will present the different forces used in our knowledge-based deformable model, by focusing on our contributions.

5.3.2 External Forces

5.3.2.1 Image Forces

External image forces attract the shape towards the desired boundaries by using image features. Given the point position x_i, the target position $\tilde{x}_i$ is sought along the normal direction n_i, such

that an image energy $E(y_i^j)$ is minimized. In practice, the energy is computed at $L = 2D + 1$ positions y_i^j, regularly spaced by a step S, symmetrically with respect to x_i (i.e. D times inside and outside the mesh). Parameters D and S are named search depth and search step, respectively. Depending on the image features and chosen strategies, different image energies with corresponding image forces can be devised.

Gradient Direction Force The gradient energy E^{g} has been already presented in a regularized and normalized form (Eq. (5.4)) in Sec. 5.2.3.3 to allow the correct summation of energy terms. A regularized and normalized gradient force $\boldsymbol{f}^{\mathrm{ng}} = \alpha^{\mathrm{ng}}(\tilde{x} - x)$ is thus simply derived from this energy. In the context of force computation, the regularized normalization can be removed as the edge strength is also a strong indicator of the presence of the structure boundaries, yielding a non-normalized gradient force $\boldsymbol{f}^{\mathrm{g}} = \alpha^{\mathrm{g}}(\tilde{x} - x)$. In both cases, the gradient direction is maintained as a force only based on gradient magnitude will not be able to distinguish the right interface between tissues (e.g., bone boundary in our MRI images is darker than surrounding muscular tissue, i.e. gradients point outward like our mesh normals).

Intensity Profile Force Based on intensity profiles (IP, Sec. 4.4), we consider two types of image energies. The first energy E^{ip} denotes the energy computed with the NCC similarity measure, recalled in Sec. 5.2.3.3. For each point x_i, this energy compares a reference profile q_i^{ref} with L candidate profiles p_i^j, $j \in [1, L]$ based on the NCC measure. The NCC has the nice property to be invariant under intensity affine changes, which confers some robustness against variations in the local appearance of structures across individuals and protocols. We chose this similarity measure as it outperformed various common measures in the context of MRI musculoskeletal segmentation [Gil07].

The second energy E^{mgd} is based on the multivariate Gaussian distribution (MGD) formulation which exploits the Mahalanobis distance as described in Sec. 4.4.3. In particular, the PCA-based Mahalanobis formulation (4.11) is preferred as it reduces the memory footprint and minimizes the influence of the noise in the data. We denote the forces based on IP and MGD as $\boldsymbol{f}^{\mathrm{ip}} = \alpha^{\mathrm{ip}}(\tilde{x} - x)$ and $\boldsymbol{f}^{\mathrm{mgd}} = \alpha^{\mathrm{mgd}}(\tilde{x} - x)$, respectively.

5.3.2.2 Constraint Point-based Force

The segmentation evolution can be controlled by using what is referred to as constraint points (CP) [Gil07]. Three types of CPs are available:

1. Internal points: the model evolves to include the points in its interior.
2. External points: the model evolves to exclude the points from its interior.
3. Frontier points: the model evolves to have the points lying exactly on its surface.

In practice, constraint points attract or repel meshes by creating forces on some vertices. An example of an internal CP P is depicted in Fig. 5.8(a), where the closest face of the mesh, here represented by two vertices P_0 and P_1, is attracted under the action of two forces $\boldsymbol{f}_{\mathbf{1}}^{\mathrm{cp}}$ and $\boldsymbol{f}_{\mathbf{2}}^{\mathrm{cp}}$ applied on P_0 and P_1, respectively. Forces become inactive when the CP is in the interior of the mesh (Fig. 5.8(b)).

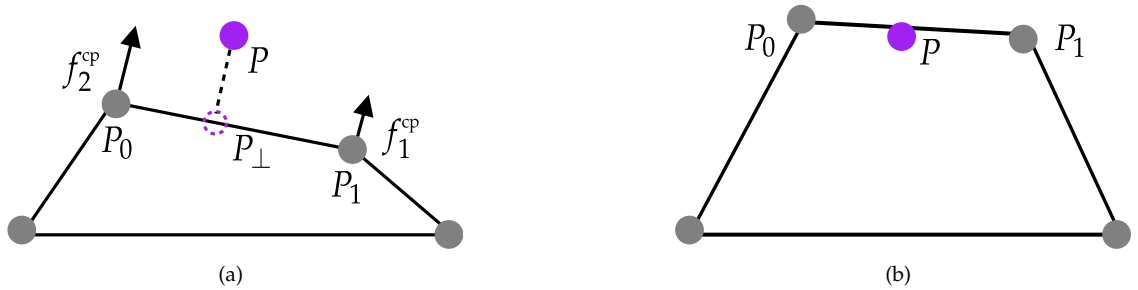


Figure 5.8: Illustration of the forces based on an internal constraint point (CP) P. (a) The closest face (P_0, P_1) to the CP P is attracted by creating 2 forces f_1^{cp} and f_2^{cp} on P_0 and P_1 respectively, whose calculation depends on P and its projection $P_\perp$ on the face. (b) End of evolution in which P is located inside the model.

5.3.2.3 Symmetric Force

Another type of prior knowledge that can be exploited in the segmentation context is the symmetry property. Indeed, a structure of interest can have a point, an axis or a plane of symmetry which can be used to regulate its deformation. For instance, the pelvis is composed of a right and left hip bones which are mostly symmetric with respect to a sagittal plane. This symmetry is commonly observed in healthy subjects.

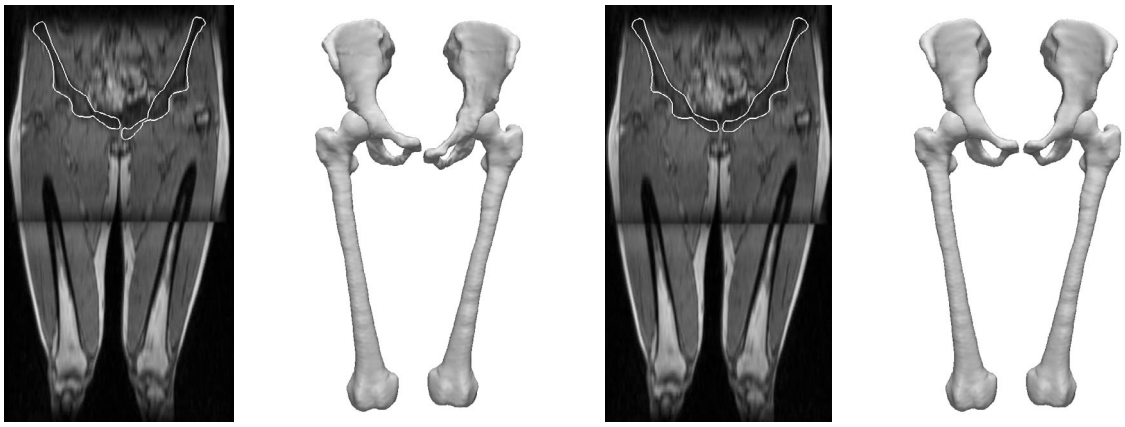
The use of two independent models to segment each hip bone generally confers a greater flexibility in the segmentation since each model can evolve freely. In fact, consider the situation where very narrow pelvic models are used as training shapes to built an SSM. In the case where a subject presents a wide pelvis, the segmentation might not be able to correctly segment the bones due to over-constraining shape priors which prevent the hip bones to separate too much from each other. In this situation, two independent deformable models of hip bones with associated SSMs would have been more adapted.

In most of our experiments, segmented hip bones appeared as symmetric at the end of the segmentation. However, in some cases asymmetry was observed as depicted in Fig. 5.9(a) and (b). The left hip bone is clearly erroneous, especially at the level of the pubic arc. Strong image artifacts in this image area could be at the origin of this deviation from the correct result.

To tackle this issue, we devised a force which enforces a symmetry between the two structures. This force has to regulate the evolution of each shape until a symmetry is reached with respect to a symmetry plane. The two shapes $S_a = \{P_1^a, \ldots, P_n^a\}$ and $S_b = \{P_1^b, \ldots, P_n^b\}$ must be in point correspondence and have the same of number of points n. These requirements are fulfilled by our hip bone shapes. Given a symmetry plane $\mathcal{P}$ with normal $\boldsymbol{n}$, the symmetry property implies two conditions:

1. the middle point C_i of the segment $[P_i^a P_i^b]$ lies on the plane $\mathcal{P}$.
2. vector $\boldsymbol{P_i^a P_i^b}$ is parallel to n.

To satisfy these conditions, we propose the following procedure to compute target positions $\widetilde{P_i^a}$ and $\widetilde{P_i^b}$:



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

Figure 5.9: Symmetric force illustration. Subfigures (a) and (b) illustrate segmented hip bones as mesh overlays and 3d models, respectively, that are not symmetric. Based on our force which enforces symmetry, the segmentation results can be corrected for this subject as shown in (c) and (d).

1. the plane $\mathcal{P}$ is estimated by performing a PCA with all current middle points $0.5(P_i^a + P_i^b)$. The principal direction with smaller variance is chosen as the normal $\boldsymbol{n}$ to the plane and an arbitrary point on the plane P_m is selected (Fig. 5.10(a)).
2. For each pair of points $(P_i^a,\ P_i^b)$, the middle point C_i is projected on the plane as $C_i^\perp$, and new points $P_i^{\star a}$ and $P_i^{\star b}$ are computed as the projections of P_i^a and P_i^b on the line passing through C_i and directed by n, respectively (Fig. 5.10(b)). These points are simply computed with the following formulas:

$$C_i^\perp = C_i + < \boldsymbol{C_i P_m} | \boldsymbol{n} > \boldsymbol{n} \tag{5.6}$$

$$P_i^{\star x} = C_i + < \boldsymbol{C_i P_i^x} | \boldsymbol{n} > \boldsymbol{n},\ \forall x \in \{a, b\} \tag{5.7}$$

3. Target points $\widetilde{P_i^a}$ and $\widetilde{P_i^b}$ are eventually computed as (Fig. 5.10(c)):

$$\widetilde{P_i^x} = P_i^x + \boldsymbol{P_i^x P^{\star x}} + \boldsymbol{C_i C_i^\perp},\ \forall x \in \{a, b\} \tag{5.8}$$

As shown in Fig. 5.10(c), the final points do satisfy the abovementioned symmetry conditions. A corrected result based on this symmetric force is visible in Figs. 5.9(c) and (d), where the left hip bone is now clearly symmetric and the segmentation is of satisfactory quality.

5.3.3 Internal Forces

Internal forces regulate the model evolution by considering smoothing and shape priors.

5.3.3.1 Smoothing Force

We used the smoothing force proposed by Gilles *et al.* [GMT10]. This force $\boldsymbol{f}^{\text{smo}}$ is based on the smoothing equation (4.1) presented in Sec. 4.2.2.2 which estimates the smoothed target position of a vertex. This smoothing is based on an improved version of the Laplacian smoothing designed to minimize the recurrent shrinking effect.

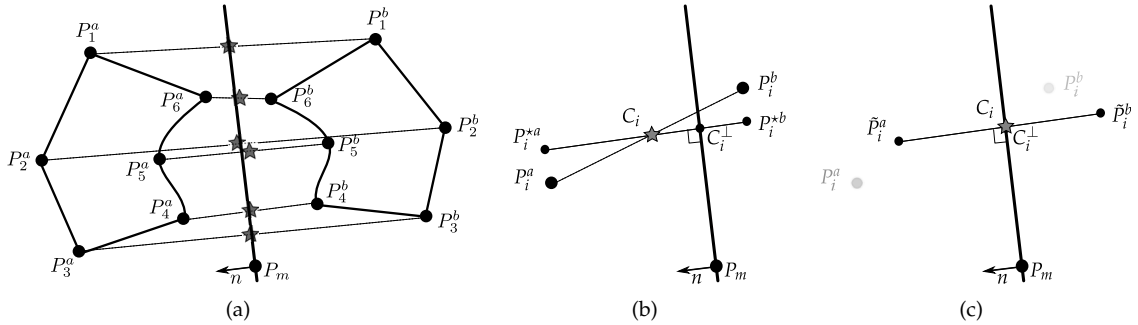


Figure 5.10: Symmetric force computation. (a) from two objects in point correspondence, a plane of symmetry is approximated from the pairs of points $(P_i^a,\ P_i^b)$. (b) For each pair of points, intermediate point positions are used to ultimately compute new target positions as shown in (c).

5.3.3.2 Shape Memory Force

Based on the simplex parameters detailed in Appendix Sec. A.3.3, a force f^{smem} referred to as "shape memory" force [Del99, GMT10] is devised. It follows the concept of strain energy, which enforces deformed structures to recover pre-defined configurations, encoded in our case by pre-defined simplex parameters. Given a triplet of reference simplex parameters associated with each shape vertex x_i, a target position $\tilde{x}_i$ is simply computed by applying the generative equation (See Fig. A.4) based on local parameters, which are expressed from the current local configuration of the shape. A reference simplex mesh is derived from a generic shape such as the mean shape of a SSM.

5.3.3.3 PDM-based Force

Our shape prior regulates the current shape $\boldsymbol{x}$ by finding its closest representation $\tilde{\boldsymbol{x}}$ given its associated SSM (Sec. 4.3.1.3) expressed as a point distribution model (PDM). This yields a force on each point x_i: $f^{\text{ssm}} = \alpha^{\text{ssm}}(\tilde{x}_i - x_i)$. Scaling constraints are chosen for the shape parameter constraints (Sec. 4.3.1.3). During the shape regularization, the pose of the shape expressed as a transform is estimated with the same alignment procedure found in the generalized Procrustes (Sec. 4.3.1.2). Since models are potentially evolving in images with small FOV, chances are that some model vertices will be out of the image FOV. To account for these points outside the FOV of the image which are not reliable, it is necessary to also weight the alignment as explained in Section 4.3.2.2. Otherwise, erroneous vertices will corrupt the alignment and ultimately affect the computation of all the PDM-based forces as illustrated in Fig. 5.11. This example depicts a shape with some erroneous points out of the image FOV (Fig. 5.11(a)). To recover a plausible shape which matches as closely as possible the non-corrupted part of the shape located in the image extents (black rectangle) a regularization was done with an SSM. Figure 5.11(c) clearly illustrates the effect of a non-robust alignment in the regularization process, while Fig. 5.11(b) depicts a robust approach.

At each mesh resolution level, a different alignment transform type can be chosen to affect the "locality" of the regularization as previously explained in Sec. 4.3.1.3. A coarse-to-fine SSM

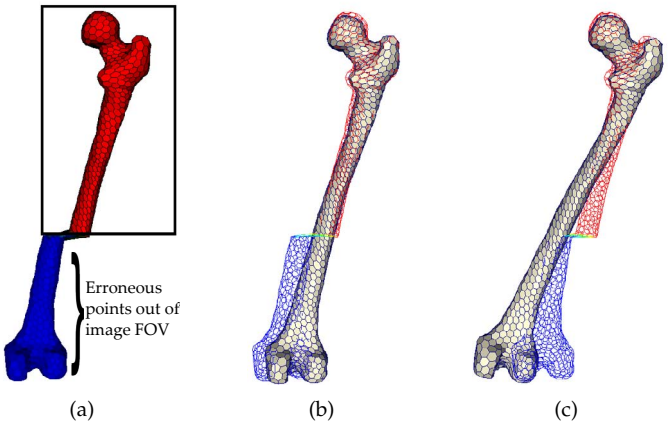


Figure 5.11: Robust alignment for PDM-based force computation. (a) Shape vertices outside the image boundaries (rectangle) are considered to be erroneous points. This shape is regularized by finding its closest shape expressed by a PDM. (b) If erroneous points are correctly taken into account, the regulated shape (filled surface representation) closely matches the correct part of the shape in the image FOV, otherwise (c) it will attempt to satisfy both inside and outside parts. This results in a non-correctly fitted shape in the image extents.

scheme is thus possible, by progressively increasing the SSM locality throughout the segmentation. Later on in the evaluation Sec. 5.5.2, we will see that in the special situation of small image FOV, a carefully chosen multi-resolution alignment scheme improves the robustness and accuracy of the segmentation.

5.3.3.4 MRF-based Force

In Sec. 4.3.3, we presented a modeling of the local shape variations based on a Markov-Gauss distribution model. In this Section, we will devise a force that exploits this local shape prior. Let us consider $\boldsymbol{d} = \{d_t, t \in \Omega\}$ which designates an observation field on a regular lattice defined by the image I. This field is related to the image information exploited in external images forces (gradient, IP, etc.). Our goal is to estimate the true (unknown) vertex positions $\boldsymbol{Y}^*$ by using a Maximum *a posteriori* (MAP, Sec. B.4.2) formulation of the random process $\boldsymbol{Y}$ from the observation $\boldsymbol{d}$ [KH98]:

$$\boldsymbol{Y}^* = \underset{\boldsymbol{Y}}{\operatorname{argmax}}\, P(\boldsymbol{Y}|\boldsymbol{d}) = \underset{\boldsymbol{Y}}{\operatorname{argmax}}\, P(\boldsymbol{d}|\boldsymbol{Y})P(\boldsymbol{Y}) = \underset{\boldsymbol{Y}}{\operatorname{argmax}}\, P(\boldsymbol{Y},\boldsymbol{d}) \tag{5.9}$$

Commonly, $P(\boldsymbol{d}|\boldsymbol{Y})$ is assumed to follow a Gibbs distribution (Sec. B.6.2), i.e. $P(\boldsymbol{d}|\boldsymbol{Y})$ is proportional to $\exp -E_d(\boldsymbol{Y},\boldsymbol{d})$. As a result, the joint distribution $P(\boldsymbol{Y},\boldsymbol{d})$ is also a Gibbs distribution: $P(\boldsymbol{Y},\boldsymbol{d}) = \frac{1}{Z}\exp - E(\boldsymbol{Y},\boldsymbol{d})$. Parameter Z is assumed to be constant and independent of Y [KH98]. The energy is expressed as:

$$\begin{aligned} E(\boldsymbol{Y},\boldsymbol{d}) &= E_d(\boldsymbol{Y},\boldsymbol{d}) + \beta \sum_{i=1}^{M} \left(\frac{1}{\eta^2} \sum_{j \in \mathcal{N}(i)} \left\| \delta_i - \delta_j \right\|^2 + \frac{1}{\sigma^2} \left\| \delta_i \right\|^2 \right) \\ &= \sum_{i=1}^{M} \left[E_{d,i}(y_i, d_{y_i}) + \beta E_{m,i}(y_i, y_{\mathcal{N}(i)}) \right] = \sum_{i=1}^{M} E_i \end{aligned} \tag{5.10}$$

where M denotes the number of shape vertices, $\delta_i = y_i - x_i$ and d_{y_i} is the image information at vertex y_i. The random process $\boldsymbol{Y}$ will be inferred from the image information and reference positions $\boldsymbol{X}$. The parameter β weights the regularization induced by the energy term based on spatial deformations with respect to the image-based energy term. When β is null, the procedure can be seen as a standard Maximum-Likelihood (ML, Sec. B.4.3) strategy that ignores geometrical considerations.

To solve this MAP problem, the energy E should be *globally* minimized which is not a trivial task. Instead, we consider E as a kind of potential energy and devise appropriate forces f to be used in our framework (i.e., $f = -\nabla E$). Each local energy E_i is derived with respect to positions y_i and y_j ($j \in \mathcal{N}(i)$), to get local contributions in forces and force derivatives. By summing up these contributions (e.g., $f_i = -\partial E_i/\partial y_i - \sum \partial E_j/\partial y_i$), forces and force derivatives can be computed for each particle in order to perform an implicit integration step. An analytic formulation of the energy $E_{m,i}$ derivatives can be expressed (See Sec. A.5 in Appendices for details), whereas the derivative of $E_{d,i}$ cannot be explicitly formulated. Finite differences approximations may also lead to incorrect results as E_d may not vary smoothly enough. The force expression of the image forces depicted in Sec. 5.3.2.1 is thus reused to get forces and force derivatives for $E_{d,i}$ ($E_{d,i}$ is not dependent on y_j). The MAP-MRF problem resolution hence yields two types of forces: $\boldsymbol{f}^{\text{im}}$ and $\boldsymbol{f}^{\text{mrf}}$ (β is then similar to a weighting parameter α^{mrf}), where $\boldsymbol{f}^{\text{im}}$ is in our case a combination of gradient- and IP-based image forces.

Different choices are possible to set $\boldsymbol{X}$, which is kept constant during the optimization of $\boldsymbol{Y}$. It can be fixed once for all during the evolution, typically after the initialization procedure as it is done in [KH98]. This assumption is strongly dependent on the quality of the initialization. Alternatively, $\boldsymbol{X}$ can be set during the segmentation when the evolving shape is near the optimal one. We decided to employ this scheme at the end of each evolution of a resolution level k (See experiment in Sec. 5.5.3.2). In fact, at this stage the shape at resolution k is usually close to the optimal result, and hence it can be locally refined with the MRF forces only, SSM-based forces being turned off. Smoothing forces are also not longer necessary as the MRF formulation achieves some smoothing via its term $\|\delta_i - \delta_j\|^2$.

5.3.4 Coupled Registration-Segmentation

5.3.4.1 Introduction

In general, segmentation yields the reconstruction of structures of interest to support diagnosis and planning, while registration allows the alignment of images to estimate motion or changes in time of a patient. As mentioned by Yezzi et al. [YZK01], these tasks can strongly benefit from each other. Their seminal paper described a way to unify registration and segmentation problems in a variational framework using deformable models. Their work gave birth to a variety of coupled registration-segmentation methods [MC03, US05, WV05, DR07, LHV07] with various extensions such as the support of more than 2 images or non-rigid transforms. Voxel-based approaches were also proposed and formulated as a MAP estimation problem [WN03,

PFL+05] or MRF-based approaches [MS10]. Some other approaches [WV05, PFL+05] made used of atlases to solve the registration-segmentation problem. Atlases might not be always available and atlas-based registration generally demands the use of non-rigid transforms which decreases the method effectiveness in case where linear transforms are expected (e.g., a rigid structure imaged in different positions).

5.3.4.2 Framework Overview

We propose a coupled registration-segmentation framework based on our deformable model, to especially segment and co-registered bony structures in multiple MRI images. However, the proposed framework is quite general and could be applied to any kind of structures and with various types of deformable model approaches as explained in the following.

Our goal is to segment and register the same structures of interest (SOIs) in N images $I^1, \ldots, I^N$. For the sake of brevity, we only present our methodology in case of a single SOI, but the proposed approach can be easily extended to multiple SOIs. We assume that the segmentation is achieved by using N deformable models (DM) represented by shapes $S^1, \ldots, S^N$, composed of a same number n of points x_j^i: $S^i = \{x_1^i, \ldots, x_n^i\}$. No restrictions are made on the nature of the DM (e.g., curves [US05], simplex meshes [Del99]), except that their shapes must share the same number of points and be in point correspondence. Similarly, DMs can adopt any type of evolution strategies (e.g., Gradient flow [YZK01], our Newtonian evolution depicted in Sec. 5.3.1) as long as they can be expressed as an update operator $\Psi(S^i)$, which returns the new state of the shape at each evolution step.

The registration consists in computing the mappings $g^{i \leftarrow j}$ to transform S^j to S^i. Similarly to [LHV07], instead of computing the N^2 possible mappings $g^{i \leftarrow j}$, we estimate the mappings g^i that transform a common shape $\bar{S}$ to each shape S^i. In case of linear transforms and shapes in point correspondence, we express these mappings g^i by using the generalized Procrustes alignment (GPA) that was presented in Sec. 2.5.2. During this procedure, mean shape $\bar{S} = \{\bar{x_1}, \ldots, \bar{x_n}\}$ and mapping g^i estimates are iteratively updated based on a least square optimization (Eq. (2.8)).

An overview of the steps involved in our coupled registration-segmentation framework is depicted in Fig. 5.12, in which segmentation and registration are used in an interleaved manner. The steps are repeated until segmentation convergence. First, shapes S^i are updated given the segmentation update operator Ψ. Then, the resulting shapes $\Psi(S^i)$ are aligned with the GPA. The final stage performs a linear "blending" at the point level between shapes $\Psi(S^i)$ and the back-transformed mean shapes $g^i(\bar{S})$ to get new shapes S'^i for the next iteration:

$$S'^i = \lambda g^i(\bar{S}) + (1 - \lambda)\Psi(S^i) \tag{5.11}$$

Parameter $\lambda \in [0, 1]$ is used as a stiffness coefficient, which expresses the degree of constraint applied by the registration to the segmentation. Theoretically, if shapes only differ by a linear transform, λ should be set to 1. However, in practice it is better to slightly relax λ at the beginning of the segmentation. This provides more freedom to the segmentation, especially when the

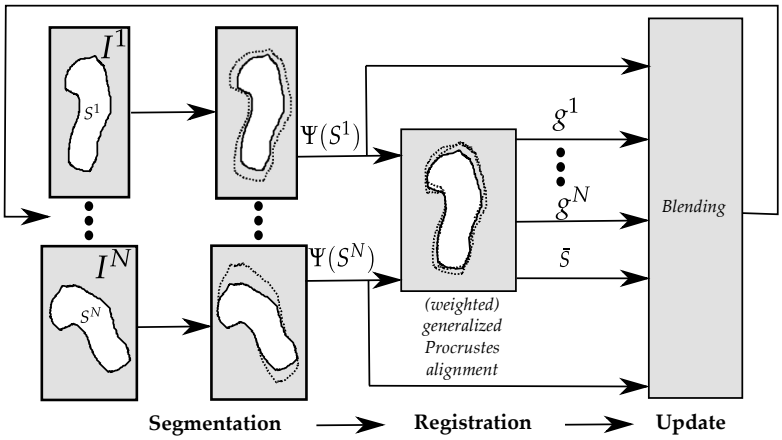


Figure 5.12: Coupled registration-segmentation step: initially, a segmentation step is executed, evolving N shapes $S^i \rightarrow \Psi(S^i)$; resulting shapes $\Psi(S^i)$ are then aligned with a (weighted) GPA to yield alignment mappings g^i and mean shape $\bar{S}$; finally, shapes $\Psi(S^i)$ and back-transformed mean shapes $g^i(\bar{S})$ are blended to produce new shapes.

shapes are not initialized closely enough and they need to be significantly deformed.

5.3.4.3 Weighting Procedure

A major issue with the aforementioned step is that all shape points are equally treated in the registration. If these points are invalid from a segmentation viewpoint (e.g. point not lying on organ boundary or out of the image FOV), they will corrupt the registration. As a result, bad estimates of the mean $\bar{S}$ and mappings g^i will be computed. Hence we apply a simple yet efficient solution that uses a weighted GPA, like it was performed in Sec. 4.3.2 to create robust SSMs.

We express the weights w_j^i as the sum of "reliability" terms $F^k(\theta(x_j^i))$: $w_j^i = \sum_{1 \leq k \leq N} F^k(\theta(x_j^i))$. The quantity $\theta(x_j^i) = \theta_j^i$ is a feature vector computed from the image I^i, the shape S^i and the point x_j^i. The *intra*-reliability term $F^i(\theta_j^i)$ accounts for the segmentation quality at x_j^i in image I^i. Similarly, *inter*-reliability terms $F^k(\theta_j^i), k \neq i$ additionally consider the coupling with other images. If we consider (i) the features vector θ_j^i as the observation of a random variable Θ, (ii) the event w "θ_j^i is the features vector obtained in case of optimal segmentation at x_j^i", and (iii) $F^k(\theta_j^i) = p^k(w|\theta_j^i)$ as a posteriori probability function, w_j^i computation is analog to a "sum rule" commonly used in multi-classifiers approaches.

The computation of weights is critical. In our framework, weights are constantly updated during the segmentation as follows: shape points out of their corresponding image FOV have null weights (lowest weight), otherwise the NCC measure is used to define the reliability terms. If rq_i^j and p_i^j denote respectively the reference and extracted profiles of point x_j^i, we define the reliability terms as follows:

$$\text{intra-reliability } F^i(\theta_j^i) = 0.5(NCC(p_j^i, rq_j^i) + 1)/N \tag{5.12}$$

$$\text{inter-reliability } F^k(\theta_j^i) = 0.5(NCC(p_j^i, q_j^k) + 1)/N, \ \forall k \in [1, N] \text{ and } k \neq i \tag{5.13}$$

Since the NCC varies within [-1, 1], weights take values in [0, 1]. By giving a higher weight to a shape point whose IP is similar to the other shapes' IPs, robust estimates of the mean shape and alignment transforms are obtained. This provides thus an effective way to express point reliability. The NCC is invariant to linear intensity changes, thus being useful for MRI images acquired with different protocols.

5.4 Interactivity and Control

In Sec. 2.5.5.3, we explained that regardless of the type of the segmentation approach, there are conditions in which the behavior of the algorithm will not be optimal, inevitably leading to segmentation errors. These conditions include e.g. pathological cases or strong image artifacts that were not accounted for by the a priori knowledge. As a consequence, users must be provided with control mechanisms which can interactively update and correct the segmentation evolution [OS01].

In Sec. 5.4.1, we will investigate the use of GPU to improve the interactivity of our deformable model-based segmentation while Sec. 5.4.2 will present our framework to concurrently and collaboratively segment the same image by multiple users.

5.4.1 Improved Interactivity with GPU Processing

5.4.1.1 Introduction

In general, segmentation approaches can be classified as low- and high-level approaches. Low-level approaches directly work on voxel information and are usually burdened by the intensive manipulation of volumetric image data. As a result, these methods naturally appeared at first as good candidates for GPU implementations. Indeed, images are regular lattices on which read access can be very efficient due to the optimized GPU texture management (e.g., caching). We can find in the literature GPU-based implementations of the watershed [SS00] and region growing methods [SHN03], along with Markov random fields [WBK+09] and graph cuts approaches [Pan09, LSS09]. Image registration (Sec. 2.3) also strongly benefited from parallel computing, giving birth to a variety of implementations based on mutual information [TEE+08], sum of squared differences [KDR+06], demons [SKSF07, MOOXS08], viscous-fluid regularization [ØNDSE+08] or regularized gradient flow [SDR04].

Higher-level approaches, such as deformable models, were also ported to GPU architectures by considering implicit deformable models as image lattices. Levelset approaches [LCW03, CLW04] (Sec. 2.4.3) became particularly popular in the GPU-segmentation community as significant speed-ups and interactive rendering were made available. Geodesic active contours [CKS97] were also efficiently implemented in GPU [UPT+08, SUP+09] by using the total variation formulation, mostly to quickly segment structures by foreground and background separation in 2D images.

Nevertheless, little work has been made in implementing explicit discrete deformable models in GPU for segmentation purposes. That is unfortunate since discrete deformable models offer

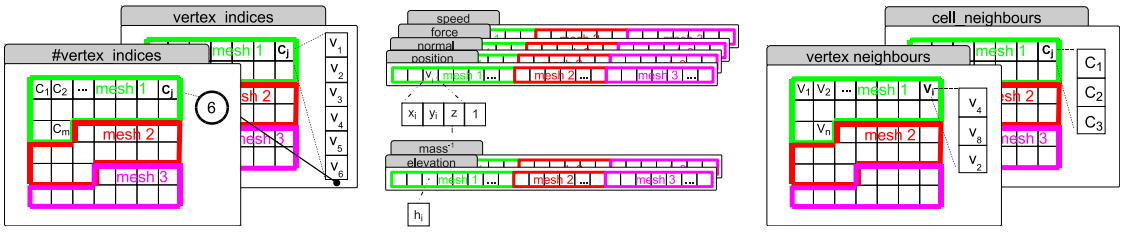
several advantages as previously stated in Chap. 2. Indeed, they provide an intuitive and more appropriate control of the shape deformations compared to implicit models. Furthermore, they are much more robust against image artifacts than most low-level approaches thanks to the use of shape regularization. In GPU, methods for implementing active contours based on gradient flow (Sec. 2.4.2) have been proposed [KB09, HK06], but they were limited to the case of 2D images. On the other hand, many works exploited physically-based surface or volumetric deformable models in GPU in other application domains, such as spring mass systems [MHS05, GW05], cloth simulation [RNS06], volumetric mesh deformation [TE05] or Finite Element Modeling [GSMY+07].

5.4.1.2 GPU Framework Architecture

In this thesis, we hence propose a flexible GPU framework for fully interactive parallel segmentation of multiple volumetric objects based on our discrete deformable models. The main goal is to increase the interactivity of the segmentation while offering an intuitive user control on the segmentation. The GPU framework uses the same segmentation methodology depicted in Sec. 5.3 in which discrete deformable models, represented as 2-simplex meshes, are dynamically evolved as a set of particles subject to internal and external forces. In our implementation, only the core segmentation features were included to validate the capability of the GPU approach.

Our GPU segmentation framework is implemented using the NVidia CUDA. To keep the focus on our contributions, we do not cover the CUDA architecture as well as the OpenGL language. We suggest interested readers to refer to the programming guides [NVI10, SWND07]. The proposed framework architecture is divided according to the major stages involved in a simulation step previously reported in Sec. 5.3: force computation, numerical integration and mesh update. In the following, each stage is described along with the designed data access layer for storage and retrieval purposes of 2-simplex mesh data.

5.4.1.3 Simplex Mesh Data Access Layer



(a) per vertex GPU arrays (b) per cell GPU arrays (c) vertex/cell neighbours GPU arrays

Figure 5.13: GPU arrays. (a) Per vertex GPU arrays which are shared by all meshes in the system; (b) Cell description held in GPU: e.g. cell C_j is composed of vertices $\{V_1, V_2, V_3, V_4, V_5, V_6\}$ and the number of vertices (6) is stored in a separated GPU array; (c) Neighbor information per vertex on the GPU: e.g. vertex V_i has 3 neighbors $\{V_4, V_8, V_2\}$. Similarly, cell C_j has neighbor cells $\{C1, C2, C3\}$.

We have designed a simple data access layer for storage and retrieval of the simplex mesh data (see Fig. 5.13). In our GPU segmentation framework, meshes are encoded as vertices within one unique array. This encoding strategy has the advantage of being very compact in terms of space

requirements and it supports a straightforward distribution of the data for parallel processing within the GPU. We store position and normal for each vertex using a CUDA one-dimensional texture sampler bound to linear memory, because we use these attributes in both the read and write modes. The same strategy is used for other local parameters like the mass or elevation h of a particle with respect to its neighbors (see Sec. A.3 of Appendices for a reminder on 2-simplex mesh geometry).

Besides, we store other *(mostly) invariant* per-vertex attributes using 2D samplers mapped to `cudaArrays` which are special data structures of CUDA, providing (i) bigger amounts of memory allocation in the GPU and (ii) fast cached reads. In order to access the desired data, we just need to compute the offset in each dimension of a 2D texture. Examples of per vertex attributes stored using 2D textures are the neighbor indices and the indices of the 3 neighbor cells sharing the same vertex. Finally, we store some per-cell information like the indices of the vertices of each cell and the number of vertices forming a cell, which is not necessarily the same for all the cells.

The volumetric image information is stored in a raw uncompressed format by using 3D textures. This is to take advantage of spatial locality to provide a fast access for read operations which are the most expensive in terms of access time. Moreover, 3D textures provide cheap trilinear interpolation that is intensively used by image-related operations. In this first implementation of the framework, moderate-size volume data sets are handled, so special compression or multi-resolution strategies are not necessary.

After each numerical integration, all meshes are processed in parallel at once by calling a CUDA kernel (i.e. a GPU-threaded function) responsible of the update of each point's position, normal, area and elevation. Cell centers are also updated since they are required by smoothing and shape prior forces (See next Section) and the visualization based on the triangular tessellation of 2-simplex mesh (Sec. A.3.2). Writing operations are then performed using data structures in global memory space, because `cudaArrays` and texture memory writing operations are not yet supported by current versions of CUDA.

5.4.1.4 Force Computation

Our GPU framework re-implements the IP- and gradient-based external image forces (Sec. 5.3.2.1), as well as the smoothing (Sec. 5.3.3.1) and shape memory (Sec. 5.3.3.2) internal forces. Once the information of all meshes has been updated, the computation of the forces is distributed in parallel by assigning each particle to a GPU thread. Each thread calls a CUDA kernel which performs the actual computation of a force. Forces can be enabled and disabled for each mesh registered in the framework. For this purpose we maintain in the GPU a simplex mesh activation register storing the weight α_i^f of the force f for the i-th mesh. Changes on this state are supported at any moment during the runtime of the segmentation algorithm.

The different forces are computed in a sequential order to avoid asynchronous updates of the total resultant force. This behavior is directly related to the availability of atomic operations,

only supported by NVIDIA graphics cards with CUDA compute capability above `1.0`[1]. In any case, forces computation time is quite heterogeneous and variable. In particular when one force is much more expensive than the others, we can assume that the total time of computation in parallel for all the forces is almost equivalent to the time of the most expensive force in terms of computation time (See later on the results on force time distributions in the experiment Sec. 5.5.6).

Most of the forces were implemented without any special GPU optimizations, since the proposed data access layer (e.g., texture samplers) and parallelism strategies (e.g. simultaneous processing of all vertices with one kernel) were already efficient. However, "interactive" forces required more attention and will be detailed in Sec. 5.4.1.7.

5.4.1.5 Numerical Integration

Once the resultant force per particle is computed, we need to update the new states of the particles by solving a series of discrete differential equations. Compared to the CPU implementation, we decided not to exploit the implicit Euler integration. This choice has been mainly driven by practical implementation choices. Indeed, iterative approaches such as the conjugate gradient are usually exploited to avoid the explicit inversion of large sparse matrices in implicit integration. These approaches require a considerable effort to be successfully implemented in the GPU, which significantly complicates the design of the GPU framework.

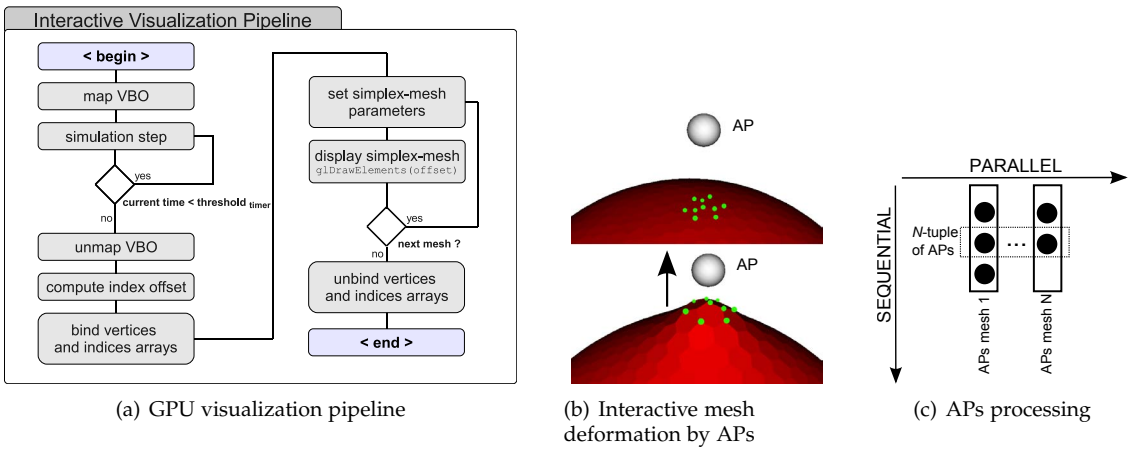
Since solving the integration with an explicit method is known to require a small stepsize in order to converge, we chose to move the calculations to the GPU in order to decrease the stepsize and reach better interactive frame rates. Among the explicit techniques implemented in the GPU, we have obtained the best results with a Verlet-based approach [Ver67]. The Verlet equation is implemented in a CUDA kernel and is computed in parallel for each particle of the simplex meshes. The final result is eventually written in the point position array.

5.4.1.6 Visualization

To ensure real-time interaction, we established a balance between simulation and visualization tasks by fixing the maximum amount of time that the simulation can spend. As long as we are within the bounds of this time interval, one or more segmentation iterations are performed. Subsequently the simplex meshes are rendered by using the Vertex Buffer Object (VBO) OpenGL extension which significantly reduces data transfers between GPU and CPU. The sequence of operations needed for updating and displaying the simplex mesh data in the VBO are shown in Fig. 5.14(a). First, all the VBO attributes array are mapped to the GPU linear memory. Next, we perform a series of simulation iterations and write the partial result in the memory used by the VBO. Meshes are then updated in parallel before the VBO is unmapped. Subsequently, all meshes are rendered by (i) calculating the entry offset in the index array, (ii) binding all VBO attribute arrays in OpenGL, (iii) setting all simplex mesh rendering parameters (e.g. color or

1. Our GPU did not support atomic operations. Moreover, atomic operations are known to be relatively slow which might affect the performance of our framework.

clipping planes) and finally calling the `glDrawElements` OpenGL command with the current index offset and bound arrays as parameters. The default OpenGL state is then restored by unbinding the vertex and index arrays. In addition, we display other useful information such as standard transversal, sagittal or coronal slices to explore the volumetric image. To render these slices, the Pixel Buffer Object (PBO) OpenGL extension was used as it supports interoperability with the CUDA architecture. More specifically, a PBO buffer is declared and the 2D slice of interest is written by using a kernel. Finally, the PBO is used to define a 2D texture which is rendered on a simple 2D quad.



(a) GPU visualization pipeline

(b) Interactive mesh deformation by APs

(c) APs processing

Figure 5.14: (a) GPU visualization pipeline. Interaction by attraction points: (b) a point and its 2-order neighborhood (small discs) of a spherical mesh are attracted by an attraction point (AP) yielding the mesh deformation. (c) In a sequential manner, a N-tuple of APs is processed in parallel to find the points on each mesh to be attracted.

5.4.1.7 Interaction

All the parameters affecting the segmentation (e.g., force contribution α_i, timestep) can be dynamically changed, providing a first level of interactivity. Such actions are however insufficient to accurately interact with the segmentation, and "pictorial-input" is preferred [OS01]. We implemented an interaction based on attraction points (AP). This interaction is similar to the constraint point force concept previously presented in Sec. 5.3.2.2. For each mesh j, different attraction points can be positioned in the 3D space by clicking on rendered slices. Given $\boldsymbol{C}_j^k$ the k-th AP associated with mesh j, we find the closest mesh point $\boldsymbol{P}_j^0$ to $\boldsymbol{C}_j^k$ and its p-order neighboring points $\boldsymbol{P}_j^i, \forall i \in [1, N_p]$ (e.g., a 2-order neighborhood is composed of the neighbors of neighbors of $\boldsymbol{P}_j^0$ as exemplified in Fig. 5.14(b)), and compute the following weighted attraction forces $\boldsymbol{f}^a$ applied on each vertex $\boldsymbol{P}_j^i$:

$$\forall i \in [0, N_p], \boldsymbol{f}^a_{\boldsymbol{P}_j^i} = \frac{w_i(\boldsymbol{C}_j^k - \boldsymbol{P}_j^0)}{\sum_{l \in [0,N_p]} w_l} \tag{5.14}$$

$$w_i = \alpha_j^a / \left\| \boldsymbol{C}_j^k - \boldsymbol{P}_j^i \right\| \tag{5.15}$$

In practice, various N-tuples of APs $\{\boldsymbol{C}_1^k, \ldots, \boldsymbol{C}_N^k\}$ are sequentially processed (Fig. 5.14(c)), where N denotes the number of meshes. Note that this N-tuple can actually have less than N elements as not all meshes have the same number of APs. Given a N-tuple of APs, the closest point $\boldsymbol{P}_j^0$ is found on each mesh j by (i) running in parallel a kernel on each mesh point $\boldsymbol{x}_j^i$ that computes the squared Euclidean distance between $\boldsymbol{C}_j^k$ and $\boldsymbol{x}_j^i$, and by (ii) applying a parallel reduction on the array of squared distances to get for each mesh j the index of the closest mesh point $\boldsymbol{P}_j^0$. Since for each mesh j of M_j points, M_j CUDA threads will access the same memory location containing the coordinates of the AP C_j^k, this information is initially loaded into *shared* memory to speed-up read-accesses. Furthermore, the possibility to interactively select, move or delete each AP is provided to users by intuitive point and grab actions, leading to a real-time control of the mesh deformation.

5.4.2 Collaborative and Concurrent Control

5.4.2.1 Introduction on Collaborative Telemedicine

Telemedicine is a rapidly developing application system in the medical domain that harnesses the advances in telecommunication and multimedia technologies [WCP06] and has lead to many healthcare benefits, such as telesurgery applications for distance-based surgery [MLG+01]. In particular, teleradiology is a telemedicine system that involves the efficient distribution and sharing of medical images from one geographical location to another. Teleradiology has evolved from video/audio/text based conferencing and shared 2D images running on imaging workstations with local area network (LAN) [ZSH+00] to current state-of-the-art systems that are capable of streaming volumetric medical images in real-time via a wireless network [PKI08]. Moreover, modern systems now support thin client/cloud computing [CKCF07] that processes (e.g., volume rendering) and stores all the medical image data in a client-server relationship, thus not requiring imaging workstations to directly access to the images.

Despite the remarkable growth in teleradiology, little progress has been made in the mechanism that enables interactive collaboration among multiple users. In teleradiology applications, collaboration tools include the usual video/audio/text for communication among the users (teleconferencing), image navigation, and image editing, which we refer to hereon as "collaborative editing". Basic image editing involves appending information to the image without changing the image's state, such as textual annotations and drawing of regions of interest (ROIs). For editing that results in a changed state of the image (e.g., brightness/contrast, lookup table) a simple "lock and release" has become the default control mechanism. In such a strict locking mechanism [PKI08, SWCHPC+05], one user requests the object and applies changes on it, while all other users cannot make changes to it until the object is "released".

Hence, these studies do not allow concurrent editing, where every user's action can be executed timely and collaboratively, because only a single user who has ownership of the object can edit it. In collaboration with a MIRALab team focused on networking and mobility, we proposed a framework for real-time and interactive collaborative segmentation based on our deformable

models. Unlike conventional approaches with strict locking mechanism, we relax the time constraint in collaborative editing in order to increase real-time and interactive performance of users: multiple users may request edits within a predefined time range, where these requests are then managed (filtered) based on image processing constraints. The filtered request is then executed and all the users are acknowledged of the processed requests.

Intentionally, this work is briefly covered in this thesis, in particular all aspects related to networking (communication substrate) or server-client architectures. We refer interested readers to our publications [SNH+09, HNS+10] for full details.

5.4.2.2 Overview of the Proposed Framework

Our collaborative editing framework consists of three main components: (i) Collaborative editing mechanism, (ii) Iterative 3D segmentation and (iii) Communication substrate for collaboration. An overview of the framework is visible in Fig. 5.15. Once a user initiates the collaboration session, i.e., loads the image and sends invitation to other users to join, multiple users can connect to the session and start the collaboration. Iterations of the segmentation are executed by a processing module (iterative 3D segmentation) which relays the results to all the users in the session through the communication substrate. The results are composed of the image data (MRI slice) and segmentation data (segmentation overlay) which are proper to each user. At any point, any user can collaboratively segment the image by adjusting the segmentation parameters via interactive image annotation (collaborative editing mechanism). These parameters are then inserted as new constraints in the segmentation and used in the next segmentation iteration.

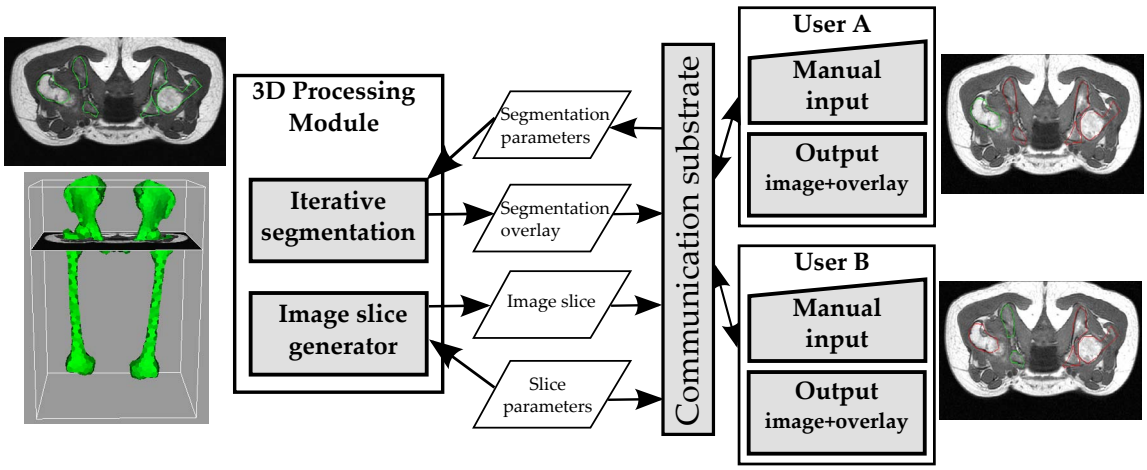


Figure 5.15: Collaborative segmentation framework. It is composed of a processing server which communicates with clients through a communication substrate. The server holds and process the segmentation data (3D meshes and image) while clients only have a 2D view corresponding to a slice of the 3D image. In this example, the server and client views are synchronized, and different clients (varying color annotations) are concurrently making changes to the segmentation. Each client can apply modification to a selected model (in green overlay on the client view).

5.4.2.3 Segmentation

To ensure the efficiency of the collaborative process, the segmentation approach used in the framework must fulfill the following requirements:

- *Iterative*: the algorithm must be iterative thereby enabling different users to amend, stop and resume the segmentation process. Most importantly, each iteration has to provide meaningful intermediate results that can be correctly interpreted by users. This provides an efficient monitoring of the algorithm evolution [OS01].
- *Multiple parameters support*: The algorithm must be able to accept multiple parameters (changes) from multiple clients. Parameters can be global (e.g., a weight coefficient) or local (e.g., local constraints). Appropriate decision rules have to be designed to fuse the various changes or constraints.
- *Locality of changes*: Changes made by users should have an immediate local influence. These changes may ultimately affect the process in a global but progressive manner.

Our discrete deformable model-based segmentation (Sec. 5.3) is hence a good candidate, as well as its GPU implementation (Sec. 5.4.1). In fact, these iterative algorithms provide intuitive visual feedback by progressively deforming shapes, whose evolution can be controlled by global parameters such as force weighting or local changes based on constraint (Sec. 5.3.2.2) and attraction (Sec. 5.4.1.7) points for the CPU and GPU versions, respectively. With our segmentation approach, the segmentation overlay (Fig. 5.15) is simply composed of the evolving model contours and the constraint points, which are overlaid on the transmitted slices.

5.4.2.4 Collaborative Client-Server Mechanism

A unique instance of the 3D image with the corresponding evolving 3D meshes is hosted and managed by the processing module (Fig. 5.15). This service provider acts as a server and is responsible of the segmentation evolution, the filtering of user requests and generation of data. The processing module processes user data between a desired number of iteration steps, relaxing time constraints and improving interactivity. Every user, who participates to the collaborative segmentation, operates through a thin client which (i) send commands to the module to notify (change in) its status (e.g., request of new slice position, addition of constraint points, change of some segmentation parameters, etc.) and (ii) receives updated information (new image and segmentation data). An efficient caching management is implemented to resend whenever possible the same information to multiple users (e.g. a slice) avoiding the recomputation of the same information.

Communications are easily and seamlessly done over a network through the communication substrate (Fig. 5.15). Clients can run on heterogeneous devices, such as those with limited resources (e.g., portable devices) since all the processing is done by the server. Views of the clients can be synchronized (2D views in Fig. 5.15) which means that all the users observe the same slice. A second mode consists in letting the users to view a slice independently to the other users in the same collaborative session. Experiments in the evaluation Sec. 5.5.7 will illustrate

some collaborative scenarios exploiting these modes.

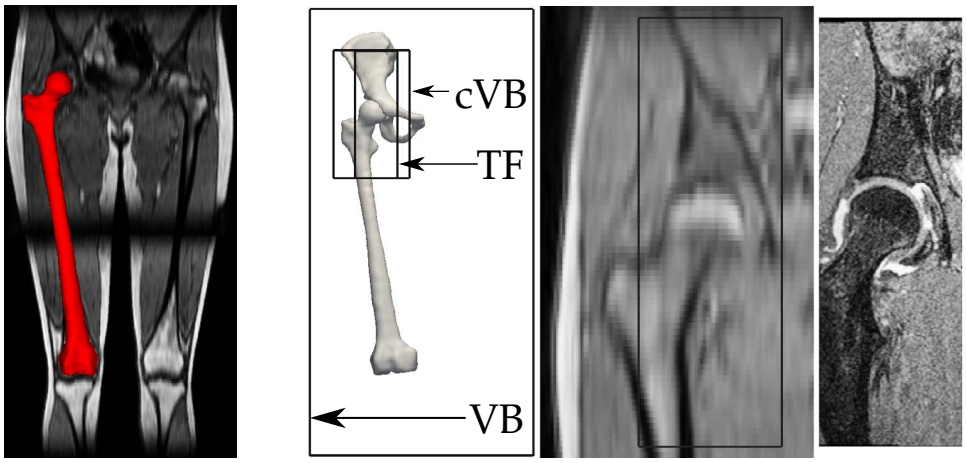
5.5 Evaluation

In this Section, we will evaluate our knowledge-based deformable models in the context of image segmentation by analyzing its main aspects such as initialization, user control, interactivity and impact of the field of view (FOV).

5.5.1 Experiment Setup

5.5.1.1 Image Material

Thirty female ballet dancers from the Grand Théâre of Geneva and 13 additional female subjects (average age of all 43 subjects: 25.1 years) gave written consent to be acquired using two different MRI protocols on a 1.5T MRI device (Philips Medical Systems). These protocols were defined by MIRALab and the radiological department of the University Hospital of Geneva [Gil07]. The first protocol VIBE (Axial 3D T1, TR/TE= 4.15/1.69 ms, FOV/Matrix= 35 cm/256 × 256, resolution= 1.367 × 1.367 × 5 mm) was used to acquire both thighs of each subject in supine position covering the FOV from the iliac crests to the beginning of the proximal tibias (Fig. 5.16(a)). This protocol had fast acquisition time resulting in low resolution images. This produced strong partial volume effect which is clearly visible in the articular area (middle of Fig. 5.16(b)).



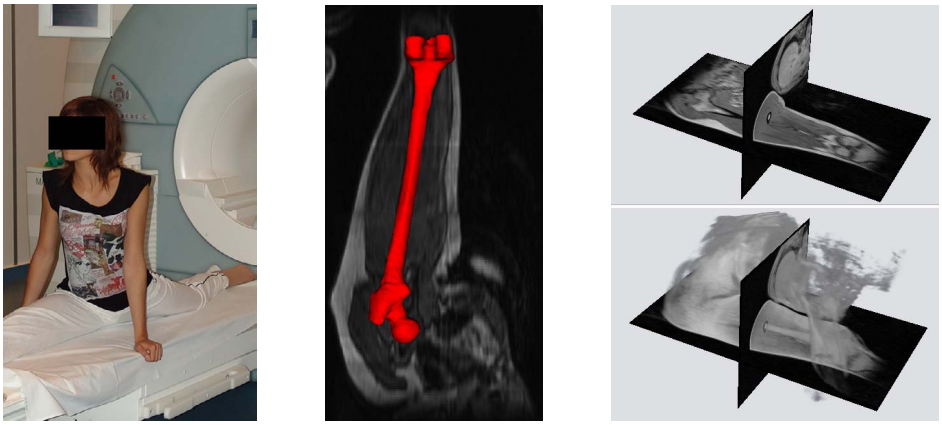
(a) (b)

Figure 5.16: MRI images acquired in supine position used in experiments. (a) Coronal slice of MRI volume acquired with VIBE protocol and with a right femur bone overlay. (b) In the left subfigure the different field of views (FOV) of the VIBE (VB), right cropped VIBE (cVB) and right TRUEFISP (TF) images can be compared. Bones are only partially visible in the TF and cVB datasets, particularly the femurs. In the middle and right subfigures, enlarged coronal slices of the hip articulation with cVB and TF protocols are respectively shown, where the black rectangle in the cVB slice corresponds to the TF FOV extents.

The second protocol TRUEFISP (Sagittal 3D T2*, TR/TE= 10.57/4.63 ms, FOV/Matrix= 20 cm/384 ×384, resolution= 0.52 × 0.52 × 0.6 mm) had a significantly smaller FOV around the right or left hip joint but it offered a better resolution and signal-to-noise ratio (right of Fig. 5.16(b)). The

left part of Fig. 5.16(b) shows the large difference between the FOV of VIBE (VB) and TRUEFISP (TF) protocols. It also illustrates an additional type of artificially created images, which was obtained by cropping the large VIBE datasets around the hip joint with slightly larger FOV than the TF. These cropped images are hereafter referred to as cropped VIBE (cVB). Volumes of cVB and TF corresponded in average to 6.2% and 3.5% of VB's volume, respectively. These images with small FOV will be used to assess the impact of FOV and SSMs builit with complete and corrupted shapes on the segmentation.

Furthermore, dancers entered the scanner in right and left split postures (Fig. 5.17(a)) and were scanned using the same protocols, producing the "split" datasets (Figs. 5.17(b) and (c)). Subjects were monitored and asked to remain as still as possible during the acquisition to avoid the formation of motion artifacts.



(a) (b) (c)

Figure 5.17: MRI images acquired in split position used in our experiments. (a) A volunteer before entering the MRI scanner in split posture. (b) A coronal slice of a split MRI image with bone overlay. (c) Volume rendering of a split MRI showing the extreme posture performed by the subject.

5.5.1.2 Shape and Appearance Priors Computation

The left and right hip joint bones (i.e., the bone "sides") which consist of the pair of hip bones and femurs (i.e. the bone "types") were semi-manually segmented as described in Sec. 4.3.1.1 of previous Chap. 4, yielding a large collection of 172 models. These models were used as reference or training shapes. For each shape chosen for training purposes, four different shape resolutions were created based on the multi-resolution scheme (Sec. 4.2.1.3). From the training dataset composed of images and shapes, shape and appearance priors were computed as follows. SSMs per shape resolution were built after point correspondence refinement (Sec. 4.3.1.2) with different multi-resolution alignment schemes (Section 4.3.1.3) that will be specified in each experiment. For the highest shape resolution level, no SSM was computed and the segmentation evolution regularization was only dictated by the smoothing constraints. This allowed the segmentation of fine details not represented by the SSM.

In the modeling of the IPs appearance, we exploited both the first order statistics by computing

an average intensity profile (IP) per vertex and a multivariate Gaussian distribution (MGD) (Sec. 4.4). For each IP, 25 intensity samples at the shape interior regularly spaced by 0.5 mm were computed. Similarly, we computed 5 samples at the shape exterior [Gil07].

5.5.1.3 Initialization and Evolution Settings

In experiments involving the constrained global initialization of Sec. 5.2.3, the HJC was manually positioned in the images with an error experimentally set as $\delta = [1,\ 1,\ 1]$ voxels, where minor changes (e.g., $\delta = [1.5,\ 1.5,\ 1.5]$) had no significant effect in the overall segmentation performance. The parameters we used for the DE minimizer were suggested in the work of [SP95]: $F = 0.8$, $C = 0.9$ and the number of individuals was set as 10 times the total number of parameters to optimize. Furthermore, we used the "RAND1BIN" crossover scheme and set the number of generations to 100. Each individual of the first generation was uniformly randomized.

In all experiments equal weight was given to image forces and energies in both the constrained initialization (i.e., $\beta = 0.5$ in Sec. 5.2.3.4) and the deformable model-based segmentation (i.e. $\alpha^{g} = \alpha^{ip} = \alpha^{mgd}$ in Sec. 5.3.2.1). All experiments used the gradient force $\boldsymbol{f}^{g}$ in conjunction with exclusively one type of IP-based force, either $\boldsymbol{f}^{ip}$ or $\boldsymbol{f}^{mgd}$. Similarly, the smoothing force (Sec. 5.3.3.1) was exclusively used with either the PDM-based force $\boldsymbol{f}^{pdm}$ (Sec. 5.3.3.3) or the shape memory force $\boldsymbol{f}^{smem}$ (Sec. 5.3.3.2), with equal weights (i.e. $\alpha^{smo} = \alpha^{pdm} = \alpha^{smem}$). When hip bones were simultaneously segmented, the weight α^{sym} of the symmetric force (Sec. 5.3.2.3) corresponded to 10% of the external forces' total weight. The total contribution of external forces was set to 30% of the total forces as proposed in [GMT10]. Constraint point-based forces (Sec. 5.3.2.2) were not used in any experiment, except in those based on collaborative segmentation (Sec. 5.5.7) since they involved user interaction.

The multi-resolution evolution scheme (Sec. 4.2.1.4) adopted an iteration schedule of 300/100/20/5 (i.e. first level had 300 iterations, second one had 100, etc.). Similarly, the search depth L used in image forces decreased according to the schedule 10/10/5/2 with a constant search step S of 0.5 mm (Sec. 5.3.2.1). Settings for the special case of MRF-based force are presented in the corresponding experiment (Sec. 5.5.3).

5.5.1.4 Description of Experiments

We used the distance measures presented in Sec. 4.5.1.1 to assess the accuracy and robustness of the segmentation based on the available ground truth. To statistically compare the results of two segmentation approaches based on the statistical significance procedure of Sec. 4.5.1.2, the ASRSD measure was selected since it is correlated to both the average and variance of the distance. A total of 6 main experiments were set up:

- **Experiment 1 – Hip joint bone segmentation** (Sec. 5.5.2): comprehensive study of femurs and hip bones segmentation in supine MRI with different FOV.
- **Experiment 2 – Robust SSM and MRF Local Modeling** (Sec. 5.5.3): evaluation in the segmentation context of SSMs built with complete and corrupted shapes, by also investigating the impact of the training size and the performance of MRF.

– **Experiment 3 – Dual-Posture MRI Segmentation** (Sec. 5.5.4): segmentation of femurs in supine and split MRI acquired with the VB protocol. This set of MRI images is denoted as "dual-posture" MRI.
– **Experiment 4 – Coupled registration-segmentation** (Sec. 5.5.5): Coupled segmentation and registration of femurs in multiple VB and TF images.
– **Experiment 5 – GPU-based Segmentation** (Sec. 5.5.6): segmentation of supine VB images with the GPU-based approach.
– **Experiment 6 – Collaborative segmentation** (Sec. 5.5.7): illustration of the collaborative segmentation.

All experiments were run on a single-core 3.40 GHz PC equipped with an NVidia GTX 8800 graphics board with 768 Mb RAM.

5.5.2 Experiment 1 – Hip Joint Bone Segmentation

In this experiment, the efficiency of hip joint bone segmentation was assessed with our collection of 43 datasets with different combinations of internal and external forces as well as image FOV. All trials followed the leave-one-out (LOO) testing procedure: first, a test image was selected and the corresponding shapes were excluded from the training dataset. Training bone shapes with various shape resolutions were then selected to build the necessary SSMs and appearance priors.

5.5.2.1 Fully Visible Bone Segmentation

We initially tested our segmentation approach with the 43 VB datasets in which the complete bone structures were visible. The bones were initialized with the landmark-based approach (Sec. 5.2.2.1) by using the landmark set defined in [Gil07]. Various combinations of the following forces were tested:

– image forces (Sec. 5.3.2.1): IP with NCC similarity measure (IP) and MGD with approximated Mahalanobis distance (MGD).
– internal forces (Sec. 5.3.3): SSM-based force with the RSS, RAA and SSS multi-alignment schemes (Sec. 4.3.1.3) and shape memory forces (SMEM).

For comparative purposes, one series of trials ("IP+SSS (NG)") was performed with the normalized gradient force f^{ng} (Sec. 5.3.2.1), whereas all other experiments used the non-normalized gradient implementation.

Results Table 5.1 presents the LOO segmentation results for all the different combinations with the VB datasets.

Table 5.1: LOO segmentation results with VB dataset (average±standard deviation). F: femurs, HB: hip bones. ASSD, ASRSD, MSD are in mm, VOE in %. Average times are in sec and for all 4 bones, and include model loading.

Bone	ASSD	ASRSD	MSD	VOE	Time
Landmark-based initialization					
F	3.13±0.66	4.49±0.94	18.38±4.82	34.44±5.48	–
HB	2.63±0.71	3.75±0.94	16.22±4.58	42.52±8.71	–

Table 5.1: LOO segmentation results with VB dataset (continued)

All	2.88±0.68	4.12±0.94	17.30±4.70	38.48±7.09	30
Fine segmentation IP + SSS					
F	1.28±0.31	2.07±0.45	9.54±2.16	16.39±2.67	–
HB	1.15±0.40	1.90±0.52	9.75±2.17	22.54±4.88	–
All	1.21±0.35	1.98±0.48	9.65±2.17	19.47±3.77	280
Fine segmentation IP + SSS (NG)					
F	1.23±0.34	2.00±0.51	9.16±2.68	16.17±2.96	–
HB	1.22±0.38	2.00±0.49	10.08±2.26	23.50±4.71	–
All	1.23±0.36	2.00±0.50	9.62±2.47	19.83±3.83	278
Fine segmentation IP + RSS					
F	1.39±0.74	2.28±1.22	10.05±4.71	17.23±5.84	–
HB	1.28±0.38	2.11±0.55	10.86±2.38	24.21±4.73	–
ALL	1.33±0.56	2.19±0.89	10.46±3.55	20.72±5.29	240
Fine segmentation IP + RAA					
F	1.39±0.74	2.29±1.27	10.08±4.98	17.20±5.91	–
HB	1.28±0.39	2.11±0.56	10.82±2.47	24.25±4.88	–
All	1.34±0.57	2.20±0.92	10.45±3.72	20.73±5.39	253
Fine segmentation IP + SMEM					
F	1.44±0.36	2.37±0.51	11.00±2.27	17.59±3.59	–
HB	1.19±0.48	2.01±0.70	10.55±3.79	22.85±6.41	–
All	1.31±0.42	2.19±0.61	10.77±3.03	20.22±5.00	186
Fine segmentation MGD + SSS					
F	1.30±0.33	2.12±0.46	9.47±2.08	16.24±3.12	–
HB	1.11±0.44	1.89±0.61	9.86±2.93	21.45±5.36	–
All	1.20±0.38	2.00±0.53	9.66±2.51	18.85±4.24	363
Fine segmentation MGD + SMEM					
F	1.39±0.38	2.29±0.52	10.20±2.43	17.28±3.93	–
HB	1.09±0.45	1.91±0.63	10.35±3.14	20.82±5.90	–
All	1.24±0.41	2.10±0.58	10.27±2.78	19.05±4.91	344

From these results, we can state that the landmark-based initialization had errors (ASSD 2.88 ± 0.68 for all bones) significantly higher than any deformable model-based fine segmentation (largest p-value=10^{-15}). Furthermore, the low standard deviations (std) of the fine segmentation results (e.g., ASSD std varying in interval [0.35-0.57]) indicate robust segmentation approaches without outliers.

The configurations which produced the best overall results were "SSS+IP" (ASSD: 1.21±0.35) and "SSS+MGD" (ASSD: 1.20±0.38). The use of the MGD appearance seemed to slightly improve the results albeit no statistical difference could be established. This was also observed between the IP+SMEM and MGD+SMEM combinations, where the use of the second order appearance statistics MGD decreased the overall ASSD to about 7%. Among the IP-based configurations with

shape priors based on PDM, the SSS multi-alignment strategy performed better than both RAA and RSS schemes, where statistical differences were only observed for hip bones (p-value < 0.048). The use of regularized and normalized gradients was statistically not more beneficial (overall ASSD of 1.23±0.36) than using the non-normalized gradient force (overall ASSD of 1.21±0.35). As a result, the non-normalized gradient was always chosen for all the other experiments.

The differences between the ASSD and ASRSD, and the large MSD values of all the fine segmentation results suggest that the error distribution over the shapes was inhomogeneous. By considering for instance the SSS+IP configuration, we computed, for each bone type and side, the point-to-surface distance d_S (Sec. 4.5.1.1) between each point of the reference shape and the segmented shape, averaged over all the trials. Visual inspection of the error distribution in Fig. 5.18 shows that large errors were mostly localized in small regions. These findings were inline with an error histogram analysis, which reported values such as 1.58 and 2.75 mm respectively for 75th and 95th percentiles of the femur. These variations in the identified small regions of the femur were expected as these regions exhibit large variations among different individuals. With respect to the image resolution ($1.36 \times 1.36 \times 5$ mm), the segmentation results of all approaches were within one voxel in error and appeared to closely match the bone structures as illustrated in Fig. 5.19. In all experiments, the segmentation appeared to be slightly more accurate with hip bones than with femurs (e.g., ASSD: 1.15±0.40 vs. 1.28±0.31 for the SSS+IP configuration). This was attributed to the greater complexity of the femur anatomical features, which made the segmentation more challenging.

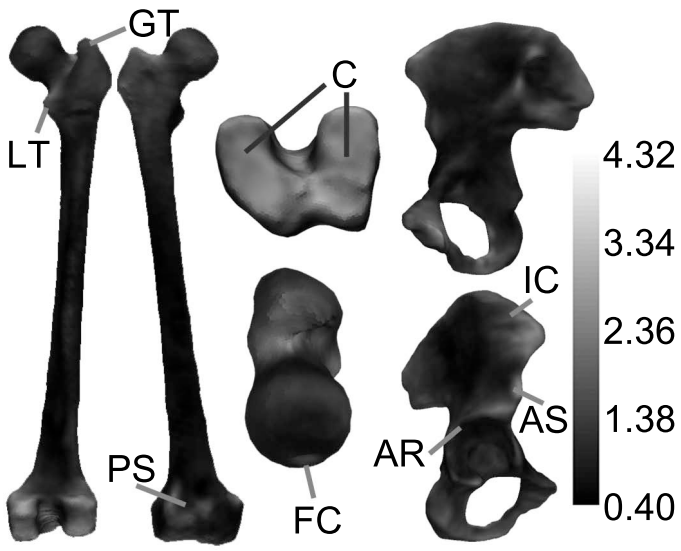


Figure 5.18: Point-to-surface distance in mm between reference and final segmentation results (SSS+IP configuration) with VB datasets mapped as grey-level values on arbitrary femur and hip bone shapes. Large errors are localized in regions such as the femur's lesser and greater trochanters (LT, GT), patellar surface (PS), condyles (C) and fovea capitis (FC) as well as hip bone's iliac crest (IC), anterior inferior spine (AS) and acetabular rim (AR).

The landmark-based initialization based on TPS interpolation took in average 30 sec including mesh loading time. This time did not include the placement of markers, which usually requires 2-3 min for an experienced user. Computation times of the different segmentation configurations

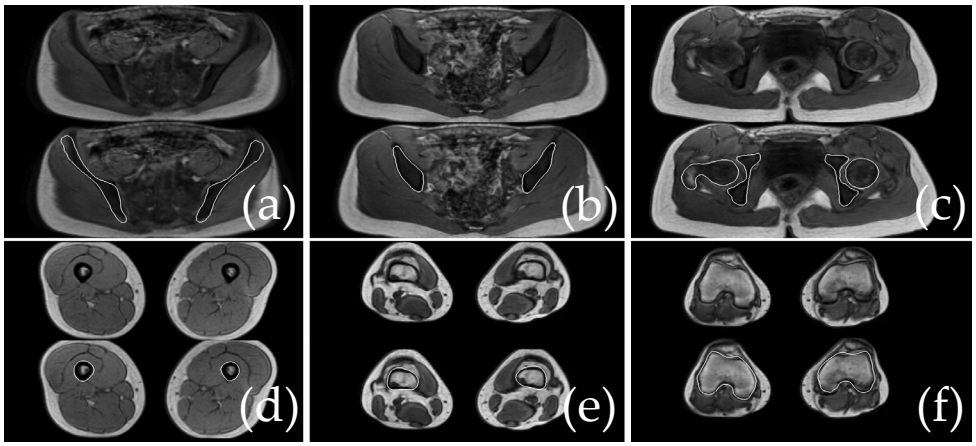


Figure 5.19: An example of a final segmentation result showing the axial slices from a VB dataset. For each subfigure a)-f), the fine segmentation results are displayed as white segmentation contours.

varied between 186 sec (IP+SMEM) and 363 sec (MGD+SSS). With respect to appearance priors, IP-based methods required less computation time than approaches based on the more complex MGD priors. Similarly, the speed of implementations based on SSM multi-alignment schemes was related to the complexity of the alignment and the number of iterations in each level. For instance, a rigid (R) alignment being cheaper to compute than a similarity (S) alignment, the RSS scheme yielded smaller segmentation times than the SSS scheme. Similarly, despite the affine alignment being more time consuming than a similarity alignment, RAA-based segmentations took less time than those using the SSS scheme since the first resolution level had the highest number of iterations (300).

5.5.2.2 Partially visible bone segmentation

The constrained global initialization (GI) introduced in Section 5.2.3 was applied on datasets where bones were only partially visible in the images (Fig. 5.16(b)). A LOO experiment as in the previous trials was first conducted on the cVB dataset and then on the TF dataset, which contained images with the smallest FOV. Three configurations were considered: IP+RAA, IP+SSS and IP+SMEM. Here, only the shape prior strategy varied.

Results with cropped VIBE (cVB) datasets Table 5.2 presents the LOO segmentation results on the cVB dataset after GI.

These results show that GI has successfully initialized the shapes with no failures (i.e., situations not recoverable with a fine segmentation, identified as results with a high error like when VOE > 60%), and that the initialization was sufficiently accurate. The accuracy is indicated by the significant decrease in the overall ASSD as exemplified in Fig. 5.20 for the RAA+IP scheme. Absence of statistical differences between the fine segmentations and the initialization was only observed with the left femur for all the RAA-, SSS- and SMEM-based combinations. This is attributed to the fact that the initialization of femurs had small errors (average ASSD 1.17±0.38) and thus no significant improvements were possible. Nevertheless, these results demonstrate that the

Bone	ASSD	ASRSD	MSD	VOE	Time
Constrained global initialization (GI)					
F	1.27±0.44	2.00±0.57	8.63±2.66	18.93±4.82	–
HB	1.72±0.49	2.89±0.76	16.94±4.33	29.05±5.40	–
All	1.49±0.46	2.44±0.66	12.79±3.49	23.99±5.11	550
Fine segmentation IP+RAA					
F	1.21±0.53	1.91±0.70	7.57±2.46	18.02±6.12	–
HB	1.03±0.38	1.79±0.52	10.39±2.65	20.06±4.77	–
All	1.12±0.46	1.85±0.61	8.98±2.55	19.04±5.44	169
Fine segmentation IP+SSS					
F	1.33±0.61	2.07±0.84	8.09±2.95	19.36±7.14	–
HB	1.04±0.36	1.80±0.48	10.00±2.48	20.22±4.48	–
All	1.19±0.49	1.93±0.66	9.05±2.72	19.79±5.81	184
Fine segmentation IP+SMEM					
F	1.77±0.59	2.61±0.72	9.75±2.66	22.09±5.57	–
HB	1.41±0.48	2.39±0.71	14.14±4.45	24.84±5.62	–
All	1.59±0.54	2.50±0.72	11.95±3.55	23.46±5.59	130

Table 5.2: LOO segmentation results with cVB dataset after constrained global initialization and fine segmentation (average±standard deviation). F: femurs, HB: hip bones. ASSD, ASRSD, MSD are in mm, VOE in %. Mean times are in sec for left and right pairs of bones (i.e. 4 bones), and include model loading.

fine segmentation based on PDMs improved with respect to the initialization, with an overall segmentation error at subvoxel accuracy (e.g., average ASSD of 1.12 mm). By comparing the different configuarions, RAA- and SSS-based configurations were statistically not distinguishable, although the RAA-based approach seemed to produce more accurate results (1.12±0.46 vs. 1.19±0.49). Conversely, both approaches based on PDM outperformed the approach based on shape memory: the overall ASSD for SMEM was 1.59±0.54 against 1.19±0.49 and 1.12±0.46 for configurations with SSS and RAA schemes, respectively. These differences were statistically significant (largest p-value: 2.5E-4) except for the left femur due to the fact that all approaches benefitted from an excellent initialization of this structure.

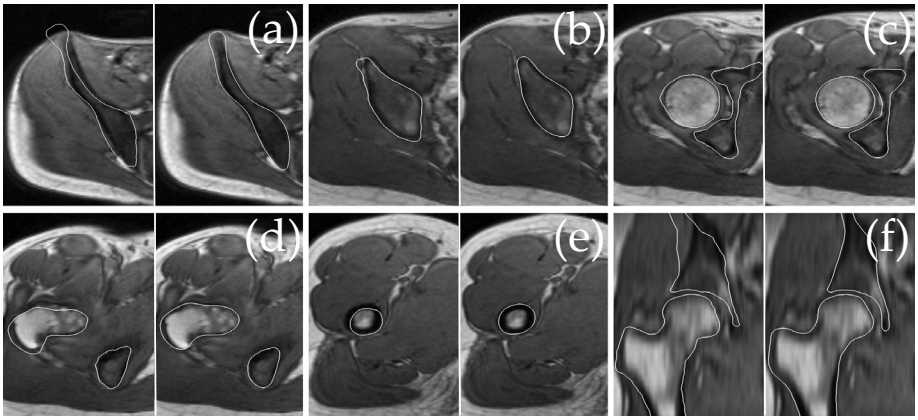


Figure 5.20: Examples of constrained initialization and segmentation results on slices of the cVB dataset with RAA+IP configuration. The subfigures a)-f) show the initialization on the left and the fine segmentation on the right. Mesh contours are overlaid in white.

An important evaluation is the performance of the algorithm in the presence of small FOV around the joint area, and more specifically, the femoral head (FH) and the acetabulum (AB) because their morphology changes may indicate potential joint disorders. We thus also computed

the average asymmetric surface distances (AASD, see Sec. 4.5.1.1) between the reference and the segmented results to assess errors in these regions. For the most accurate combination (i.e. RAA+IP), AASD for both sides were computed with the results of 1.35 ± 0.50 mm for AB and 1.45 ± 0.51 mm for FH. This difference between the hip bone and femur segmentation was also observed in a global manner (e.g., the overall ASSD for hip bones was 1.03 mm against 1.21 mm for femurs with the IP+RAA mode).

Table 5.2 also listed the execution times to initialize and segment a left and right pairs of bones (although a cVB trial only segmented one side (left or right) at a time). These time results show that the GI initialization was the most time consuming operation, as about 10 min were in average necessary to initialize the 4 bones. Furthermore, times for fine segmentation were similar in proportion to the experiments with VB datasets: IP+SMEM was the fastest approach, while IP+RAA was less time consuming in average than IP+SSS. However, a clear difference in execution times between trials with fully and partially visible bones was observed and was related to the different image FOVs of VB and cVB images.

Results with TRUEFISP (TF) datasets Table 5.3 presents the LOO segmentation results on the TF dataset after GI. No failed cases were reported in the GI procedure which produced initializations of satisfactory accuracy (ASSD of 1.80±0.56). For approaches based on PDM, improvements between the initialization and the fine segmentation were observed (e.g., 27.6% improvment for IP+RAA with TF) whereas the SMEM-based segmentation was clearly not able to yield improved and accurate results (overall ASSD jumped from 1.80±0.56 to 3.09±1.29). The best approach in terms of accuracy and robustness was the IP+RAA configuration (overall ASSD of 1.31±0.44) which statistically performed better than the IP+SSS combination (overall ASSD of 1.46±0.62) in segmenting hip bones (p-value < 0.023). The approach based on SMEM forces returned extremely poor results with errors statistically greater than those of the IP+RAA (largest p-value: 3E-17) and IP+SSS (largest p-value: 1.9E-7) approaches, regardless of the structure type.

Bone	ASSD	ASRSD	MSD	VOE	Time
Constrained global initialization (GI)					
F	1.61±0.71	2.28±1.15	11.19±5.73	19.42±7.20	–
HB	1.99±0.41	2.90±0.63	18.08±4.52	30.63±5.47	–
All	1.80±0.56	2.59±0.89	14.63±5.10	25.03±6.32	510
Fine segmentation IP+RAA					
F	1.31±0.61	1.80±1.03	8.67±5.12	16.31±6.05	–
HB	1.30±0.27	1.89±0.39	12.77±3.06	21.48±3.86	–
All	1.31±0.44	1.84±0.70	10.72±4.09	18.89±4.93	125
Fine segmentation IP+SSS					
F	1.28±0.57	1.73±0.96	8.18±5.06	15.93±5.64	–
HB	1.64±0.67	2.37±0.96	15.24±5.45	26.66±9.98	–
All	1.46±0.62	2.05±0.95	11.71±5.25	21.30±7.79	154
Fine segmentation IP+SMEM					
F	2.33±2.25	3.12±2.98	12.19±8.66	24.49±12.15	–
HB	3.84±0.62	6.92±1.46	38.51±6.02	45.68±5.36	–
All	3.09±1.29	5.02±2.01	25.35±7.14	35.09±8.55	62

Table 5.3: LOO segmentation results with TF dataset after constrained global initialization and fine segmentation (average±standard deviation). F: femurs, HB: hip bones. ASSD, ASRSD, MSD are in mm, VOE in %. Mean times are in sec and for left and right pairs of bones (i.e. 4 bones), and include model loading.

Despite the excellent results of the IP+RAA coupling (Fig. 5.21), the overall segmentation error was worse than that of cVB (e.g., ASSD of 1.31±0.44 for TF against 1.12±0.46 for cVB), especially when comparing the image resolution ($0.52 \times 0.52 \times 0.6$ mm) and quality (better signal-to-noise ratio) of TF images. This observation supports our expectation that a smaller FOV is more challenging as discussed later in Sec. 5.6.1.1. Nevertheless, the robustness of the IP+RAA segmentation was considered high since the overall ASSD standard deviation (0.44 mm) was below the average voxel size ($\approx$ 0.54 mm). The overall results were within ASSD error of 2 voxels, the ASSD errors of the acetabulum (AB) and femoral head (FH) regions were 0.9±0.22 and 1.18±0.45, respectively, and small errors had minor influences on the segmentation results in these regions as illustrated in Fig. 5.21(c) and (d). Regarding the execution times, the results exhibited the same patterns as previous experiments, where segmentations based on PDM were more time consuming than the SMEM-based approach.

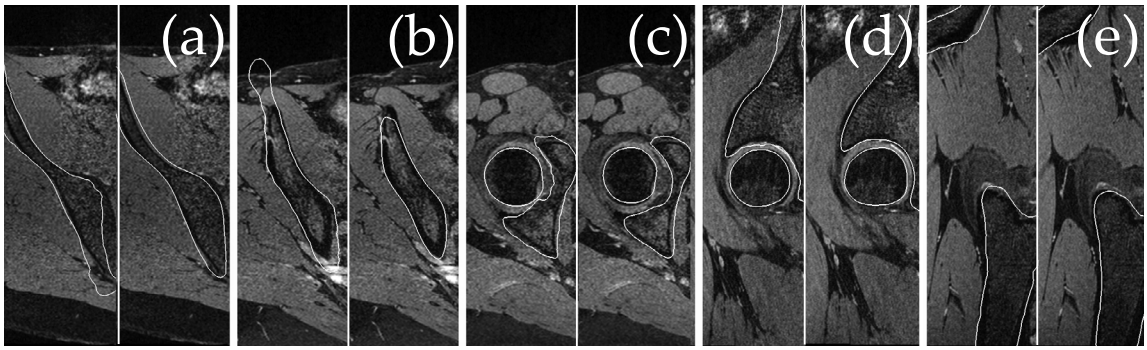


Figure 5.21: Examples of constrained initialization/segmentation results on slices of the TF dataset with RAA+IP configuration. For each subfigure a)-e), the initialization is on the left and the fine segmentation on the right. Mesh contours are overlaid in white.

5.5.3 Experiment 2 – Robust SSM and MRF Local Modeling

5.5.3.1 Robust SSM Performance

Skočaj *et al.* [SLB07] demonstrated that a robust PCA outperformed conventional PCA when using corrupted data, as the latter was unable to handle erroneous data. However, many of the tests they conducted only underpinned the good Specificity of the robust SSM. The performance of an SSM in the context of segmentation cannot be explained only by the statistical model, as there are many factors, such as appearance priors and evolution strategies, which come into play.

We therefore devised this experiment to investigate whether the mixing of non-corrupted S° and corrupted $S^{\bullet}$ shapes into a robust SSM (Section 4.3.2) would result in better segmentation results compared to those obtained with an SSM only built from S°. Ten VB datasets were randomly chosen as the testing data, and the remaining 33 datasets were used to build 3 different SSMs:

- A "full-SSM" built using all 33 reference shapes.
- A "sub-SSM" built from 12 shapes randomly chosen among the 33 reference shapes.
- A "mixed-SSM" created from the 12 reference shapes of the sub-SSM and 21 corrupted shapes.

The corrupted shapes were obtained from the segmentation of the cVB datasets in the previous experiment (Section 5.5.2.2), in which all points out of FOV were not reliable. These shapes were aligned using the weighted alignment described in Sec. 4.3.2.2.

For the mixed-SSM, we chose the mdEM robust PCA approach as it produced the best results in our evaluation experiments (Sec. 4.5.4). The initialization process described in Sec. 4.3.2.2 was used as a start to the robust PCA. The configuration IP+SSS was chosen since it exhibited excellent performance in the previous experiment (Sec. 5.5.2.1) with VB datasets. Furthermore, exactly the same landmark-based initialization was used, i.e. with identical landmarks placement. This provided a fair comparison with the previous experiment results.

Table 5.4 compares the performances of the segmentation on the VB dataset using full-, mixed- and sub-SSM performed after the landmark-based initialization. The table shows that the segmentation results with 33 complete shapes (ASSD 1.29±0.32 mm) were slightly worse than those of previous experiment (Sec. 5.5.2.1) in which 43 full shapes were considered (ASSD 1.21±0.35 mm, see Table 5.1). Although no statistical difference was established, this observation is consistent with the idea that larger training sets usually improve the SSM's segmentation efficiency. This idea is also supported by the results of the sub-SSM where the use of a smaller training set (12 shapes) produced significantly worse results compared to the full-SSM (p-value$<$ 0.04 for each bone type and side). As expected, the Generality of the sub-SSM was too weak to efficiently segment the shapes. The most interesting result is the comparison of mixed- and full-SSMs results which are not differentiable from a statistical viewpoint. This strongly suggests that the proposed combination of complete and corrupted shapes in the construction of SSM can be very robust and efficient.

Bone	ASSD	ASRSD	MSD	VOE
Full-SSM				
F	1.35±0.26	2.13±0.34	9.71±1.81	17.42±2.11
HB	1.23±0.38	2.00±0.51	10.15±2.02	23.52±4.57
All	1.29±0.32	2.07±0.43	9.93±1.91	20.47±3.34
Mixed-SSM				
F	1.33±0.27	2.10±0.35	9.16±1.79	17.33±2.41
HB	1.24±0.32	2.03±0.46	10.42±2.86	23.61±3.81
All	1.29±0.30	2.06±0.41	9.79±2.32	20.47±3.11
Sub-SSM				
F	1.53±0.29	2.39±0.38	10.47±1.87	19.73±2.48
HB	1.46±0.41	2.33±0.58	11.97±3.21	27.11±4.69
All	1.49±0.35	2.36±0.48	11.22±2.54	23.42±3.59

Table 5.4: Segmentation results on VB dataset with full-, mixed- and sub-SSM after landmark-based initialization on 10 datasets (average±standard deviation). F: femurs, HB: hip bones. ASSD, ASRSD, MSD are in mm, VOE in %.

5.5.3.2 MRF Local Modeling Evaluation

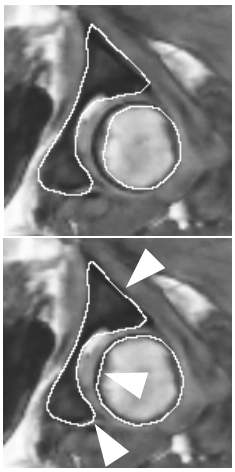
We studied the performance of forces based on the local shape variation modeling by MRF as presented in Sec. 5.3.3.4. Two situations were investigated. In the first case, the MRF prior was coupled with a PDM trained with all available training samples (except the current one being

tested) whose efficiency was demonstrated in IP+SSS experiments. For comparison purposes, the IP+SSS configuration without MRF was run on the same data with identical conditions (initialization, training data, etc.). In the second case, we used the previous sub-SSM only composed of 12 training samples, whose segmentation results were mitigated due to the relative small size of the training set. All experiments were conducted with the same 10 subjects of the previous experiment evaluating the robust SSM performance.

As presented in Sec. 5.3.3.4, the exploitation of MRF priors is equivalent to a MAP problem which can be solved with a dynamic evolution using image f^{im} and regularization f^{mrf} forces. When MRF priors are used it is no longer necessary to simultaneously use shape (f^{smem} or f^{pdm}) or smoothing forces, since the force f^{mrf} already achieves smoothing and shape regularization. In our experiments, we used MRF-based forces in an interleaved manner with the SSM multi-resolution approach: after all the scheduled iterations of a resolution level, the current model state was set as the reference shape $\boldsymbol{X}$ and only MRF forces f^{mrf} and image forces (IP-based and gradient forces as in the IP+SSS configuration) were used to drive the model evolution (Sec. 5.3.3.4). The iteration schedule for the MRF evolution was set to 50/20/5. The last resolution level did not rely on MRF priors to better capture fine details not expressed by the priors. Values of MRF parameters (Sec. 5.3.3.4) were $1/\eta^2 = 1$ and $1/\sigma^2 = 0.25$ as suggested in the work of Kervrann and Heitz [KH98]. The weighting parameter $\beta = \alpha^{\text{mrf}}$ was progressively modified to give an increasing importance to local variations based on the shape resolution, by adopting the schedule 1.5/2/3. Results are reported in Table 5.5(a) for both scenarios.

Bone	ASSD	ASRSD	MSD	VOE	Time
IP + SSS (SSM built with 42 shapes)					
F	1.30±0.26	2.03±0.33	9.02±1.78	17.16±2.11	–
HB	1.25±0.42	2.06±0.57	10.64±2.54	23.75±5.19	–
All	1.28±0.34	2.05±0.45	9.83±2.16	20.46±3.65	277
IP + SSS (SSM built with 42 shapes) + MRF					
F	1.29±0.21	2.03±0.27	8.65±1.46	17.12±1.87	–
HB	1.29±0.31	2.09±0.44	9.60±2.11	20.53±3.50	–
All	1.29±0.26	2.06±0.35	9.13±1.79	18.83±2.68	366
IP + SSS (SSM built with 12 shapes) + MRF					
F	1.36±0.25	2.10±0.32	8.65±1.15	18.03±2.26	–
HB	1.34±0.33	2.18±0.47	10.65±2.22	25.33±4.30	–
All	1.35±0.29	2.14±0.39	9.65±1.68	21.68±3.28	361

(a)



(b)

Table 5.5: (a) Segmentation results on 10 VB datasets with MRF-based priors (average±standard deviation). F: femurs, HB: hip bones. ASSD, ASRSD, MSD are in mm, VOE in %. Mean times are in sec for left and right pairs of bones (i.e. 4 bones), and include model loading time. (b) An example of before (top) and after (bottom) MRF-based evolution, where white arrows indicate a better delineation of fine details.

In the first scenario involving a large training shape set, the use of MRF-based priors does not seem not to be beneficial as a sensible increase of the errors was observed (e.g., ASSD: 1.28±0.34 to 1.29±0.26) although this difference could not be statistically proven. However, in the case

where the training dataset was not very big (12 shapes), the MRF-based evolution significantly improved the results. In fact, the ASSD for both bone types decreased from 1.49±0.35 (Sub-SSM results in Table 5.4) to 1.35±0.29, where statistical difference was only detected for femurs (p-value < 3.8E-3). It was noticed that MRF often refined the delineation of anatomical features, such as the femoral head or the acetabulum (Fig. 5.5(b)), thanks to its local modeling of variations.

Timing results showed a higher computational cost of the MRF-based approaches which was mostly explained by a higher number of iterations and the additional computation of energy terms related to the MRF local modeling.

5.5.4 Experiment 3 – Dual-Posture MRI Segmentation

5.5.4.1 Description

In this experiment, we exploit the MRI protocol, denoted as "dual-posture" MRI, which acquires image pairs of lower limbs with VIBE protocol consisting of normal (supine) and extreme (split) postures, where the split posture reveals the maximum extent of the dancer's vertical lower limb motion as presented in previous Sec. 5.5.1.1. Such imaging data provide rare insights into the movement of the dancers and enable new observations from the effects of extreme motion and its impact on surrounding anatomical structures [CAVMT09].

We aim to segment the femur at the two postures. The supine posture can benefit from a priori knowledge in terms of appearance (intensity distribution within and around tissues of interest) and organ poses, which facilitates the use of automated methods of bone segmentation. However, due to the largely varying orientations and morphologies between the subjects, and the strong soft tissue deformations evident around the joint areas, the split posture cannot benefit from existing algorithms that are optimized for the standard supine acquisition.

We propose an automatic method to segment the dual-posture MRI where the results from the segmentation of the supine MRI are used to assist the segmentation of the split MRI of the same subject. Our approach exploits the characteristics of the femur being a hard structure unaffected by changes in postures between the supine and the split. Our proposed segmentation methodology works in two phases – Supine and Split postures – as shown in Fig. 5.22.

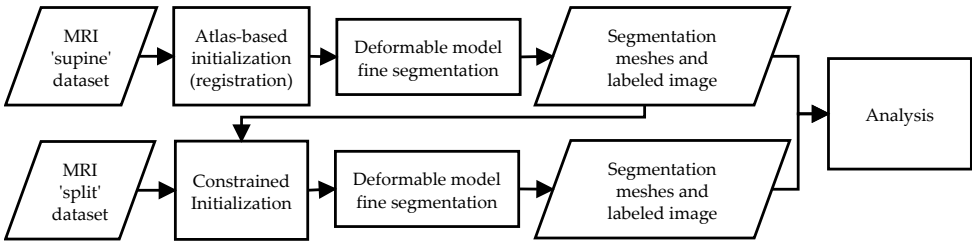


Figure 5.22: Dual-posture MRI segmentation overview.

For the supine segmentation, the atlas-based registration (Sec. 5.2.2.2) was used as an initialization for our deformable model method. The derived result was then used for the split MRI

segmentation as an initialization parameter in the constrained global initialization (Sec. 5.2.3). In this initialization of the split MRI, shape parameters were not varied since the shape femur was the segmentation result of the supine MRI acquired with the same subject. Nevertheless, a deformable model was again used for further refinements after the initialization because the supine femur might not have been perfectly delineated. During the deformable-based segmentation, the IP+SSS scheme of the first experiment (Sec. 5.5.2.1) was once again selected.

5.5.4.2 Results

Our proposed algorithm was evaluated with 10 arbitrarily chosen dancer datasets. The atlas was randomly chosen among the remaining 33 datasets. SSMs were built with all dancer models except those of the model being tested. All results were qualitatively and quantitatively measured against the ground truth.

Supine MRI segmentation Averaged results of the supine posture segmentation are shown in Table 5.6 for both the registration (initialization) and deformable model-based fine segmentation. All the data were registered using the same atlas and registration parameters defined in Sec. 5.2.2.2. The registration errors were high and inconsistent (e.g., average ASSD of 5.86±2.15 mm for both sides), however, the results in all cases were sufficient for its use as an initialization of the subsequent fine segmentation. Indeed, the fine segmentation resulted in significant improvements, e.g. the VOE of 50.13% dropped to 17.98%.

Bone	**ASSD**	**ASRSD**	**MSD**	**VOE**
Atlas-based registration				
F	5.86±2.15	7.90±3.07	26.78±9.32	50.13±9.71
Fine segmentation				
F	1.48±0.36	2.36±0.59	10.37±2.23	17.98±2.66

Table 5.6: Supine overall registration and fine segmentation errors (average±standard deviation) averaged for left and right femurs (F) from 10 subjects. ASSD, ASRSD, MSD in mm. VOE in %.

The errors between the registration and fine segmentation results were in average improved by 75%, and the results were significantly different (p-value=1E-4 for left and right femurs). More importantly, the variance between the subjects had even larger improvements of 83% thus suggesting the consistency and the reliability of the fine segmentation, even with not so accurate initializations.

The quantitative findings were confirmed by the visual results, where Fig. 5.23 illustrates the accuracy of the proposed dual-posture segmentation, exemplified by subject 1 for which the registration error was the highest (ASSD of 10.29 mm) and subject 10 who had the lowest initialization errors (ASSD of 3.73 mm). The visual differences between the fine segmentation and the ground truth are minor for both subjects. The overall performance of the fine segmentation results was ASSD of 1.48±0.36 mm. Comparative examples between the fine segmentation and the ground truth are shown in Fig. 5.24.

In order to study the largest error range, we compared all the segmentation models to their

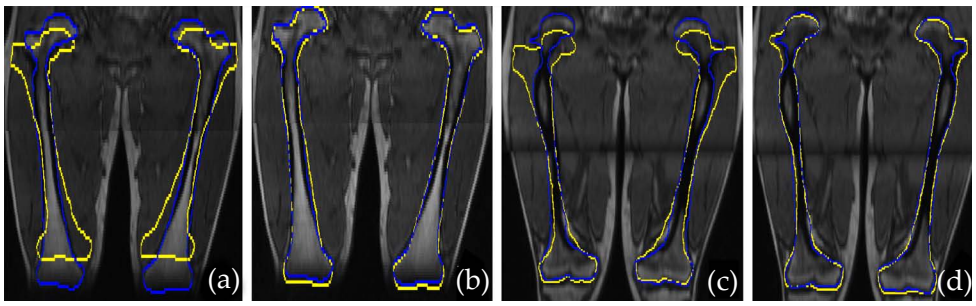


Figure 5.23: Supine registration vs. ground truth. (a), (c): registration (yellow) vs. ground truth (blue) results for subject 1 and 10 respectively in supine MRI. (b), (d): fine segmentation (yellow) vs. ground truth (blue) results for subject 1 and 10 respectively.

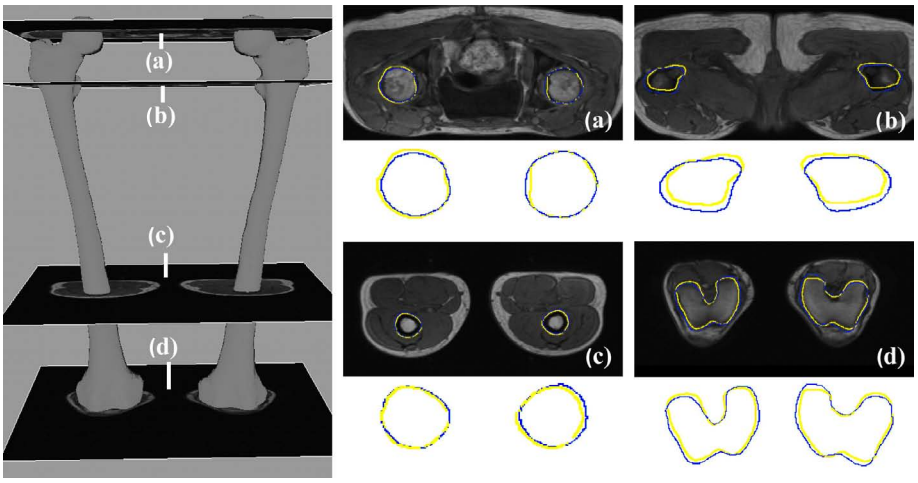


Figure 5.24: Comparison between the ground truth (light yellow) and the fine segmentation (dark blue) result in a supine MRI data with zoomed overlays.

ground truth counterparts using the point-to-surface distance d_S as we did in experiment 1 (see Fig. 5.18). This produced an error distribution over the bone, which was color-mapped for quantitative inspection (Fig. 5.25(a)). Largest errors were localized in the regions of the lesser (LT) and greater (GT) trochanters, and particularly the patellar surface (PS) and the condyles (C). Nevertheless, an error distribution analysis showed that 80% of errors were below 2 mm meaning that large errors did not excessively cover large portions of the bone.

Split MRI segmentation The results of split segmentation are reported in Table 5.7 with visual comparisons in Fig. 5.26.

Bone	**ASSD**	**ASRSD**	**MSD**	**VOE**
Constrained global initialization (GI)				
F	2.22±1.10	3.24±1.43	12.28±4.29	24.80±9.21
Fine segmentation				
F	1.30±0.26	2.07±0.35	9.53±1.75	16.83±2.62

Table 5.7: Split overall registration and fine segmentation errors (average±standard deviation) averaged for left and right femurs (F) from 10 subjects. ASSD, ASRSD, MSD in mm. VOE in %.

These results highlight the capabilities of the constrained global initialization (GI) to yield initial

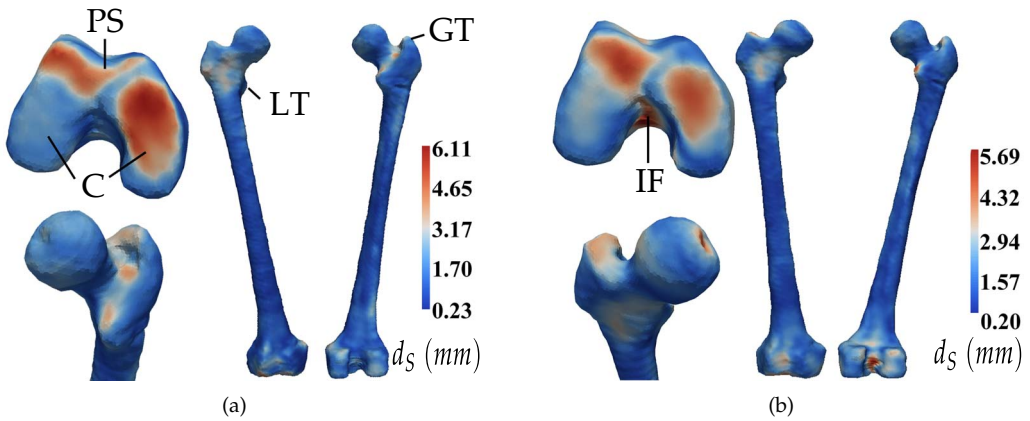


Figure 5.25: Average point-to-surface distance d_S error distribution over a standard right femur shape. A color mapping of the error on the shape reveals the largest errors on the regions of the patellar surface (PS), the greater (GT) and lesser trochanters (LT), condyles (C) and intercondylar fossa (IF) for the (a) supine and (b) split final results.

models that matched closely to the ground truth (averaged ASSD of 2.22±1.10 mm). As in the supine data, the fine segmentation again significantly improved the overall ASSD, from 2.22±1.10 to 1.30±0.26 mm (p-value = 0.024). As with the supine results, the split segmentation resulted in satisfactory agreement to the ground truth, in low standard deviation (0.26 mm), which highlighted the precision of the approach, and in a similar error distribution (Fig. 5.25(b)). Compared to the supine results, the overall accuracy was better but some high errors were observed in other highly localized regions, such as the intercondylar fossa (IF, Fig. 5.25(b)).

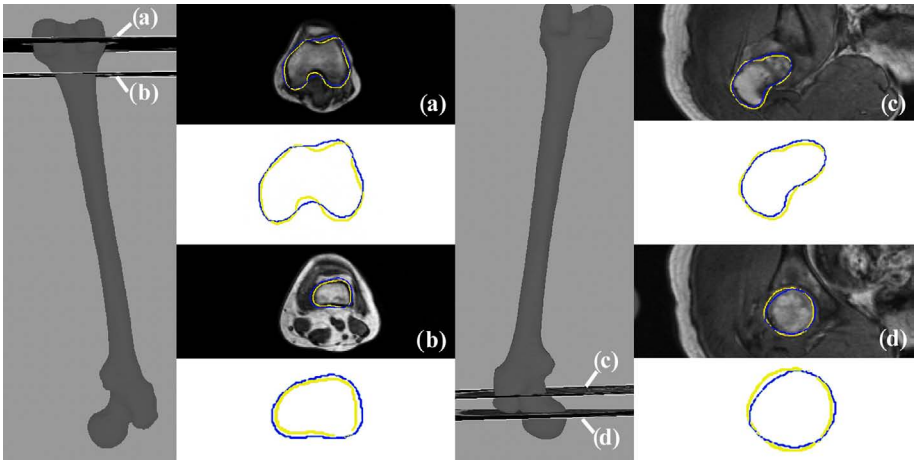


Figure 5.26: Comparison between the ground truth (light yellow) and the fine segmentation (dark blue) in split MRI with zoomed overlays. (a)-(b) and (c)-(d) are respectively from a right and left split segmentations.

Execution Time The proposed algorithm required in average 11 min to process the dual-posture MRI (segmenting both the left and the right femurs in supine and split images). The measured time encompassed both the registration (non-rigid and model-to-image) and deformable model-based segmentations.

5.5.5 Experiment 4 – Coupled Registration-Segmentation

5.5.5.1 Description

In our context of musculoskeletal imaging and analysis, the acquisition of MRI images of the hip joint at multiple levels of detail (LOD) is often used to identify bone changes and soft tissue abnormalities, which might explain the development of arthritis. In this protocol, the low resolution VB MRI that covers a large field of view (FOV) is acquired to image the whole hip joint structures whereas the high resolution TF MRI with limited FOV and centered at the hip joint is acquired to reveal the fine details of the soft tissues (See Fig. 5.16 in Sec. 5.5.1.1). In this experiment, we name this acquisition protocol "multiple LOD MRI" and focus our study on the femur.

Since VB and TF images may not be accurately aligned, the use of the coupled registration-segmentation approach (Sec. 5.3.4) is appropriate to segment and register the bone structure. Furthermore, the use of a "richer" intra-subject information expressed by multiple aligned images is expected to improve the segmentation. Compared to experiment 1 (Sec. 5.5.2), the segmentation of both the TF and VB images is done simultaneously. In our experiments, the femur is a rigid structure and all the data is intra-subject, therefore the unknown mappings g_i to compute (Sec. 5.3.4) are rigid transforms. The blending parameter λ, which interpolates segmentation with alignment results, was linearly increased from 0.7 to 1 throughout the process.

Fifteen subjects were randomly selected from the 43 subjects database. For each subject, three different types of multiple LOD MRIs were processed with our proposed coupled registration-segmentation framework: (i) "supine VB + supine TF images", (ii) "supine VB + split TF images" and (iii) "supine VB + supine TF + split TF images". For comparative analysis, the VB was also solely segmented using the same deformable model as in the coupled approach. This is referred to as a *single segmentation*. Femur models were initialized in supine VB images with the landmark-based approach (Sec. 5.2.2.1), while the globally constrained initialization (GI) (Sec. 5.2.3) was applied on supine and split TF datasets. Based on the results of experiment 1 (Sec. 5.5.2), the IP+SSS and IP+RAA strategies were selected for the deformable model-based segmentation of the VB and TF images, respectively.

5.5.5.2 Results

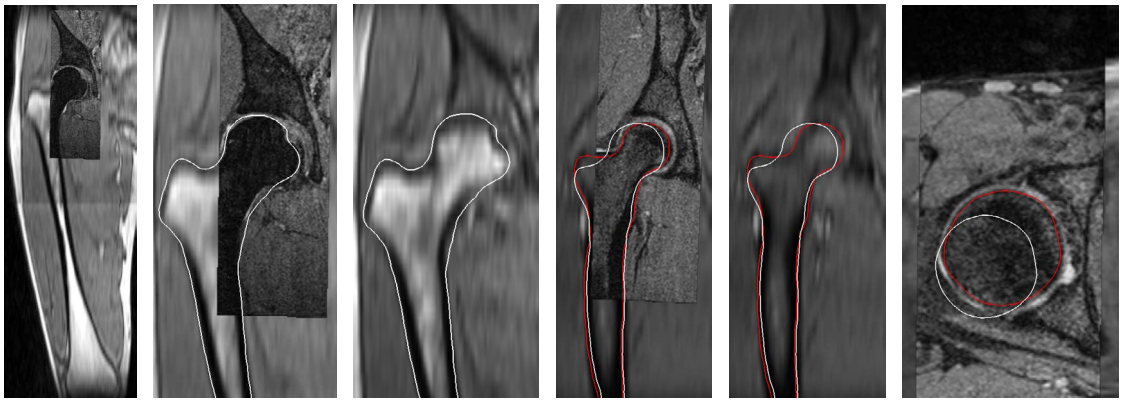
Results are reported in Table 5.8 for all the different multiple LOD MRIs. All segmentation results were compared to the ground truth in supine VB images.

In the first multiple LOD dataset, subjects were in supine position and exhibited low flexion/abduction/adduction amplitudes of the thigh. An example of result of our coupled approach is shown in Fig. 5.27(a) and (b)), where both supine VB and TF images have been automatically aligned (superimposed together) and segmented, respectively. Here, we can see that the alignment appears to be satisfactory and the segmentation shape correctly contours the bone in both VB and TF MRIs. The surrounding soft tissues, i.e. muscles, were slightly misaligned,

Bone	ASSD	ASRSD	MSD	VOE	Time
Supine VB + supine TF					
F	1.05±0.21	1.79±0.30	7.99±2.09	13.90±1.87	156
Supine VB + split TF					
F	1.07±0.20	1.81±0.28	8.07±2.15	14.02±1.89	160
Supine VB + supine TF + split TF					
F	1.05±0.20	1.79±0.28	8.11±2.12	13.78±1.80	245
Supine VB single segmentation					
F	1.21±0.60	2.02±0.89	8.84±3.98	15.19±4.77	60

Table 5.8: Coupled registration-segmentation accuracy and timing results with multiple LOD MRIs (average±standard deviation). F: femurs. ASSD, ASRSD, MSD are in mm, VOE in %. Mean times are in sec and averaged for left and right femurs.

although visibly accurate. This was expected since the alignment was based on a rigid mapping of the femur bone and thus did not take into account the soft tissue deformations. In Fig. 5.27(c), the segmentation result using only the VB is presented. Visual differences to the coupled approach results of Fig. 5.27(b) are very subtle.



(a) (b) (c) (d) (e) (f)

Figure 5.27: Coupled registration-segmentation results with supine VB and TF MRIs. An example of alignment resulting from our coupled approach is shown in (a) with corresponding white segmentation contours overlaid in (b). In comparison, single segmentation results of the VB image are shown in (c). In another example, subfigures (d) and (e) show the overlay of segmentation results of the single (white) and coupled (red) approaches in the VB MRI, with and without the superimposed aligned TF image, respectively. Subfigure (f) is in the axial direction and clearly shows in this example the large errors of the single segmentation.

However, as reported in Table 5.8, the overall errors of the coupled approach (ASSD of 1.05±0.21) were lower than those of the single segmentation (ASSD of 1.21±0.60), although no statistical difference was established. In fact, the coupled registration-segmentation approach was in some cases much more effective as depicted in Fig. 5.27(d)–(f), where the single segmentation yielded very poor results compared to those of the coupled approach. In general, the coupled approach was markedly more accurate in the segmentation of the finer details that were only available in the TF (Fig. 5.28(a)–(c)). The point-to-surface distance error was calculated between the ground truth and the segmentation results for both the single (Fig. 5.28(d)) and the coupled (Fig. 5.28(e)) approaches, averaged among all the subjects, and finally mapped on an arbitrary reference shape.

These distance errors better reveal that the finer information from the TF contributed to a better segmentation (e.g., fine details of the fovea capitis).

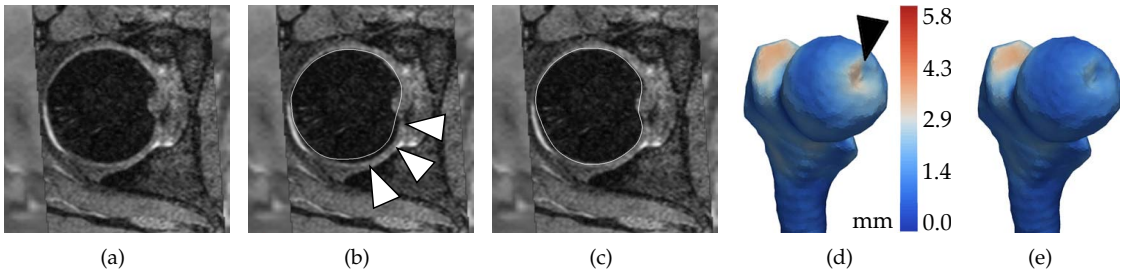


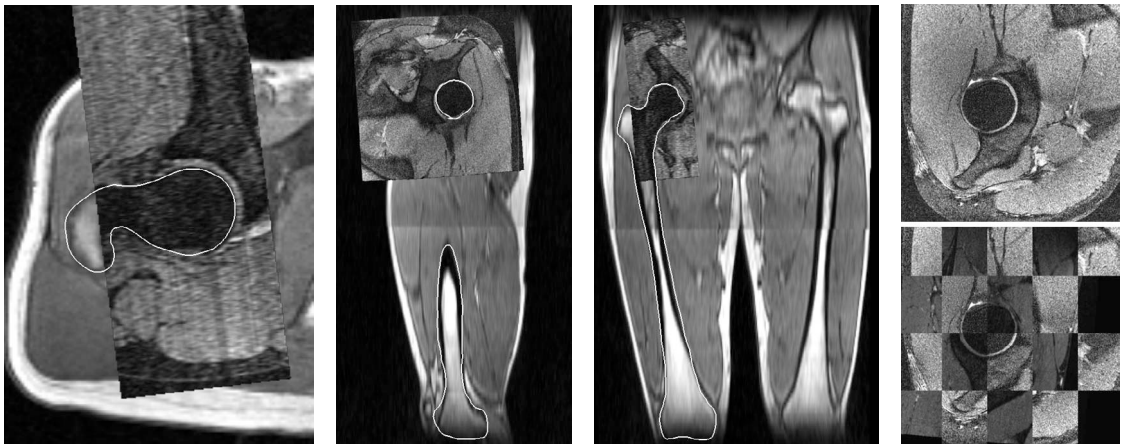
Figure 5.28: Fine detail delineation by coupled registration-segmentation. (a) A close-up transversal view of a femoral head visible in the TF image. Results from (b) single and (c) coupled segmentation are overlaid as white contours. The white arrows indicate poorly segmented areas in (b) which are markedly corrected in (c) using the additional information from the TF MRI. Average point-to-surface distance errors computed from the single (d) and coupled (e) segmentation results are mapped on an arbitrary femur shape. The black arrow indicates the Fovea Capitis for which larger error differences are observed.

In the presented segmentation trials, the position of the leg was neutral for both the VB and the TF MRIs. However, in clinical practice, it often happens that stronger alignment differences between the image acquisitions are observed due to patient movement between the scans. Large leg rotation patterns make the alignment more challenging and locally affect the intensity distribution around the bones. In order to analyze the capability of our coupled framework in the presence of large movement between the scans, we tested it with pairs of supine VB and split TF images for each of the 15 subjects.

Figures 5.29(a)–(c) illustrate the strong misalignment that needed to be recovered between the bone extracted in the VB and the TF MRIs. The registration of the femur performed well, where we can clearly see the shape outline of the femur between the VB and the TF. The checkerboard visualization (Fig. 5.29(d)) also highlights the accuracy of this registration. Similarly to the previous results, the quantitative measures of the femur segmentation were almost identical as reported in Table 5.8 (ASSD of 1.21±0.60 and 1.07±0.20 for single and coupled approaches respectively). This suggested that our framework is robust in regards to the mis-alignment orientation of the LOD MRIs as long as these MRIs share a common SOI.

We further conducted tests to evaluate our coupled approach in the case where more than two MRI images were available. When our framework was applied to multiple LOD MRIs consisting of a VB and two TF (supine and split) MRIs, all three images were successfully aligned and the segmentation accuracy was in line with all the other experiments (single and coupled results with ASSD of 1.21±0.60 and 1.05±0.20, respectively). In this case, the use of an additional TF image did not yield any significant improvement of the accuracy, the obtained error being already excellent with the previous VB+TF configuration.

In any case, the coupled approach seemed to produce better results than those of single segmentation approaches reported in previous experiments, although the number of test datasets varied. For instance, the best ASSD for femurs in VB images in experiment 1 (Sec. 5.5.2) and



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

Figure 5.29: Segmentation and alignment result of a multiple LOD MRI composed of supine VB and split TF images shown in (a) coronal (b) transversal and (c) sagittal view slices. In this example, the subject performed a large rotation movement of the leg (split). (d) Top: slice of the TF image centered on the femoral head; bottom: the VB image is resampled into the TF space and overlaid with checkerboard visualization to illustrate the quality of the registration.

experiment 3 (Sec. 5.5.4) were respectively 1.21±0.53 (43 test datasets) and 1.48±0.38 (10 test datasets), whereas the coupled approach had an error of 1.05±0.20 (15 test datasets).

The segmentation of a single VB image only required around 60 sec. As expected, the coupled registration-approach was more time consuming. The coupled registration and segmentation of 2 femur models took in average 158 sec for the scenarios involving one supine VB and one TF image (split or supine). For comparison, the simultaneous but independent segmentation of one VB and one TF image was estimated at 90 sec. This significant difference was due to the additional operations involved in the coupled registration-segmentation approach: the generalized Procrustes alignment (Sec. 5.3.4.2), the computation of inter-reliability terms (Sec. 5.3.4.3) and the blending performed after each segmentation iteration. As a result, the configuration involving one VB supine image with two TFs, split and supine, presented the largest computation time (245 sec).

5.5.6 Experiment 5 – GPU-based Segmentation

5.5.6.1 Description

In this experiment, we compare the accuracy and speed of our GPU-framework (Sec. 5.4.1) against our CPU-based implementation of discrete deformable models. For comparison purposes, only the features common to both implementations were used, thus collision detection and sophisticated shape priors (i.e. statistical shape model and MRF-based modeling) were omitted from the CPU implementation. We used a subset of 28 VB MRI images from our database in the supine position. Compared to previous CPU-based experiments, only the first three shape resolutions were used (denoted here as coarse, medium and fine mesh resolutions) and the iteration schedule was hence set to 300/100/20, respectively. Both left and right pairs of femur and hip

bone were simultaneously segmented, i.e. 4 fully visible bones as in experiment 1 (Sec 5.5.2.1).

5.5.6.2 Results

Results are reported in Table 5.9, with computation times that were evaluated per mesh resolution and did not account for loading, saving and rendering of the meshes.

	Coarse	**Medium**	**Fine**
$ASSD_{GPU}$ (mm)	1.93±0.50	1.66±0.43	1.62±0.44
$ASSD_{CPU}$ (mm)	2.07±0.70	1.60±0.63	1.58±0.63
$time_{GPU}$ (s)	1.85 (0.006)	1.27 (0.012)	0.42 (0.021)
$time_{CPU}$ (s)	44.9 (0.15)	57.2 (0.57)	29.9 (1.49)
#iterations	300	100	20
#vertices	2656	10624	42496

Table 5.9: GPU- vs. CPU-based MRI Hip joint bone segmentation results for each mesh resolution level: accuracy error ($ASSD$ in mm) for GPU and CPU; times for GPU and CPU with time/iteration in parenthesis; #iterations denotes the number of iterations and #vertices is the total number of vertices for all meshes.

The GPU approach was consistently about $25-70\times$ faster than the CPU version to execute a single time step. As expected, this difference became more significant when the number of vertices increased, as e.g. the CPU needed 1.49 sec to process 42K vertices while 21 ms were only necessary for the GPU implementation. These figures also show that the parallelization was not fully performed at the vertex level as an increase of the vertices yielded a non-negligible increase of the time to perform an iteration. Nevertheless, this time did not scale linearly with the number of vertices (it seems to double when the number of vertices is quadrupled) which highlights the presence and good performance of the underlying parallelization. Most importantly, the update frequencies of the GPU approach were consistently above the minimum 10Hz of refresh rates required for interactivity.

A time distribution analysis of the different steps of the simulation is given in Fig. 5.30(a). It reveals that the force computation accounted for approximatively 99% of total time per iteration, while mesh update and numerical integration contributions remained below 1%. Similarly, the time distribution per force type, illustrated in Fig. 5.30(b), shows that image forces (IP- and gradient-based forces) were the most time consuming forces and that IP-based force clearly dominated with its 93% of the total force computation time. Furthermore, the computation time of image forces was dependent on the image size as shown in Fig. 5.30(c) and 5.30(d), where the simulation was performed with three different image sizes ($257\times258\times140$, $514\times516\times280$ and $771\times774\times420$). The size $257\times258\times140$ was the average original size of the images used in the experiment, while the other two sizes were artifically obtained by supersampling the original datasets. For both types of forces, the bigger the image was, the longer the computation took.

In terms of accuracy, both CPU and GPU approaches returned similar results. Indeed, the final ASSD error of the CPU segmentation was 1.58 ± 0.63 against 1.62 ± 0.44 for the GPU implementation. Similarly, errors between both approaches were very close at the end of each iteration sequence which exploited a given mesh resolution. These similar accuracy results highlight the

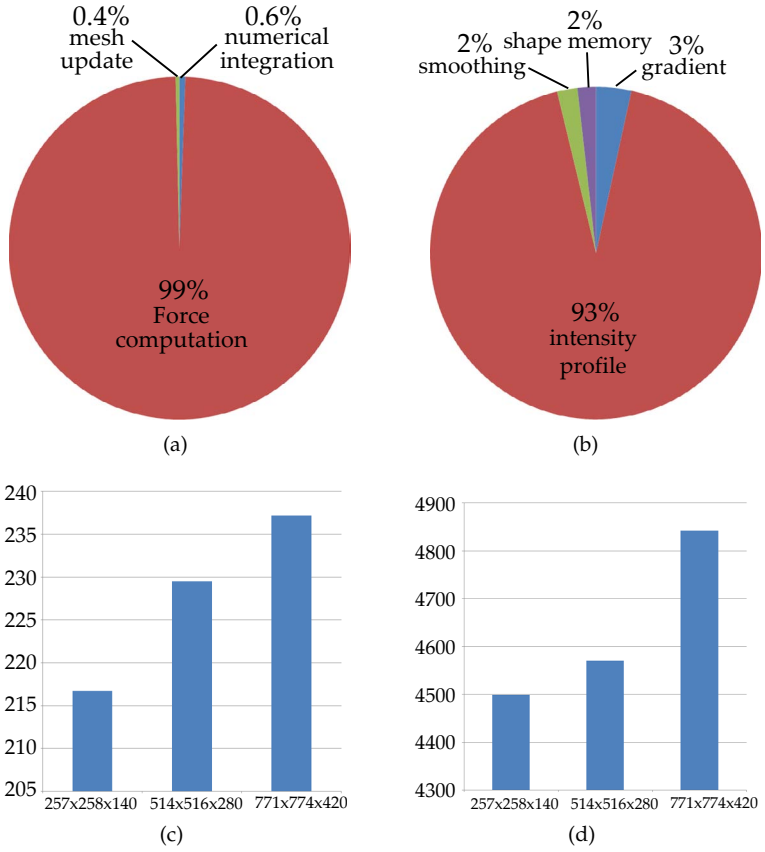


Figure 5.30: GPU time distributions and results. (a) Average time distributions per type of simulation step. (b) Average time distributions per force type. Subfigures (c) and (d) depict the total time in sec spent by gradient- and IP-based forces in function of image size ($W \times H \times D$), respectively.

correct implementation of the segmentation into the GPU formalism. A visual inspection (Fig. 5.31) confirmed a satisfactory segmentation in most bony regions.

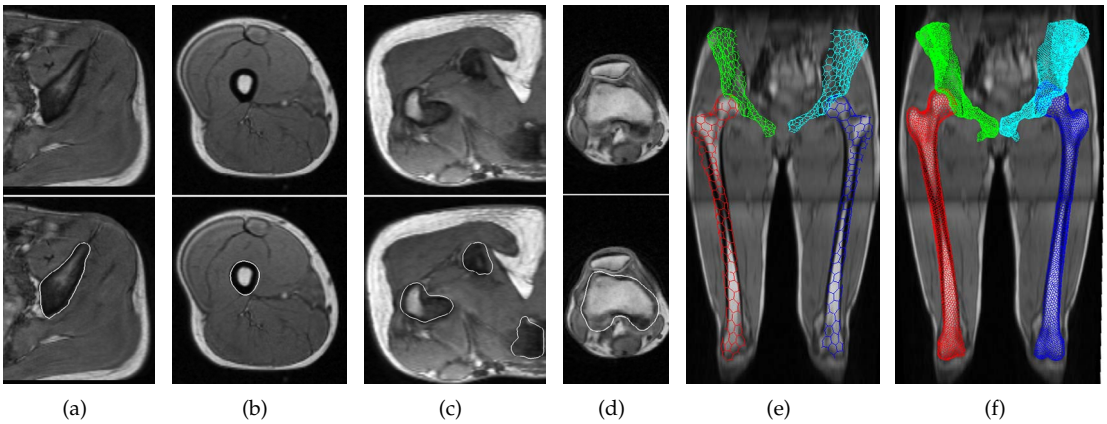


Figure 5.31: GPU-based MRI hip joint bone segmentation examples: from (a) to (d) axial slices are shown without and with white mesh overlays on top and bottom of the subfigure, respectively. In (e), a coarse mesh is initialized at the beginning of the segmentation, and in (e) the final result is shown with meshes at their higher resolution.

Nevertheless, the current GPU approach suffered from mesh inter-penetrations observed in the joint area (Fig. 5.32(a)). However, thanks to the new possibilities of higher update frequencies of the GPU approach, attraction points (Sec. 5.4.1.7) could be easily added in real-time to correct bad segmentation results as shown in Fig. 5.32(b), where inter-penetrations were avoided improving the ASSD about 9% (1.52 to 1.40 mm) for this particular subject.

(a) (b) (c)

Figure 5.32: Correction of bone segmentation by attraction points (APs). (a) Result without APs where the red arrow points at mesh inter-penetrations. (b) Corrected result with APs, which are displayed as (c) small spheres in a 3d view with a transparent hip bone.

5.5.7 Experiment 6 – Collaborative Segmentation

5.5.7.1 Description

In order to illustrate the concepts of the proposed medical collaborative system presented in Sec. 5.4.2, two application scenarios were designed and evaluated:

1. **"Teacher-student" scenario**: a teacher shows to students the reading of radiological images and the concept of semi-automatic segmentation for ROI annotation on the images as illustrated in Fig. 5.33. In this scenario, all the students observe the segmentation evolution as the teacher initiates the segmentation and changes the viewing slice across the dataset. Then the teacher shows how to manually and locally correct the segmentation evolution by inserting various constraint points. Once this simple principle is explained, students are invited to interact with the segmentation process. The teacher is assigned greater weight (i.e. weight α^{cp} in the constraint point-based force of Sec . 5.3.2.2) to the constraint points than the students, thus, enabling the teacher to override students' constraints.
2. **"Expert-expert" scenario**: two (or more) experts are segmenting the same dataset. Each of them can monitor and modify the segmentation in different parts of the dataset hence expediting the segmentation process. As with the "teacher-students" scenario, when experts work in the same slice, weights may be allocated based on the experience of the experts, thus giving more priority on the segmentation constraints to the more experienced user.

In the first scenario, user views were synchronized with the teacher view, while asynchronous navigation was granted to experts in the second scenario (Sec. 5.4.2.4). The first scenario was

Figure 5.33: Illustration of the teacher-student collaborative segmentation scenario. A teacher (center) uses her laptop to monitor and guide the two students in a collaborative segmentation session, where one student is using a portable device (left), and the other is using a conventional workstation (right).

evaluated in the muscle segmentation context [GMT10], where 21 muscles were simultaneously segmented. In this experiment, 20 users joined in a collaborative session and the teacher-student behavioral pattern was simulated, i.e. most of the time one user explored and interacted with the segmentation process while the others only observed. The evaluation of the expert-expert mode was performed with a hip joint bone segmentation experiment based on our deformable model, in which two experienced users collaboratively and concurrently segmented the 4 bones in the image.

For each experiment, raw image slices (i.e. without compression optimizations) were used in the transmission for testing purposes, and for performance measurements, only orthogonal image slices of the volume were available (i.e. no oblique viewing plane of the MRI dataset). We measured the average number of updates per second (ups in Hz), averaged over 100 segmentation iterations, at different levels:

- Segmentation: ups of an iterative segmentation step.
- Client update: ups of the user client update. A client update was triggered every time a new slice with corresponding overlay (mesh contours and constraint points) was received and prepared for rendering. The time to receive an answer to a request from the server was hence also considered here. Since the rendering was performed in another thread, the ups on the rendering was not representative of the client responsiveness and was not measured at this level.

The server was run on a dedicated PC, while clients were uniformly distributed over 4 other workstations (hence 5 clients per PCs in the teacher-student scenario). All PCs were equipped of a single core Pentium 4 running at 3.4 GHz with 2 GB or RAM.

5.5.7.2 Results

Simulation results are shown in Table 5.10.

These results show that our framework did not expose any significant degradation in system responsiveness and preserved satisfactory interactive performance of users in varying conditions. Measured ups at the client level was acceptable ($\approx$ 4 Hz) but not always optimal for collaborative editing. This was related to the speed of the segmentation process which bound the overall ups. The complexity of the segmentation, in terms of memory consumption and execution speed

	Teacher-student	Expert-expert
3D image size	$260 \times 511 \times 242$	$483 \times 358 \times 270$
Number of models	21	4
Segmentation iteration ups	3.8	4.6
Client update ups	3.5	4.2
Number of clients	20	2

Table 5.10: Results of the collaborative segmentation experiments reporting the ups in Hz with varying number of users, models and image size ($W \times H \times D$ voxels), averaged over 100 iterations.

was mostly related to the size of the segmented image and the number of models that were simultaneously segmented. These factors were expected to also have an impact on the system as the image data and the model contours (overlaid on the slice at the client side) needed to be sent over the network. Additionally, the number of clients introduced additional load at the server side. Faster version of the segmentation, such as our highly interactive GPU approach (Sec. 5.5.6), coupled with more powerful computational resources are expected to greatly improve the interactivity. Finally, the client was very light in terms of memory usage (< 1 MB) and CPU consumption (< 1%) which would greatly facilitate its usage on portable devices.

5.6 Discussion

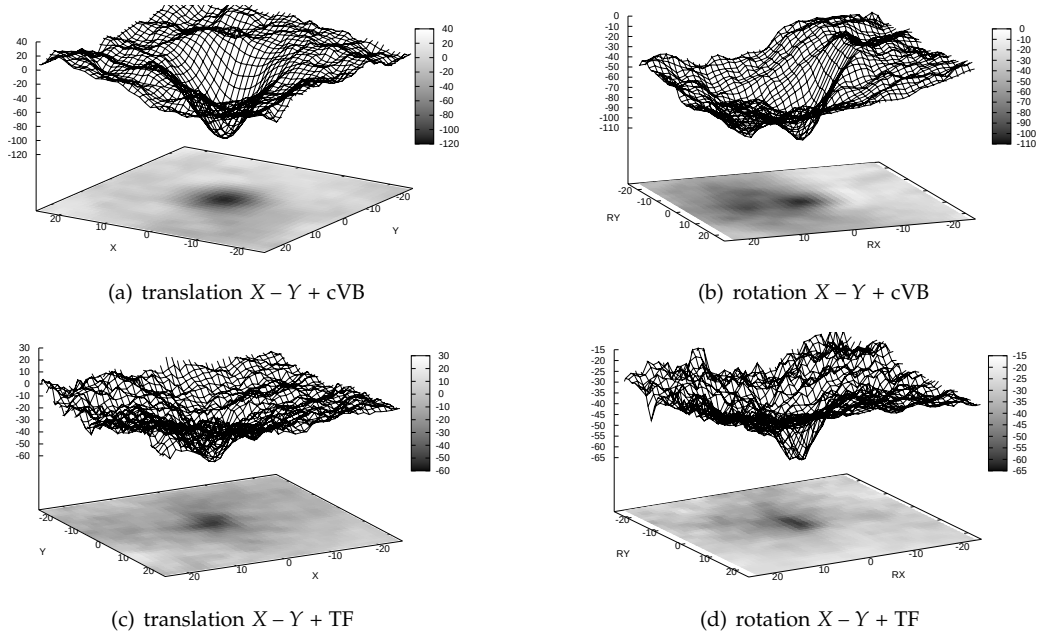
5.6.1 Impact of Image Characteristics and Knowledge-based Strategies

5.6.1.1 Image Field of View

In this section, we will discuss the impact of the image FOV on the different aspects of segmentation and SSM construction. As previously mentioned in Sec. 4.3.2.1, a reduced FOV makes the structure to be segmented only partially visible. The amount of image cues is thus reduced and we strongly suggest that this ultimately affects the performance of segmentation approaches, especially those based on deformable models and SSMs.

Initialization Model initialization is affected by small FOVs as we already discussed in Sec. 5.2.3.1. In fact, the landmark-based approach (Sec. 5.2.2.1) becomes hardly applicable at all and the atlas-based registration (Sec. 5.2.2.2) faces strong difficulties when the atlas and test image FOVs are very different. Hence, it was necessary to devise an adapted initialization approach. As reported in experiment 1 (Sec. 5.5.2), our constrained global initialization (GI) (Sec. 5.2.3) proved to be very robust (no failures were reported in all experiments) and accurate as depicted in Fig. 5.20 with images having a small FOV. The cost function C_f however presented various local minima as shown in Fig. 5.34 and was not always smooth, especially in case of the TF dataset (Fig. 5.34(c) and 5.34(d)).

More precisely, the function appeared to be smooth around the desired minimum, but the capture range was not very large (e.g., ≈2 and 1 cm in $X-Y$ translation, respectively for the cVB and TF cases). With small FOVs, we observed that various shape instances generated by the transform and shape parameters were more likely to return a strong fit with the wrong structures in the



(a) translation $X-Y$ + cVB

(b) rotation $X-Y$ + cVB

(c) translation $X-Y$ + TF

(d) rotation $X-Y$ + TF

Figure 5.34: Example of a constrained initialization's cost function plot around optimum on cVB ((a)-(b)) and TF ((c)-(d)) datasets of the right femur. Ground truth reference shape was translated in the $X-Y$ direction (mm) in the first column, while it was rotated around the X and Y axes (degrees) in the second column.

image. In this context, the usage of local optimizers was seriously limited. Since satisfactory bone size and pose estimates cannot be usually predicted in advance, finding a reasonable initialization to a local minimizer is an issue. Additional experiments showed that while a local minimizer initialized with the mean shape of the SSM sometimes returned acceptable initializations, there were many occasions on which it failed with the cVB and TF datasets (e.g., in cVB initialization, more than 10 trials produced results with VOE > 60%). Conversely, differential evolution was able to escape local minima, and was also very simple to implement and parameterize since it is a self-adaptive approach.

Evolution The image FOV undoubtedly had an impact on the performance of our deformable model-based segmentation. In some cases, the impact of poor initialization was diminished if the image FOV was sufficiently large. For example in experiment 1 (Sec. 5.5.2), the initialization of the landmark-based approach on the VB images was much less accurate than the GI on cVB images (ASSD of 2.88±0.68 vs. 1.49±0.46) but the fine segmentation was still able to produce good segmentation of the full bones. Similarly, in experiment 3 (Sec. 5.5.4), the atlas-based initialization was absolutely not accurate (ASSD of 5.86±2.15), but still the deformable model-based segmentation decreased the overall ASSD to 1.48±0.36. We suggest that this is partially due to the fully visible bones within the VB datasets, which guided the evolution of the segmentation by providing sufficient image cues.

In experiment 1, the final segmentation results of the experiments on the cVB and VB datasets demonstrated that the overall accuracy of the cVB (ASSD: 1.12±0.46) was better than the VB

(ASSD: 1.21±0.35), despite the latter having a much larger FOV. This contradicts our assumption that a larger FOV improves the segmentation. Although it appears that the better initialization of the cVB may explain this observation, the cVB metrics were actually better due to a "favored" error computation, which did not include the distal parts of the segmented femurs, as these parts were not present in the cVB images. In fact, in the VB images the segmentation of the femur condyles in the distal part performed poorly (Fig. 5.18) as these structures were (i) located almost parallel to the axial plane (and were thus subject to PVE), (ii) very close to the image extremities (thereby having poor signal-to-noise ratio), and (iii) in proximity to the knee articulation (hence exhibiting fuzzy boundaries). This statement was confirmed by computing the metrics of the segmentation with VB images only within the image extents of the cVB: the overall ASSD of the VB segmentation dropped from 1.21±0.35 to 1.17±0.42 and was statistically indistinguishable from the overall ASSD of the cVB results, 1.12±0.46.

The combination of moderate initialization with small FOVs is thus a serious handicap for segmentation. This explains why TF images were generally more difficult to segment in experiment 1. The additional reduction of the FOV of the TF images is expected to significantly degrade the segmentation performance. To illustrate this assumption, we conducted additional experiments on TF images with much smaller FOVs, in which the initialization and averaged IPs were identical to the original TF experiments (Sec. 5.5.2.2) in order to remove bias from possible different initializations and appearance priors. The smaller FOV was computed by reducing each side length of the original TF volume by 80% (i.e. reduction of ≈ 51% in volume). The overall ASSD error significantly increased (p-value < 0.002) from 1.31±0.44 to 1.65±0.81, highlighting the negative impact of small FOVs on the segmentation evolution.

5.6.1.2 Shape and Appearance Priors

The use of prior knowledge expressed as statistical modeling of shape and appearance variations has brought many benefits in all the key stages of deformable models.

Initialization The constrained initialization has greatly benefited from the generative property of SSM which instantiates "valid" shapes from few control parameters (Eq. (4.3) in Sec. 4.3.1.3). This shape parameterization, coupled with geometrical constraints derived from the hip joint kinematics (Sec. 5.2.3.3), yielded an efficient constrained initialization which drastically reduced the search space of the global minimization (Sec. 5.2.3.5). As a result, the robustness of the initialization was improved which explains the good results of the GI in experiment 1 (Sec. 5.5.2.2).

Evolution To attenuate the limitations from small FOVs, the use of SSM in the deformable models was of significant importance. By bringing robustness in the shape regularization process, SSM-based segmentation was found to outperform conventional deformable models. In fact, the deformable model based on weak shape priors expressed as shape memory (SMEM) constraints (Sec. 5.3.3.2) performed very poorly with the cVB and TF images in experiment 1 (Sec. 5.5.2.2) as overall ASSDs of 1.59±0.54 and 3.09±1.29 were computed, respectively. Furthermore, while large

FOV may alleviate the negative effects of poor initializations, we are convinced that the use of SSM remains essential. In fact, we observed that the deformable model based on the shape memory prior was incapable to cope with the mediocre atlas-based initialization (ASSD of 5.86±2.15) in experiment 3 (Sec. 5.5.4.2).

Nevertheless, special care was required when choosing the multi-resolution SSM scheme (Sec. 4.3.1.3) in order to yield optimal results. The SSS scheme reported the best results with the VB datasets, while RAA scheme performed better with cVB and TF images. More exactly, we only observed statistical differences between SSS and RAA schemes when segmenting the hip bones in the TF datasets as depicted in Fig. 5.35. We observed that with insufficient accuracy in the initialization, combined with the absence of sufficient image cues or a poor appearance model, the SSS-based evolution was impeded. The use of a rigid alignment with the coarsest resolution in the RAA scheme prevented shapes from "over-stretching" when trying to fit the shape with the image, as exemplified in Fig. 5.35(c) and (g). From our observation, femurs were not subject to this problem, likely due to better image features, such as the strong image gradients around the shaft (Fig. 5.20(e)), and also better initializations.

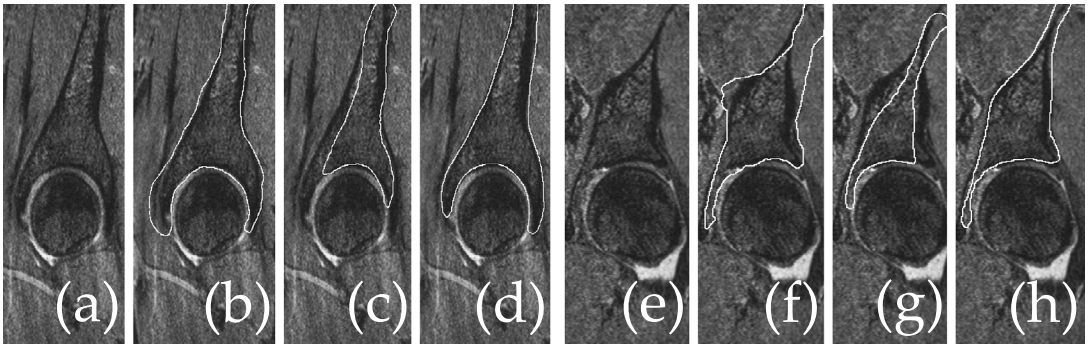


Figure 5.35: Right hip bone comparative results on TF dataset. (a) coronal slice with (b) constrained global initialization, (c) SSS scheme-based and (d) RAA scheme-based fine segmentation results overlaid in white. (e)-(h) Comparative results on a sagittal slice. The RAA scheme improved the initialization and appears as better than the SSS scheme which tended to over-stretch the hip bone as depicted in (c) and (g).

Regarding the use of appearance priors (Sec. 4.4) in our experiments, the coupling of multivariate Gaussian (MGD) priors with the Mahalanobis distance did not significantly outperform the use of reference profiles with the normalized cross correlation (NCC). As we explained, the robustness of the NCC measure with respect to linear intensity changes was a real asset. In fact, the possible absence of these intensity variations in the training dataset would prevent the MGD model from capturing them, affecting thus its efficiency. In all cases, appearance priors were complementary to the gradient features, especially when edge information was missing or non-reliable (e.g., fuzzy boundaries between articulations). As a result, they were essential in the segmentation approach and were thus always exploited in conjunction with the gradient-based forces. The NCC-based approach was in particular preferred to MGD priors for its lower complexity, which yielded faster segmentation algorithms (see timing results in Table 5.1 of experiment 1).

SSM construction and enlargement The limitation of creating and using incomplete shapes based on the image FOV has been efficiently addressed in this thesis by devising a robust evolution and SSM construction approaches that were able to cope with corrupted shapes. This improved automation and thus made possible the enlargement of a unique SSM per shape, which remained efficient in the segmentation context. Indeed, our evaluation results of experiment 2 (Sec. 5.5.3.1) demonstrated that the segmentation results from the images with small FOV could be efficiently re-used to improve the SSM performance in subsequent segmentations. In situations where the Generality of the SSM was poor (e.g., our sub-SSM experiment with 12 shapes), our methodology to enrich the SSM with corrupted shapes was shown to significantly improve the segmentation performance. Standard SSM performance measures (e.g. Compactness and Specificity presented in Sec. 4.5.1.3) were not used to compare sub- with mixed- or full-SSMs, as there were different numbers of models in the individual SSMs. Moreover, these measures do not necessarily reflect the segmentation performance and were also criticized by some studies [MDW08].

MRF-based Local Modeling The modeling of local variations based on MRF (Sec. 5.3.3.4) was beneficial when the Generality property of the SSM was not satisfactory due to the training dataset being too small. In experiment 2 (Sec. 5.5.3.2), a multi-resolution refinement based on the MRF evolution improved the overall ASSD error of the sub SSM (trained with only 12 shapes) to about 15% (i.e. from 1.59 to 1.35 mm). This improvement in accuracy was achieved at the expense of an increase of 22% in the computation time. The MRF was suitable to delineate fine details with the condition of being close to the optimal solution. In fact, the MRF modeling required to define the reference shape $\boldsymbol{X}$ (Sec. 4.3.3) to control the amplitude of the local deformations δ_i. Our solution consisted in setting the reference shape as the shape state at the end of the iterations of each resolution level.

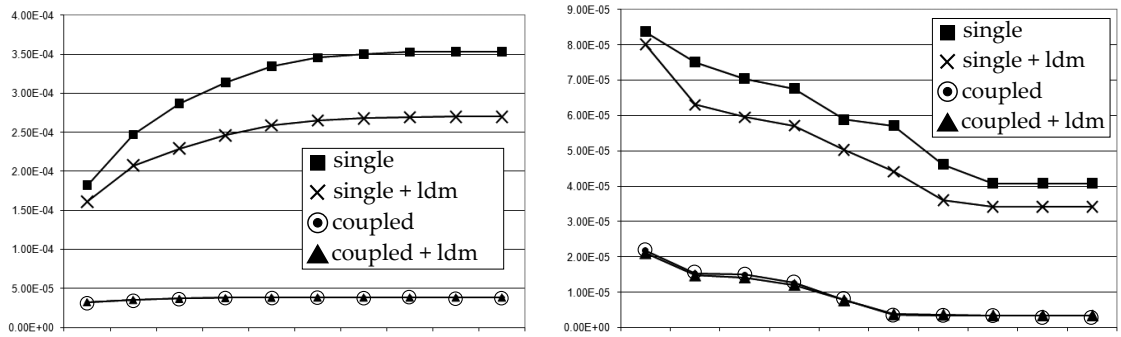
A difficulty that we faced with the MRF-based forces was the setting of the parameters. In [KH98], no rules of thumb were provided to select these parameters and we needed to perform several trials to find appropriate values of β and the number of iterations. If there was a way to statistically model the reference shape $\boldsymbol{X}$, a promising solution would consist in estimating by statistical inference these unknown parameters based on e.g. an EM algorithm.

5.6.1.3 Coupled Registration-Segmentation

The presented coupled registration-segmentation approach evaluated in experiment 4 (Sec. 5.5.5) appeared to notably improve segmentation results while simultaneously co-registering multiple images. The use of this knowledge expressed by the coupling of registration and segmentation tasks had an impact on the speed of the segmentation approach.

Furthermore, some preliminary experiments showed that a coupled registration-segmentation has also a positive impact on shape correspondence. Indeed, we designed an experiment involving femurs to be co-registered and segmented from 10 images. An affine mapping was used since all images came from different subjects and the blending parameter λ was set to a low

value of 0.05 to not penalize the segmentation (Sec. 5.3.4.2). We ran the experiment in coupled (i.e. with the coupled registered-segmentation approach) and single (i.e. 10 independent segmentations) modes. We also applied the correspondence refinement based on landmark sliding (Sec. 4.3.1.2) on the results of these two modes for comparison purposes. We compared the SSM correspondence metrics (Sec. 4.5.1.3) with the single and coupled modes. Examples of the Compactness and Generality curves are shown in Fig. 5.36(a) and 5.36(b), respectively. These curves clearly show that (i) the coupled approach yielded much better performance metrics than a single approach and that (ii) the correspondence refinement based on landmark sliding could not improve the results of the coupled mode since those were already excellent.



(a) Compactness (b) Generality

Figure 5.36: Impact of coupled registration-segmentation on shape correspondence illustrated with femurs. Curves indicate (a) Compactness and (b) Generality measures. Axis denotes the number of modes and "ldm" refers to a correspondence optimized by landmark sliding.

These results show a very promising usage of the coupled registration-segmentation approach. Intuitively, these observations make sense since the coupled registration-segmentation approach enforces point correspondence across shapes via its registration and blending stages. However its similar performance with respect to the correspondence optimization using the landmark sliding is striking in this preliminary experiment. The coupled approach may hence be a good alternative (or at least a complement) to the correspondence refinement by landmark sliding.

However, despite the accuracy of the coupled approach (ASSD of 1.58±0.18) was better than in the single mode (ASSD of 1.65±0.36), like it was previously observed in experiment 4, this accuracy was not very satisfactory when compared to the other experiment results. This was due to the fact that the affine mapping could not cope with non-rigid inter-subject variations. A promising solution would be hence to use an extension of the Procrustes alignment to non-rigid transforms (e.g., [CP10]). In any case, these are preliminary results and a more thorough evaluation is necessary.

5.6.2 Overall Assessment of Segmentation Approach

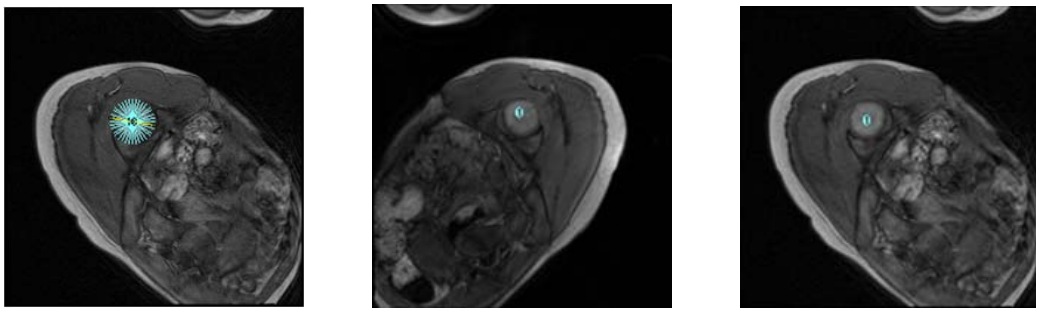
5.6.2.1 Initialization

Constrained global initialization The constrained global initialization based on the differential evolution (DE) optimization (Sec. 5.2.3.5) proved to be very robust and versatile. In experiment 1 (Sec. 5.5.2.2) both shape and transform parameters were simultaneously optimized to initialize models in images with small FOV, while in experiment 3 (Sec. 5.5.4) the transform between supine and split postures was computed. The robustness of the approach is obtained with a sufficient number of iterations to correctly explore the search space or escape local minima, which is the main strength of global minimization approaches. The number of iterations was empirically set to 100 based on several trials. In some situations, it was observed that the DE quickly converged to the right result and as a consequence several iterations were unnecessary. It would be hence interesting to explore optimization approaches that design more clever stopping criteria [ZL07] to save some computations. However as reported by Nearchou [Nea08], it is not possible to guarantee that an optimal solution has been obtained after a number of iterations due the stochastic behavior of DE. Alternatively, we found that the number of iterations could be decreased if the search space was more restricted based on prior knowledge. For instance, in the initialization of the split dataset (experiment 3 in Sec. 5.5.4), rotation angles could be more tightly bound in a small interval (full rotation space was used in our experiments).

Furthermore, DE is known to be potentially subject to the phenomenon of stagnation which indicates a premature convergence of the algorithm despite of having a diverse and unconverged population [LZ00]. We did not observe this particular case in our experiments but it could be addressed by using more sophisticated parameter settings [Zah02, Tvr06, HQS06].

Eventually, the placement of the hip joint center (HJC) (Sec. 5.2.3.2) was not very demanding, as in practice we noticed that small errors of 1-2 pixels (1-3 mm) were well handled by the DE in which translation parameters were allowed to vary in the error interval δ. Nevertheless, a more thorough sensitivity analysis of the HJC should be done. Furthermore, methods to automatically detect the HJC could be investigated. A collaboration with Kim [KSMT10] underpinned the non-efficiency of automatic methods based on edge detectors, such as Hough transforms [vVBV02, CYDL06] which failed to provide adequate results. This was due to the low image resolution and weak gradients at the femoral head boundaries in VB images. Hence a robust method of line profile matching (Fig. 5.37(a)) coupled with density clustering [ABKS99] was proposed to group the profiles that best represented the HJC. On a test database of 23 supine and split VB images, the algorithm was able to detect 81.52% of the HJC, with further 9.78% detected, but not as the best estimate, to a combined total of 91.3%, as illustrated in Fig. 5.37(b) and 5.37(b).

Atlas-based initialization The low accuracy in the atlas-based registration of experiment 3 (Sec. 5.5.4) was attributed to the use of an atlas which had quite large differences in intensity distributions and anatomical variations when compared to subject datasets. Furthermore, the registration attempted to register MRI images that were mostly composed of soft tissues, thus



(a) (b) (c)

Figure 5.37: Automatic HJC detection [KSMT10]. (a) Illustration of line profile extraction (blue segments) for detection of HJC candidates. Examples of automatically detected HJCs (blue discs) are visible in (b) and (c) for a left and a right split images, respectively.

the registration emphasized the similarity among the soft tissues rather than the bony structures. Nevertheless, the atlas used was based on a single subject scan and was able to produce a satisfactory initialization for this experiment. This was possible because all the subject scans were from the same population group, sharing same gender, age group, and similar physical attributes (e.g., height). We anticipate that with the addition of statistical properties in the construction of the atlas, e.g. atlas averaging via clustering of registered data [ZHR06], the initialization of the segmentation should be improved, thus leading to further improvements in the segmentation.

The segmentation of a split posture could theoretically follow the same procedures as in the supine segmentation, i.e., registration to an atlas for initialization. However, in our assessment, a selection of a subject for use as the atlas in split postures failed due to the ubiquitous inter-subject variations between the split images (unpredictable variations of femur rotation angles and surrounding soft tissue deformations), which strongly affected the registration process. Similarly, an intra-subject registration between the supine and split MRI remained even more problematic due to the larger rotation variations between the two postures.

The greatest advantage of using the atlas-based registration is that it is a fully automatic process as long as appropriate conditions are met (i.e. similar FOVs and subject positions). Coupled with our robust SSM-based evolution, we can thus expedite the bone segmentation of supine VB images in an automatic and reliable way.

5.6.2.2 Accuracy and Robustness

MRI Images acquired with the VB protocol are difficult to segment automatically due to poor image characteristics such as large slice thickness, strong bone intensity inhomogeneities, and diffused and unclear boundaries at joint interface. In this study, we obtained highly accurate and robust results (e.g., ASSD of 1.21±0.35 in experiment 1), by harnessing previously developed methods [GMT10] with robust multi-resolution SSMs. The segmentation on TF images also performed well in some regions of major interest despite having an overall segmentation error that was higher than the segmentation error on the VB dataset. For instance, the segmentation of iliac crests caused some trouble (e.g., Fig 5.21(a) in experiment 1) but the challenging areas

of the AB and FH were accurately segmented (ASSD 1.04±0.33, see Figs. 5.21(c) and (d) in Sec. 5.5.2.2). The final results with the TF datasets were good in the majority of the subjects, but the overall error was moderate despite the better image quality (higher resolution, better signal-to-noise ratio). As discussed in Section 5.6.1.1, the initialization accuracy suffered from a more challenging cost function minimization (not smooth, presence of many local minima as shown in Fig. 5.34(c) and 5.34(c)) which affected the fine segmentation. The use of higher order statistics or estimated probability distribution functions for the IPs is expected to improve these segmentation results. However, due to the large size of IPs (25 samples) with respect to the relatively low number of training datasets, effective approaches based on e.g. IP multi-modal clustering [CD09] or *k*-NN classifiers [HMMW07] should be investigated, to minimize the effects of the curse of dimensionality (Sec. 2.5.6).

Another interesting observation is that in all the experiments, the hip bones were equally or better segmented than the femurs. These results are consistent with other studies (e.g., [KLZH09]) which exploited SSMs to segment the hip joint bones and reported greater difficulties in the segmentation of the femurs. In this study, we attribute these results to the greater complexity of features found in the femur morphology. In many experiments we noticed that the inter-subject variability of these features was harder to capture by the SSMs (e.g., Fig. 5.18 in experiment 1 and Fig. 5.25 in experiment 3), as it was reported in the study of Barratt *et al.* [BCE+08].

From a global MRI bone segmentation perspective, we devised an approach that is very promising for computer-aided surgical applications, regardless of the size of the FOV or the resolution. In fact, Livyatan *et al.* [LYJ03] reported that a target reconstruction error (TRE) of approximatively 1.5 mm (2 to 3 mm worst case) was acceptable for bone computer-aided surgery applications. Since the TRE is also seen as an error of a reconstructed surface model in comparison to its associated ground truth [ZSG10] and our overall ASSD measures from all experiments were all below 1.5 mm independently of the image type, our approach has a great potential clinical practice. In comparison, Gilles *et al.* reported larger errors (1.7 mm in [GMT10] and 1.6 mm in [Gil07]), despite the authors jointly used multiple images to segment the bones. In our case, when multiple images were used in our coupled registration-segmentation experiment (Sec. 5.3.4), a very good accuracy of 1.07 mm was obtained with femurs.

The GPU-based approach evaluated in experiment 5 produced satisfactory results with ASSD of 1.62±0.44 (Sec. 5.5.6.2) being marginally over the recommended 1.5 mm value. The articular area presented segmentation errors mainly due to the bone proximity and the presence of fuzzy boundaries. We discussed the possibility to tackle these errors in the articular region by interactively using attraction points (Fig. 5.32). While this interactive action is efficient in most of the cases, it is preferred to exploit additional corrective techniques to minimize or simply remove the user interaction. Firstly, collision response and detection techniques could be implemented to prevent mesh inter-penetrations and hence bring additional robustness. Furthermore, the exploitation of our sophisticated shape priors such as SSMs will undoubtedly improve the performance of the GPU implementation, given the excellent results obtained with these priors in the other experiments. However, the implementation of SSM-based forces in GPU may de-

mand complex numerical tools. For instance, Singular Value Decomposition (SVD) is required to achieve the alignment of shapes [Ume91] in order to perform the SSM regularization. Many of these numerical tools are common but demand additional effort to be efficiently implemented into GPU formalism. Thankfully, they are progressively ported to CUDA, like the SVD implementation of Lahabar and Narayanan [LN09]. However, their integration into an existent framework is not always straightforward as special data structures are often required. These structures do not necessarily satisfy the requirements of our simulation, and as a result this may demand a careful (re-)design of the GPU framework.

5.6.2.3 Speed and Interactivity

Results of experiments showed that the speed of the segmentation was dependent on different factors. The image FOV had for example an impact as some computations (e.g., image-based metrics) could be avoided for mesh vertices falling out of the FOV, which accelerated the process (e.g, compare the TF vs. VB timing results of experiment 1 in Sec. 5.5.2). The speed of image-based forces in the GPU implementation was also affected by the image size in terms of number of voxels, as reported in Fig. 5.30(c) and 5.30(d) in experiment 5. Similarly, the complexity of forces or strategies influenced the speed of the segmentation. Indeed a SSM-based evolution was more time consuming than a shape memory-based segmentation in experiment 1 (Sec. 5.5.2), or the coupled-registration approach was slower than a single segmentation as reported in experiment 4 (Sec. 5.5.5).

The different segmentation approaches that were proposed (robust and local shape priors, coupled registration-segmentation, etc.) aimed at improving the accuracy and robustness of the segmentation as reported in Sec. 5.6.2.2. As a result, the focus was not in preserving the interactivity of the segmentation. In fact, timing measurements in experiment 5 (Sec. 5.5.6) showed that the CPU version was unable to fully support interactivity, since update frequencies were about 0.6 – 6.7Hz. Similarly, the collaborative framework in experiment 6 (Sec. 5.5.7) exhibited an average update per seconds of 4 Hz.

Conversely, a significant improvement of the interactivity was obtained with the GPU-based implementation of our discrete deformable models. The GPU approach was consistently about 25 – 70× faster than the CPU version to execute a single time step. This was expected since the GPU implementation benefited from parallel processing while the CPU-based approach operated in a sequential manner. In all cases, the time taken by a single GPU iteration was perfectly matching with interactivity constraints. Indeed, update frequencies for the GPU were about 47 – 162Hz, thus easily supporting the minimum 10Hz of refresh rates required for interactivity. Having full support to interactivity paves the way to novel segmentation approaches, in which the user is fully able to interact with the segmentation loop.

We have measured that, in the GPU implementation, force computation accounted for approximatively 99% of total time per iteration, mesh update and numerical integration contributions being below 1% (Fig. 5.30(a)). Hence, an additional level of parallelization could be achieved by

computing all forces in parallel. This force parallelism could be achieved with atomic operations which would sum up the force contributions in parallel in the force accumulation array. Since atomic operators do not necessarily improve performances [2] and their availability depends on the CUDA compute capability of the GPU, an alternative approach would consist in using a proper array for each force to avoid memory writing conflicts and thus the use of atomic operations. However, such an implementation is obtained at the expense of a larger memory consumption.

In our segmentation context, effort should be instead spent in speeding up image-based forces, and especially those using intensity profile. In fact, the IP-based force clearly dominated with its 93% of the total force computation time (Fig. 5.30(b)). Despite the use of an efficient computation of the NCC, this force remained very expensive. The main reason is that during the search for the optimal target position $\boldsymbol{y}^*$ with the lowest image energy, we compute the IP-based energy L times in a sequential manner (Sec. 5.3.2.1). Hence, once again we could parallelize the computation of the energy for each position y_i, $i \in [1,\ L]$. However, before considering such a parallelization, special attention should be paid in preserving a good tradeoff between memory usage, speed and implementation complexity. Last but not least, the CUDA implementation performance should be more throroughly evaluated and optimized if necessary [NVI10] (e.g., fine tuning of grid and block sizes, avoidance of possible bank conflicts, reduction of number of cycles per instruction and memory latency, etc.).

From a global MRI bone segmentation perspective, the speed of the different CPU-based segmentations remained very satisfactory, given that most of the approaches were automatic or required very little user intervention. The GPU-based approach showed that the use of parallelization techniques, coupled with modern hardware (our PC specifications were quite modest as reported in Sec. 5.5.1) supporting multiple cores, are expected to substantially speed-up the CPU-based automatic and collaborative segmentation approaches. This would support the clinical integration and acceptance of our deformable model-based segmentation.

5.6.2.4 Control

Regardless of the interactivity of the segmentation approach, we investigated efficient way to control the evolution of the deformable models. Constraint points from [Gil07] were cleverly exploited in the collaborative segmentation to yield the concurrent segmentation of the same image by multiple users. A user satisfaction study [HNS+10] revealed that most users were strongly or moderately satisfied with the collaborative approach. Most importantly, participants felt more comfortable to collaborate with others with the proposed algorithm because more freedom was given compared to an approach based on the strict locking mechanism. Nevertheless, additional studies should be conducted with end users (e.g., clinicians) to clearly assess the performance of our collaborative framework.

The use of attraction points (Sec. 5.4.1.7) in the GPU framework were very useful in circumventing some of the weaknesses of the approach (i.e. absence of collisions management and strong

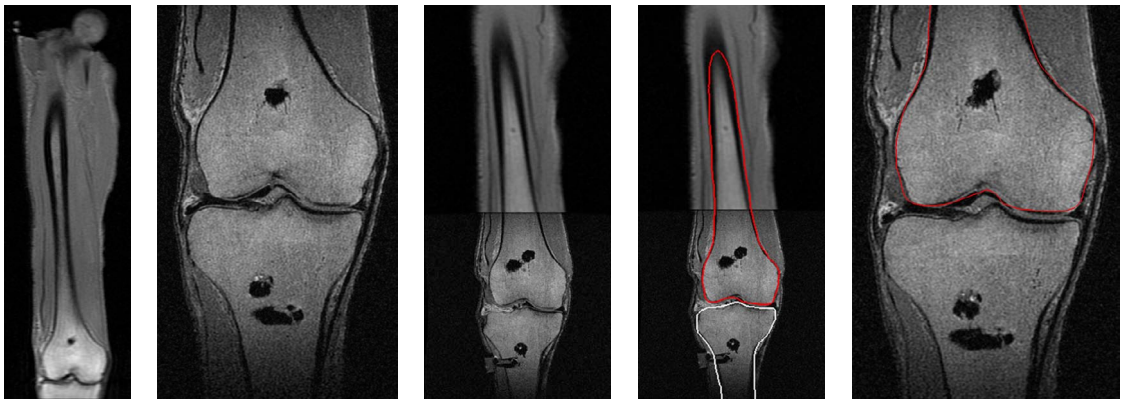
2. http://www.strobe.cc/cuda_atomics/cuda_atomics.pdf

shape priors) as illustrated in Fig. 5.32. Thanks to the efficient implementation of this interaction force (e.g., use shared memory and parallel reduction) into GPU formalism, no slowdowns were observed even in the presence of hundreds of attraction points.

5.6.2.5 Generality, Applicability and Extensibility

Generality In this work, only the constrained global initialization (Sec. 5.2.3) is specific to the hip segmentation and therefore our proposed robust SSM evolution (Sec. 5.3.3.3 and 5.3.3.4) and construction processes (Sec. 4.3) can be applied without any restriction for the segmentation of other structures. In fact, the methodology to compute shape and intensity priors in a robust manner can be re-used for other structures or image modalities. Furthermore, by running our deformable approach, the resulting segmentation results could be used to update or compute an SSM, independently of the shape nature (complete or corrupted shapes).

Applicability As discussed in the introduction Chapter, our work has great potentials and can be exploited in various medical and biomechanical studies. Many of our segmented models were used at MIRALab to conduct dynamic simulations of the hip joint and investigate the effects of (extreme) motion on the dancer population [ACS+09, CSKC+09, CAVMT09, CMTB+10]. Furthermore, our segmentation algorithms were used in the 3D Anatomical Human project [3DA10]. For instance, a partner acquired a cadaver leg to carry out some mechanical experiments to assess the stability of the knee joint. Different MRI images of the limb with varying characteristics (e.g., high resolution image of the knee with small FOV or low resolution of the thigh with large FOV) were obtained based on MR imaging. Based on our SSM-based deformable model and coupled registration-segmentation algorithms, images were co-registered and bones were successfully segmented as depicted in Fig. 5.38.

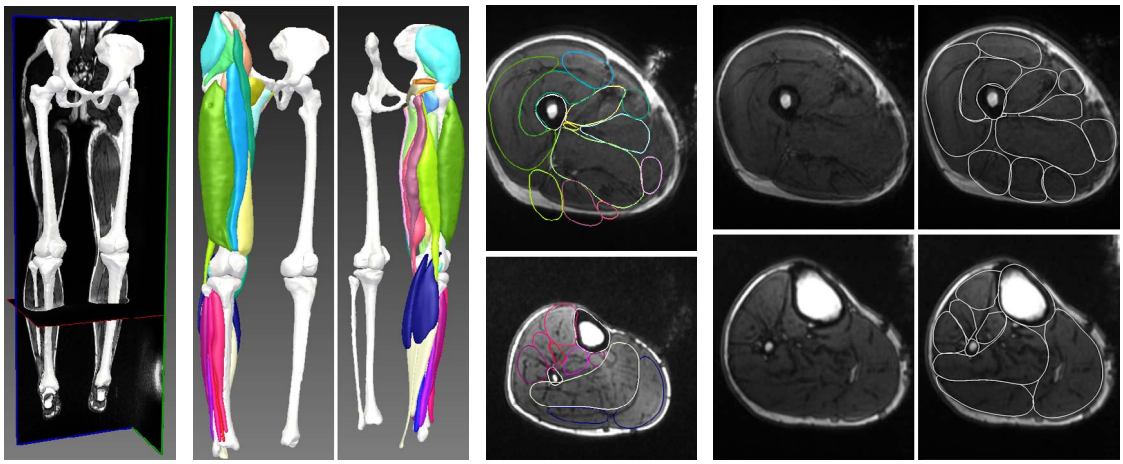


(a) (b) (c) (d) (e)

Figure 5.38: MRI imaging of the thigh and knee of a cadaver specimen are visible in (a) and (b), respectively. Based on our coupled registration-segmentation approach, images can be jointly registered and segmented as shown in (c) and (d), respectively. Red and white overlays represent the segmented femur and tibia, respectively. Subfigure (e) shows a slice with femur segmentation overlay of the knee image.

Similarly, we exploited our automatic bone segmentation and atlas-based initialization to segment muscles from an MRI image of the full lower limb, preprocessed with the guidelines given

in Sec. 5.2.1. The atlas was constructed with our generic model construction explained in Sec. 4.2 and exemplified in Sec. 4.5.2.1. From the segmented bones (Fig. 5.39(a)), we computed an affine alignment between the atlas and the test images to initialize the atlas-based registration. The transformed atlas muscle models from the estimated non-rigid transformation are visible in Fig. 5.39(b), while some mesh overlays are shown in Fig. 5.39(c). Based on the deformable model approach of [GMT10], an acceptable segmentation of the muscular structures of both the thigh and the leg was obtained as depicted in Fig. 5.39(d).



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

Figure 5.39: Segmentation of the lower limb. (a) Segmented bone 3D models. (b) Initialized muscle models from a non-rigid atlas-based registration overlaid on transversal slices in (c). (d) Muscle overlays after a deformable model-based segmentation on transversal slices of the thigh (top) and leg (bottom).

This patient-specific modeling of the lower limb supported various advanced neuromuscular studies as reported in [SSC+09, MSP+09]. As an illustrative example, the segmented knee model presented in Sec. 4.5.2.2 was used by Sandholm *et al.* [SSP+10] to compare different knee modelings to be applied in neuromuscular simulations [DRC+06, DAA+07]. In particular, a geometrically-based knee model fitted ellipsoids on the condyles of the segmented models (Fig. 5.40), in order to derive a joint kinematical model which constrained the adduction-abduction rotation and distration-compression translation of the knee.

Extensibility In this thesis, we did not have access to a sufficient amount of training data of soft organs such as the muscles. As a result, we could not build and experiment SSMs for soft structures. Nevertheless, our overall methodology should remain valid in this context. We also envision possible extensions of our kinematics-inspired initialization to other articulations. For instance, the shoulder (approximated as a ball-and-socket joint) could directly benefit from this concept, whereas the knee joint initialization could exploit a hinge-like motion constraint for which the user should specify only two points for the axis direction. Eventually, additional prior knowledge could be included in the statistical models such as the pose of structures [KLZH09] or some pathological conditions [YOT+09] to improve the Generality aspect of the SSMs. In fact, the proposed deformable models were robust against some abnormal cases such as knee surgery

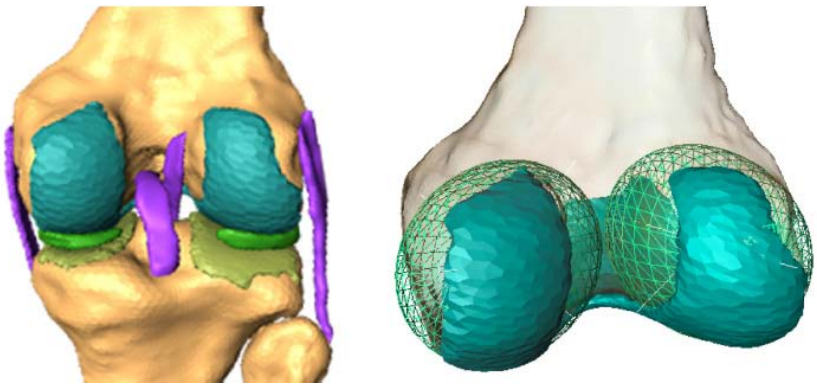
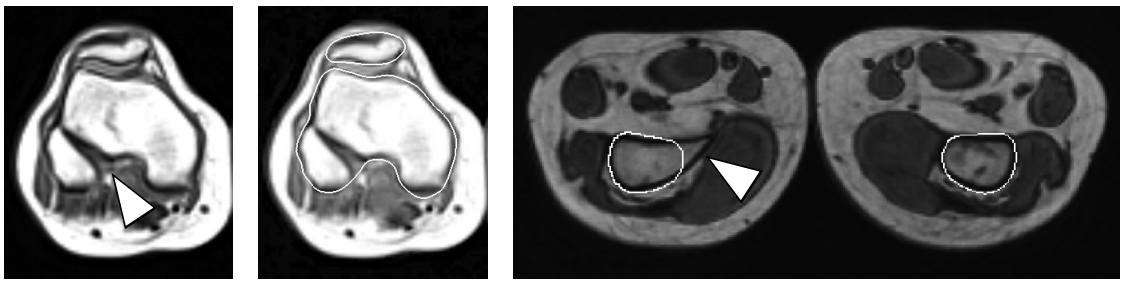


Figure 5.40: Use of segmented knee models to support neuromuscular simulation. Our segmented knee model (left) is used to define geometrical constraints based on the fitting of ellipsoids to the condyles (right). Images courtesy and used by permissions from [SSP+10].

exemplified in Fig. 5.41(a) and (b), but other large deviations from healthy subjects were more difficult to tackle as shown in Fig. 5.41(c).

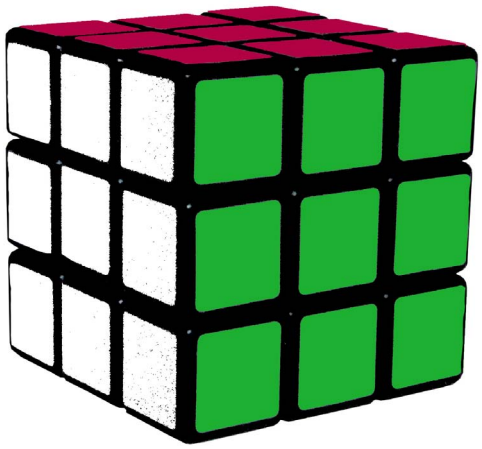


(a) (b) (c)

Figure 5.41: Examples of post-surgical and pathological segmentations. In (a) the subject underwent a knee surgery resulting in bone loss as pointed by the white arrow. Our deformable model manages to correctly segment the distal femur despite the presence of this abnormality as indicated by white mesh overlays in (b). In the second example (c) the segmentation failed in accurately delineating an abnormal deformation due to a benign osteochondrome (cartilage tumor) at the right distal femoral metaphysis (white arrow).

Chapter 6

Conclusion



6.1 Discussion

The work presented in this thesis has addressed the important and challenging problem of subject-specific musculoskeletal model reconstruction based on segmentation of MRI medical images. Image segmentation is a core necessity of many medical computer-assisted technologies and often represents a major bottleneck in widespread use of computer-assisted solutions in clinical practices. Musculoskeletal modeling is attracting a lot of interest due to the staggering prevalence of musculoskeletal pathologies and related disorders in aging societies.

As a result, the research conducted in this thesis was stemmed from the increasing demand of appropriate musculoskeletal modeling solutions and from the many technical challenges faced in image segmentation. Indeed, an efficient segmentation approach is primarily characterized by its accuracy and robustness. Automation is also essential to expedite tedious and time consuming manual segmentations, but interactive control mechanisms remain necessary to monitor and correct, if necessary, the segmentation evolution during its execution.

In this thesis, the diversity of MRI imaging protocols with varying image (image resolution, field of view) and subject (posture, amplitude of movement) characteristics required the use of prior knowledge to design efficient segmentation approaches. Prior knowledge, expressed as statistical priors or specialized strategies (e.g., coupled registration-segmentation), has shown to bring robustness and sub-voxel accuracy. We paid special attention to the construction of these priors by studying the issues related to shape correspondence and mixing of training shapes of different types (i.e., complete, corrupted). Moreover, a step towards a fully automated method has been made by investigating initialization with no or little user intervention, while the interactive control process has been studied under the perspectives of speed improvement (by the use of GPU) and collaborative interactive segmentation (i.e. concurrent editing of a same image by multiple users).

Nevertheless, additional work is needed for segmentation algorithms to be routinely employed in clinical practice, whereby offering decision making support and improving the diagnostic as well as the surgical efficiency. Similarly, more effort is required to produce subject-specific anatomical models that will be seamlessly integrated in simulation studies.

6.2 Contributions and Benefits

The main contribution of this thesis is a complete methodology to train and employ knowledge-based deformable models for the purposes of medical image analysis and subject-specific modeling. Our work has demonstrated the efficient coupling of shape and appearance priors with discrete deformable models. Contributions were made in all the key stages of statistical deformable models leading to the following benefits:

– **Accuracy**: with an average bone segmentation error of 1.21 mm, our approach has a great potential for bone segmentation and its use for computer-aided applications.

- **Robustness**: the good performance against heterogeneous acquisition protocols (different image fields of view, image artifacts, unconventional subject postures, multiple non-registered scans, etc.) underpinned the robustness of our deformable model-based segmentation.
- **Speed and Automation**: depending on the acquisition context, computation time varied between 3 to 11 min in average to segment hip joint structures, requiring no or very little operator intervention.
- **Interactivity and Control**: by offering highly interactive segmentation solutions based on GPU programming along with mechanisms for concurrent and collaborative editing, the segmentation process is more efficiently monitored.
- **Generality**: the proposed methodology for the construction of generic, statistical and subject-specific models makes our approach quite versatile in terms of type of modeled structures.

Our segmentation methodology has been successfully exploited in different application scenarios at MIRALab and by partners of the 3D Anatomical Human project [3DA10]:

- **Simulation of the hip joint** [ACS+09, CSKC+09, CAVMT09, CMTB+10]: based on datasets of the dancers and a control group of asymptomatic subjects, we segmented bone models with our approach along with soft structures using [GMT10]. Subsequently, these models were processed and used in simulations involving anatomy, kinematics and soft tissue deformations. This work was part of a clinical study whose aim is to better understand and explain signs of osteoarthritis observed in dancers' hips.
- **Subject-specific neuromuscular and soft-tissue simulations** [SSC+09, MSP+09]: based on our musculoskeletal modeling, subject-specific models including segmented bones, ligaments and muscles with corresponding attachments, were exploited in neuromuscular (lower limb) and soft-tissue (muscle) simulations.
- **Interactive 3D medical visualization** [ABG+09]: segmented models of a full lower limb were interactively visualized by direct volume rendering on a novel holographic display. The examination of the segmented musculoskeletal structures, which presented a complex spatial organization, was a good case study of medical visualization.
- **Anatomical atlases for teaching and training** [CKM+10]: an anatomical browser was built on top of an ontology of the lower limb, composed of our segmented models, for teaching and training purposes.

6.3 Limitations and Future Work

By focusing on very specific research directions, a thesis cannot always explore in full details the rich variety of available techniques that may circumvent some of the limitations of the proposed work. Similarly, the evaluation of the proposed methodology can always benefit from more testing and application scenarios. Below is a list of future research to improve and pursue the work initiated in this thesis.

6.3.1 Prior Knowledge Expression

Prior knowledge proved to be very efficient in improving the robustness and regulating the segmentation based on deformable models. However, we also demonstrated the negative impact of poor priors on the segmentation. Hence, the construction and the nature of the priors could be further improved.

- One of the main limitations of Principal Component analysis (PCA) is to assume the Gaussianity and linearity of the observed variations. This assumption may be invalid in some situations, hence the exploration of less restrictive approaches (e.g., kernel PCA, Independent Component Analysis) should be considered.
- From a global viewpoint, the Generality of the statistical shape models (SSM) could be improved by enlarging the training shape database whenever possible. Furthermore, training shapes should span a larger variety of subject characteristics (gender, age, pathological conditions).
- Regarding the Generality of SSMs built with soft musculoskeletal structures like muscles, artificial shape variations could be created based on soft tissue simulations that would generate artificial yet realistic organ deformations.
- Alternative appearance features could be considered to improve the Specificity of the appearance models. Region- and histogram-based features are of interest but special care should be taken in preserving reasonable computational resources (e.g., memory issues).
- Robust PCA approaches could benefit from robust statistical estimators (e.g., M-estimators) which simultaneously identify and weight non-reliable features. This would allow for instance to account for vertices of training models that are within the image extents but that are not perfectly on the structure boundaries (i.e. due to an incorrect local segmentation).

6.3.2 Deformable Model Initialization

The initialization being very critical in deformable model-based approaches, initialization approaches may be improved to yield a better subsequent segmentation.

- The image energies in the constrained global initialization could account for richer appearance information, such as image regions (e.g., model interior), and be based on more sophisticated metrics (e.g., mutual information).
- Atlas-based initialization has a great potential and should be further investigated. A more careful selection of the atlas or the construction of a probabilistic atlas are promising research directions. In future studies, a more thorough analysis of the non-rigid registration parameter tuning (e.g., number of levels, grid size, fluid and viscous regularization, use of masks, etc.) should be investigated.
- Image pre-processing techniques, such as thresholding or feature detection, could be exploited to coarsely initialize the model. This would reduce the search space in the constrained global initialization (GI) and improve the success rate of the registration-based approach.
- A multi-resolution image pyramid could be used in the GI to improve its robustness and

possibly reduce the overall number of iterations.
- Similarly, parallelization techniques should significantly speed up the GI. In fact, we envision two levels of parallelization: a first at the vertex level to compute the image energy (like it is done in the GPU-based evolution) and a second at the individual level in the differential evolution (DE). Furthermore, more sophisticated stopping criteria of the DE algorithm could be exploited to reduce the number of generations.

6.3.3 Deformable Model Evolution

Despite the very good performance of the segmentation, there is still room for improvements to reach the same levels of accuracy and robustness that of manual segmentations. Some limitations are related to the quality of the priors but also to the fundamental mechanisms of the segmentation evolution.

- Target positions of vertices in image-based forces could be globally optimized (e.g., via graph cut techniques). This would increase the robustness of the evolution against local image artifacts.
- Similarly, adaptive weighting of forces could be done at the vertex level based on e.g. the quality of the image features (i.e., if it is poor, we decrease and increase the weights of image and internal forces, respectively).
- Alternative shape prior-free internal forces, more sophisticated than mere smoothing, should be studied to cope with the cases where shape priors are too specific and cannot adapt to configurations which were not present in the training database (e.g., pathological cases).
- All CPU-based segmentation techniques should be ultimately ported to GPU or parallelized with multiple cores since significant speedups are expected. This would moreover improve the interactivity of the segmentation and allow the use of more time consuming but more sophisticated image-based strategies. In case of fully automatic approaches, fast techniques would yield the effective processing of large subject databases, which include more and more scans per subject (different acquisitions in time, various body areas imaged at multiple scales, etc.).

6.3.4 Evaluation and Applications

In addition to technical improvements, the evaluation procedure should be enriched to identify possible unknown limitations. Furthermore, a thoroughly validated framework in the clinical environment would expedite its acceptance by physicians as well as its integration into medical software. Hence, this would ultimately lead to our research being translated to real world application by bringing direct benefits to the patients, which remains the core motivation of our study.

- Segmentation methods should be evaluated with and adapted to a larger variety of clinical images acquired with different modalities (CT, X-Rays, etc.) and image protocols. Pathological (e.g., necrotic femoral heads, fractures) and surgical conditions (e.g., bone cutting) should be

also included in the testing datasets. As previously mentioned, the characteristics of the subjects should be also varied. This heterogeneous material will eventually improve the robustness of the algorithms and ensure a better clinical integration.
- Multiple users should (semi-)manually segment the same datasets to account for inter-observer variability. Moreover, this will minimize potential errors introduced in the training data.

Musculoskeletal models should be eventually extended to account for multiple anatomical (genetics, cell, organ, etc.), physiological (e.g., muscle activity) and biomechanical (e.g., soft tissue mechanics) scales in a consistent manner. This will strongly support the research on the virtual physiological human in the following aspects:

- Aided-diagnosis: the automatic segmentation of musculoskeletal structures help in the localization and identification of abnormalities (e.g., tumors, signs of arthritis, articulation impingements) and in the decision making process (proposition of possible treatment).
- Predictive medicine: simulations are run based on subject-specific models to assess e.g. the outcomes of virtual treatments (e.g., gait analysis after prosthesis placement) or the evolution of pathologies over time. This also supports the planning of surgeries.
- Assisted-surgery: based on pre-operative planning, virtual representations of the patient structures of interest are adapted intra-operatively (e.g., shape of soft organs is changed to account for intra-patient discrepancies or physiological variations such as breathing) to provide guidance during the surgery (e.g., augmented reality of structures in the endoscopic view).
- Data collection and mining: created models are gathered into common databases and decorated with semantic information (anonymized clinical records, protocol and experiment settings, etc). This centralized knowledge yields to data mining (e.g., statistical analysis) and provides validated training and testing material to the medical and research communities.

Appendix A

Acronyms, Mathematical Notations and Technical Details

A.1 Acronyms

Acronyms used throughout the thesis and their definition:

AAM: Active Appearance Model
ASM: Active Shape Model
CPU: Central Processing Unit
GPU: Graphics Processing Unit
CT: Computed Tomography
CDF: Cumulative Distributive Function
DRR: Digitally Reconstructed Radiograph
EA: Evolutionary Algorithm
EM: Expectation-Maximization
EVD: Eigenvalue Decomposition
FAI: Femoroacetabular Impingement
FEM: Finite Element Model
FFD: Free Form Deformation
FOV: Field Of View
ICA: Independent Component Analysis
ICP: Iterative Closest Point
KDE: Kernel Density Estimation
kPCA: Kernel Principal Component Analysis
LOO: Leave-one-out
MAP: Maximum *a posteriori*
MI: Mutual Information
MSD: Musculoskeletal Disorder
MRI: Magnetic Resonance Imaging
NMI: Normalized Mutual Information
OA: Osteroarthritis
PCA: Principal Component Analysis

PDF: Probability Density Function
PDM: Point Distribution Model
PPCA: Probabilistic Principal Component Analysis
RMS: Root Mean Square
RV: Random Variable
SVD: Singular Value Decomposition
SVM: Support Vector Machine
SAM: Statistical Appearance Model
SDF: Signed Distance Function
SDM: Statistical Deformation Model
SSM: Statistical Shape Model
TPS: Thin Plate Spline

A.2 Mathematical Notations

Vectors are in **bold**, e.g. $\boldsymbol{x}$.
Matrices are in *italics*, e.g. X. A matrix X can be defined as a set of N column vectors $\boldsymbol{x_i}$ as $X = \{\boldsymbol{x_1}, \ldots, \boldsymbol{x_N}\}$. Similarly, the element of a matrix M at the ith row and jth column is M_{ij}.
A $N \times N$ diagonal matrix D whose diagonal elements are $d_1, \ldots, d_N$ is written as $D = diag(d_1, \ldots, d_N)$.
The identity matrix of dimension $N \times N$ is written as Id_N.
The dot product between two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ is written as $\boldsymbol{x} \bullet \boldsymbol{y}$ or $< \boldsymbol{x} | \boldsymbol{y} >$.

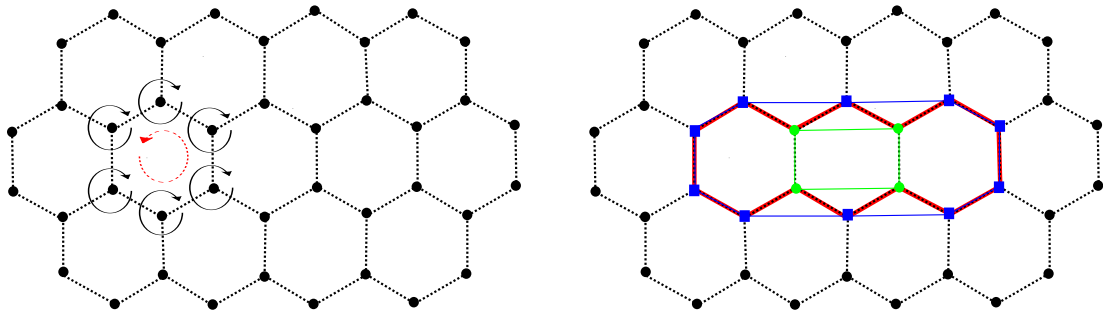
A.3 Simplex Meshes

A.3.1 Definition

A k-simplex mesh of $\mathbb{R}^d$, $k < d$, is a k manifold which presents vertices that have the constant number of $k+1$ distinct neighbors. A k-simplex mesh is generally composed of p-cells, $0 \leq p \leq k$, which are defined with a recursive rule [Del99]. In this thesis, we focus on 2-simplex meshes which can represent arbitrary surfaces with any kind of topology. In that case, a 2-simplex mesh is hence composed of vertices (0-cell), edges (1-cell) and faces (2-cell).

An orientation can be attributed to vertices based on an arbitrary order chosen for the sharing edges (e.g., the right hand rule) as shown in Fig. A.1(a). Similarly, a face orientation can be defined based on its vertices. A *contour* of a 2-simplex mesh $\mathcal{M}$ is a 1-simplex mesh whose vertices belong to $\mathcal{M}$. A contour is always a closed polygon. When a contour is exclusively composed of the vertices of a face, it is denoted as a *border* and its orientation is the same as the face orientation. In general, contours do not have a preferred orientation and it can be thus arbitrarily chosen. Each vertex of a contour can be classified into two groups, whether the orientation of its two adjacent edges belonging to the contour is compatible or not with the orientation of all three edges defined at this vertex as illustrated in Fig. A.1(b).

Simplex meshes follow the Euler-Poincaré formula which is simplified for this type of mesh due to the



(a) (b)

Figure A.1: Examples of 2-simplex mesh orientations and contours. (a) Orientations of a vertex and a face are depicted as filled black and dashed red curved arrows, respectively. (b) A contour is defined as a closed polygon composed of some vertices of the mesh (red thick polygon). Given a chosen contour orientation, contour vertices can be classified into two different groups (blue squares and green circles) based on consecutive edges orientation, resulting into two new contours (green and blue filled polygons).

constant vertex arity. For a 2-simplex mesh, the formula is [Del94a]:

$$\begin{array}{rcl} F - \frac{V}{2} & = & 2(1-g) - H \\ E & = & \frac{3V}{2} \end{array} \tag{A.1}$$

where V, E and F denote the number of vertices, edges and faces, respectively. The genus g corresponds to the number of handles of the surface (e.g. 0 for a sphere and 1 for a torus), while H denotes the number of holes and ends (e.g. a cylinder without caps at its extremities has a genus 0 and 2 ends).

For our work, we will only deal with *complete* 2-simplex meshes. A k-simplex mesh is complete if [Del94a]:

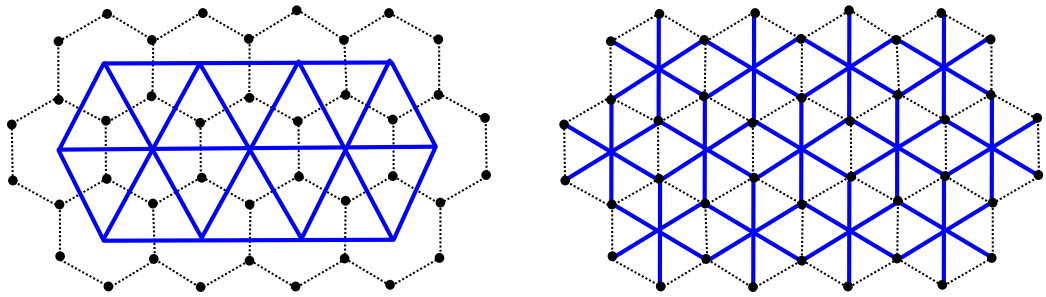
– it is oriented
– it respects specific adjacency rules between edges, vertices and faces
– two k-cells intersect along one (and only one) $(k-1)$-cell

For 2-simplex meshes, this implies in practice that two adjacent faces must solely share one unique edge.

A.3.2 Duality

An interesting property of 2-simplex meshes is their *topological* duality with triangulations. This property is generalizable for k-simplex meshes, which are the topological duals of "k-triangulations" [Del99]. It is important to understand that geometric duality of meshes is not ensured since geometric bijections between simplex meshes and triangulations do not exist. This implies that various tessellation approaches are possible to convert a simplex mesh to a triangular mesh and that there might be some geometric differences (e.g., curvature) between the two types of meshes. However, since triangulations are more commonly used (e.g., for GPU rendering) it is very useful to be able to convert a triangulation to a simplex mesh and vice versa.

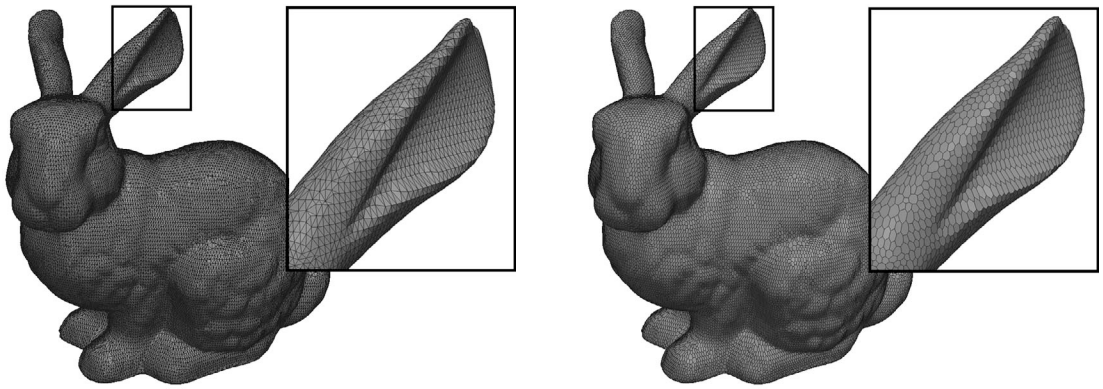
A triangular mesh is easily converted to a 2-simplex mesh by joining the center of triangles which share a same vertex to create a simplex face as depicted in Fig. A.2(a). This procedure becomes handy to convert widely available triangular models, such as the Stanford bunny [TL94] (Fig. A.3), or models produced by classic reconstruction approaches (e.g., Marching Cubes [LC87]). Conversely, a simplex mesh can be converted to a triangular one by using the same process, i.e. connecting the center of simplex faces to yield triangles. To better preserve the geometric quality of simplex meshes, an alternative tessellation to



(a) (b)

Figure A.2: Simplex/triangular mesh duality. (a) The classical approach convert vertices to faces and vice versa; thick and dotted lines depict the triangular and simplex meshes, respectively. (b) An alternative 2-simplex mesh triangular tessellation is obtained by joining the p vertices of a simplex face with its center to produce p triangles. Dotted lines depict the simplex mesh, while the union of dotted and thick lines forms the edges of the triangulation.

triangles [Gil07] is reported in Fig. A.2(b), where the vertices of the simplex mesh are included in the triangulation. In the following, we will adopt this triangular tessellation scheme every time a simplex mesh needs to be converted to a triangular mesh.



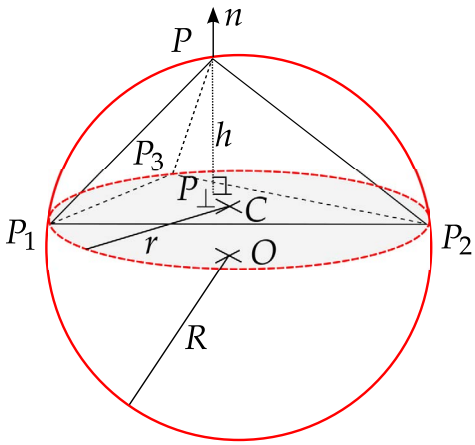
(a) (b)

Figure A.3: An example of (a) triangular to (b) simplex mesh conversion. The famous Stanford bunny is easily converted to a 2-simplex mesh.

A.3.3 Local Geometry

A local geometry can be defined to express in a unique manner the position of a vertex P with respect to its neighboring vertices P_i. Figure A.4 shows that P is computed from three parameters ϵ_1, ϵ_2 and ϕ [Del94b]. Terms ϵ_i are the barycentric coefficients (also referred to as metric parameters in [Del94b]) of the projection point $P_\perp$ on the triangle $(P_1P_2P_3)$. Therefore we have $\epsilon_1 + \epsilon_2 + \epsilon_3 = 1$. The simplex angle ϕ is a function of the distance $\|OC\|$ and the circumscribed circle and sphere with respective radii r and R.

Parameters ϵ_1, ϵ_2 and ϕ, referred to as *simplex parameters*, are independent and invariant under similarity transforms. B. Gilles [Gil07, GMT10] demonstrated that these parameters uniquely defined P from P_i only in the situation where the projection $P_\perp$ laid inside the circumscribed circle. Such a constraint is problematic since mesh configurations need to be restricted in order to ensure the uniqueness of the simplex parameters. As a result, Gilles instead proposed to directly compute and use the elevation



ϕ = angle between $\mathbf{P_1P}$ and the triangle (P_3P_2P)

$\mathbf{n} = \frac{\mathbf{P_1P_3} \wedge \mathbf{P_1P_2}}{\|\mathbf{P_1P_3} \wedge \mathbf{P_1P_2}\|}$

$$P(\epsilon_1, \epsilon_2, \phi) = \epsilon_1 P_1 + \epsilon_2 P_2 + (1 - \epsilon_1 - \epsilon_2) P_3 + h(\phi)\mathbf{n}$$

Figure A.4: Simplex mesh local geometry [Del94b]. A point P is uniquely defined with three independent and similarity invariant parameters ϵ_1, ϵ_2 and ϕ.

$h = \|PP_\perp\|$. The elevation replaces thus the simplex angle as the third simplex parameter, but since it is not a similarity invariant quantity, the scaled elevation $h_n = hS^{1/\alpha}$ is preferred, where S indicates the surface of the triangle $(P_1P_2P_3)$ and α is an invariance parameter. When $\alpha = 2$, the simplex parameter h_n becomes similarity invariant but alternative values for α can be chosen to consider non linear local scaling changes. In [Gil07], a value of $\alpha = 4$ was found as satisfactory to be used with shape memory forces that are presented in Sec. 5.3.3.2.

A.3.4 Local Descriptors

The constant vertex arity $k+1$ of k-simplex meshes yields efficient computation of local descriptors. For instance, the normal $\boldsymbol{n}$ at a point P can be easily approximated from the three neighbor positions P_1, P_2 and P_3 as depicted in Fig. A.4. Another important local quantity that is easily computed is the curvature at a vertex. Curvature computation has been intensively studied in the past, especially for triangular meshes [PP93, DMSB99, SMS+03]. H. Delingette proposed a discrete curvature that approximated the mean curvature K at P, the mean curvature being the average of the two principal curvatures [Del94b]:

$$K = \frac{sin(\phi)}{r} = \pm\frac{1}{R} \tag{A.2}$$

This discrete curvature converges to the real mean curvature when the mesh density increases.

A.3.5 Mesh Transformation and Multi-resolution Scheme

A.3.5.1 Basic Operators

In [Del94a, Del99], H. Delingette described four basic independent operators to transform a k-simplex mesh. The first two operators, T_1^k and T_2^k, delete and insert p-cells ($p \leq k$). These operators are coined *Eulerian* because they do not change the mesh topology (i.e. genus remains unchanged). Conversely, operators T_3^k and T_4^k can decompose the mesh into pieces or change its genus, and are thus referred to as *meta*-operators.

This is illustrated for example in Fig. A.1(b) where from a contour defined on the mesh, two new contours can be defined to split the mesh into two pieces. In this case, this corresponds to the T_3^2

operator. This operator is only valid with complete 2-simplex meshes [Del94a] which are used in this thesis.

A.3.5.2 Multi-resolution Scheme

Various resolutions of a shape can be generated with a multi-resolution scheme. There is a wide variety of multi-resolution schemes [RLB+02] that produce different lower resolutions of a mesh via simplification techniques [RM09] such as vertex removal [CCMS97], edge collapse [Hop96], vertex clustering [LT97] and vertex-pair collapse [GH97]. In the case of simplex meshes, the multi-resolution scheme has to preserve the vertex connectivity. H. Delingette [Del94b] proposed a scheme illustrated in Fig. A.5(a) which triples the number of points at each resolution increase. However, as reported in [Gil07], low resolution vertices are not conserved in the higher resolutions with this scheme. As a result some shape features might be lost, especially when faces are not planar or strong curvatures are formed between adjacent faces. The author proposed thus an alternative scheme, depicted in Fig. A.5(b), which preserves the vertices of the lower resolutions.

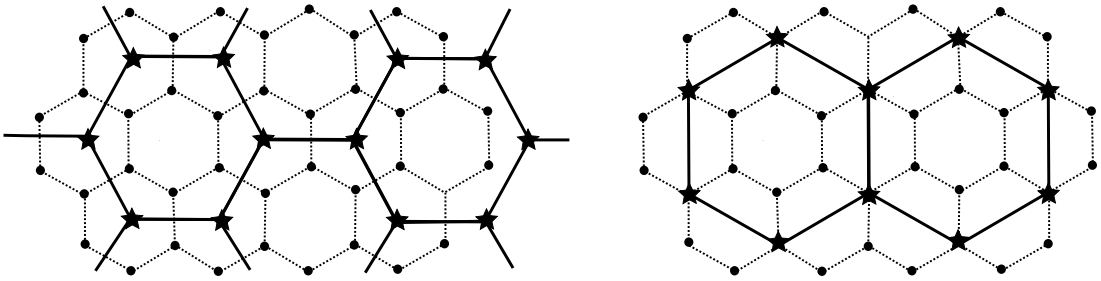


Figure A.5: Simplex mesh multi-resolution schemes. A 2-simplex mesh at resolution k ($\star$ vertices and thick edges) yields a new 2-simplex mesh at resolution $k+1$ ($\bullet$ vertices and dotted edges). (a) Classic multi-resolution scheme [Del94b]. (b) Conservative multi-resolution scheme [Gil07].

Given a mesh $\boldsymbol{X^k}$ at resolution k with N^k vertices x_j^k, we have: $x_j^{k+1} = x_j^k, \forall j \in [1, N^k]$. This scheme is hence capable of a better preservation of the features through resolution increase as vertices from lower resolution are retained, but at the expense of a higher memory footprint (number of points is multiplied by 4 between consecutive resolutions, i.e. $N_{k+1} = 4N_k$).

A.4 Shape Correspondence by Landmark Sliding

In Chapter 4, we use a correspondence method based on landmark sliding. In this Section, a more detailed description is provided. The following notations are taken from the work of Dalal *et al.* [DMW+07] which describes the correspondence method based on landmark sliding.

In such approach, an arbitrary shape with landmarks $U_L = \{u_i\},\ i \in [1, n]$ is chosen as the template shape. For each remaining target shape with landmarks $V_L = \{v_i\},\ i \in [1, n]$, the correspondence method consists in sliding the landmarks v_i over its surface by sliding distances computed by the minimization of a shape-correspondence error $\Phi(U_L, V_L)$. After each landmark sliding, the target landmark $\boldsymbol{v_i}$ is moved to $\boldsymbol{v_i} + \alpha_i \boldsymbol{p_i} + \beta_i \boldsymbol{q_i}$, where α_i and β_i are the sliding distances along two unit-length tangent vectors $\boldsymbol{p_i}$ and $\boldsymbol{q_i}$ at v_i, respectively. These local tangent directions $\boldsymbol{p_i}$ and $\boldsymbol{q_i}$ are computed by a local eignenanalysis computed by PCA. Given n target landmarks v_i, the correspondence by landmarks sliding consists to find hence the values of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ such that the function Φ is minimized:

$$\begin{array}{l}\Phi(U_L, V_L; \boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{\star\in\{x,y,z\}} (\boldsymbol{v}_\star + P_\star\boldsymbol{\alpha} + Q_\star\boldsymbol{\beta})^\top L(\boldsymbol{v}_\star + P_\star\boldsymbol{\alpha} + Q_\star\boldsymbol{\beta}) \\ +\lambda \sum_{i=1}^{n} \|v_i + \alpha_i p_i + \beta_i q_i - u_i\|^2\end{array} \tag{A.3}$$

where

$$\begin{array}{lcl}\boldsymbol{\alpha} & = & (\alpha_1, \ldots, \alpha_n)^\top \\ \boldsymbol{\beta} & = & (\beta_1, \ldots, \beta_n)^\top \\ P_\star & = & diag(p_{1\star}, \ldots, p_{n\star}) \\ Q_\star & = & diag(q_{1\star}, \ldots, q_{n\star}),\ \star \in \{x, y, z\}\end{array} \tag{A.4}$$

with the constraints

$$\begin{array}{lcl}|\alpha_i| & \leq & \epsilon \\ |\beta_i| & \leq & \epsilon,\ i = 1, \ldots, n\end{array} \tag{A.5}$$

These constraints prevent the landmarks from sliding too far at each step. This shape correspondence error uses the symmetric matrix L which expresses the required thin-plate bending energy [Boo89] to deform the template into the target shape surface.

The values of α_i and β_i can be *globally* computed with an efficient quadratic programming (QP) that accounts for constraints. In fact the minimization problem is a convex QP problem with bound constraints. Such a problem has for general form [GW03]:

$$\boldsymbol{x} = \underset{\boldsymbol{x}}{\operatorname{argmin}} \frac{1}{2}\boldsymbol{x}^\top Q\boldsymbol{x} + \boldsymbol{c}^\top\boldsymbol{x} \tag{A.6}$$

subject to

$$x_i \geq l_i,\ i \in \mathcal{L},\ x_i \leq t_i,\ i \in \mathcal{U}, \tag{A.7}$$

where Q is a $m \times m$ symmetric positive semi-definite matrix, $\boldsymbol{c}$ is a vector of dimension m, $\mathcal{L}$ and $\mathcal{U}$ are subsets of $\{1, \ldots, m\}$.

With some algebraic rewriting, it can be shown that our landmarks-sliding problem is equivalent to a QP problem with bound constraints where:

$$\boldsymbol{x} = (\boldsymbol{\alpha}^\top|\boldsymbol{\beta}^\top)^\top \tag{A.8}$$

$$m = 2n,\ l_i = -\epsilon,\ t_i = \epsilon,\ \mathcal{L} = \mathcal{U} = [1, 2n] \tag{A.9}$$

$$Q = 2\sum_{\star\in\{x,y,z\}} H_\star \tag{A.10}$$

$$H_\star = \left(\begin{array}{c|c} P_\star^\top(L + \lambda Id_n)P_\star & P_\star^\top(L + \lambda Id_n)Q_\star \\ \hline Q_\star^\top(L + \lambda Id_n)P_\star & Q_\star^\top(L + \lambda Id_n)Q_\star \end{array}\right) \tag{A.11}$$

$$c = 2\left(\begin{array}{c} \sum_{\star\in\{x,y,z\}} P_\star(L + \lambda Id_n)\boldsymbol{v}_\star - \lambda P_\star\boldsymbol{u}_\star \\ \sum_{\star\in\{x,y,z\}} Q_\star(L + \lambda Id_n)\boldsymbol{v}_\star - \lambda Q_\star\boldsymbol{u}_\star \end{array}\right) \tag{A.12}$$

To solve this QP problem for each pair of target and template models, we used the OOQP library of Gertz and Wright [GW03].

A.5 Internal Forces based on MRF Formulation

In Sec. 5.3.3.4, we described how to translate the MRF shape priors into forces. The aim of these forces is to minimize the local energy E_i at vertex y_i:

$$E_i = E_{d,i}(y_i, d_{y_i}) + \beta E_{m,i}(y_i, y_{\mathcal{N}(i)})$$

As previously discussed, only forces and force derivatives of the model energy $E_{m,i}$ have an analytic formulation, which is detailed here. $E_{m,i}$ is expressed as:

$$E_{m,i} = \sum_{i=1}^{M} \left(\frac{1}{\eta^2} \sum_{j \in \mathcal{N}(i)} \left\| \delta_i - \delta_j \right\|^2 + \frac{1}{\sigma^2} \left\| \delta_i \right\|^2 \right) \tag{A.13}$$

where $\delta_i = y_i - x_i$. The total model force $f_{m,i}$ for a particle y_i is the contribution of local forces:

$$f_{m,i} = -\frac{\partial E_{m,i}}{\partial y_i} - \sum_{j \in \mathcal{N}(i)} \frac{\partial E_{m,j}}{\partial y_i} = F^i_{m,i} + \sum_{j \in \mathcal{N}(i)} F^i_{m,j} \tag{A.14}$$

where:

$$F^i_{m,i} = \frac{2}{\eta^2} \left[\sum_{j \in \mathcal{N}(i)} (\delta_j - \delta_i) \right] - \frac{2}{\sigma^2} \delta_i \tag{A.15}$$

$$\forall j \in \mathcal{N}(i),\ F^i_{m,j} = \frac{2}{\eta^2} (\delta_i - \delta_j) \tag{A.16}$$

These expansions derive from the rule $\partial \left\| x - u \right\|^2 / \partial x = 2(x-u)\partial(x-u)/\partial x$. Similarly, force derivatives expressed as 3×3 matrices can be computed from the following quantities:

$$\frac{\partial F^i_{m,j}}{\partial y_k} = \begin{cases} -(\frac{6}{\eta^2} + \frac{2}{\sigma^2}) Id_3 & \text{if} \quad i = j = k \\ \frac{2}{\eta^2} Id_3 & \text{if} \quad i = j \text{ and } k \in \mathcal{N}(i) \\ \frac{2}{\eta^2} Id_3 & \text{if} \quad j \in \mathcal{N}(i) \text{ and } k = i \\ -\frac{2}{\eta^2} Id_3 & \text{if} \quad j \in \mathcal{N}(i) \text{ and } k \in \mathcal{N}(i) \end{cases} \tag{A.17}$$

Appendix B

Technical Background on Statistics and Probability Theory

B.1 Introduction

The use of knowledge into medical image analysis techniques essentially relies on fundamental principles of statistics and probability theory. As a result, a technical background on statistics and the probability theory is given in this appendix.

B.2 Probability Theory

Probability theory deals with the study and understanding of random, or chancy, phenomena. It is also the logical development of the behavior of some mathematical entities, called *random variables*, that will be presented in Sec. B.3. It is an ubiquitous theory widely spread in scientific areas such as data mining, information theory, signal and image processing, quantum mechanics, machine learning and economics.

An event E is composed of *outcomes*, which are the results of *random experiments*. These random events will exhibit in the "long run" some certain statistical patterns which can be predicted. Initially, these events were mainly discrete and were studied with combinatorial methods. Then continuous aspects of the events were taken in account in the modern probability theory, whose most of the foundations were laid by Andrey Nikolaevich Kolmogorov. He especially presented his notion of sample space Ω which denotes the set of all possible outcomes.

B.2.1 Probability

An event E is a set of outcomes and thus a subset of the sample space ($E \subset \Omega$). To each event a probability function $P(.)$ can be associated, which satisfies the following probability (Kolmogorov) axioms:

$$\forall E \subseteq \Omega,\ 0 <= P(E) <= 1 \tag{B.1}$$

$$P(\Omega) = 1 \tag{B.2}$$

$$\forall E, F,\ E \cap F = \varnothing,\ P(E \cup F) = P(E) + P(F) \tag{B.3}$$

Intuitively, a probability represents the "chances" that an event has to occur when the random experiment is performed. Two kinds of interpretation prevail: frequentist and Bayesian. In the first one, the probability is seen as as the long-run expected frequency of occurrence. It differs from the Bayesian (or subjective) interpretation in which the probability is related to degree of belief. This last interpretation is particularly legitimate when you consider experiments that cannot be repeated, a thing problematic for the frequentist approach.

B.2.2 Conditional Probability

The conditional probability $P(E|F)$ expresses "the probability of E if F is known to occur". The following equation illustrates the conditional probability:

$$P(E|F) = \frac{P(E \cap F)}{P(F)} \tag{B.4}$$

If $\{E_1, ..., E_n\}$ denotes a family of events such that

1. the events are mutually exclusive ($E_i \cap E_j = \varnothing$) and
2. their union is the sample space Ω ($\bigcup_{i=1}^{n} E_i = \Omega$),

we got the *total probability formula*:

$$\forall F \in \Omega,\ P(F) = \sum_{i=1}^{n} P(F|E_i)P(E_i) \tag{B.5}$$

B.2.3 Bayes' Rule

By using the previous conditional probability equation B.4, the *Bayes' rule* can be expressed:

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{P(F \cap E)}{P(F)} = \frac{P(F|E)P(E)}{P(F)} \tag{B.6}$$

where $P(E)$ is called the *prior probability* of E, $P(F|E)$ denotes the *likelihood* of F given E, and $P(E|F)$ is the *posterior probability*. Sometimes $P(F)$ is designated as the *evidence* and in some cases it is more easily computed with the total probability formula B.5. The Bayes' rule is at the origin of Bayesian inference which may convert a difficult problem (i.e. estimation of the posterior probability) to an easier one. Many segmentation approaches expresses the problem based on Bayesian rules and inference. The notion of inference is illustrated in Sec. B.4.

B.3 Random Variable

Most of the time, things we want to study can be described by numbers and the classic set theory applied to outcomes is not anymore the most appropriate. The notion of *random variable* (RV) is hence introduced to mimic outcomes that are the result of "chance".

B.3.1 Definition

A random variable X is a measurable function from the sample space Ω to another measurable space S called the *state space*: $X : \Omega \to \mathcal{S}$. If Ω is uncountable, events must be chosen into a suitable space $\mathcal{F}$ (σ-algebra) in order to be able to produce a probability for their outcomes. For instance, the Borel algebra of Ω is a common choice. The probability function P is thus a function from $\mathcal{F}$ to [0,1] and

the combination (Ω, F, P) is denoted as a *probability space*. In many practical cases, $\mathcal{S} = \mathbb{R}$ and X is a *real*-valued RV if $\{w \in \Omega : X(w) \leq x\} \in \mathcal{F},\ \forall x \in \mathbb{R}$.

In this manuscript, the probability of $\{w \in \Omega : X(w) \leq x\}$ is written as $P(X \leq x)$ where uppercase letters denote RV (e.g. X) whereas lowercase letters represent the values assigned by the RV (e.g. x). If S is a part of the real line $\mathbb{R}$, the RV X is termed as a *continuous* RV whereas if Ω maps into the integers, we speak about a *discrete* RV.

B.3.2 Probability Density Function

$P(X \leq x)$ can be interpreted as a function of x, the *cumulative distributive function* (or CDF) Φ_X, which satisfies the following properties:

$$1)\ 0 \leq \Phi_X(x) \leq 1 \tag{B.7}$$

$$2)\ \lim_{x \to -\infty} \Phi_X(x) = 0 \text{ and } \lim_{x \to +\infty} \Phi_X(x) = 1 \tag{B.8}$$

$$3)\ \Phi_X(x+h) \geq \Phi_X(x) \text{ for } h \geq 0 \qquad (\Phi_X \text{ is non-decreasing}) \tag{B.9}$$

$$4)\ \lim_{h \to 0^+} \Phi_X(x+h) = \Phi_X(x) \qquad (\Phi_X \text{ is right continuous}) \tag{B.10}$$

When the CDF can be derived with respect to x, the *probability density function* ϕ_X for a RV X can be defined:

$$\phi_X(x) = \frac{d\Phi(x)}{dx} \tag{B.11}$$

Probability density functions (PDF) must satisfy:

$$1)\ \phi_X(x) \geq 0 \text{ for all } x \tag{B.12}$$

$$2)\ \int_{Supp_X} \phi_X(x)dx = 1, \text{ where } Supp_X = \{x \in \mathbb{R}, \phi_X(x) > 0\} \text{ (support of } \phi_X) \tag{B.13}$$

Another notation for stating that X is distributed with a PDF ϕ is: $X \sim \phi$. It is also possible to define a sort of probability distribution f for discrete RV X by using the set $\{p_i\} = \{P(X = x_i), i = 1, 2, ...\}$, the PDF being defined through Dirac functions:

$$f_X(x) = \sum_i p_i \delta(x - i) \tag{B.14}$$

Such function is called *probability mass function* (PMF) to make a clear distinction with respect to classical continuous RV PDFs. However sometimes it is more convenient to use an unified notation, so in in this manuscript PDF and PMF may be written as $P_X(x)$. Note that for a continuous RV X, $P(X = x)$ is null for all x. Probability distribution and probability density function will be considered as synonymous notions in this thesis.

Table B.1 illustrates some of the main discrete and continuous distributions. For example, let us consider $X \sim \mathcal{B}(n, p)$. The second column tell us that the distribution of the Binomial law is: $P(X = k) = C_k^n p^k (1-p)^{n-k}$, $k \in \{0, ..., n\}$. This law can be also thought as the sum of n independent RVs $X_i \sim \mathcal{B}(1, p)$ (X_i is also said to follow a *Bernoulli*'s law). One of the major continuous distribution is the *Normal* distribution $\mathcal{N}(\mu, \sigma)$ (also named *Gaussian* distribution). Its PDF is:

$$\mathcal{N}(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{B.15}$$

Many physiological and physical phenomena can be approximated by this distribution. Moreover the *Central Limit Theorem* states that the sum of independent RVs X_i defined on the same probability space and sharing the same distribution can be approximated by a Normal distribution. More precisely, the RV $S_n = X_1 + ... + X_n$ will converge toward a Normal distribution when n approaches $+\infty$. The Normal distribution is thus often used as an approximation of some phenomena due to its simplicity and numerous nice properties.

Law **Parameter(s)**	**Type**	**Density** **or Probabilities**	**Expectation**	**Variance**
Binomial $\mathcal{B}(n,p)$ $n \in \mathbb{N}^*, p \in]0,1[$	discrete	$C_k^n p^k (1-p)^{n-k}$ $k \in \{0,...,n\}$	np	$np(1-p)$
Negative Binomial $n \in \mathbb{N}^*, p \in]0,1[$	discrete	$C_{k-1}^{n-1} p^n (1-p)^{k-n}$ $k \in \{n, n+1, ...\}$	$\frac{n}{p}$	$n\frac{1-p}{p^2}$
Poisson $\mathcal{P}(\lambda)$ $\lambda \in \mathbb{R}^*$	discrete	$e^{-\lambda}\lambda^k/k!$ $k \in \mathbb{N}$	λ	λ
Geometric $\mathcal{G}(p)$ $p \in]0,1[$	discrete	$p(1-p)^{k-1}$ $k \in \mathbb{N}^*$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Uniform $\mathcal{U}(a,b)$ $a \in \mathbb{R}, b \in \mathbb{R}, a < b$	continuous	$\frac{1}{(b-a)}\mathbb{1}_{]a,b[}(x)$ $x \in \mathbb{R}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normal $\mathcal{N}(\mu,\sigma)$ $\mu \in \mathbb{R}, \sigma \in \mathbb{R}^{+*}$	continuous	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $x \in \mathbb{R}$	μ	σ^2
Exponential $\mathcal{E}(\lambda)$ $\lambda \in \mathbb{R}^{+*}$	continuous	$\lambda e^{-\lambda x}\mathbb{1}_{\mathbb{R}^+}(x)$ $x \in \mathbb{R}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

Table B.1: Classic probability distributions

B.3.3 Expectation and Variance

Sometimes it can be interesting to summarize certain attributes that describe some behaviors of the PDF. It must be stressed that these attributes are not RV although they result from operations applied on the

distributions. A classical attribute is the *expectation* or *mean* $E[X]$ of a RV X:

$$E[X] = \int_{Supp_X} x\phi_X(x)dx \qquad \text{(continuous formulation)} \tag{B.16}$$

$$E[X] = \sum_i x_i p_i \qquad \text{(discrete formulation)} \tag{B.17}$$

The expectation can be seen as the center of gravity or the central value of a distribution. The spread or variability of a distribution can be expressed by the variance $Var(X)$:

$$Var(X) = \int_{Supp_X} (x - E[X])^2 \phi_X(x)dx \qquad \text{(continuous formulation)} \tag{B.18}$$

$$Var(X) = \sum_i (x_i - E[X])^2 p_i \qquad \text{(discrete formulation)} \tag{B.19}$$

$$Var(X) = E[(X - E[X])^2] \qquad \text{(general formulation)} \tag{B.20}$$

$$Var(X) = E[X^2] - E[X]^2 \qquad \text{(Huygens formula)} \tag{B.21}$$

$\sqrt{Var}$ is referred to as the *standard deviation*. Additional characteristics of a probability distribution are given by higher-order moments of the density function such as the *skewness* which measures the asymmetry of a distribution. Table B.1 reports the variance and expectation of some classical distributions. All listed distributions in Table B.1, expect the Uniform $\mathcal{U}$ distribution, belong to the *exponential* distributions family and are fully described by their expectation and variance statistics. More generally, distributions of the exponential family are fully explained by a finite number of *sufficient* statistics.

B.3.4 Independence and Correlation

Two RVs X and Y are *not* correlated if and only if $E[XY] - E[X]E[Y] = 0$. The correlation expresses in some ways the strength and direction of a linear relationship between two RVs. Independence can be expressed in terms of conditional probabilities or expectation. X and Y are independent if and only if:

$$E[g_1(X)g_2(Y)] - E[g_1(X)][g_2(Y)] = 0 \text{ for } \textit{any} \text{ measurable functions } g_1, g_2 \text{ or,} \tag{B.22}$$

$$P(X|Y) = P(X) \tag{B.23}$$

Independence implies thus much more restrictive conditions to be satisfied compared to uncorrelation. That is why in many real and practical cases, independence is hardly satisfied. However the independence consequences and properties make it an extremely appealing condition because it can really simplify problems which would have been not tractable without independence assumption. As a consequence, independence is assumed to be verified in some problems although it is not the case. This may lead to incorrect results but it may in some cases not have drastic consequences. Some important properties of independence and correlation are as follows:

$$X, Y \text{ independent} \Rightarrow X, Y \text{not correlated (from eq. B.22)}, \tag{B.24}$$

$$X, Y \text{ independent} \Leftrightarrow P(X, Y) = P(X)P(Y), \tag{B.25}$$

$$\text{“}X, Y \text{ independent} \Leftrightarrow X, Y \text{ not correlated”} \Leftrightarrow X, Y \sim \mathcal{N} \tag{B.26}$$

$$X_i, \ldots, X_n \text{ independent} \Rightarrow Var(X_1 + \ldots + X_n) = Var(X_1) + \ldots + Var(X_n) \tag{B.27}$$

The second property B.25 can be reformulated for continuous RVs by using their PDFs: X and Y are independent if their joint PDF is the product of the marginal PDFs:

$$\phi_{X,Y}(x,y) = \phi_X(x)\phi_Y(y) \tag{B.28}$$

B.4 Statistical Inference

In real cases, only the observations of a physical process are available rather than the underlying RVs. This means that the collection and analysis of data will be used to estimate the form of the probability distributions of RVs as well as the *parameters* that model the process. Statistics come into play and the estimation procedure is denoted as statistical *inference*.

In Statistics, the observed data is a *sample* of the *population* that represents the elements about which we want to infer. It is then a subset of the population data, the difficulty consists thus in learning about the population from a sample. The observed data result from a physical process which can be modeled by a RV X. A random sample $\{X_1, \ldots, X_n\}$ is composed of n mutually independent RVs X_i that are identically distributed (i.e. $\forall i \in [1,n], X \sim X_i \sim \mathcal{L}$). The independent observations $\{x_1, \ldots, x_n\}$ are values assigned by these RVs X_i. The purpose of the inference is to learn something about X by analyzing the observed data.

B.4.1 Estimators

Let us suppose the RV $X = \{X_1, \ldots, X_n\}$ has a PDF $\phi_X(.;\theta)$, where θ is a parameters vector characterizing the distribution $\mathcal{L}$. An estimator $\hat{\theta}_n = \hat{\theta}(X_1, \ldots, X_n)$ is a RV that will estimate θ. The fact that the form of the distribution law of X is *known* must be stressed because an estimator will try to estimate the parameters *not* how X is distributed. By definition, $\hat{\theta}$ is an estimator converging toward θ if and only if $\hat{\theta}$ converges *in distribution* toward θ:

$$\lim_{n\to\infty} \hat{\theta}(X_1, \ldots, X_n) = \theta \text{ for any series of independent RVs } X_i \text{ identically distributed} \tag{B.29}$$

This kind of convergence is also referred to as convergence in law or *weak* convergence. It is a mode of convergence implied by all other kind of RV convergences.

B.4.1.1 Consistency

A more restrictive mode of convergence is the convergence *in probability*: $\hat{\theta}_n$ converges in probability to θ if for every $\epsilon > 0$,

$$\lim_{n\to\infty} P(|\hat{\theta}_n - \theta| > \epsilon) = 0 \tag{B.30}$$

If an estimator converges in probability toward the parameter θ to estimate, the estimator is said to be *consistent*. Consistency can be interpreted as the need to have progressively better estimates of the sought parameter with growing sample sizes.

B.4.1.2 Unbiasness

A highly desirable property for an estimator is to be *unbiased*:

$$\forall n,\ b(\hat{\theta}_n) = E[\hat{\theta}_n] - \theta = 0 \tag{B.31}$$

which means that the expected value of the estimator must be equal to the true value of the parameter. b is called the *bias* of the estimator. Biasness is not necessarily a disadvantage, biased estimators may have desirable properties. For instance the variance $Var(\hat{\theta}_n)$ of a biased estimator can converge faster to 0 that the variance of an unbiased estimator of the same parameter. As a consequence, if we consider the *Bienaymé-Tchebicheff inequality*

$$P(|\hat{\theta}_n - \theta| \geq \epsilon) \leq \frac{1}{\epsilon^2} Var(X) \tag{B.32}$$

and the fact that the variances of the biased and unbiased estimator converge to 0: (i) the consistency property is verified but most importantly (ii) the biased estimator tends to be "consistent in a faster way" than the unbiased one. This can be interesting in real applications where reaching large values of n is almost impossible, the biased estimator can prove hence to be a more useful and powerful statistic.

B.4.1.3 Examples

Traditional estimators are presented in this paragraph. Let us consider first the following estimator:

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{B.33}$$

The *weak law of large numbers* states that for every $\epsilon > 0$,

$$\lim_{n \to \infty} P(|\bar{X}_n - \mu| > \epsilon) = 0 \tag{B.34}$$

where $\mu = E[X_i] = E[X]$, $\forall i$. That is the estimator $\bar{X}_n$ converges in probability toward the expectation of RV X, so it is consistent. Moreover it is easy to show that $E[\bar{X}_n] = \frac{1}{n} \sum_{i=1}^{n} E[X_i] = \frac{1}{n} \sum_{i=1}^{n} E[X] = E[X]$, this estimator is hence additionally unbiased. Let us consider now the following estimator:

$$S_n^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu)^2 \tag{B.35}$$

It is a consistent estimator of the variance $Var[X]$ but is a biased because $E[S_n^2] = Var[X] - E[(E[X] - \mu)^2] = \frac{n-1}{n} Var[X] < Var[X]$. An unbiased estimator of the variance is thus

$$S'^2_n = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \mu)^2 \tag{B.36}$$

As discussed in the previously, the biased version of the variance estimator is not always a bad choice because its lower variance can yield greater statistical power. Note that if μ is replaced by the true expectation $E[X]$ in Eq. (B.35), the estimator S_n^2 is *not* biased.

B.4.2 Maximum *a posteriori* (MAP)

The MAP estimation relies on a simple but powerful principle: given the observation sample $x = \{x_1, \ldots, x_n\}$, the best parameters vector θ is the one which maximizes the probability of having θ given the observation sample x. This kind of estimation is possible when we have access to (or suppose to know) the form of the law of the RV X which dictates the behavior of the observation RVs X_i. The MAP estimation looks for an estimate $\hat{\theta}_n$ such as

$$\hat{\theta}_n = \underset{\theta}{\operatorname{argmax}}\, P(\theta|x) \tag{B.37}$$

It may happen that such a posterior probability may be hard to compute, the Bayes' rule B.6 may be then preferred for the MAP formulation:

$$\hat{\theta}_n = \underset{\theta}{\operatorname{argmax}}\ P(\theta|x) = \underset{\theta}{\operatorname{argmax}}\ \frac{P(x|\theta)P(\theta)}{P(x)} = \underset{\theta}{\operatorname{argmax}}\ P(x|\theta)P(\theta) \tag{B.38}$$

the last equality results from the fact that x is supposed to be independent of θ. We will see later that it is a formulation commonly used for problems such as recognition (e.g. [JDM00]) and segmentation (e.g. [PFG+06, Cua03, LGF00]).

B.4.3 Maximum Likelihood (ML)

R.A. Fisher introduced and developed the concept of ML estimation in the 1920s. The basic idea is that the ML estimation will select a parameters estimate that makes the data more likely to occur. The *likelihood function* for the parameter θ is defined as:

$$L(\theta|x) = P(x_1, \ldots, x_n|\theta) \tag{B.39}$$

$$= \phi_X(x_1, \ldots, x_n|\theta) \qquad \text{(continuous)} \tag{B.40}$$

$$= P(X_1 = x_1, \ldots, X_n = x_n|\theta) \qquad \text{(discrete)} \tag{B.41}$$

The *maximum likelihood estimator* (MLE) of θ finds an estimate $\hat{\theta}_n$ by maximizing the likelihood function:

$$\hat{\theta}_n = \underset{\theta}{\operatorname{argmax}}\ L(\theta) \tag{B.42}$$

The ML estimation can be also thought as a special case of the MAP estimation where the prior $P(\theta)$ is omitted by considering that it follows a uniform law. When the RVs X_i, from which the x_i are drawn, are independent the MLE is rewritten as a product of marginals PDFs or PMFs:

$$\hat{\theta}_n = \underset{\theta}{\operatorname{argmax}}\ \prod_{i=1}^{n} P_{X_i}(x_i|\theta) \tag{B.43}$$

In general, the MLE may not be unique or may not even exist. In case of independence of the RVs X_i the estimator is *asymptotically* unbiased, i.e. its bias tends to zero when $n \to +\infty$.

B.4.4 Expectation Maximization (EM) Method

The Expectation-Maximization algorithm was introduced by Dempster *et al.* [DLR77]. The MLE being is some cases intractable, the EM-algorithm proposes a way to compute the MLE based on an iterative procedure. Remember that the objective here is to find $\hat{\theta}$ which maximizes $L(\theta)$ (Eq. (B.42)) or, for convenience, the log likelihood function $L'(\theta) = lnP(x|\theta)$. The idea of EM is to introduce an hidden observation z with RV Z that will make the maximization of the likelihood function $L(\theta)$ much simpler. Consider the following function $l(\theta|\theta_n)$:

$$l(\theta|\theta_n) = L'(\theta) + \sum_z P(z|x, \theta_n) ln\left(\frac{P(x|z, \theta)P(z|\theta)}{P(x|z, \theta_n)P(x|\theta_n)}\right) \tag{B.44}$$

It can be proven [Bor04] that $l(\theta|\theta_n)$ is bounded above by the log likelihood function $L'(\theta)$ and that for $\theta = \theta_n$, functions $l(\theta|\theta_n)$ and $L'(\theta)$ are equal. The EM algorithm iteratively updates θ_n so that $l(\theta_n|\theta_{n-1})$

increases and so does $L'(\theta_n)$. Formally:

$$\begin{aligned} \theta_{n+1} &= \underset{\theta}{\operatorname{argmax}}\, l(\theta|\theta_n) \\ &= \underset{\theta}{\operatorname{argmax}}\, \textstyle\sum_z P(z|x,\theta_n) lnP(x,z|\theta) \\ &= \underset{\theta}{\operatorname{argmax}}\, E_{Z|x,\theta_n}[lnP(x,z|\theta)] \end{aligned} \tag{B.45}$$

where $E_{Z|x,\theta_n}[lnP(x,z|\theta)]$ is the conditional expectation. EM steps are thus:
– E-step: compute $E_{Z|x,\theta_n}[lnP(x,z|\theta)]$
– M-step: maximize it with respect to θ
While this seems to complicate the problem, the EM algorithm becomes very handy when observations x are distributed according to an exponential family. For instance, in case where x is distributed according a multivariate Gaussian distribution, closed forms exist to compute at each iteration step the new value of θ_n.

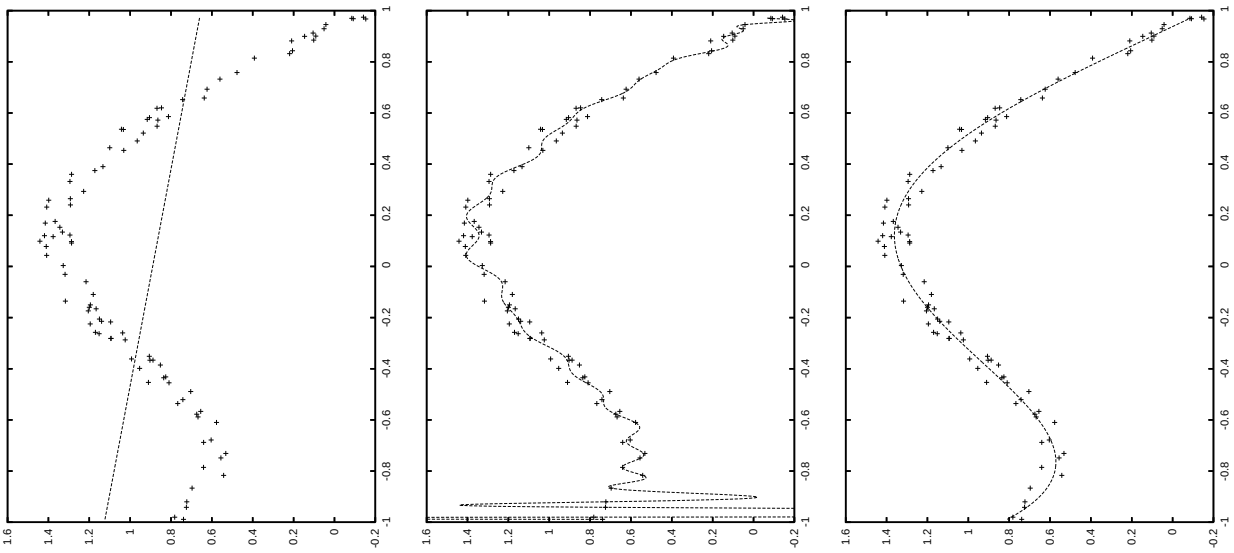
B.4.5 Minimum Description Length (MDL)

Introduced by Jorma Rissanen in 1978, the MDL principle is a powerful concept that can be sometimes considered as a form of statistical inference. According to Grünwald [Grü04], statistical inference can be seen as finding regularity in the observed data. Indeed, if we are able to compress the data, i.e. describe it with fewer symbols than those used in a literal description of the data, we are in a way *learning* about it.

In the MDL principle, a trade-off between the goodness-of-fit and the model complexity is sought. The MDL criterion tries to satisfies the Occam's Razor concept in *automatically* choosing a model $\mathcal{M}$ that expresses as close as possible the data $\mathcal{D}$ with the lowest complexity ("simplest" model formulation with as less as possible parameters). The MDL principle stems from an idea of data transmission: the goal is to transmit the shortest message embedding the data model description and the parameters associated with it. Let us consider for example the Fig. B.1 where three different polynomials are considered to express the relationship (the model) between the x-coordinate and y-coordinate of a series of 100 points. These polynomials belong to a family of models. The first one (Fig. B.1(a)) is a simple line, it is assumed that the relationship is linear. The result is a very poor approximation. The second polynomial of degree 40 (Fig. B.1(b)) performs a good approximation with a sum squares error (SSE) value of 0.168. Although this model description approximates accurately the *observed* data, it may approximate very poorly *new* data generated by the real underlying model. In fact, some point positions may have been corrupted by noise. This model is too complex and tends to be *overfitted*, it performs well for the training data but have great chances to become error-prone with new data. On the contrary, the last polynomial of degree 4 (Fig. B.1(c)) offers a less good fitting (SEE=0.26) but it intuitively appears as a more adequate model to describe the data. The MDL procedure will specifically attempt to find such model by balancing complexity and goodness-of-fit.

From a more formal viewpoint, a *language* $\mathcal{L}$ is used to describe properties of the data. This language must be a general-purpose language with which a program can be devised to output the data and to halt. The shorter is the program, the better is the compression and more we learn from the data. The MDL strategy consists then in finding such *shortest* program. But this is an uncomputable problem: a general procedure that given an arbitrary sequence of data finds the shortest program and stops does not exist.

A *Two-stage* MDL approach is then used in practice: let us consider the languages $\mathcal{L}_m(\mathcal{M})$ and $\mathcal{L}_c(\mathcal{D}|\mathcal{M})$



(a) (b) (c)

Figure B.1: MDL principle illustration: (a) simple, (b) complex and (c) trade-off polynomials (dashed curves) to approximate observed data (crosses).

that describe the model $\mathcal{M}$ and the data $\mathcal{D}$ given the model $\mathcal{M}$ respectively. Let us consider the operator $|.|$ as the length in bits of a description. This MDL strategy will then choose a model $\mathcal{M}^*$ such as [Grü04, Lec89]:

$$\mathcal{M}^* = \underset{\mathcal{M}}{\operatorname{argmin}} \, |\mathcal{L}_c(\mathcal{D}|\mathcal{M})| + |\mathcal{L}_m(\mathcal{M})| \tag{B.46}$$

We can consider that the term $|\mathcal{L}_c(\mathcal{D}|\mathcal{M})|$ refers to a notion of *coverage* (i.e. "fit at best the data") whereas the term $|\mathcal{L}_m(\mathcal{M})|$ to a notion of *conciseness* (i.e. "use a simple model"). An interesting point is that MDL can be equivalent to a MAP formulation when we have some prior probabilities on the model and the data. Indeed, Rissanen [Ris83] showed that optimal descriptive languages $\mathcal{L}_m^*$ and $\mathcal{L}_c^*$ exist and that they satisfy the following properties [Lec89, HY01]:

$$|\mathcal{L}_m^*(\mathcal{M})| = -\log_2 P(\mathcal{M}) \tag{B.47}$$

$$|\mathcal{L}_c^*(\mathcal{D}|\mathcal{M})| = -\log_2 P(\mathcal{D}|\mathcal{M}) \tag{B.48}$$

So the MDL strategy consists in finding the model $\mathcal{M}$ that minimizes the quantity $-\log_2 P(\mathcal{M}) - \log_2 P(\mathcal{D}|\mathcal{M})$, which is thus equivalent to a MAP estimation. Albeit both estimation strategies are equivalent, the practical choice of a strategy depends from the context. Sometimes a descriptive language can be more easily computed than some prior probabilities or vice-versa. Furthermore, it may happen that prior probabilities are better suited to describe the coverage while a description length expresses more naturally the conciseness (e.g. [Lec89]), resulting in a mix of both strategies.

To conclude, MDL strategy can be used for parametric estimation in which a model is selected among a family of models and its parameters are computed. Concretely speaking, the two-stage strategy is adopted [HY01]:

- Select a model, compute its parameters with an estimator $\hat{\theta}_n$ and the description length $|\mathcal{L}_m^*(\hat{\theta}_n)|$
- Express the data given the selected model and compute its description length $|\mathcal{L}_c^*(\mathcal{D}|\hat{\theta}_n)|$
- Find the model which minimizes the quantity $|\mathcal{L}_m^*| + |\mathcal{L}_c^*|$

B.5 Nonparametric Estimation

In the previous section, the statistical inference dealt mostly with *parametric estimation*. Indeed, we saw how estimators could be used to find parameters associated with a given distribution model. It was assumed that the model was *known*. Different strategies were thus presented to infer the parameters from the data and the known model.

In many real cases however, the distribution law that governs a process is not accessible, parameters inference is no longer possible. This section will hence present some *nonparametric estimation* techniques that try to cope with this absence of information. Basically, different kinds of approaches exist: (i) those which try to build a distribution model from the data and estimate the model parameters, and (ii) those that will directly estimate a PDF or a PMF from the data.

B.5.1 Histogram Approach

The histogram approach is one of the simplest method which derives essentially from the frequentist approach of probability (Sec. B.2.1). Let us assume that the data belongs to a N-dimensional space and that it is restricted to a finite support D. This support is partitioned into a fixed grid of m non overlapping bins B_k (an uniform distribution is not mandatory). Let v_k denotes the volume of each bin B_k (the volume is the product of the N side lengths of a bin). Let consider our usual observation sample $\{x_1, \dots, x_n\}$ and a function I_{B_k} such as:

$$I_{B_k}(x) = \begin{cases} 1, & \text{if } x \text{ falls into bin } B_k \\ 0, & \text{otherwise} \end{cases} \tag{B.49}$$

Letting $n_k = \sum_{i=1}^{n} I_{B_k}(x_i)$ be the number of sample values falling into B_k, we can estimate the N-dimensional PDF of X as:

$$P_X(x) = \sum_{k=1}^{m} \frac{n_k}{n v_k} I_{B_k}(x), \ \forall x \in D \text{ (0 otherwise)} \tag{B.50}$$

The quality of the histogram estimation depends on the number of bins (or on the bin volume, since both quantities are related) and the data sample size n. Usually, more the number of bins is high, more the estimation becomes accurate. This however remains valid in cases where the sample has an "acceptable" size. An infinite number of bins will change nothing if the sample size is too small. This constraint is even more inevitable in high dimensions, i.e. when N becomes significantly large. This phenomenon is called the *Curse of the Dimensionality* and it will be studied more in details in the next Sec. B.5.3.

B.5.2 Kernel Density Estimation (KDE)

Kernel density estimation provides a way to estimate by interpolation techniques an unknown PDF from observed samples $x_1, \dots, x_n$. In KDE, a kernel function is generated around each sample x_i and an estimate P_n of the PDF is found by summing up and normalizing the kernel contributions:

$$P_n(x) = \frac{1}{C} \sum_{i=1}^{n} K(x_i)) \tag{B.51}$$

where K is the kernel function in the d-dimensional space and C is a normalization constant so that the integral of P_n over $\mathbb{R}^d$ is equal to 1.

B.5.2.1 Parzen Windowing

The most common KDE approach is the Parzen windowing [Par62]. In this approach, the kernel is a Gaussian function, i.e. $K(x) = \mathcal{N}(x; x_i, \Sigma)$, where Σ denotes the covariance matrix. It is assumed that this covariance matrix is the same for all samples. Moreover, it is common to assume that $\Sigma = \sigma^2 I$ which implies that the standard deviation σ is the same along each dimension. The standard deviation is also called the bandwidth. It can be proved that when $n \to \infty$, P_n converges to the true PDF [Par62]. With a finite number of samples, the bandwidth has to be carefully selected otherwise the PDF estimate will differ too much from the real PDF. Various approaches are possible, the most popular method is to use an ML estimator coupled with cross validation to regularize the PDF estimate (otherwise a naive ML would return $\sigma = 0$)[Awa06].

B.5.2.2 Other approaches

In [PV03], three approaches were studied: the Parzen windowing, the k–NN KDE and the k–neighbor Parzen (k–NP) method which combines both approaches. k–NN KDE considers different Σ_i for each sample x_i in order to better express local distributions and uses the $k-1$ closest neighbors to x_i to estimate the covariance matrix. One of the main issues with KDE is their high computational burden, especially in high dimension [GM03]. As a result, various approaches were proposed to reduce the complexity while preserving at best the accuracy. Gurwicz and Lerner [GL05] proposed a spline-approximated KDE which consisted in smoothing the density by splines. Splines being manipulated with fewer coefficients, a significant reduction of the amount of data to be estimated was achieved, improving thus the KDE speed. Gray and Moore [GM03] presented various optimization strategies, especially based on space-partioning trees. They proposed in particular their Dual-Tree method which was later on compared in [RDG07] with an improved version of the fast Gauss transform. The fast Gauss transform (FGT) [GS91] is an approximation of the sum of N Gaussians at M d–dimensional points, which already achieves an important speed-up by reducing the complexity from $\mathcal{O}(NM)$ to $\mathcal{O}(N+M)$. The improved version of Raykar *et al.* [RDG07] proposed to extend the FGT by adopting a clustering of the samples based on the k-means algorithm. The idea is that when the PDF is estimated at x, instead of using all samples it appears as more appropriate to only use the samples present in the same cluster of x, thus reducing the number of computations. In the same philosophy, Zhang *et al.* [ZTK05] improved the mean shift algorithm by only considering neighbor information in order to devise a fast KDE approach. Eventually, effort can be also put into using adapted hardware and parallelization techniques as illustrated by Ohmer *et al.* [OMB05] who implemented KDE methods on the GPU.

B.5.3 Curse of the Dimensionality

To understand the *curse of dimensionality* [Bel57], consider the histogram approach presented in Sec. B.5.1. In this approach, the number of bins increases exponentially with the dimensionality. In practice, this means that the number of samples should be sufficiently large with respect to the features dimension. Otherwise, the histogram kernel density estimation quality degrades. In case of density estimation, this phenomenon is also called *empty space phenomenon* [PP97], which in case of histograms translates as an empty bin problem. In the domain of classifier, Jain *et al.* [JDM00] discussed in particular the effects of the curse of dimensionality which leads to the *peaking phenomenon* in classifiers design.

Hopefully, some cases exist where statistical inference is not totally cursed by the absence of a large quantity of data. In fact, various works observed that in some situations features of interest of high dimension actually lie on *low-dimensional* manifolds, and as a consequence any knowledge on this sub-

space might compensate for the lack of large training dataset [Awa06]. For instance, Bengio and Vincent [BV04] improve the Parzen windowing KDE by devising an algorithm able to capture the local structure of the underlying manifold. Popat and Picard [PP97] proposed the use of clustering approaches to address the problem of the curse of dimensionality.

B.6 Markov Random Fields

B.6.1 Definition

A random field is a family of RVs $X = \{X_t\}_{t\in\mathcal{T}}$ where each RV X_t is defined on a sample space Ω [Awa06]. We can associate a neighborhoods family $\mathcal{N} = \{\mathcal{N}_t\}_{t\in\mathcal{T}}$ with the index set $\mathcal{T}$ such that [Awa06]:

$$\begin{aligned} &\mathcal{N}_t \subset \mathcal{T} \\ &t \notin \mathcal{N}_t \\ &(u \in \mathcal{N}_t) \Leftrightarrow (t \in \mathcal{N}_u) \end{aligned} \tag{B.52}$$

$\mathcal{N}_t$ denotes the indices of the neighborhood of index t. We can now define a Markov Random Field (MRF) $X(\Omega, \mathcal{N})$ as:

$$P(x_t) > 0,\ \forall t \rightarrow P(x_1, \ldots, x_{|\mathcal{T}|}) > 0 \qquad \text{(positivity)} \tag{B.53}$$

$$\forall t,\ P(X_t|\{X_u\}_{u\in\{\mathcal{T}\setminus\{t\}\}}) = P(X_t|\{x_t\}_{t\in\mathcal{N}_t}) \qquad \text{(Markovianity)} \tag{B.54}$$

The Markovianity property implies that given the neighborhood $\mathcal{N}_t$ of an RV X_t, the remaining RVs do not provide additional information about this RV. This property allows many inference problems to become tractable as the very high complexity of computing $P(X_t|\{X_u\}_{u\in\{\mathcal{T}\setminus\{t\}\}})$ is reduced to a subset of RVs defined in the neighborhood. Based on the Bayes' rule (B.6), this moreover simplifies the computation of $P(X_1, \ldots, X_{|\mathcal{T}\|})$.

B.6.2 Gibbs Distribution

Despite the Markovianity property reduces the complexity of inference, the computation of $P(x)$ can remain very difficult. Besag [Bes74] presented thus the link between MRF and Gibbs random field (GRF). Before defining a GRF, we must define the notion of clique. A clique c is a subset of the index set $\mathcal{T}$ if every distinct pair of indices in c are neighbors given a neighborhood $\mathcal{N}_t$. We denote $\mathcal{C}$ the collection of all possible cliques for the neighborhood family $\mathcal{N}$.

A GRF follows a Gibbs distribution [Awa06]:

$$\begin{aligned} P(x) &= \frac{1}{Z} e^{-\frac{U(x)}{\tau}} \\ U(x) &= \sum_{c\in\mathcal{C}} V_c(x) \\ Z &= \int_x e^{-\frac{U(x)}{\tau}} dx \end{aligned} \tag{B.55}$$

where τ is the temperature, U is the energy function, V_c denotes a clique-potential function and Z refers to as the partition function. The Hammersley-Clifford theorem [Bes74] states that a MRF is equivalent to a GRF. This provides thus a clear formulation of the MRF distribution. Nevertheless, inference schemes are necessary to estimate the unknown parameters of the Gibbs distribution.

A Gauss-Markov process is a continuously-valued random vector following a multivariate Gaussian distribution $\mathcal{N}(\mu, \Sigma)$. The Markovianity property will produce a spare inverse covariance matrix Σ^{-1}

where each variable interacts with a few others through the quadratic energy function U; index i is neighbor of j if $\Sigma^{-}1_{ij} = \Sigma^{-1}_{ji} \neq 0$ [Per98].

Bibliography

[3DA10] "3d anatomical human project," http://3dah.miralab.ch/, 2006-2010.

[AA04] M. Alexa and A. Adamson, "On normals and projection operators for surfaces defined by point sets," in *Proc. Eurographics Symp. Point-based Graphics*, M. Alexa, M. Gross, H. Pfister, and S. Rusinkiewicz, Eds. Eurographics Association, 2004, pp. 149–156.

[AASK08] V. Athitsos, J. Alon, S. Sclaroff, and G. Kollios, "Boostmap: An embedding method for efficient nearest neighbor retrieval," *IEEE Trans Pattern Anal Mach Intell*, vol. 30, no. 1, pp. 89–104, 2008.

[AB02] S. Aylward and E. Bullitt, "Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction," *IEEE Trans Med Imag*, vol. 21, no. 2, pp. 61–75, 2002.

[ABG+09] M. Agus, F. Bettio, A. Giachetti, E. Gobbetti, J. Iglesias Guitián, F. Marton, J. Nilsson, and G. Pintore, "An interactive 3D medical visualization system based on a light field display," *Vis Comput*, vol. 25, pp. 883–893, 2009.

[ABKS99] M. Ankerst, M. Breunig, H. Kriegel, and J. Sander, "OPTICS: ordering points to identify the clustering structure," in *Proc. SIGMOD: Management of data*. ACM Press, 1999, pp. 49–60.

[Ack95] M. J. Ackermann, "The visible human project," http://www.nlm.nih.gov, 1995.

[ACK01] L. A. Alexandre, A. C. Campilho, and M. Kamel, "On combining classifiers using sum and product rules," *Pattern Recogn Lett*, vol. 22, no. 12, pp. 1283–1289, 2001.

[ACS+09] L. Assassi, C. Charbonnier, J. Schmid, P. Volino, and N. Magnenat-Thalmann, "From MRI to Anatomical Simulation of the Hip Joint," *Computer Animation and Virtual Worlds Journal, Special Issue on Physiological Human*, vol. 20, no. 1, pp. 53–66, 2009.

[AFP00] M. A. Audette, F. P. Ferrie, and T. M. Peters, "An algorithmic overview of surface registration techniques for medical imaging," *Med Image Anal*, vol. 4, no. 3, pp. 201–217, 2000.

[AHH06] U. Ascher, E. Haber, and H. Huang, "On effective methods for implicit piecewise smooth surface recovery," *SIAM J Sci Comput*, vol. 28, no. 1, pp. 339–358, 2006.

[AK02] F. M. Alkoot and J. Kittler, "Moderating k-NN Classifiers," *Pattern Anal Appl*, vol. 5, no. 3, pp. 326–332, 2002.

[AMN+98] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu, "An optimal algorithm for approximate nearest neighbor searching fixed dimensions," *J ACM*, vol. 45, no. 6, pp. 891–923, 1998.

[And09] M. Andrecut, "Parallel GPU Implementation of Iterative PCA Algorithms," *J Comput Biol*, vol. 16, no. 11, pp. 1593–1599, 2009.

[AOC07] J.-H. Ahn, J.-H. Oh, and S. Choi, "Learning principal directions: Integrated-squared-error minimization," *Neurocomputing*, vol. 70, pp. 1372–1381, 2007.

[AS95] D. Adalsteinsson and J. Sethian, "A fast level set method for propagating interfaces," *J Comput Phys*, vol. 118, no. 2, pp. 269–277, 1995.

[AT01] A. Aubel and D. Thalmann, "Interactive modeling of the human musculature," in *Proc. CA*, H.-S. Ko, Ed. IEEE Computer Society, 2001, pp. 167–174.

[Awa06] S. P. Awate, "Adaptive, nonparametric markov models and information-theoretic methods for image restoration and segmentation," Ph.D. dissertation, University of Utah, 2006.

[BAGD07] S. Blemker, D. Asakawa, G. Gold, and S. Delp, "Image-based musculoskeletal modeling: Applications, advances, and future opportunities," *J Magn Reson Imag*, vol. 25, pp. 441–451, 2007.

[BB03] L. Ballerini and L. Bocchi, "Multiple genetic snakes for bone segmentation," in *Applications of Evolutionary Computing*, vol. LNCS 2611. Springer, 2003, pp. 346–356.

[BBLS05] R. Beichel, H. Bischof, F. Leberl, and M. Sonka, "Robust active appearance models and their application to medical image analysis," *IEEE Trans Med Imag*, vol. 24, no. 9, pp. 1151–1169, 2005.

[BCA96] E. Bardinet, L. Cohen, and N. Ayache, "Tracking and motion analysis of the left ventricle with deformable superquadrics," *Med Image Anal*, vol. 1, no. 2, pp. 129–149, 1996.

[BCE+08] D. C. Barratt, C. S. Chan, P. J. Edwards, G. P. Penney, M. Slomczykowski, T. J. Carter, and D. J. Hawkes, "Instantiation and registration of statistical shape models of the femur and pelvis using 3D ultrasound imaging," *Med Image Anal*, vol. 12, no. 3, pp. 358–374, 2008.

[BCM05] A. Buades, B. Coll, and J. Morel, "A Non-Local Algorithm for Image Denoising," in *Proc. CVPR*, vol. 2. IEEE Computer Society, 2005, pp. 60–65.

[BCP+06] J. Boisvert, F. Cheriet, X. Pennec, N. Ayache, and H. Labelle, "A novel framework for the 3D analysis of spine deformation modes," in *Research into Spinal Deformities*, ser. Studies in Health Technology and Informatics, D. Uyttendaele and P. Dangerfield, Eds., vol. 123. IOS Press, 2006, pp. 176–182.

[BD05] S. Blemker and S. Delp, "Three-dimensional representation of complex muscle architectures and geometries," *Ann Biomed Eng*, vol. 33, no. 5, pp. 661–673, 2005.

[BDBS02] K. Baek, B. Draper, J. Beveridge, and K. She, "PCA vs. ICA:A Comparison on the FERET Data Set," in *Proc. JCIS*. JCIS / Association for Intelligent Machinery, 2002, pp. 824–827.

[BE+84] J. Bezdek, R. Ehrlich *et al.*, "FCM: The fuzzy c-means clustering algorithm," *Comput Geosci*, vol. 10, no. 2-3, pp. 191–203, 1984.

[Bel57] R. Bellman, *Dynamic programming*. Princeton University Press, 1957.

[Bes74] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J Roy Stat Soc B*, vol. 36, no. 2, pp. 192–236, 1974.

[BFS+07] P. Bourgeat, J. Fripp, P. Stanwell, S. Ramadan, and S. Ourselin, "MR image segmentation of the knee bone using phase information," *Med Image Anal*, vol. 11, no. 4, pp. 325–335, 2007.

[BGK95] C. Brechbuhler, G. Gerig, and O. Kubler, "Parametrization of closed surfaces for 3-D shape description," *Comput Vis Image Understand*, vol. 61, no. 2, pp. 154–170, 1995.

[BGNV03] M. Bruijne, B. Ginneken, W. Niessen, and M. Viergever, "Active shape model segmentation using a non-linear appearance model: application to 3D AAA segmentation," Utrecht University: Information and Computing Sciences, Tech. Rep., 2003.

[BH95] A. Baumberg and D. Hogg, "An adaptive eigenshape model," in *Proc. BMVC*, vol. 1. BMVA Press, 1995, pp. 87–96.

[BHTR90] M. Bomans, K. Hohne, U. Tiede, and M. Riemer, "3-D segmentation of MR images of the head for 3-D display," *IEEE Trans Med Imag*, vol. 9, no. 2, pp. 177–183, 1990.

[BJ85] P. J. Besl and R. C. Jain, "Three-dimensional object recognition," *ACM Comput Surv*, vol. 17, no. 1, pp. 75–145, 1985.

[BKS+99] M. Backhaus, T. Kamradt, D. Sandrock, D. Loreck, J. Fritz, K. Wolf, H. Raber, B. Hamm, G. Burmester, and M. Bollow, "Arthritis of the finger joints: a comprehensive approach comparing conventional radiography, scintigraphy, ultrasound, and contrast-enhanced magnetic resonance imaging," *Arthritis Rheum*, vol. 42, no. 6, pp. 1232–1245, 1999.

[BLS+07] C. Bass, S. Lucas, R. Salzar, M. Oyen, C. Planchak, B. Shender, and G. Paskoff, "Failure properties of cervical spinal ligaments under fast strain rate deformations," *Spine*, vol. 32, no. 1, p. E7, 2007.

[Blu67] H. Blum, "A Transformation for Extracting New Descriptors of Shape," in *Models for the Perception of Speech and Visual Form*, W. Wathen-Dunn, Ed. Cambridge: MIT Press, 1967, pp. 362–380.

[BM97] W. Barrett and E. Mortensen, "Interactive live-wire boundary extraction," *Med Image Anal*, vol. 1, no. 4, pp. 331–341, 1997.

[BM05] S. Bianchi and S. Marcelis, "Musculoskeletal Sonography," in *Musculoskeletal Diseases: Diagnostic Imaging and Interventional Techniques*, C. von Schulthess, G.K. Zollikofer, Ed. Springer, 2005, pp. 156–166.

[BM06] V. Brest, J. and Žumer and M. Maučec, "Self-adaptive differential evolution algorithm in constrained real-parameter optimization," in *Proc. CEC*. IEEE Computer Society, 2006, pp. 919–926.

[BML+02] J. Bosch, S. Mitchell, B. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. Reiber, "Fully automated endocardial contour detection in time sequences of echocardiograms by three-dimensional active appearance models," *Progress In Biomedical Optics and Imaging*, vol. 3, no. 22, pp. 452–462, 2002.

[BMS02] M. Bartlett, J. Movellan, and T. Sejnowski, "Face recognition by independent component analysis," *IEEE Trans. Neural Netw.*, vol. 13, no. 6, pp. 1450–1464, 2002.

[BMVS02] G. Behiels, F. Maes, D. Vandermeulen, and P. Suetens, "Evaluation of image features and search strategies for segmentation of bone structures in radiographs using active shape models," *Med Image Anal*, vol. 6, no. 1, pp. 47–62, 2002.

[Boo89] F. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Trans Pattern Anal Mach Intell*, vol. 11, no. 6, pp. 567–585, 1989.

[Bor04] S. Borman, "The Expectation Maximization Algorithm A short tutorial," Online: http://www.seanborman.com/publications, 2004.

[Bou06] C. Bouveyron, "Modélisation et classification des données de grande dimension - application à l'analyse d'images," Ph.D. dissertation, Université Joseph Fourier, 2006.

[BR05] J. Beltran and M. Recht, "Shoulder," in *Musculoskeletal Diseases: Diagnostic Imaging and Interventional Techniques*, C. von Schulthess, G.K. Zollikofer, Ed. Springer, 2005, pp. 3–6.

[Bre65] J. Bresenham, "Algorithm for computer control of a digital plotter," *IBM Systems journal*, vol. 4, no. 1, pp. 25–30, 1965.

[BS97] D. Bhattacharya and S. Sinha, "Invariance of stereo images via the theory of complex moments," *Pattern Recogn*, vol. 30, no. 9, pp. 1373–1386, 1997.

[BS00] M. Brejl and M. Sonka, "Object localization and border detection criteria design inedge-based image segmentation: automated learning from examples," *IEEE Trans Med Imag*, vol. 19, no. 10, pp. 973–985, 2000.

[BSB+93] I. Barrodale, D. Skea, M. Berkley, R. Kuwahara, and R. Poeckert, "Warping digital images using thin plate splines," *Pattern Recogn*, vol. 26, no. 2, pp. 375–376, 1993.

[BSP+07] M. Barth, M. Smith, I. Pedrosa, R. Lenkinski, and N. Rofsky, "Body MR Imaging at 3.0 T: Understanding the Opportunities and Challenges1," *Radiographics*, vol. 27, no. 5, pp. 1445–1462, 2007.

[BT00a] A. Brett and C. Taylor, "A method of automated landmark generation for automated 3D PDM construction," *Image Vis Comput*, vol. 18, no. 9, pp. 739–748, 2000.

[BT00b] A. Brett and C. Taylor, "Construction of 3d shape models of femoral articular cartilage using harmonic maps," in *Proc. MICCAI*, vol. LNCS 1935, Springer. Springer, 2000, pp. 41–92.

[BV04] Y. Bengio and P. Vincent, "Manifold Parzen Windows," CIRANO, CIRANO Working Papers 2004s-30, 2004.

[BW98] D. Baraff and A. Witkin, "Large steps in cloth simulation," in *Computer Graphics Proc., Annu. Conf. Series*. SIGGRAPH, 1998, pp. 43–54.

[CACC05] P. Coyte, C. Asche, R. Croxford, and B. Chan, "The economic cost of musculoskeletal disorders in Canada," *Arthritis Rheum*, vol. 11, no. 5, pp. 315–325, 2005.

[Cae10] C. Cael, *Functional Anatomy: Musculoskeletal Anatomy, Kinesiology, and Palpation for Manual Therapists*. Wolters Kluwer - Lippincott Williams & Wilkins, 2010.

[Can86] J. F. Canny, "A computational approach to edge detection," *IEEE Trans Pattern Anal Mach Intell*, vol. 8, pp. 679–698, 1986.

[Cat66] R. Cattell, "The Scree Test For The Number Of Factors," *Multivariate Behav Res*, vol. 1, no. 2, pp. 245–276, 1966.

[CAVMT09] C. Charbonnier, L. Assassi, P. Volino, and N. Magnenat-Thalmann, "Motion study of the hip joint in extreme postures," *Vis Comput*, vol. 25, no. 9, pp. 873–882, 2009.

[CC07] P. Campadelli and E. Casiraghi, "Liver segmentation from CT scans: A survey," in *Proc. WILF*, vol. LNCS 4578. Springer, 2007, pp. 520–528.

[CCB+09] T. Cresson, R. Chav, D. Branchaud, L. Humbert, B. Godbout, B. Aubert, W. Skalli, and J. De Guise, "Coupling 2D/3D registration method and statistical model to perform 3D reconstruction from partial x-rays images data," in *Proc. EMBC*. IEEE Computer Society, 2009, pp. 1008–1011.

[CCMS97] A. Ciampalini, P. Cignoni, C. Montani, and R. Scopigno, "Multiresolution decimation based on global error," *Vis Comput*, vol. 13, no. 5, pp. 228–246, 1997.

[CCP09] G. Costache, P. Corcoran, and P. Puslecki, "Combining PCA-based datasets without retraining of the basis vector set," *Pattern Recogn Lett*, vol. 30, pp. 1441–1447, 2009.

[CD09] F. Chung and H. Delingette, "Multimodal Prior Appearance Models based on Regional Clustering of Intensity Profiles," in *Proc. MICCAI*, vol. LNCS 5762(2). Springer, 2009, pp. 1051–1058.

[CDNA07] M. J. Costa, H. Delingette, S. Novellas, and N. Ayache, "Automatic Segmentation of Bladder and Prostate Using Coupled 3D Deformable Models," in *Proc. MICCAI*, vol. LNCS 4791. Springer, 2007, pp. 252–260.

[CDS02] M. Collins, S. Dasgupta, and R. Schapire, "A generalization of principal component analysis to the exponential family," *Adv Neural Inform Process Syst*, vol. 1, pp. 617–624, 2002.

[CET98a] T. Cootes, G. Edwards, and C. Taylor, "A comparative evaluation of active appearance model algorithms," in *Proc. BMVC*. BMVA Press, 1998, pp. 680–689.

[CET98b] T. Cootes, G. Edwards, and C. Taylor, "Active appearance models," in *Proc. ECCV*, vol. LNCS 1407. Springer, 1998, pp. 484–498.

[CFS+07] J. E. Cates, P. T. Fletcher, M. A. Styner, M. E. Shenton, and R. T. Whitaker, "Shape Modeling and Analysis with Entropy-Based Particle Systems," in *Proc. IPMI*, vol. LNCS 4584. Springer, 2007, pp. 333–345.

[CG07] R. Cangelosi and A. Goriely, "Component retention in principal component analysis with application to cdna microarray data," *Biol Direct*, vol. 2, no. 2, 2007, online.

[CH00] C. Croux and G. Haesbroeck, "Principal component analysis based on robust estimators of the covariance or correlation matrix: Influence functions and efficiencies," *Biometrika*, vol. 87, no. 3, pp. 603–618, 2000.

[CHA96] "Charm website, esprit project 9036," http://ligwww.epfl.ch/~maurel/CHARM/, 1993-1996.

[Che02] H. Chen, "Principal Component Analysis with Missing Data and Outliers," Electrical and Computer Engineering Department, Rutgers University, Piscataway, NJ, 08854, Tech. Rep., 2002.

[Cho04] S. Choi, "Sequential EM learning for subspace analysis," *Pattern Recogn Lett*, vol. 25, no. 14, pp. 1559–1567, 2004.

[Cho07] S. Choi, "Wake-Sleep PCA," in *Proc. IJCNN*. IEEE Computer Society, 2007, pp. 2432–2435.

[Chr94] G. E. Christensen, "Deformable shape models for anatomy," Ph.D. dissertation, Washington University Sever Institute Of technology, 1994.

[Chr99] G. Christensen, "Consistent linear-elastic transformations for image matching," in *Proc. IPMI*, vol. LNCS 1613. Springer, 1999, pp. 224–237.

[CHTH93] T. F. Cootes, A. Hill, C. J. Taylor, and J. Haslam, "The Use of Active Shape Models for Locating Structures in Medical Images," in *Proc. IPMI*, vol. LNCS 687. Springer-Verlag, 1993, pp. 33–47.

[CJ96] J. Cadima and I. Jolliffe, "Size-and shape-related principal component analysis," *Biometrics*, vol. 52, no. 2, pp. 710–716, 1996.

[CJ01] G. Christensen and H. Johnson, "Consistent image registration," *IEEE Trans Med Imag*, vol. 20, no. 7, pp. 568–582, 2001.

[CJSW01] H. Cheng, X. Jiang, Y. Sun, and J. Wang, "Color image segmentation: advances and prospects," *Pattern Recogn*, vol. 34, no. 12, pp. 2259–2281, 2001.

[CKCF07] L. Constantinescu, J. Kim, C. Chan, and D. Feng, "Automatic Mobile Device Synchronization and Remote Control System for High-Performance Medical Applications," in *Proc. EMBS*. IEEE Computer Society, 2007, pp. 2799–2802.

[CKM+10] I. Ciuciu, H. Kang, R. Meersman, J. Schmid, N. Magnenat-Thalmann, Iglesias Guitián, and E. Gobbetti, "Collaborative semantic content management: an ongoing case study for imaging applications," in *Proc. ECKM*. Academic Publishing, 2010.

[CKN02] T. Cootes and P. Kittipanya-Ngam, "Comparing variations on the active appearance model algorithm," in *Proc. BMVC*, vol. 2. BMVA Press, 2002, pp. 837–846.

[CKS97] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *Int J Comput Vis*, vol. 22, no. 1, pp. 61–79, 1997.

[CKS03] D. Cremers, T. Kohlberger, and C. Schnörr, "Shape statistics in kernel space for variational image segmentation," *Pattern Recogn*, vol. 36, no. 9, pp. 1929–1943, 2003.

[CKSS97] V. Caselles, R. Kimmel, G. Sapiro, and C. Sbert, "Minimal surfaces based object segmentation," *IEEE Trans Pattern Anal Mach Intell*, vol. 19, no. 4, pp. 394–398, 1997.

[CL96] R. Casey and E. Lecolinet, "A survey of methods and strategies in character segmentation," *IEEE Trans Pattern Anal Mach Intell*, vol. 18, no. 7, pp. 690–706, 1996.

[CLMT08] C. Charbonnier, E. Lyard, and N. Magnenat-Thalmann, "Analysis of extreme hip motion in professional ballet dancers," in *Proc. Int. Symp. 3D Analysis of Human Movement*, 2008.

[CLW04] J. E. Cates, A. E. Lefohn, and R. T. Whitaker, "Gist: an interactive, gpu-based level set segmentation tool for 3d medical images," *Med Image Anal*, vol. 8, no. 3, pp. 217–231, 2004.

[CMT+04] T. F. Cootes, S. Marsland, C. J. Twining, K. Smith, and C. Taylor, "Groupwise Diffeomorphic Non-rigid Registration for Automatic Model Building," in *Proc. ECCV*, vol. LNCS 3024. Springer, 2004, pp. 316–327.

[CMTB+10] C. Charbonnier, N. Magnenat-Thalmann, C. Becker, P. Hoffmeyer, and J. Menetrey, "An integrated platform for hip joint osteoarthritis analysis: Design, implementation and results," *Int J Comput Assist Radiol Surg*, vol. 5, no. 4, pp. 351–358, 2010.

[Coh91] L. Cohen, "On active contour models and balloons," *CVGIP: Image Understanding*, vol. 53, no. 2, pp. 211–218, 1991.

[COM09] "Come swiss national project," http://co-me.ch/, 2001-2009.

[Coo01] T. Cootes, "Statistical Models of Appearance for Computer Vision. Working report," 2001.

[Cov68] T. Cover, "Estimation by the nearest neighbor rule," *IEEE Trans Inform Theor*, vol. 14, no. 1, pp. 50–55, 1968.

[CP10] B. Combès and S. Prima, "An efficient em-icp algorithm for symmetric consistent non-linear registration of point sets," in *Proc. MICCAI*, vol. LNCS 6362. Springer, 2010, pp. 594–601.

[CRD07] D. Cremers, M. Rousson, and R. Deriche, "A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape," *Int J Comput Vis*, vol. 72, no. 2, pp. 195–215, 2007.

[Cre06] D. Cremers, "Dynamical statistical shape priors for level set-based tracking," *IEEE Trans Pattern Anal Mach Intell*, vol. 28, no. 8, pp. 1262–1273, 2006.

[CRM94] G. Christensen, R. Rabbitt, and M. Miller, "3D brain mapping using a deformable neuroanatomy," *Phys Med Biol*, vol. 39, pp. 609–618, 1994.

[CS05] C. Chung and L. Steinbach, "Magnetic Resonance Imaging of the Elbow," in *Musculoskeletal Diseases: Diagnostic Imaging and Interventional Techniques*, C. von Schulthess, G.K. Zollikofer, Ed. Springer, 2005, pp. 7–12.

[CSKC+09] C. Charbonnier, J. Schmid, F. Kolo-Christophe, N. Magnenat-Thalmann, C. Becker, and P. Hoffmeyer, "Virtual Hip Joint: from Computer Graphics to Computer-Assisted Diagnosis," in *Eurographics 2009 - First Medical Prize*. Eurographics Association, 2009, pp. 1–4.

[CT94] T. Cootes and C. Taylor, "Using Grey-Level Models to Improve Active Shape Model Search," in *Proc. ICPR*, 1994, pp. A:63–67.

[CT95] T. Cootes and C. Taylor, "Combining point distribution models with shape models based on finite element analysis," *Image Vis Comput*, vol. 13, no. 5, pp. 403–409, 1995.

[CT01] T. F. Cootes and C. J. Taylor, "Constrained Active Appearance Models," in *Proc. ICCV*, vol. 1. Los Alamitos, CA, USA: IEEE Computer Society, 2001, p. 748.

[CT04] T. Cootes and C. Taylor, "Anatomical statistical models and their role in feature extraction," *The British Journal Of Radiology*, vol. 77, pp. 133–139, 2004.

[CTBT08] T. Cootes, C. Twining, K. Babalola, and C. Taylor, "Diffeomorphic statistical shape models," *Image Vis Comput*, vol. 26, no. 3, pp. 326–332, 2008.

[CTT+02] Y. Chen, H. D. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K. S. Gopinath, R. W. Briggs, and E. A. Geiser, "Using Prior Shapes in Geometric Active Contours in a Variational Framework," *Int J Comput Vis*, vol. 50, no. 3, pp. 315–328, 2002.

[Cua03] M. B. Cuadra, "Atlas-based segmentation and classification of magnetic resonance brain images," Ph.D. dissertation, Faculté Sciences et Techniques de Lingénieur, Institut de traitement des signaux, Ecole Polytechnique Fédérale de Lausanne, 2003.

[CVC+95] L. Clarke, R. Velthuizen, M. Camacho, J. Heine, and M. Vaidyanathan, "MRI segmentation: methods and

applications," *Magn Reson Imag*, vol. 13, no. 3, pp. 343–368, 1995.

[CYDL06] M. Cao, C. Ye, O. Doessel, and C. Liu, "Spherical parameter detection based on hierarchical Hough transform," *Pattern Recogn Lett*, vol. 27, no. 9, pp. 980–986, 2006.

[CYP+08] P. Coupé, P. Yger, S. Prima, P. Hellier, C. Kervrann, and C. Barillot, "An Optimized Blockwise Non Local Means Denoising Filter for 3D Magnetic Resonance Images," *IEEE Trans Med Imag*, vol. 27, no. 4, pp. 425–441, 2008.

[DAA+07] S. Delp, F. Anderson, A. Arnold, P. Loan, A. Habib, C. John, E. Guendelman, and D. Thelen, "OpenSim: open-source software to create and analyze dynamic simulations of movement," *IEEE Trans Biomed Eng*, vol. 54, no. 11, pp. 1940–1950, 2007.

[Dav97] C. Davatzikos, "Spatial transformation and registration of brain images using elastically deformable models," *Comput Vis Image Understand*, vol. 66, no. 2, pp. 207–222, 1997.

[Dav02] R. H. Davies, "Learning Shape: Optimal Models for Analysing Natural Variability," Ph.D. dissertation, Division of Imaging Science and Biomedical Engineering, University of Manchester, 2002.

[DAWW07] R. Dalvi, R. Abugharbieh, D. Wilson, and D. Wilson, "Multi-Contrast MR for Enhanced Bone Imaging and Segmentation," in *Proc. EMBC*. IEEE Computer Society, 2007, pp. 5620–5623.

[dBGVN03] M. de Bruijne, B. Ginneken, M. Viergever, and W. Niessen, "Adapting active shape models for 3D segmentation of tubular structures in medical images," in *Proc. IPMI*, vol. LNCS 2732. Springer, 2003, pp. 136–147.

[dBN04] M. de Bruijne and M. Nielsen, "Image Segmentation by Shape Particle Filtering," in *Proc. ICPR*, vol. 3. Washington, DC, USA: IEEE Computer Society, 2004, pp. 722–725.

[Del94a] H. Delingette, "Modélisation, Déformation et Reconnaissance d'objets tridimensionnels à l'aide de maillages simplexes," Ph.D. dissertation, Ecole Centrale de Paris, 1994.

[Del94b] H. Delingette, "Simplex meshes: a general representation for 3D shape reconstruction," INRIA, Tech. Rep. 2214, 1994.

[Del99] H. Delingette, "General Object Reconstruction Based on Simplex Meshes," *Int J Comput Vis*, vol. 32, no. 2, pp. 111–146, 1999.

[DFP08] E. Dam, P. Fletcher, and S. Pizer, "Automatic shape model building based on principal geodesic analysis bootstrapping," *Med Image Anal*, vol. 12, no. 2, pp. 136–151, 2008.

[DG96] M. Desbrun and M. Gascuel, "Smoothed particles: A new paradigm for animating highly deformable bodies," in *Proc. Computer Animation and Simulation*, vol. 96. Springer Verlag, 1996, pp. 61–76.

[DGS01] L. Ding, A. Goshtasby, and M. Satter, "Volume image registration by template matching," *Image Vis Comput*, vol. 19, no. 12, pp. 821–832, 2001.

[DGZ07] X. Dong, A. M. González Ballester, and G. Zheng, "Automatic Extraction of Femur Contours from Calibrated X-Ray Images using Statistical Information," *J Multimed*, vol. 2, no. 5, pp. 46–54, 2007.

[Dic45] L. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[DJM+09] P. Dalal, L. Ju, M. McLaughlin, X. Zhou, H. Fujita, and S. W. 3D, "Open-Surface Shape Correspondence for Statistical Shape Modeling: Identifying Topologically Consistent Landmarks," in *Proc. ICCV*. IEEE Computer Society, 2009.

[DLH07] F. Ding, W. Leow, and T. Howe, "Automatic Segmentation of Femur Bones in Anterior-Posterior Pelvis X-Ray Images," in *Proc. CAIP*, vol. LNCS 4673. Springer, 2007, pp. 205–212.

[DLR77] A. Dempster, N. Laird, and D. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *J Roy Stat Soc B*, vol. 39, no. 1, pp. 1–38, 1977.

[DlTB01] F. De la Torre and M. Black, "Robust principal component analysis for computer vision," in *Proc. ICCV*, vol. 1. IEEE Computer Society, 2001, pp. 362–369.

[DLvB+10] J. Dworzak, H. Lamecker, J. von Berg, T. Klinder, C. Lorenz, D. Kainmüller, H. Seim, H. Hege, and S. Zachow, "3D reconstruction of the human rib cage from 2D projection images using a statistical shape model," *Int J Comput Assist Radiol Surg*, vol. 5, no. 2, pp. 111–124, 2010.

[DMSB99] M. Desbrun, M. Meyer, P. Schröder, and A. Barr, "Implicit fairing of irregular meshes using diffusion and curvature flow," in *Proc. SIGGRAPH*. ACM Press, 1999, pp. 317–324.

[DMW+07] P. Dalal, B. Munsell, S. Wang, J. Tang, K. Oliver, H. Ninomiya, X. Zhou, and H. Fujita, "A Fast 3D Correspondence Method for Statistical Shape Modeling," in *Proc. CVPR*. IEEE Computer Society, 2007, pp. 1–8.

[DP95] C. Davatzikos and J. Prince, "An active contour model for mapping the cortex," *IEEE Trans Med Imag*, vol. 14, no. 1, pp. 65–80, 1995.

[DR07] M. Droske and M. Rumpf, "Multiscale joint segmentation and registration of image morphology," *IEEE Trans Pattern Anal Mach Intell*, vol. 29, no. 12, pp. 2181–2194, 2007.

[DRC+06] M. Damsgaard, J. Rasmussen, S. Christensen, E. Surma, and M. de Zee, "Analysis of musculoskeletal systems in the AnyBody Modeling System," *Simulat Model Pract Theor*, vol. 14, no. 8, pp. 1100–1111, 2006.

[dRKD00] D. de Ridder, J. Kittler, and R. Duin, "Probabilistic pca and ica subspace mixture models for image segmentation," in *Proc. BMVC*. BMVA Press, 2000, pp. 112–121.

[DRL+06] R. Donner, M. Reiter, G. Langs, P. Peloschek, and H. Bischof, "Fast active appearance model search using canonical correlation analysis," *IEEE Trans Pattern Anal Mach Intell*, vol. 28, no. 10, pp. 1690–1694, 2006.

[DRT06] S. Dambreville, Y. Rathi, and A. Tannenbaum, "Shape-based approach to robust image segmentation using kernel pca," in *Proc. CVPR*. IEEE Computer Society, 2006, pp. 977–984.

[DS97] N. Duta and M. Sonka, "Segmentation and interpretation of MR brain images using an improved knowledge-based active shape model," in *Proc. IPMI*, vol. LNCS 1230. Springer, 1997, pp. 375–380.

[DTC+02] R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor, "3D Statistical Shape Models Using Direct Optimisation of Description Length," in *Proc. ECCV*, vol. LNCS 2352. Springer, 2002, pp. 3–20.

[DXD03] C. Davatzikos, T. Xiaodong, and S. Dinggang, "Hierarchical active shape models, using the wavelet transform," *IEEE Trans Med Imag*, vol. 22, pp. 414–423, 2003.

[Ehr90] H. Ehricke, "Problems and approaches for tissue segmentation in 3-D MR imaging," in *Proc. SPIE*, M. Loew, Ed., vol. 1233. SPIE, 1990, pp. 128–137.

[EHW+00] J. Ehrhardt, H. Handels, T. Wegner, B. Strathmann, W. Plotz, and S. Poppl, "An Anatomical Atlas to Support the Virtual Planning of Hip Operations," *Stud Health Tech Informat*, vol. 77, pp. 1226–1230, 2000.

[EK82] H. T. Eastment and W. J. Krzanowski, "Cross-validatory choice of the number of components from a principal component analysis," *Technometrics*, vol. 24, no. 1, pp. 73–77, 1982.

[EKW09] M. Erdt, M. Kirschner, and S. Wesarg, "Simultaneous Segmentation and Correspondence Establishment for Statistical Shape Models," in *Modelling the Physiological Human: Proc. 3DPH*, vol. LNCS 5903. Springer, 2009, pp. 25–35.

[EPS+08] O. Ecabert, J. Peters, H. Schramm, C. Lorenz, J. Von Berg, M. Walker, M. Vembar, M. Olszewski, K. Subramanyan, G. Lavi, and J. Weese, "Automatic model-based segmentation of the heart in CT images," *IEEE Trans Med Imag*, vol. 27, no. 9, pp. 1189–1201, 2008.

[ER00] R. Everson and S. Roberts, "Inferring the Eigenvalues of Covariance Matrices From Limited, Noisy Data," *IEEE Trans Acoust Speech Signal Process*, vol. 48, no. 7, pp. 2083–2091, 2000.

[EW04] D. Elentuck and P. E. William, "Direct Magnetic Resonance Arthrography," *Eur Radiol*, vol. 14, no. 11, pp. 1956–1967, 2004.

[FA96] J. Feldmar and N. Ayache, "Rigid, affine and locally affine registration of free-form surfaces," *International journal of computer vision*, vol. 18, no. 2, pp. 99–119, 1996.

[FB81] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[FBCO07] J. Fripp, P. Bourgeat, S. Crozier, and S. Ourselin, "Segmentation of the Bones in MRIs of the Knee Using Phase, Magnitude, and Shape Information," *Academic Radiology*, vol. 14, no. 10, pp. 1201–1208, 2007.

[FBGI00] S. J. Ferguson, J. T. Bryant, R. Ganz, and K. Ito, "The influence of the acetabular labrum on hip joint cartilage consolidation: a poroelastic finite element model," *J Biomech*, vol. 33, no. 8, pp. 953–960, 2000.

[FCWO07] J. Fripp, S. Crozier, S. Warfield, and S. Ourselin, "Automatic segmentation of the bone and extraction of the bone-cartilage interface form magnetic resonance images of the knee," *Phys Med Biol*, vol. 52, pp. 1617–1631, 2007.

[Fel00] F. Feldman, "Musculoskeletal Radiology: Then and Now," *Radiology*, vol. 216, no. 2, pp. 309–316, 2000.

[Fer95] L. Ferré, "Selection of components in preincipal component analysis: A comparison of methods," *Comput Stat Data Anal*, vol. 19, no. 6, pp. 669–682, 1995.

[FGXL10] Y. Fu, W. Gao, Y. Xiao, and J. Liu, "A framework for automatic construction of 3D PDM from segmented volumetric neuroradiological data sets," *Comput Meth Programs Biomed*, vol. 97, no. 3, pp. 199–210, 2010.

[FLJ99] M. Fleute, S. Lavallée, and R. Julliard, "Incorporating a statistically based shape model into a system for computer-assisted anterior cruciate ligament surgery," *Med Image Anal*, vol. 3, no. 3, pp. 209–222, 1999.

[FLJ03] P. Fletcher, C. Lu, and S. Joshi, "Statistics of shape via principal geodesic analysis on Lie groups," in *Proc. CVPR*, vol. 1. IEEE Computer Society, 2003, pp. 95–101.

[Flu92] J. Flusser, "An adaptive method for image registration," *Pattern Recogn*, vol. 25, no. 1, pp. 45–54, 1992.

[FM96] N. Foster and D. Metaxas, "Realistic animation of liquids," *Graph Model Image Process*, vol. 58, no. 5, pp. 471–483, 1996.

[FM08] B. Fischer and J. Modersitzki, "Ill-posed medicine—an introduction to image registration," *Inverse Problems*, vol. 24, no. 3, pp. 4008–4028, 2008.

[FMT+04] J. Fernandez, P. Mithraratne, S. Thrupp, M. Tawhai, and P. Hunter, "Anatomically based geometric modelling of the musculo-skeletal system and other organs," *Biomech Model Mechanobiol*, vol. 2, no. 3, pp. 139–155, 2004.

[FOP+05] J. Folkesson, O. Olsen, P. Pettersen, E. Dam, and C. Christiansen, "Combining binary classifiers for automatic cartilage segmentation in knee mri," in *Proc. CVBIA*, vol. LNCS 3765. Springer, 2005, pp. 230–239.

[FPAB04] Y. Fridman, S. Pizer, S. Aylward, and E. Bullitt, "Extracting branching tubular object geometry via cores," *Med Image Anal*, vol. 8, no. 3, pp. 169–176, 2004.

[FRSN02] A. Frangi, D. Rueckert, J. Schnabel, and W. Niessen, "Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling," *IEEE Trans Med Imag*, vol. 21, no. 9, pp. 1151–1166, 2002.

[FS06] K. Fritscher and R. Schubert, "3D image segmentation by using statistical deformation models and level sets," *Int J Comput Assist Radiol Surg*, vol. 1, no. 3, pp. 123–135, 2006.

[GEG+04] L. Galois, S. Etienne, L. Grossin, A. Watrin-Pinzano, C. Cournil-Henrionnet, D. Loeuille, P. Netter, D. Mainard, and P. Gillet, "Dose-response relationship for exercise on severity of experimental osteoarthritis in rats: a pilot study," *Osteoarthritis Cartilage*, vol. 12, no. 10, pp. 779–786, 2004.

[GH97] M. Garland and P. Heckbert, "Surface simplification using quadric error metrics," in *Proc. SIGGRAPH*. ACM Press, 1997, p. 216.

[Gil07] B. Gilles, "Anatomical and Kinematical Modelling of the Musculoskeletal System from MRI," Ph.D. dissertation, Université de Genève, 2007.

[GKB+96] P. Gingins, P. Kalra, P. Beylot, N. Magnenat-Thalmann, and J. Fasel, "Using VHD to build a Comprehensive Human Model," in *Proc. Visible Human Project Conf (CD-ROM)*, 1996.

[GL05] Y. Gurwicz and B. Lerner, "Bayesian network classification using spline-approximated kernel density estimation," *Pattern Recogn Lett*, vol. 26, no. 11, pp. 1761–1771, 2005.

[Gla93] C. Glasbey, "An analysis of histogram-based thresholding algorithms," *Graph Model Image Process*, vol. 55, no. 6, pp. 532–537, 1993.

[GM03] G. Gray and A. Moore, "Nonparametric density estimation: toward computational tractability," in *Proc. SDM*. SIAM, 2003.

[GMB06] R. Gross, I. Matthews, and S. Baker, "Active appearance models with occlusion," *Image Vis Comput*, vol. 24, no. 6, pp. 593–604, 2006.

[GMMT06] B. Gilles, L. Moccozet, and N. Magnenat-Thalmann, "Anatomical modelling of the musculoskeletal system from MRI," in *Proc. MICCAI*, vol. LNCS 4190. Springer, 2006, pp. 289–296.

[GMT10] B. Gilles and N. Magnenat-Thalmann, "Musculoskeletal MRI segmentation using multi-resolution simplex meshes with medial representations," *Med Image Anal*, vol. 14, no. 3, pp. 291–302, 2010.

[Gow75] J. Gower, "Generalized Procrustes analysis," *Psychometrika*, vol. 40, pp. 33–51, 1975.

[GP08] B. Gilles and D. Pai, "Fast musculoskeletal registration based on shape matching," in *Proc. MICCAI*, vol.

LNCS 5242. Springer, 2008, pp. 822–829.

[GPB+03] R. Ganz, J. Parvizi, M. Beck, M. Leunig, H. Nötzli, and K. Siebenrock, "Femoroacetabular impingement: a cause for osteoarthritis of the hip," *Clin Orthop Relat Res*, vol. 417, pp. 112–120, 2003.

[Gra18] H. Gray, *20th U.S. edition of Gray's Anatomy of the Human Body*, W. Lewis, Ed., 1918.

[GRB93] J. Gee, M. Reivich, and R. Bajcsy, "Elastically deforming 3D atlas to match anatomical brain images," *J Comput Assist Tomo*, vol. 17, no. 2, pp. 225–236, 1993.

[GRL+98] S. Gold, A. Rangarajan, C. Lu, S. Pappu, and E. Mjolsness, "New algorithms for 2D and 3D point matching pose estimation and correspondence," *Pattern Recogn*, vol. 31, no. 8, pp. 1019–1031, 1998.

[Grü04] P. Grünwald, "A Tutorial Introduction to the Minimum Description Length Principle," online, arXiv:math/0406077v1 [math.ST], 2004.

[GS91] L. Greengard and J. Strain, "The fast Gauss transform," *SIAM J Sci Stat Comput*, vol. 12, no. 1, pp. 79–94, 1991.

[GSJS86] T. Gasser, L. Sroka, and C. Jennen-Steinmetz, "Residual variance and residual pattern in nonlinear regression," *Biometrika*, vol. 73, no. 3, pp. 625–633, 1986.

[GSMY+07] D. Göddeke, R. Strzodka, J. Mohd-Yusof, P. McCormick, S. Buijssen, M. Grajewski, and S. Turek, "Exploring weak scalability for FEM calculations on a GPU-enhanced cluster," *Parallel Computing*, vol. 33, no. 10-11, pp. 685–699, 2007.

[GVL96] G. Golub and C. Van Loan, *Matrix computations*, 3rd ed. Johns Hopkins Univ Pr, 1996.

[GVSS00] I. Guskov, K. Vidimče, W. Sweldens, and P. Schröder, "Normal meshes," in *Proc. SIGGRAPH*. ACM Press, 2000, pp. 95–102.

[GW03] E. M. Gertz and S. J. Wright, "Object-Oriented Software for Quadratic Programming," *ACM Trans Math Software*, vol. 29, pp. 58–81, 2003.

[GW05] J. Georgii and R. Westermann, "Mass-Spring Systems on the GPU," *Simulat Model Pract Theor*, vol. 13, pp. 693–702, 2005.

[GWC+03] X. Gu, Y. Wang, T. Chan, P. Thompson, and S. Yau, "Genus zero surface conformal mapping and its application to brain surface mapping," in *Proc. IPMI*, vol. LNCS 2732. Springer, 2003, pp. 172–184.

[GXDX10] P. Gamage, S. Xie, P. Delmas, and W. Xu, "Segmentation of radiographic images under topological constraints: application to the femur," *Int J Comput Assist Radiol Surg*, vol. 5, pp. 425–435, 2010. [Online]. Available: http://dx.doi.org/10.1007/s11548-009-0399-6

[GZ79] K. Gabriel and S. Zamir, "Lower rank approximation of matrices by least squares with any choice of weights," *Technometrics*, vol. 21, no. 4, pp. 489–498, 1979.

[HAGM04] G. Hamarneh, R. Abu-Gharbieh, and T. McInerney, "Medial Profiles for Modeling Deformation and Statistical Analysis of Shape and their use in Medical Image Segmentation," *Int J Shape Model*, vol. 10, no. 2, pp. 187–209, 2004.

[HE02] M. Hauth and O. Etzmuss, "A High Performance Solver for the Animation of Deformable Objects using Advanced Numerical Methods," *Comput Graph Forum*, vol. 20, no. 3, pp. 319–328, 2002.

[Hei08] T. Heimann, "Statistical Shape Models For 3d Medical Image Segmentation," Ph.D. dissertation, University Of Heidelberg, 2008.

[HES03] M. Hauth, O. Etzmuß, and W. Straßer, "Analysis of numerical methods for the simulation of deformable models," *The Visual Computer*, vol. 19, no. 7, pp. 581–600, 2003.

[HG04] S. Ho and G. Gerig, "Profile scale-spaces for multiscale image match," in *Proc. MICCAI*, vol. LNCS 3216(1). Springer, 2004, pp. 176–183.

[HH97] T. Heap and D. Hogg, "Improving Specificity in PDMs using a Hierarchical Approach," in *Proc. BMVC*. BMVA Press, 1997.

[HHD+99] M. Holden, D. Hill, E. Denton, J. Jarosz, T. Cox, and D. Hawkes, "Voxel similarity measures for 3D serial MR brain image registration," in *Proc. IPMI*, vol. LNCS 1613. Springer, 1999, pp. 472–477.

[HHL07] M. S. Hansen, M. F. Hansen, and R. Larsen, "Diffeomorphic Statistical Deformation Models," in *Proc. ICCV*. IEEE Computer Society, 2007, pp. 1–8.

[HJ08] A. Hurvitz and L. Joskowicz, "Registration of a CT-like atlas to fluoroscopic X-ray images using intensity correspondences," *Int J Comput Assist Radiol Surg*, vol. 3, no. 6, pp. 493–504, 2008.

[HK06] Z. He and F. Kuester, "GPU-Based Active Contour Segmentation Using Gradient Vector Flow," *Advances in Visual Computing*, vol. 4291, pp. 191–201, 2006.

[HKR93] D. Huttenlocher, G. Klanderman, and W. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Trans Pattern Anal Mach Intell*, vol. 15, no. 9, pp. 850–863, 1993.

[HLM07] Y. Huang, Q. Liu, and D. Metaxas, "A component based deformable model for generalized face alignment," in *Proc. ICCV*. IEEE Computer Society, 2007, pp. 1–8.

[HLSB04] M. Hadwiger, C. Langer, H. Scharsach, and K. Bühler, "State of the art report 2004 on gpu-based segmentation," VRVis, Tech. Rep., 2004.

[HM02] C. Hoad and A. Martel, "Segmentation of MR images for computer-assisted surgery of the lumbar spine," *Phys Med Biol*, vol. 47, no. 19, pp. 3503–3517, 2002.

[HM06] E. Haber and J. Modersitzki, "Intensity gradient based registration and fusion of multi-modal images," in *Proc. MICCAI*, vol. LNCS 4191. Springer, 2006, pp. 726–733.

[HM09] T. Heimann and H.-P. Meinzer, "Statistical shape models for 3D medical image Segmentation: A review," *Med Image Anal*, vol. 13, no. 4, pp. 543–563, 2009.

[HMK+02] P. Heinze, D. Meister, R. Kober, J. Raczkowsky, and H. Worn, "Atlas-based segmentation of pathological knee joints," *Stud Health Tech Informat*, vol. 85, pp. 198–203, 2002.

[HMM02] P. Hall, D. Marshall, and R. Martin, "Adding and subtracting eigenspaces with eigenvalue decomposition and singular value decomposition," *Image Vis Comput*, vol. 20, no. 13-14, pp. 1009–1016, 2002.

[HMMW07] T. Heimann, S. Munzing, H. Meinzer, and I. Wolf, "A Shape-Guided Deformable Model with Evolutionary Algorithm Initialization for 3D Soft Tissue Segmentation," in *Proc. IPMI*, N. Karssemeijer and B. Lelieveldt, Eds., vol. LNCS 4584. Springer, 2007, pp. 1–12.

[HNS+10] S. Han, N. Nijdam, J. Schmid, J. Kim, and N. Magnenat-Thalmann, "Collaborative telemedicine for interactive multiuser segmentation of volumetric medical images," *The Visual Computer*, vol. 26, no. 6-8, pp. 639–648, 2010.

[HO97] A. Hyvarinen and E. Oja, "A Fast Fixed-Point Algorithm for Independent Component Analysis," *Neural Comput*, vol. 9, no. 7, pp. 1483–1492, 1997.

[Hol75] J. Holland, "Adaptation in natural and artificial systems," *Ann Arbor, MI: University of Michigan Press*, vol. 31, 1975.

[Hop96] H. Hoppe, "Progressive meshes," in *Proc. SIGGRAPH*. ACM Press, 1996, pp. 99–108.

[Hor87] B. Horn, .K.P., "Closed-form solution of absolute orientation using unit quaternions," *J Opt Soc Am A*, vol. 4, no. 4, pp. 629–642, 1987.

[HP76] S. Horowitz and T. Pavlidis, "Picture segmentation by a tree traversal algorithm," *J ACM*, vol. 23, no. 2, pp. 368–388, 1976.

[HPE+07] H. Hufnagel, X. Pennec, J. Ehrhardt, H. Handels, and N. Ayache, "Point-Based Statistical Shape Models with Probabilistic Correspondences and Affine EM-ICP," in *Proc. BVM*. Springer, 2007, pp. 434–438.

[HQS06] V. Huang, A. Qin, and P. Suganthan, "Self-adaptive differential evolution algorithm for constrained real-parameter optimization," in *Proc. CEC*. IEEE Press, 2006, pp. 17–24.

[HSV+05] A. Heemskerk, G. Strijkers, A. Vilanova, M. Drost, and K. Nicolay, "Determination of mouse skeletal muscle architecture using three-dimensional diffusion tensor imaging," *Magn Reson Med*, vol. 53, no. 6, pp. 1333–1340, 2005.

[HT92] A. Hill and C. Taylor, "Model-based image interpretation using genetic algorithms," *Image Vis Comput*, vol. 10, no. 5, pp. 295–300, 1992.

[HvGS+09] T. Heimann, B. van Ginneken, M. Styner, Y. Arzhaeva, V. Aurich, C. Bauer, A. Beck, C. Becker, R. Beichel, G. Bekes *et al.*, "Comparison and Evaluation of Methods for Liver Segmentation from CT datasets," *IEEE Trans Med Imag*, vol. 28, pp. 1251–1265, 2009.

[HW79] J. Hartigan and M. Wong, "A K-means clustering algorithm," *J Roy Stat Soc C Appl Stat*, vol. 28, pp. 100–108, 1979.

[HWM07] T. Heimann, I. Wolf, and H. Meinzer, "Automatic generation of 3D statistical shape models with optimal landmark distributions," *Meth Inform Med*, vol. 46, no. 3, pp. 275–281, 2007.

[HWWM05] T. Heimann, I. Wolf, T. Williams, and H. Meinzer, "3D active shape models using gradient descent

optimization of description length," in *Proc. IPMI*, vol. LNCS 3565. Springer, 2005, pp. 566–577.

[HY01] M. Hansen and B. Yu, "Model Selection and the Principle of Minimum Description Length," *J Am Stat Assoc*, vol. 96, pp. 746–774, 2001.

[Hyv99] A. Hyvärinen, "Survey on independent component analysis," *Neural Comput Surv*, vol. 2, pp. 94–128, 1999.

[INT10] "Intermedia project," http://intermedia.miralab.ch/, 2006-2010.

[ISNC05] L. Ibanez, W. Schroeder, L. Ng, and J. Cates, *The ITK Software Guide*, 2nd ed., Kitware, Inc. ISBN 1-930934-15-7, http://www.itk.org/ItkSoftwareGuide.pdf, 2005.

[Jac93] D. Jackson, "Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches," *Ecology*, vol. 74, no. 8, pp. 2204–2214, 1993.

[JC02] H. Johnson and G. Christensen, "Consistent landmark and intensity-based image registration," *IEEE Trans Med Imag*, vol. 21, no. 5, pp. 450–461, 2002.

[JDJG04] S. Joshi, B. Davis, M. Jomier, and G. Gerig, "Unbiased diffeomorphic atlas construction for computational anatomy," *NeuroImage*, vol. 23, pp. 151–160, 2004.

[JDM00] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Trans Pattern Anal Mach Intell*, vol. 22, no. 1, pp. 4–37, 2000.

[JMG97] S. Joshi, M. Miller, and U. Grenander, "On the geometry and shape of brain sub-manifolds," *Int J Pattern Recogn Artif Intell*, vol. 11, no. 8, pp. 1317–1344, 1997.

[JNRS10] M. Journée, Y. Nesterov, P. Richtárik, and R. Sepulchre, "Generalized Power Method for Sparse Principal Component Analysis," *J Mach Learn Res*, vol. 11, pp. 517–553, 2010.

[Jol02] I. Jolliffe, *Principal component analysis*. Springer verlag, 2002.

[JPF+02] S. Joshi, S. Pizer, P. T. Fletcher, P. Yushkevich, A. Thall, and J. S. Marron, "Multiscale deformable model segmentation and statistical shape analysis using medial descriptions," *IEEE Trans Med Imag*, vol. 21, no. 5, pp. 538–550, 2002.

[JZDJ98] A. K. Jain, Y. Zhong, and M.-P. Dubuisson-Jolly, "Deformable template models: a review," *Signal Process*, vol. 71, no. 2, pp. 109–129, 1998.

[KA03] J. Kittler and F. M. Alkoot, "Sum Versus Vote Fusion in Multiple Classifier Systems," *IEEE Trans Pattern Anal Mach Intell*, vol. 25, no. 1, pp. 110–115, 2003.

[KA05] H. Knutsson and M. Andersson, "Morphons: Segmentation using elastic canvas and paint on priors," in *Proc. ICIP*, vol. 2. IEEE Computer Society, 2005, pp. 1226–1229.

[KB09] E. Kienel and G. Brunnett, "Tile-based Image Forces for Active Contours on GPU," in *Proc. Eurographics*. Eurographics Association, 2009, pp. 89–92.

[KBG+95] P. Kalra, P. Beylot, P. Gingins, N. Magnenat-Thalmann, P. Volino, P. Hoffmeyer, J. Fasel, and F. Terrier, "Topological Modeling Of Human Anatomy Using Medical Data," in *Proc. Comput Anim*. IEEE Computer Society, 1995, pp. 172–180.

[KBG+98] T. Kapur, P. Beardsley, S. Gibson, W. Grimson, and W. Wells, "Model-based segmentation of clinical knee MRI," in *Proc. IEEE Int. Work. Model-Based 3D Image Analysis*. IEEE Computer Society, 1998, pp. 97–106.

[KDR+06] A. Köhn, J. Drexl, F. Ritter, M. König, and H. Peitgen, "GPU accelerated image registration in two and three dimensions," in *Proc. BVM*. Springer, 2006, pp. 261–265.

[KE06] J. Karlsson and A. Ericsson, "A ground truth correspondence measure for benchmarking," in *Proc. ICPR*, vol. 3. IEEE Computer Society, 2006, pp. 568–573.

[KEK03] Y. Kang, K. Engelke, and W. Kalender, "A new accurate and precise 3-D segmentation method for skeletal structures in volumetric CT data," *IEEE Trans Med Imag*, vol. 22, no. 5, pp. 586–598, 2003.

[KGG+03] C. Kauffmann, P. Gravel, B. Godbout, A. Gravel, G. Beaudoin, J. Raynauld, J. Martel-Pelletier, J. Pelletier, and J. De Guise, "Computer-aided method for quantification of cartilage thickness and volume changes using MRI: validation study using a synthetic model," *IEEE Trans Biomed Eng*, vol. 50, no. 8, pp. 978–988, 2003.

[KH98] C. Kervrann and F. Heitz, "A Hierarchical Markov Modeling Approach for the Segmentation and Tracking of Deformable Shapes," *Graph Model Image Process*, vol. 60, no. 3, pp. 173–195, 1998.

[KHC05] I. Kardos, L. Hajder, and D. Chetverikov, "Bone Surface Reconstruction From CT/MR Images Using Fast

Marching and Level Set Methods," in *Proc. HACIPPR*. Austrian Computer Society, 2005, pp. 41–48.

[KHSJ01] H. Kargupta, W. Hunag, K. Sivakumar, and E. Johnson, "Distributed clustering using collective principal component analysis," *Knowledge and Information Systems*, vol. 3, no. 4, pp. 422–448, 2001.

[Kie97] H. A. Kiers, "Weighted least squares fitting using ordinary least squares algorithms," *Psychometrika*, vol. 62, no. 2, pp. 251–266, 1997.

[Kir84] S. Kirkpatrick, "Optimization by simulated annealing: Quantitative studies," *J Stat Phys*, vol. 34, no. 5, pp. 975–986, 1984.

[KK08] J. Kratky and J. Kybic, "Three-dimensional segmentation of bones from CT and MRI using fast level sets," in *Proc. SPIE Medical Imaging*, vol. 6914. SPIE, 2008, pp. 47–57.

[KL04] J. Koikkalainen and J. Lötjönen, "Reconstruction of 3-D Head Geometry from Digitized Point Sets: An Evaluation Study," *IEEE Trans Inform Tech Biomed*, vol. 8, no. 3, pp. 377–386, 2004.

[Kla07] O. Klar, "Interactive GPU-based Segmentation of Large Medical Volume Data with Level-Sets," in *Web Proc. CESCG*, 2007.

[KLZH08] D. Kainmüller, H. Lamecker, S. Zachow, and H.-C. Hege, "Coupling Deformable Models for Multi Object Segmentation," in *Proc. ISBMS*, vol. LNCS 5104. Springer, 2008, pp. 69–78.

[KLZH09] D. Kainmüller, H. Lamecker, S. Zachow, and H.-C. Hege, "An Articulated Statistical Shape Model for Accurate Hip Joint Segmentation," in *Proc. EMBC*. IEEE Computer Society, 2009, pp. 6345–6351.

[KMA+00] K. Krissian, G. Malandain, N. Ayache, R. Vaillant, and Y. Trousset, "Model Based Detection of Tubular Structures in 3D Images," *Comput Vis Image Understand*, vol. 80, no. 2, pp. 130–171, 2000.

[KMB+03] S. Klim, S. Mortensen, B. Bodvarsson, L. Hyldstrup, and H. Thodberg, "More active shape model," *Image Vis Comput*, vol. 3, pp. 26–28, 2003.

[KMG+10] S. Khallaghi, P. Mousavi, R. Gong, S. Gill, J. Boisvert, G. Fichtinger, D. Pichora, D. Borschneck, and P. Abolmaesumi, "Registration of a statistical shape model of the lumbar spine to 3d ultrasound images," in *Proc. MICCAI*, vol. LNCS 6362. Springer Berlin / Heidelberg, 2010, pp. 68–75.

[KMH+04] R. Keiser, M. Müller, B. Heidelberger, M. Teschner, and M. Gross, "Contact Handling for Deformable Point-Based Objects," in *Proc. VMV*, B. Girod, M. Magnor, and H.-P. Seidel, Eds. IOS Press, 2004, pp. 315–322.

[KNO+09] R. Kurazume, K. Nakamura, T. Okada, Y. Sato, N. Sugano, T. Koyama, Y. Iwashita, and T. Hasegawa, "3d reconstruction of a femoral shape using a parametric model and two 2d fluoroscopic images," *Comput Vis Image Understand*, vol. 113, no. 2, pp. 202–211, 2009.

[Koi06] J. Koikkalainen, "Image Databases in Medical Applications," Ph.D. dissertation, Helsinki University of Technology, Finland, 2006.

[KPL+03] M. Kaus, V. Pekar, C. Lorenz, R. Truyen, S. Lobregt, and J. Weese, "Automated 3-D PDM construction from segmented images using deformable models," *IEEE Trans Med Imag*, vol. 22, no. 8, pp. 1005–1013, 2003.

[KPL+04] J. Koikkalainen, M. Pollari, J. Lötjönen, S. Kivistö, and K. Lauerma, "Segmentation of Cardiac Structures Simultaneously from Short- and Long-Axis MR Images," in *Proc. MICCAI*, vol. LNCS 3216. Springer, 2004, pp. 427–434.

[KSG98] A. Kelemen, G. Szekely, and G. Gerig, "Three-dimensional model-based segmentation of brain mri," in *Proc. WBIA*. IEEE Computer Society, 1998, pp. 4–13.

[KSG99] A. Kelemen, G. Szekely, and G. Gerig, "Elastic model-based segmentation of 3-d neuroradiological data sets," *IEEE Trans Med Imag*, vol. 18, no. 10, pp. 828–839, 1999.

[KSMT10] J. Kim, J. Schmid, and N. Magnenat-Thalmann, "Automatic Detection of Hip Joint Center for the Visualization of Extreme Leg Motion," in *Proc. 2nd 3DAH Summer School*, 2010.

[KSP07] S. Klein, M. Staring, and P. Pluim, J., "Evaluation of Optimization Methods for Nonrigid Medical Image Reg-istration using Mutual Information and B-splines," *IEEE Trans Image Process*, vol. 16, no. 12, pp. 2879 – 2890, 2007.

[KSWM09] I. Kress, K. Siebenrock, S. Werlen, and T. Mamisch, "Sports-related alterations of the hip joint and correlation to typical findings in femuro-acetabular impingement: Mr findings in 50 elite karate fighters," in *European Congress on Radiology*, vol. Volume 19, no. 1, 2009, p. 64.

[KT98] A. Kotcheff and C. Taylor, "Automatic construction of eigenshape models by direct optimization," *Med Image Anal*, vol. 2, no. 4, pp. 303–314, 1998.

[KvdH04] B. L. Kaptein and F. C. T. van der Helm, "Estimating muscle attachment contours by transforming geometrical bone models," *J Biomech*, vol. 37, no. 3, pp. 263–273, 2004.

[KVP+10] P. Krekel, E. Valstar, F. Post, P. Rozing, and C. Botha, "Combined surface and volume processing for fused joint segmentation," *Int J Comput Assist Radiol Surg*, vol. 5, no. 3, pp. 263–273, 2010.

[KWT88] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int J Comput Vis*, vol. 1, no. 4, pp. 321–331, 1988.

[LAL+05] J. Lötjönen, K. Antila, E. Lamminmäki, J. Koikkalainen, M. Lilja, and T. Cootes, "Artificial enlargement of a training set for statistical shape models: Application to cardiac images," in *Proc. FIMH*, vol. LNCS 3504. Springer, 2005, pp. 92–101.

[Lam02] J. Lampinen, "A constraint handling approach for the differential evolution algorithm," in *Proc. CEC*. IEEE Computer Society, 2002, pp. 1468–1473.

[LB00] N. Lane and J. Buckwalter, "Exercise and osteoarthritis," *Curr Opin Orthop*, vol. 11, no. 1, pp. 62–65, 2000.

[LC87] W. Lorensen and H. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," in *Proc. SIGGRAPH*, M. Stone, C., Ed. ACM Press, 1987, pp. 163–169.

[LCW03] A. E. Lefohn, J. E. Cates, and R. T. Whitaker, "Interactive, GPU-Based Level Sets for 3D Segmentation," in *Proc. MICCAI*, vol. LNCS 2878. Springer, 2003, pp. 564–572.

[LE01] R. Larsen and H. Eiriksson, "Robust and resistant 2D shape alignment," Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, Tech. Rep., 2001.

[Lec89] Y. G. Leclerc, "Constructing simple stable descriptions for image partitioning," *Int J Comput Vis*, vol. 3, no. 1, pp. 73–102, 1989.

[LEHE97] A. Losch, F. Eckstein, M. Haubner, and K. Englmeier, "A non-invasive technique for 3-dimensional assessment of articular cartilage thickness based on MRI part 1: development of a computational method," *Magn Reson Imag*, vol. 15, no. 7, pp. 795–804, 1997.

[Lev00] M. E. Leventon, "Statistical models in medical image analysis," Ph.D. dissertation, Massachusetts Institute of Technology. Dept. of Electrical Engineering and Computer Science, 2000.

[Lev03] D. Levin, "Mesh-independent surface interpolation," in *Geometric Modeling for Scientific Visualization*, Brunnett, Hamann, and Mueller, Eds. Springer-Verlag, 2003, vol. 3, pp. 37–49.

[LFG+98] L. M. Lorigo, O. D. Faugeras, W. E. L. Grimson, R. Keriven, and R. Kikinis, "Segmentation of Bone in Clinical Knee MRI Using Texture-Based Geodesic Active Contours," in *Proc. MICCAI*, vol. LNCS 1496. Springer-Verlag, 1998, pp. 1195–1204.

[LGF00] M. E. Leventon, W. E. L. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," in *Proc. CVPR*, vol. 1. IEEE Computer Society, 2000, pp. 316–323.

[LH03] B. Luo and E. Hancock, "A unified framework for alignment and correspondence," *Comput Vis Image Understand*, vol. 92, no. 1, pp. 26–55, 2003.

[LHV07] N. Lord, J. Ho, and B. Vemuri, "USSR: A unified framework for simultaneous smoothing, segmentation, and registration of multiple images," in *Proc. ICCV*. IEEE Computer Society, 2007, pp. 1–6.

[LI05] Y. Li and W. Ito, "Shape Parameter Optimization for AdaBoosted Active Shape Model," in *Proc. ICCV*, vol. 1. IEEE Computer Society, 2005, pp. 251–258.

[LJT00] Z. Lin, J. Jin, and H. Talbot, "Unseeded region growing for 3D image segmentation," in *ACM International Conference Proceeding Series*, vol. 9. ACM Press, 2000, pp. 31–37.

[LK00] C. Lorenz and N. Krahnstöver, "Generation of point-based 3d statistical shape models for anatomical objects," *Comput Vis Image Understand*, vol. 77, no. 9, pp. 175–191, 2000.

[LKHW05] A. E. Lefohn, J. M. Kniss, C. D. Hansen, and R. T. Whitaker, "A streaming narrow-band algorithm: interactive computation and visualization of level sets," in *Proc. SIGGRAPH Courses*. New York, NY, USA: ACM Press, 2005, p. 243.

[LLPS05] J. Lee, J. Lee, M. Park, and S. Song, "An extensive comparison of recent classification tools applied to microarray data," *Comput Stat Data Anal*, vol. 48, no. 4, pp. 869–885, 2005.

[LLS02] H. Lamecker, T. Lange, and M. Seebass, "A Statistical Shape Model for the Liver," in *Proc. MICCAI*, vol. LNCS 2489. Springer, 2002, pp. 421–427.

[LM99] J.-O. Lachaud and A. Montanvert, "Deformable meshes with automated topology changes for coarse-to-fine three-dimensional surface extraction," *Med Image Anal*, vol. 3, no. 2, pp. 187–207, 1999.

[LMW+05] K. Li, S. Millington, X. Wu, D. Chen, and M. Sonka, "Simultaneous segmentation of multiple closed surfaces using optimal graph searching," in *Proc. IPMI*, vol. LNCS 3565. Springer, 2005, pp. 406–417.

[LN09] S. Lahabar and P. J. Narayanan, "Singular value decomposition on gpu using cuda," in *Proc. IPDPS*. IEEE Computer Society, 2009, pp. 1–10.

[LRN+08] L. Liu, D. Raber, D. Nopachai, P. Commean, D. Sinacore, F. Prior, R. Pless, and T. Ju, "Interactive separation of segmented bones in CT volumes using graph cut," in *Proc. MICCAI*, vol. LNCS 5241. Springer, 2008, pp. 296–304.

[LS97] L. Lam and S. Suen, "Application of majority voting to pattern recognition: an analysisof its behavior and performance," *IEEE Trans Syst Man Cybern Syst Hum*, vol. 27, no. 5, pp. 553–568, 1997.

[LS06] C. Lewis and S. Sahrmann, "Acetabular labral tears," *Phys Ther*, vol. 86, no. 1, pp. 110–121, 2006.

[LSHD04] H. Lamecker, M. Seebaß, H.-C. Hege, and P. Deuflhard, "A 3D statistical shape model of the pelvic bone for segmentation," in *Proc. SPIE*, vol. 5370. SPIE, 2004, pp. 1341–1351.

[LSS09] J. Liu, J. Sun, and H. Shum, "Paint selection," *ACM Trans Graph*, vol. 28, no. 3, pp. 1–7, 2009.

[LT97] K. Low and T. Tan, "Model simplification using vertex-clustering," in *Proc. I3D*. ACM Press, 1997, pp. 75–81.

[LvdGM+02] B. Lelieveldt, R. van der Geest, S. Mitchell, J. Bosch, M. Sonka, and J. Reiber, "3D active appearance models: Fully automatic detection of endo-and epicardial contours in short-axis cardiac MR data," in *Proc. ISMRM*, vol. 2. MIRA, 2002, p. 1668.

[LWH06] H. Lamecker, T. Wenckebach, and H. Hege, "Atlas-based 3D-Shape Reconstruction from X-Ray Images," in *Proc. ICPR*, vol. 1. IEEE Computer Society, 2006, pp. 371–374.

[LWTX06] L. Liang, F. Wen, X. Tang, and Y. Xu, "An integrated model for accurate shape alignment," in *Proc. ECCV*, vol. LNCS 3954. Springer, 2006, pp. 333–346.

[LXMJ03] Y. Li, L. Xu, J. Morphett, and R. Jacobs, "An integrated algorithm of incremental and robust PCA," in *Proc. ICIP*, vol. 1. IEEE Computer Society, 2003, pp. 245–248.

[LYJ03] H. Livyatan, Z. Yaniv, and L. Joskowicz, "Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT," *IEEE Trans Med Imag*, vol. 22, no. 11, pp. 1395–1406, 2003.

[LZ00] J. Lampinen and I. Zelinka, "On stagnation of the differential evolution algorithm," in *Proc. MENDEL*, 2000, pp. 76–83.

[LZJ04] S. Li, L. Zhu, and T. Jiang, "Active shape model segmentation using local edge structures and AdaBoost," in *Proc. MIAR*, vol. LNCS 3150. Springer, 2004, pp. 121–128.

[LZZ+00] J. Lynch, S. Zaim, J. Zhao, A. Stork, C. Peterfy, and H. Genant, "Cartilage segmentation of 3D MRI scans of the osteoarthritic knee combining user knowledge and active contours," in *Proc. SPIE*, vol. 3979. SPIE, 2000, p. 925.

[MBL+91] J. V. Miller, D. E. Breen, W. E. Lorensen, R. M. O'Bara, and M. J. Wozny, "Geometrically deformed models: a method for extracting closed geometric models form volume data," *ACM Comput Graph*, vol. 25, no. 4, pp. 217–226, 1991.

[MC03] M. Moelich and T. Chan, "Joint segmentation and registration using logic models," UCLA, Tech. Rep. CAM Report 03-06, 2003.

[MD98] J. Montagnat and H. Delingette, "Globally constrained deformable models for 3D object reconstruction," *Signal Process*, vol. 71, no. 2, pp. 173–186, 1998.

[MD05] J. Montagnat and H. Delingette, "4D deformable models with temporal constraints: application to 4D cardiac image segmentation," *Med Image Anal*, vol. 9, no. 1, pp. 87–100, 2005.

[MDA01] J. Montagnat, H. Delingette, and N. Ayache, "A review of deformable surfaces: Topology, geometry and deformation," *Image Vis Comput*, vol. 19, pp. 1023–1040, 2001.

[MDW07] B. Munsell, P. Dalal, and S. Wang, "A New Benchmark for Shape Correspondence Evaluation," in *Proc. MICCAI*, N. Ayache, S. Ourselin, and A. Maeder, Eds., vol. LNCS 4791. Springer-Verlag Berlin Heidel-

berg, 2007, pp. 307–314.

[MDW08] B. Munsell, P. Dalal, and S. Wang, "Evaluating shape correspondence for statistical shape analysis: A benchmark study," *IEEE Trans Pattern Anal Mach Intell*, vol. 30, no. 11, pp. 2023–2039, 2008.

[MH80] D. Marr and E. Hildreth, "Theory of edge detection," *Proc Roy Soc Lond B Biol Sci*, pp. 187–217, 1980.

[MH06] C. McIntosh and G. Hamarneh, "Vessel Crawlers: 3D Physically-based Deformable Organisms for Vasculature Segmentation and Analysis," in *Proc. CVPR*, vol. 1. IEEE Computer Society, 2006, pp. 1084–1091.

[MHS05] J. Mosegaard, P. Herborg, and T. Sorensen, "A GPU accelerated spring mass system for surgical simulation," *Stud Health Tech Informat*, vol. 111, pp. 342–348, 2005.

[MI05] K. McLeish and P. Irarrazaval, "Scan time reduction with an adaptive field of view," *Magn Reson Imag*, vol. 23, no. 1, pp. 47 – 52, 2005.

[MKE03] J. Mezger, S. Kimmerle, and O. Etzmuß, "Hierarchical Techniques in Collision Detection for Cloth Animation," *J. WSCG*, vol. 11, no. 2, pp. 322–329, 2003.

[MKP+05] A. McPherson, J. Kärrholm, V. Pinskerova, A. Sosna, and S. Martelli, "Imaging knee position using MRI, RSA/CT and 3D digitisation," *J Biomech*, vol. 38, no. 2, pp. 263–268, 2005.

[MLG+01] J. Marescaux, J. Leroy, M. Gagner, F. Rubino, D. Mutter, M. Vix, S. Butner, and M. Smith, "Transatlantic robot-assisted telesurgery," *Nature*, vol. 413, pp. 379–380, 2001.

[MLvdG+00] S. Mitchell, B. Lelieveldt, R. van der Geest, J. Schaap, J. Reiber, and M. Sonka, "Segmentation of cardiac mr images: An active appearance model approach," in *Proc. SPIE*, vol. 3979. SPIE, 2000, p. 224.

[MLZ+10] J. Ma, L. Lu, Y. Zhan, X. Zhou, M. Salganicoff, and A. Krishnan, "Hierarchical Segmentation and Identification of Thoracic Vertebra Using Learning-Based Edge Detection and Coarse-to-Fine Deformable Model," in *Proc. MICCAI*, vol. LNCS 6361. Springer, 2010, pp. 19–27.

[MNPV99] E. Meijering, W. Niessen, J. Pluim, and M. Viergever, "Quantitative comparison of sinc-approximating kernels for medical image interpolation," in *Proc. MICCAI*, vol. LNCS 1679. Springer, 1999, pp. 210–217.

[MNV01] E. Meijering, W. Niessen, and M. Viergever, "Quantitative evaluation of convolution-based methods for medical image interpolation," *Med Image Anal*, vol. 5, no. 2, pp. 111–126, 2001.

[MOOXS08] P. Muyan-Ozcelik, J. Owens, J. Xia, and S. Samant, "Fast deformable registration on the GPU: A CUDA implementation of demons," in *Proc. ICCSA*. IEEE Computer Society, 2008, pp. 223–233.

[Mor77] J. Moré, "The Levenberg-Marquardt algorithm: implementation and theory," *Numer Anal*, vol. 630, pp. 105–116, 1977.

[Mor95] B. S. Morse, "Computation of object cores from grey-level images," Ph.D. dissertation, University of North Carolina, Chapel Hill, Chapel Hill, NC, USA, 1995.

[MS10] D. Mahapatra and Y. Sun, "Joint Registration and Segmentation of Dynamic Cardiac Perfusion Images Using MRFs," in *Proc. MICCAI*, ser. Lecture Notes in Computer Science, T. Jiang, N. Navab, J. P. Pluim, and M. A. Viergever, Eds., vol. LNCS 6361. Springer, 2010, pp. 493–501.

[MSP+09] X. Maurice, A. Sandholm, N. Pronost, R. Boulic, and D. Thalmann, "A subject-specific software solution for the modeling and the visualization of muscles deformations," *Vis Comput*, vol. 25, pp. 835–842, 2009.

[MSV95] R. Malladi, J. Sethian, and B. Vemuri, "Shape Modeling with Front Propagation: A Level Set Approach," *IEEE Trans Pattern Anal Mach Intell*, vol. 17, pp. 158–175, 1995.

[MT96] T. McInerney and D. Terzopoulos, "Deformable models in medical images analysis: a survey," *Med Image Anal*, vol. 1, no. 2, pp. 91–108, 1996.

[MT99] W. Maurel and D. Thalmann, "A case study on human upper limb modelling for dynamic simulation," *Comput Meth Biomech Biomed Eng*, vol. 2, no. 1, pp. 65–82, 1999.

[MT00] T. McInerney and D. Terzopoulos, "T-Snakes: Topology Adaptive Snakes," *Med Image Anal*, vol. 4, no. 2, pp. 73–91, 2000.

[MTCS08] N. Magnenat-Thalmann, C. Charbonnier, and J. Schmid, "Multimedia Application to the Simulation of Human Musculoskeletal System: A Visual Lower Limb Model from Multimodal Captured Data," in *Proc. MMSP*. IEEE Publisher, 2008, pp. 520–525.

[MTYCS04] N. Magnenat-Thalmann, L. Yahia-Cherif, and H. Seo, "Modeling Anatomical-Based Humans," in *Proc. ICIG*. IEEE Publisher, 2004, pp. 476–480.

[MV98] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Med Image Anal*, vol. 2,

no. 1, pp. 1–36, 1998.

[MVDEV96] J. Maintz, P. Van Den Elsen, and M. Viergever, "Evaluation of ridge seeking operators for multimodality medical image matching," *IEEE Trans Pattern Anal Mach Intell*, vol. 18, no. 4, pp. 353–365, 1996.

[MWA06] B. Moghaddam, Y. Weiss, and S. Avidan, "Spectral bounds for sparse PCA: Exact and greedy algorithms," *Adv Neural Inform Process Syst*, vol. 18, pp. 915–922, 2006.

[MWMTT98] W. Maurel, Y. Wu, N. Magnenat-Thalmann, and D. Thalmann, *Biomechanical Models for Soft Tissue Simulation*, ser. Basic Research. New York: Springer-Verlag, 1998.

[MY01] M. Miller and L. Younes, "Group actions, homeomorphisms, and matching: A general framework," *International Journal of Computer Vision*, vol. 41, no. 1, pp. 61–84, 2001.

[NB06] J. Noble and D. Boukerroui, "Ultrasound image segmentation: A survey," *IEEE Trans Med Imag*, vol. 25, no. 8, pp. 987–1010, 2006.

[Nea08] A. C. Nearchou, "A differential evolution approach for the common due date early/tardy job scheduling problem," *Comput Oper Res*, vol. 35, no. 4, pp. 1329–1343, 2008.

[NL98] A. Neumann and C. Lorenz, "Statistical shape model based segmentation of medical images," *Comput Med Imag Graph*, vol. 22, no. 2, pp. 133–143, 1998.

[NL05] V. D. M. Nhat and S. Lee, "Improvement on PCA and 2D PCA Algorithms for Face Recognition," in *Proc. CIVR*, Springer-Verlag, Ed., vol. LNCS 3568, 2005, pp. 568–577.

[NLBA07] N. Nguyen, D. Laurendeau, and A. Branzan-Albu, "A new segmentation method for MRI images of the shoulder joint," in *Proc. CRV*. IEEE Computer Society, 2007, pp. 329–338.

[NM65] J. Nelder and R. Mead, "A simplex method for function minimization," *Comput J*, vol. 7, no. 4, p. 308, 1965.

[NM05] J. Nascimento and J. Marques, "Adaptive snakes using the EM algorithm," *IEEE Trans Image Process*, vol. 14, no. 11, pp. 1678–1686, 2005.

[NMK+06] A. Nealen, M. Muller, R. Keiser, E. Boxerman, and M. Carlson, "Physically based deformable models in computer graphics," *Comput Graph Forum*, vol. 25, no. 4, pp. 809–836, 2006.

[Nob88] J. Noble, "Finding corners," *Image Vis Comput*, vol. 6, no. 2, pp. 121–128, 1988.

[NSC+05] Z. Niu, S. Shan, X. Chen, B. Ma, and W. Gao, "Enhance ASMs Based on AdaBoost-Based Salient Landmarks Localization and Confidence-Constraint Shape Modeling," in *Advances in Biometric Person Authentication*, vol. LNCS 3781. Springer, 2005, pp. 9–14.

[NSP+04] S. Nicolau, J. Schmid, X. Pennec, L. Soler, and N. Ayache, "An augmented reality & virtuality interface for a puncture guidance system: Design and validation on an abdominal phantom," in *Proc. MIAR*, vol. LNCS 3150. Springer, 2004, pp. 302–310.

[NT98] L. P. Nedel and D. Thalmann, "Real time muscle deformations using mass-spring systems," in *Proc. CGI*. IEEE Computer Society, 1998, pp. 156–165.

[NTHF02] V. Ng-Thow-Hing and E. Fiume, "Application-specific muscle representations," in *Proc. GI*, W. Sturzlinger and M. McCool, Eds. AK Peters Ltd, 2002, pp. 107–116.

[NVI10] NVIDIA Corporation, *NVIDIA CUDA Programming Guide, ver. 3.1*, May 2010.

[OBB+04] S. Ordas, L. Boisrobert, M. Bossa, M. Huguet, M. Laucelli, S. Olmos, and A. Frangi, "Grid-enabled automatic construction of a two-chamber cardiac PDM from a large database of dynamic 3D shapes," in *Proc. ISBI*. IEEE Computer Society, 2004, pp. 416–419.

[OEvJ99] H. Overhoff, S. Ehrich, and U. von Jan, "Reliable identification of sphere-shaped femoral heads in 3-D image data," in *Proc. SPIE*, vol. 3661. SPIE, 1999, pp. 1377–1387.

[OLU+03] C. Oost, B. Lelieveldt, M. Üzümcü, H. Lamb, J. Reiber, and M. Sonka, "Multi-view active appearance models: Application to X-ray LV angiography and cardiac MRI," in *Proc. IPMI*, vol. LNCS 2732. Springer, 2003, pp. 234–245.

[OMB05] J. Ohmer, F. Maire, and R. Brown, "Implementation of Kernel Methods on the GPU," in *Proc. DICTA*. IEEE Computer Society, 2005, pp. 78–86.

[OMSA09] K. Oberhofer, K. Mithraratne, N. Stott, and I. Anderson, "Anatomically-based musculoskeletal modeling: prediction and validation of muscle deformation during walking," *Vis Comput*, vol. 25, no. 9, pp. 843–851, 2009.

[ØNDSE+08] K. Østergaard Noe, B. De Senneville, U. Elstrøm, K. Tanderup, and T. Sørensen, "Acceleration and validation of optical flow based deformable registration for image-guided radiotherapy," *Acta Oncol*, vol. 47, no. 7, pp. 1286–1293, 2008.

[OS88] S. Osher and J. Sethian, "Fronts propagating with curvature dependent speed: Algorithms based on hamilton-jacobi formulation," *J Comput Phys*, vol. 79, pp. 12–49, 1988.

[OS01] S. Olabarriaga and A. Smeulders, "Interaction in the segmentation of medical images: A survey," *Med Image Anal*, vol. 5, no. 2, pp. 127–142, 2001.

[Ots79] N. Otsu, "A Threshold selection Method from Gray-Level Histograms," *IEEE Trans. Syst. Man Cybern.*, vol. 9, no. 1, pp. 62–66, 1979.

[Pan09] W. Pan, "Improving Interactive Image Segmentation via Appearance Propagation," in *Proc. Eurographics*, P. Alliez and M. Magnor, Eds. Eurographics Association, 2009, pp. 93–96.

[Par62] E. Parzen, "On the estimation of a probability density function and mode," *Ann Math Stat*, vol. 33, no. 3, pp. 1065–1076, 1962.

[Par91] J. Parker, "Gray level thresholding in badly illuminated images," *IEEE Trans Pattern Anal Mach Intell*, vol. 13, no. 8, pp. 813–819, 1991.

[PB07] J. Pons and J. Boissonnat, "Delaunay deformable models: Topology-adaptive meshes based on the restricted Delaunay triangulation," in *Proc. CVPR*. IEEE Computer Society, 2007, pp. 1–8.

[PBHM05] H. Park, P. Bland, A. Hero, and C. Meyer, "Least biased target selection in probabilistic atlas construction," in *Proc. MICCAI*, vol. LNCS 3750. Springer, 2005, pp. 419–426.

[PBTM06] A. Pitiot, E. Bardinet, P. Thompson, and G. Malandain, "Piecewise affine registration of biological images for volume reconstruction," *Med Image Anal*, vol. 10, no. 3, pp. 465–483, 2006.

[Pea01] K. Pearson, "On Lines and Planes of Closest Fit to Systems of Points in Space," *Phil Mag*, vol. 2, no. 6, pp. 559–572, 1901.

[PEFM98] S. M. Pizer, D. Eberly, D. S. Fritsch, and B. S. Morse, "Zoom-Invariant Vision of Figural Shape: The Mathematics of Cores," *Comput Vis Image Understand*, vol. 69, no. 1, pp. 055–071, 1998.

[Pen90] A. Pentland, "Automatic extraction of deformable part models," *Int J Comput Vis*, vol. 4, no. 2, pp. 107–126, 1990.

[Per98] P. Perez, "Markov random fields and images," *CWI Quarterly*, vol. 11, no. 4, pp. 413–437, 1998.

[PEW05] J. Peters, O. Ecabert, and J. Weese, "Feature optimization via simulated search for model-based heart segmentation," *International Congress Series*, vol. 1281, pp. 33–38, 2005.

[PFG+06] K. M. Pohl, J. Fisher, W. E. L. Grimson, R. Kikinis, and W. M. Wells, "A Bayesian model for joint segmentation and registration," *Neuroimage*, vol. 31, no. 1, pp. 228–239, 2006.

[PFJ+03] S. Pizer, P. Fletcher, S. Joshi, A. Thall, J. Chen, Y. Fridman, D. Fritsch, A. Gash, J. Glotzer, M. Jiroutek *et al.*, "Deformable m-reps for 3d medical image segmentation," *Int J Comput Vis*, vol. 55, no. 2, pp. 85–106, 2003.

[PFL+05] K. M. Pohl, J. W. Fisher, J. J. Levitt, M. E. Shenton, R. Kikinis, W. E. L. Grimson, and W. M. Wells, "A Unifying Approach to Registration, Segmentation, and Intensity Correction," in *Proc. MICCAI*, vol. LNCS 3749. Springer, 2005, pp. 310–318.

[PFS+06] K. Pohl, J. Fisher, M. Shenton, R. McCarley, W. Grimson, R. Kikinis, and W. Wells, "Logarithm odds maps for shape representation," in *Proc. MICCAI*, vol. LNCS 4191. Springer, 2006, pp. 955–963.

[PFTV92] W. H. Press, B. P. Flannery, S. A. Teukolosky, and W. T. Vetterling, *Numerical Recipes in C, Second ed.* Cambridge University Press, 1992.

[PJBB05] T. Pock, C. Janko, R. Beichel, and H. Bischof, "Multiscale Medialness for Robust Segmentation of 3D Tubular Structures," in *Proc. CVWW*, A. Hanbury and H. Bischof, Eds. PRIP, 2005, pp. 93–102.

[PKA+05] M. Pauly, R. Keiser, B. Adams, P. Dutre, M. Gross, and L. Guibas, "Meshless animation of fracturing solids," *ACM Trans Graph*, vol. 24, no. 3, pp. 957–964, 2005.

[PKB06] J. Pettersson, H. Knutsson, and M. Borga, "Automatic hip bone segmentation using non-rigid registration," in *Proc. ICPR*, vol. 3. IEEE Computer Society, 2006, pp. 946–949.

[PKI08] S. Park, W. Kim, and I. Ihm, "Mobile collaborative medical display system," *Comput Meth Programs Biomed*, vol. 89, no. 3, pp. 248–260, 2008.

[PLNS04] D. Pantazis, R. Leahy, T. Nichols, and M. Styner, "Statistical surface-based morphometry using a non-parametric approach," in *Proc. ISBI: Nano to Macro*. IEEE Computer Society, 2004, pp. 1283–1286.

[PM90] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans Pattern Anal Mach Intell*, vol. 12, no. 7, pp. 629–639, 1990.

[PMD+06] C. W. A. Pfirrmann, B. Mengiardi, C. Dora, F. Kalberer, M. Zanetti, and J. Hodler, "Cam and Pincer Femoroacetabular Impingement: Characteristic MR Arthrographic Findings in 50 Patients," *Radiology*, vol. 240, no. 3, pp. 778–785, 2006.

[PMV03] J. Pluim, J. Maintz, and M. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Trans Med Imag*, vol. 22, no. 8, pp. 986–1004, 2003.

[PP93] U. Pinkall and K. Polthier, "Computing discrete minimal surfaces and their conjugates," *Exp Math*, vol. 2, no. 1, pp. 15–36, 1993.

[PP97] K. Popat and R. W. Picard, "Cluster based probability model and its application to image and texture processing." *IEEE Trans Image Process*, vol. 6, no. 2, pp. 268–284, 1997.

[PS91] A. Pentland and S. Sclaroff, "Closed-form solutions for physically based shape modeling and recognition," *IEEE Trans Pattern Anal Mach Intell*, vol. 13, no. 7, pp. 715–729, 1991.

[PSS01] E. Praun, W. Sweldens, and P. Schröder, "Consistent mesh parameterizations," in *Proc. SIGGRAPH*. ACM Press, 2001, pp. 179–184.

[PTB+03] S. Premože, T. Tasdizen, J. Bigler, A. Lefohn, and R. Whitaker, "Particle-based simulation of fluids," *Comput Graph Forum*, vol. 22, no. 3, pp. 401–410, 2003.

[PV03] C. Peters and F. Valafar, "Comparison of Three Nonparametric Density Estimation Techniques Using Baye's Classifier Applied to Mircoarray Data Analysis," in *Proc. METMBS*, 2003, pp. 119–125.

[PVC+01] Q. Pham, F. Vincent, P. Clarysse, P. Croisille, I. Magnin, C. CNRS, and V. INSA, "A FEM-based deformable model for the 3D segmentation and trackingof the heart in cardiac MRI," in *Proc. ISPA*. IEEE Standards, 2001, pp. 250–254.

[PXP00] D. Pham, C. Xu, and J. Prince, "Current Methods In Medical Image Segmentation," *Annual Review of Biomedical Engineering*, vol. 2, no. 1, pp. 315–337, 2000.

[R D10] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2010, ISBN 3-900051-07-0.

[RBH+00] H. Rifai, I. Bloch, S. Hutchinson, J. Wiart, and L. Garnero, "Segmentation of the skull in mri volumes using deformable model and taking the partial volume effect into account," *Med Image Anal*, vol. 4, no. 3, pp. 219–233, 2000.

[RC78] T. Ridler and S. Calvard, "Picture thresholding using an iterative selection method," *IEEE Trans Syst Man Cybern Cybern*, vol. 8, no. 8, pp. 630–632, 1978.

[RDG07] V. Raykar, R. Duraiswami, and N. Gumerov, "Fast computation of sums of gaussians," *J Mach Learn Res*, vol. submitted, 2007.

[Reg02] J. Reginster, "The prevalence and burden of arthritis," *Rheumatology*, vol. 41, no. 12, pp. 3–6, 2002.

[RFS01] D. Rueckert, A. Frangi, and J. Schnabel, "Automatic construction of 3-D statistical deformation models using non-rigid registration," in *Proc. MICCAI*, W. Niessen and M. Viergever, Eds., vol. LNCS 2208. Springer, 2001, pp. 77–84.

[RFS03] D. Rueckert, A. Frangi, and J. Schnabel, "Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration," *IEEE Trans Med Imag*, vol. 22, no. 8, pp. 1014–1025, 2003.

[RG02] M. Rogers and J. Graham, "Robust active shape model search," in *Proc. ECCV*, vol. LNCS 2353. Springer, 2002, pp. 517–530.

[RHA+98] L. Reeva, C. Helmick, F. Arnett, R. Deyo, D. Felson, E. Giannini, P. Stephen, R. Hirsch, C. Marc, G. Gene *et al.*, "Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States," *Arthritis Rheum*, vol. 41, no. 5, pp. 778–799, 1998.

[Ris83] J. Rissanen, "A Universal Prior for Integers and Estimation by Minimum Description Length," *Ann Stat*, vol. 11, no. 2, pp. 416–431, 1983.

[RLB+02] J. Ribelles, A. López, O. Belmonte, I. Remolar, and M. Chover, "Multiresolution modeling of arbitrary polygonal surfaces: a characterization," *Comput Graph*, vol. 26, no. 3, pp. 449–462, 2002.

[RM09] O. E. Ripollés Mateu, "Towards a Multiresolution Model on GPU," Ph.D. dissertation, Universitat Jaume I, 2009.

[RNS06] J. Rodriguez-Navarro and A. Susin, "Non structured meshes for Cloth GPU simulation using FEM," in *Proc. VRIPHYS*, C. Mendoza and I. Navazo, Eds. Eurographics Association, 2006, pp. 1–7.

[Rog01] M. Rogers, "Exploiting Weak Constraints on Object Structure and Appearance for Segmentation of 2-D Images," Ph.D. dissertation, University of Manchester, 2001.

[Roh01] K. Rohr, *Landmark-based image analysis: using geometric and intensity models*. Springer Netherlands, 2001.

[Row98] S. Roweis, "EM algorithms for PCA and SPCA," in *Proc. NIPS*, M. I. Jordan, M. J. Kearns, and S. A. Solla, Eds., vol. 10. The MIT Press, 1998, pp. 626–632.

[RP02] M. Rousson and N. Paragios, "Shape priors for level set representations," in *Proc. ECCV*, vol. LNCS 2351. Springer, 2002, pp. 78–92.

[RPR+00] A. Roche, X. Pennec, M. Rudolph, D. Auer, G. Malandain, S. Ourselin, L. Auer, and N. Ayache, "Generalized correlation ratio for rigid registration of 3D ultrasound with MR images," in *Proc. MICCAI*, vol. LNCS 1935. Springer, 2000, pp. 203–220.

[RSGP06] M. Russell, K. Shivanna, N. Grosland, and D. Pedersen, "Cartilage contact pressure elevations in dysplastic hips: a chronic overload model," *J Orhtop Surg Res*, vol. 1, no. 6, pp. 169–177, 2006.

[RSH+99] D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, and D. Hawkes, "Nonrigid registration using free-form deformations: application to breast MR images," *IEEE Trans Med Imag*, vol. 18, no. 8, pp. 712–721, 1999.

[RSS+01] K. Rohr, H. Stiehl, R. Sprengel, T. Buzug, J. Weese, and M. Kuhn, "Landmark-based elastic registration using approximating thin-plate splines," *IEEE Trans Med Imag*, vol. 20, no. 6, pp. 526–534, 2001.

[RST+07] K. T. Rajamani, M. A. Styner, H. Talib, G. Zheng, L. P. Nolte, and M. A. González Ballester, "Statistical deformable bone models for robust 3D surface extrapolation from sparse data," *Med Image Anal*, vol. 11, no. 2, pp. 99–109, 2007.

[Rux06] G. Ruxton, "The unequal variance t-test is an underused alternative to Student's t-test and the Mann-Whitney U test," *Behavioral Ecology*, vol. 17, no. 4, pp. 688–690, 2006.

[S+05] S. Standring *et al.*, *Gray's anatomy, 39th Edition*. Churchill Livingstone, Elsevier, 2005.

[SBPC07a] J. Stough, R. Broadhurst, S. Pizer, and E. Chaney, "Regional appearance in deformable model segmentation," in *Proc. IPMI*, vol. LNCS 4584. Springer, 2007, pp. 532–543.

[SBPC07b] J. Stough, R. Broadhurst, S. Pizer, and E. Chaney, "Clustering on local appearance for deformable model segmentation," in *Proc. ISBI: From Nano to Macro*. IEEE Computer Society, 2007, pp. 960–963.

[SBSG02] M. Styner, C. Brechbuhler, G. Szckely, and G. Gerig, "Parametric estimate of intensity inhomogeneities applied to MRI," *IEEE Trans Med Imag*, vol. 19, no. 3, pp. 153–165, 2002.

[SC04] F. Shellock and J. Crues, "MR procedures: Biologic effects, safety, and patient care," *Radiology*, vol. 232, pp. 635–652, 2004.

[SCD+03] M. Straka, A. L. Cruz, L. Dimitrov, M. Srámek, D. Fleischmann, and M. E. Gröller, "Bone Segmentation in CT-Angiography Data Using a Probabilistic Atlas," in *Proc. VMV*. IOS Press, 2003, pp. 505–512.

[Sch05] F. Schick, "Whole-body MRI at high field: technical limits and clinical potential," *Eur Radiol*, vol. 15, no. 5, pp. 946–959, 2005.

[SCT03] I. Scott, T. Cootes, and C. Taylor, "Improving appearance model matching using local image structure," in *Proc. IPMI*, vol. LNCS 2732. Springer, 2003, pp. 258–269.

[SCT07] O. Sadowsky, G. Chintalapani, and R. Taylor, "Deformable 2D-3D registration of the pelvis with a limited field of view, using shape statistics," in *Proc. MICCAI*, vol. LNCS 4792. Springer, 2007, pp. 519–526.

[SD96] L. Staib and J. Duncan, "Model-based deformable surface finding for medical images," *IEEE Trans Med Imag*, vol. 15, no. 5, pp. 720–731, 1996.

[SDR04] R. Strzodka, M. Droske, and M. Rumpf, "Image registration by a regularized gradient flow. a streaming implementation in dx9 graphics hardware," *Computing*, vol. 73, no. 4, pp. 373–389, 2004.

[Set99] J. Sethian, *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*. Cambridge University Press, 1999.

[SFK+08] P. Steininger, K. Fritscher, G. Kofler, B. Schuler, M. Hanni, K. Schwieger, and R. Schubert, "Comparison of

Different Metrics for Appearance-Model-Based 2D/3D-registration with X-ray Images," *Bildverarbeitung Für Die Medizin 2008: Algorithmen-Systeme-Anwendungen*, pp. 122–126, 2008.

[SGS07] E. Syrkina, M. A. González Ballester, and G. Székely, "Correspondence Establishment in Statistical Modeling of Shapes with Arbitrary Topology," in *Proc. ICCV*. IEEE Computer Society, 2007, pp. 1–7.

[SH08] H. Shen and J. Huang, "Sparse principal component analysis via regularized low rank matrix approximation," *J Multivariate Anal*, vol. 99, no. 6, pp. 1015–1034, 2008.

[SHH99] C. Studholme, D. Hill, and D. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recogn*, vol. 32, no. 1, pp. 71–86, 1999.

[Shl03] J. Shlens, "A tutorial on principal component analysis," Online: http://www.snl.salk.edu/shlens/notes.html, 2003.

[SHN03] A. Sherbondy, M. Houston, and S. Napel, "Fast volume segmentation with simultaneous visualization using programmable graphics hardware," in *Proc. VIS*. Washington, DC, USA: IEEE Computer Society, 2003, pp. 171–176.

[SJL+07] L. Scheys, I. Jonkers, D. Loeckx, A. Van Campenhout, A. Spaepen, and P. Suetens, "Estimation Of Hip-muscle Geometry Using Automated Non-rigid Atlas-based Registration Of MR Images," in *Annual Meeting of American Society of Biomechanics*. ASB, 2007, pp. 191–192.

[SK05] Y. Shi and W. Karl, "A fast level set method without solving pdes," in *Proc. ICASSP*, vol. 2. IEEE Computer Society, 2005, pp. 97–100.

[SKH+08] H. Seim, D. Kainmüller, M. Heller, H. Lamecker, S. Zachow, and H.-C. Hege, "Automatic Segmentation of the Pelvic Bones from CT Data Based on a Statistical Shape Model," in *Proc. Eurographics Workshop VCBM*. Eurographics Association, 2008, pp. 93–100.

[SKSF07] G. Sharp, N. Kandasamy, H. Singh, and M. Folkert, "GPU-based streaming architectures for fast cone-beam CT image reconstruction and demons deformable registration," *Phys Med Biol*, vol. 52, no. 19, pp. 5771–5784, 2007.

[SL96] R. Szeliski and S. Lavallée, "Matching 3-D anatomical surfaces with non-rigid deformations using octree-splines," *Int J Comput Vis*, vol. 18, no. 2, pp. 171–186, 1996.

[SL03] D. Skočaj and A. Leonardis, "Weighted and Robust Incremental Method for Subspace Learning," in *Proc. ICCV*, vol. 2. IEEE Computer Society, 2003, pp. 1494–1501.

[SLB07] D. Skočaj, A. Leonardis, and H. Bischof, "Weighted and robust learning of subspace representations," *Pattern Recogn*, vol. 40, no. 5, pp. 1556–1569, 2007.

[SLS+09] L. Scheys, D. Loeckx, A. Spaepen, P. Suetens, and I. Jonkers, "Atlas-based non-rigid image registration to automatically define line-of-action muscle models: A validation study," *J Biomech*, vol. 42, no. 5, pp. 565–572, 2009.

[SM99] L. V. Safont and E. M. Marroquín, "3D Reconstruction of Third Proximal Femur (31-) with Active Contours," in *Proc. ICIAP*. Washington, DC, USA: IEEE Computer Society, 1999, pp. 458–463.

[SMFF07] J. Salvi, C. Matabosch, D. Fofi, and J. Forest, "A review of recent range image registration methods with accuracy evaluation," *Image Vis Comput*, vol. 25, no. 5, pp. 578–596, 2007.

[SMRW06] M. Seise, S. McKenna, J., I. Ricketts, W., and C. Wigderowitz, A., "Probabilistic Segmentation of the Knee Joint from X-ray Images," in *Web Proc. MIUA*, N. T. J. Graham and T. Cootes, Eds., vol. 1, 2006, pp. 110–114.

[SMS+03] T. Surazhsky, E. Magid, O. Soldea, G. Elber, and E. Rivlin, "A comparison of gaussian and mean curvatures estimation methods on triangular meshes," in *Proc. ICRA*, vol. 1. IEEE Computer Society, 2003, pp. 1021–1026.

[SNH+09] J. Schmid, N. Nijdam, A. Han, J. Kim, and N. Magnenat-Thalmann, "Interactive Segmentation of Volumetric Medical Images for Collaborative Telemedicine," in *Modelling the Physiological Human, Proc. 3DPH*, vol. LNCS 5903. Springer, 2009, pp. 13–24.

[SNTH03] W. Shao and V. Ng-Thow-Hing, "A general joint component framework for realistic articulation in human characters," in *Proc. I3D*. ACM Press, 2003, pp. 11–18.

[Som89] K. M. Somers, "Allometry, isometry and shape in principal components analysis," *Syst Zool*, vol. 38, no. 2, pp. 169–173, 1989.

[SP95] R. Storn and K. Price, "Differential evolution-a simple and efficient adaptive scheme for global optimization over continuous spaces," International Computer Science Institute, Tech. Rep. TR-95-012, 1995.

[SP07a] A. Seth and M. Pandy, "A neuromusculoskeletal tracking method for estimating individual muscle forces in human movement," *J Biomech*, vol. 40, no. 2, pp. 356–366, 2007.

[SP07b] A. Sharma and K. K. Paliwal, "Fast principal component analysis using fixed-point algorithm," *Pattern Recogn Lett*, vol. 28, no. 10, pp. 1151–1155, 2007.

[SPCM97] F. Scheepers, R. E. Parent, W. E. Carlson, and S. F. May, "Anatomy-based modeling of the human musculature," in *Proc. SIGGRAPH*. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1997, pp. 163–172.

[Spi93] L. Spirkovska, "A Summary of Image Segmentation Techniques," NASA, Ames Research Center, Moffett Field, California, USA, Technical Memorandum 104022, 1993.

[SRN+03] M. Styner, K. Rajamani, L.-P. Nolte, G. Zsemlye, G. Székely, C. Taylor, and R. Davies, "Evaluation of 3D Correspondence Methods for Model Building," *Inform Process Med Imag*, vol. 2732, pp. 63–75, 2003.

[SS99] D. E. T. Shepherd and B. B. Seedhom, "Thickness of human articular cartilage in joints of the lower limb," *Ann Rheum Dis*, vol. 58, no. 1, pp. 27–34, 1999.

[SS00] S. Stoev and W. Straβer, "Extracting regions of interest applying a local watershed transformation," in *Proc. Conf. Visualization*. IEEE Computer Society, 2000, pp. 21–28.

[SS05] J. Schäfer and K. Strimmer, "A Shrinkage Approach to Large-Scale Covariance Matrix Estimation and Implications for Functional Genomics," *Statistical Applications in Genetics and Molecular Biology*, vol. 4, no. 1, 2005.

[SSC+09] J. Schmid, A. Sandholm, F. Chung, D. Thalmann, H. Delingette, and N. Magnenat-Thalmann, "Musculoskeletal simulation model generation from MRI datasets and motion capture data," in *Recent Advances in the 3D Physiological Human*. Springer-Verlag, 2009, pp. 3–19.

[SSG+03] I. Schiffer, W. Schreiber, R. Graf, E. Schreiber, D. Jung, D. Rose, M. Hehn, S. Gebhard, J. Sagemüller, H. Spieß *et al.*, "No influence of magnetic fields on cell cycle progression using conditions relevant for patients during MRI," *Bioelectromagnetics*, vol. 24, no. 4, pp. 241–250, 2003.

[SSL06a] K. Sjöstrand, M. Stegmann, and R. Larsen, "Sparse principal component analysis in medical shape modeling," in *Proc. SPIE Medical imaging*, vol. 6144. SPIE, 2006, pp. 1579–1590.

[SSL06b] M. Stegmann, K. Sjöstrand, and R. Larsen, "Sparse modeling of landmark and texture variability using the orthomax criterion," in *Proc. SPIE*, vol. 6144. SPIE, 2006, pp. 485–496.

[SSM98] B. Schölkopf, A. Smola, and K. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Comput*, vol. 10, no. 5, pp. 1299–1319, 1998.

[SSP+10] A. Sandholm, C. Schwartz, N. Pronost, M. de Zee, M. Voigt, and D. Thalmann, "Evaluation of a geometry-based knee joint compared to a planar knee joint," *Vis Comput*, vol. DOI:10.1007/s00371-010-0538-7, 2010.

[SSWC88] P. K. Sahoo, S. Soltani, A. K. Wong, and Y. C. Chen, "A survey of thresholding techniques," *Comput Vis Graph Image Process*, vol. 41, no. 2, pp. 233–260, 1988.

[ST92] R. Szeliski and D. Tonnesen, "Surface modeling with oriented particle systems," *Computer Graphics*, vol. 26, no. 2, pp. 185–194, 1992.

[STA98] G. Subsol, J. Thirion, and N. Ayache, "A scheme for automatically building three-dimensional morphometric anatomical atlases: application to a skull atlas," *Med Image Anal*, vol. 2, no. 1, pp. 37–60, 1998.

[Stä05] A. Stäbler, "Bone Marrow Disorders," in *Musculoskeletal Diseases: Diagnostic Imaging and Interventional Techniques*, C. von Schulthess, G.K. Zollikofer, Ed. Springer, 2005, pp. 73–82.

[STCK03] T. Sebastian, H. Tek, J. Crisco, and B. Kimia, "Segmentation of carpal bones from CT images using skeletally coupled deformable models," *Med Image Anal*, vol. 7, no. 1, pp. 21–45, 2003.

[Ste02] M. Stegmann, B., "Analysis and segmentation of face images using point annotations and linear subspace techniques," Technical University of Denmark, Tech. Rep. IMM-REP-2002-22, 2002.

[Ste04] M. B. Stegmann, "Generative interpretation of medical images," Ph.D. dissertation, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 2004.

[STZ06] Y. Sun, E. Teo, and Q. Zhang, "Discussions of knee joint segmentation," in *Proc. ICBPE*. IEEE Computer Society, 2006.

[Sun89] P. Sundberg, "Shape and size-constrained principal components analysis," *Syst Zool*, vol. 38, no. 2, pp. 166–168, 1989.

[SUP+09] J. Santner, M. Unger, T. Pock, C. Leistner, A. Saffari, and H. Bischof, "Interactive Texture Segmentation using Random Forests and Total Variation," in *Proc. BMVC*. London, UK: BMVA Press, 2009.

[SVG02] J. Snel, H. Venema, and C. Grimbergen, "Deformable triangular surfaces using fast 1-D radial Lagrangian dynamics-segmentation of 3-D MR and CT images of the wrist," *IEEE Trans Med Imag*, vol. 21, no. 8, pp. 888–903, 2002.

[SWCHPC+05] F. Simmross-Wattenberg, N. Carranza-Herrezuelo, C. Palacios-Camarero, P. Casaseca-de-la Higuera, M. Martín-Fernández, S. Aja-Fernández, J. Ruiz-Alzola, C. Westin, and C. Alberola-López, "Group-Slicer: A collaborative extension of 3D-Slicer," *J Biomed Informat*, vol. 38, no. 6, pp. 431–442, 2005.

[SWND07] D. Shreiner, M. Woo, J. Neider, and T. Davis, *OpenGL Programming Guide: The official Guide to Learning OpenGL, Version 2.1*, 6th ed. Addison-Wesley Professional, 2007.

[SXESG07] M. Styner, S. Xu, M. El-Sayed, and G. Gerig, "Correspondence evaluation in local shape analysis and structural subdivision," in *Proc. ISBI*. IEEE Computer Society, 2007, pp. 1192–1195.

[TAD03] D. Thelen, F. Anderson, and S. Delp, "Generating dynamic simulations of movement using computed muscle control," *J Biomech*, vol. 36, no. 3, pp. 321–328, 2003.

[TB99] M. Tipping and C. Bishop, "Probabilistic principal component analysis," *J Roy Stat Soc B*, vol. 61, no. 3, pp. 611–622, 1999.

[TBA+09] P. Tresadern, H. Bhaskar, S. Adeshina, C. Taylor, and T. Cootes, "Combining Local and Global Shape Models for Deformable Object Matching," in *Web Proc. BMVC*. BMVA Press, 2009.

[TC05] K. Tan and S. Chen, "Adaptively weighted sub-pattern pca for face recognition," *Neurocomputing*, vol. 64, pp. 505–511, 2005.

[TE05] E. Tejada and T. Ertl, "Large steps in GPU-based deformable bodies simulation," *Simulat Model Pract Theor*, vol. 13, no. 8, pp. 703–715, 2005.

[TEE+08] M. Teßmann, C. Eisenacher, F. Enders, M. Stamminger, and P. Hastreiter, "GPU Accelerated Normalized Mutual Information and B-Spline Transformation," in *Proc. Eurographics Workshop VCBM*, C. Botha, G. Kindlmann, W. Niessen, and B. Preim, Eds. Delft, The Netherlands: Eurographics Association, 2008, pp. 117–124.

[Thi98] J. Thirion, "Image matching as a diffusion process: an analogy with Maxwell's demons," *Med Image Anal*, vol. 2, no. 3, pp. 243–260, 1998.

[Tho03] H. Thodberg, "Minimum description length shape and appearance models," in *Proc. IPMI*, vol. LNCS 2732. Springer, 2003, pp. 51–62.

[TICB09] C. Tejos, P. Irarrazaval, and A. Cárdenas-Blanco, "Simplex Mesh Diffusion Snakes: Integrating 2D and 3D Deformable Models and Statistical Shape Knowledge in a Variational Framework," *Int J Comput Vis*, vol. 85, no. 1, pp. 19–34, 2009.

[TKLL06] T. Tölli, J. Koikkalainen, K. Lauerma, and J. Lötjönen, "Artificially enlarged training set in image segmentation," in *Proc. MICCAI*, R. Larsen, M. Nielsen, and J. Sporring, Eds., vol. LNCS 4190. Springer, 2006, pp. 75–82.

[TL94] G. Turk and M. Levoy, "Zippered Polygon Meshes from Range Images," in *Proc. SIGGRAPH*. ACM Press, 1994, pp. 311–318.

[TM91] D. Terzopoulos and D. Metaxas, "Dynamic 3D models with local and global deformations: Deformable superquadrics," *IEEE Trans Pattern Anal Mach Intell*, vol. 13, no. 7, pp. 703–714, 1991.

[TMM78] C. Tipton, R. Matthes, and R. Martin, "Influence of age and sex on the strength of bone-ligament junctions in knee joints of rats," *J Bone Joint Surg*, vol. 60, no. 2, p. 230, 1978.

[TO03] H. Thodberg and H. Olafsdottir, "Adding curvature to minimum description length shape models," in *Proc. BMVC*, vol. 5. BMVA Press, 2003, pp. 14–16.

[Toh02] J. Tohka, "Surface Extraction from Volumetric Images Using Deformable Meshes: A Comparative Study," in *Proc. ECCV*, vol. LNCS 2352. Springer, 2002, pp. 350–364.

[TP91] M. Turk and A. Pentland, "Face recognition using eigenfaces," in *Proc. CVPR*, vol. 591. IEEE Computer Society, 1991, pp. 586–591.

[TPBF87] D. Terzopoulos, J. Platt, A. Barr, and K. Fleischert, "Elastically Deformable Models," *ACM Comput Graph*, vol. 21, no. 4, pp. 205–214, 1987.

[TSB+05] J. Teran, E. Sifakis, S. S. Blemker, V. Ng-Thow-Hing, C. Lau, and R. Fedkiw, "Creating and simulating skeletal muscle from the visible human data set," *IEEE Trans Visual Comput Graph*, vol. 11, no. 3, pp. 317–328, 2005.

[TT97] P. Thompson and A. Toga, "Detection, visualization and animation of abnormal anatomic structure with a deformable probabilistic brain atlas based on random vector field transformations," *Med Image Anal*, vol. 1, no. 4, pp. 271–294, 1997.

[Tvr06] J. Tvrdík, "Differential Evolution: Competitive Setting of Control Parameters," in *Proc. IMCSIT*. PTI, 2006, pp. 207–213.

[TW88] D. Terzopoulos and A. Witkin, "Deformable models," *IEEE Comput Graph Appl Mag*, vol. 8, no. 6, pp. 41–51, 1988.

[TWK88] D. Terzopoulos, A. Witkin, and M. Kass, "Constraints on deformable models: Recovering 3D shape and nonrigid motion," *Artificial Intelligence*, vol. 36, no. 1, pp. 91–123, 1988.

[TWT+04] A. Tsai, W. Wells, C. Tempany, E. Grimson, and A. Willsky, "Mutual information in coupled multi-shape model for medical image segmentation," *Med Image Anal*, vol. 8, no. 4, pp. 429–445, 2004.

[TYW+03] A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, W. E. Grimson, and A. Willsky, "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE Trans Med Imag*, vol. 22, no. 2, pp. 137–154, 2003.

[ÜFS+03] M. Üzümcü, A. Frangi, M. Sonka, J. Reiber, and B. Lelieveldt, "ICA vs. PCA Active Appearance Models: Application to Cardiac MR Segmentation," in *Proc. MICCAI*, vol. LNCS 2878. Springer, 2003, pp. 451–458.

[Ume91] S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns," *IEEE Trans Pattern Anal Mach Intell*, vol. 13, no. 4, pp. 376–380, 1991.

[UPB08] M. Unger, T. Pock, and H. Bischof, "Continuous Globally Optimal Image Segmentation with Local Constraints," in *Proc. CVWW*. PRIP, 2008.

[UPT+08] M. Unger, T. Pock, W. Trobin, D. Cremers, and H. Bischof, "Tvseg - interactive total variation based image segmentation," in *Proc. BMVC*. BMVA Press, 2008.

[US05] G. Unal and G. Slabaugh, "Coupled PDEs for non-rigid registration and segmentation," in *Proc. CVPR*, vol. 1. IEEE Computer Society, 2005, p. 168.

[VADF+06] H. Van Assen, M. Danilouchkine, A. Frangi, S. Ordás, J. Westenberg, J. Reiber, and B. Lelieveldt, "SPASM: A 3D-ASM for segmentation of sparse and arbitrarily oriented cardiac MRI data," *Med Image Anal*, vol. 10, no. 2, pp. 286–303, 2006.

[VBA+04] F. M. Vos, P. W. d. Bruin, J. G. M. Aubel, G. J. Streekstra, M. Maas, L. J. v. Vliet, and A. M. Vossepoel, "A Statistical Shape Model without Using Landmarks," in *Proc. ICPR*, vol. 3. Washington, DC, USA: IEEE Computer Society, 2004, pp. 714–717.

[Ver67] L. Verlet, "Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules," *Physical Review*, vol. 159, no. 1, pp. 98–103, 1967.

[VGDBLV03] B. Van Ginneken, M. De Bruijne, M. Loog, and M. Viergever, "Interactive shape models," in *Proc. SPIE Medical Imaging*, vol. 5032. SPIE, 2003, pp. 1206–1216.

[VGFS+02] B. Van Ginneken, A. Frangi, J. Staal, B. ter Haar Romeny, and M. Viergever, "Active shape model segmentation with optimal features," *IEEE Trans Med Imag*, vol. 21, no. 8, pp. 924–933, 2002.

[VLQ99] S. Valle, W. Li, and S. Qin, "Selection of the Number of Principal Components: The Variance of the Reconstruction Error Criterion with a Comparison to Other Methods," *Ind Eng Chem Res*, vol. 38, no. 11, pp. 4389–4401, 1999.

[VMT05] P. Volino and N. Magnenat-Thalmann, "Implicit midpoint integration and adaptive damping for efficient cloth simulation: Collision detection and deformable objects," *Comput Animat Virtual Worlds*, vol. 16, no. 3-4, pp. 163–175, 2005.

[VMTF09] P. Volino, N. Magnenat-Thalmann, and F. Faure, "A simple approach to nonlinear tensile stiffness for accurate cloth simulation," *ACM Trans Graph*, vol. 28, no. 4, pp. 105–116, 2009.

[VT00] P. Volino and N. M. Thalmann, "Implementing Fast Cloth Simulation with Collision Response," in *Proc. CGI*. Washington, DC, USA: IEEE Computer Society, 2000, pp. 257–266.

[VTL+04] T. Vrtovec, D. Tomaževic, B. Likar, L. Travnik, and F. Pernuš, "Automated Construction of 3D Statistical Shape Models," *Image Anal Stereol*, vol. 23, pp. 111–120, 2004.

[vVBV02] M. van der Glas, F. Vos, C. Botha, and A. Vossepoel, "Determination of position and radius of ball joints," in *Proc. SPIE Medical Imaging*, vol. 4684. SPIE, 2002, pp. 1571–1577.

[VW97] P. Viola and W. Wells, "Alignment by maximization of mutual information," *Int J Comput Vis*, vol. 24, no. 2, pp. 137–154, 1997.

[War96] S. Warfield, "Fast *k*-NN Classification for Multichannel Image Data," *Pattern Recogn Lett*, vol. 17, no. 7, pp. 713–721, 1996.

[WBK+09] J. Walters, V. Balu, S. Kompalli, V. Chaudhary, and N. Buffalo, "Evaluating the use of GPUs in liver image segmentation and HMMER database searches," in *Proc. IPDPS*. IEEE Computer Society, 2009, pp. 1–12.

[WBPR+04] K. Walker-Bone, K. Palmer, I. Reading, D. Coggon, and C. Cooper, "Prevalence and impact of musculoskeletal disorders of the upper limb in the general population," *Arthritis Care Res*, vol. 51, no. 4, pp. 642–651, 2004.

[WCP06] R. Wootton, J. Craig, and V. Patterson, *Introduction to telemedicine*. The Royal Society of Medicine Press Ltd, Second edition, London, 2006.

[WEG87] S. Wold, K. Esbensen, and P. Geladi, "Principal Component Analysis," *Chemometr Intell Lab Syst*, vol. 2, pp. 37–52, 1987.

[WGE06] L. Wang, M. Greenspan, and R. Ellis, "Validation of Bone Segmentation and Improved 3-D Registration Using Contour Coherency in CT Data," *IEEE Trans Med Imag*, vol. 25, no. 3, pp. 324–334, 2006.

[WGH+98] R. Woods, S. Grafton, C. Holmes, S. Cherry, and J. Mazziotta, "Automated image registration: I. General methods and intrasubject, intramodality validation," *J Comput Assist Tomo*, vol. 22, no. 1, pp. 141–154, 1998.

[WHA+91] S. Woo, J. Hollis, D. Adams, R. Lyon, and S. Takai, "Tensile properties of the human femur-anterior cruciate ligament-tibia complex," *Am J Sports Med*, vol. 19, no. 3, p. 217, 1991.

[Whi98] R. Whitaker, "A level-set approach to 3D reconstruction from range data," *Int J Comput Vis*, vol. 29, no. 3, pp. 203–231, 1998.

[Wil05] A. Wilson, "Radiology of Hand and Wrist Injuries," in *Musculoskeletal Diseases: Diagnostic Imaging and Interventional Techniques*, C. von Schulthess, G.K. Zollikofer, Ed. Springer, 2005, pp. 13–16.

[Wir07] O. Wirjadi, "Survey of 3D image segmentation methods," Fraunhofer ITWM, Tech. Rep., 2007.

[WK07] D. Withey and Z. Koles, "Medical image segmentation: Methods and software," in *Proc. NFSI-ICFBI*. IEEE Computer Society, 2007, pp. 140–143.

[WKR04] S. Wang, T. Kubota, and T. Richardson, "Shape correspondence through landmark sliding," in *Proc. CVPR*. IEEE Computer Society, 2004, pp. 143–150.

[WN03] P. Wyatt and J. Noble, "MAP MRF joint segmentation and registration of medical images," *Med Image Anal*, vol. 7, no. 4, pp. 539–552, 2003.

[WOG+86] S. Woo, C. Orlando, M. Gomez, C. Frank, and W. Akeson, "Tensile properties of the medial collateral ligament as a function of age," *J Orthop Res*, vol. 4, no. 2, pp. 133–141, 1986.

[WPS00] Y. Wang, B. S. Peterson, and L. H. Staib, "Shape-based 3d surface correspondence using geodesics and local geometry," in *Proc. CVPR*, vol. 2. Los Alamitos, CA, USA: IEEE Computer Society, 2000, p. 2644.

[WPS03] Y. Wang, B. S. Peterson, and L. H. Staib, "3D Brain surface matching based on geodesics and local geometry," *Comput Vis Image Understand*, vol. 89, no. 2-3, pp. 252–271, 2003.

[WQL05] S. Wang, F. Qi, and H. Li, "Add cartesian differential invariants to minimum description length shape models," in *Proc. IAPR MVA*, 2005, pp. 253–256.

[WS00] Y. Wang and L. Staib, "Boundary finding with prior shape and smoothness models," *IEEE Trans Pattern Anal Mach Intell*, vol. 22, no. 7, pp. 738–743, 2000.

[WTMT95] Y. Wu, D. Thalmann, and N. Magnenat-Thalmann, "Deformable surfaces using physically based particle

systems," in *Proc. CGI*, vol. 95. IEEE Computer Society, 1995, pp. 205–215.

[WV05] F. Wang and B. Vemuri, "Simultaneous Registration and Segmentation of Anatomical Structures from Brain MRI," in *Proc. MICCAI*, vol. LNCS 3749. Springer, 2005, pp. 17–25.

[XP97] C. Xu and J. Prince, "Gradient vector flow: A new external force for snakes," in *Proc. CVPR*. IEEE Computer Society, 1997, pp. 66–71.

[YAGE05] W. Yao, P. Abolmaesumi, M. Greenspan, and R. Ellis, "An estimation/correction algorithm for detecting bone edges in ct images," *IEEE Trans Med Imag*, vol. 24, no. 8, pp. 997–1010, 2005.

[YB89] S. Yanowitz and A. Bruckstein, "A new method for image segmentation," *Comput Vis Graph Image Process*, vol. 46, no. 1, pp. 82–95, 1989.

[YCGMMT03] L. Yahia-Cherif, B. Gilles, L. Moccozet, and N. Magnenat-Thalmann, "Individualized bone modeling from mri: Application to the human hip," in *Proc. CARS*. Springer, 2003, p. 1306.

[YDB02] W. Yambor, B. Draper, and R. Beveridge, "Analyzing PCA-based Face Recognition Algorithms: Eigenvector Selection and Distance Measures," in *Empirical Evaluation Methods in Computer Vision*, H. Christensen and J. Phillips, Eds. World Scientific Press, 2002, pp. 39–60.

[YOT+09] F. Yokota, T. Okada, M. Takao, N. Sugano, Y. Tada, and Y. Sato, "Automated Segmentation of the Femur and Pelvis from 3D CT Data of Diseased Hip Using Hierarchical Statistical Shape Model of Joint Structure," in *Proc. MICCAI*, G.-Z. e. a. Yang, Ed., vol. LNCS 5762. Springer-Verlag, 2009, pp. 811–818.

[Ypm95] T. J. Ypma, "Historical development of the newton–raphson method," *SIAM Review*, vol. 37, no. 4, pp. 531–551, 1995.

[YWT+07] T. Yamazaki, T. Watanabe, T. Tomita, K. Sugamoto, M. Ogasawara, Y. Sato, H. Yoshikawa, and S. Tamura, "3D kinematics of normal knee using X-ray fluoroscopy and CT images," in *Proc. World Congress Medical Physics Biomedical Engineering*. Springer, 2007, pp. 2793–2796.

[YZG05] P. A. Yushkevich, H. Zhang, and J. C. Gee, "Statistical modeling of shape and appearance using the continuous medial representation," in *Proc. MICCAI*, vol. LNCS 3750. Springer-Verlag, 2005, pp. 725–732.

[YZK01] A. Yezzi, L. Zollei, and T. Kapur, "A variational framework for joint segmentation and registration," in *Proc. MMBIA*. IEEE Computer Society, 2001, pp. 44–51.

[Zah02] D. Zaharie, "Critical values for the control parameters of differential evolution algorithms," in *Proc. MENDEL*, 2002, pp. 62–67.

[ZF03] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image Vis Comput*, vol. 21, no. 11, pp. 977–1000, 2003.

[ZFG08] H. Zhang, J. Fritts, and S. Goldman, "Image segmentation evaluation: A survey of unsupervised methods," *Comput Vis Image Understand*, vol. 110, no. 2, pp. 260–280, 2008.

[Zha08] Z. Zhang, "Adaptive region intensity based rigid ultrasound and CT image registration," in *Proc. CVPR*. IEEE Computer Society, 2008, pp. 1–7.

[ZHG+08] D. Zikic, M. Hansen, B. Glocker, A. Khamene, R. Larsen, and N. Navab, "Computing minimal deformations: application to construction of statistical shape models," in *Proc. CVPR*. IEEE Computer Society, 2008, pp. 1–8.

[ZHR06] L. Zhang, E. Hoffman, and J. Reinhardt, "Atlas-driven lung lobe segmentation in volumetric X-ray CT images," *IEEE Trans Med Imag*, vol. 25, no. 1, pp. 1–16, 2006.

[ZHT06] H. Zou, T. Hastie, and R. Tibshirani, "Sparse principal component analysis," *J Comput Graph Stat*, vol. 15, no. 2, pp. 265–286, 2006.

[ZL07] K. Zielinski and R. Laur, "Stopping Criteria for a Constrained Single-Objective Particle Swarm Optimization Algorithm," *Informatica*, vol. 31, pp. 51–59, 2007.

[ZSG10] G. Zheng, S. Schumann, and M. A. González Ballester., "An integrated approach for reconstructing a surface model of the proximal femur from sparse input data and a multi-resolution point distribution model: an in vitro study," *Int J Comput Assist Radiol Surg*, vol. 5, no. 1, pp. 99–107, 2010.

[ZSH+00] J. Zhang, J. Stahl, H. Huang, X. Zhou, S. Lou, and K. Song, "Real-Time Teleconsultation with High-Resolution and Large-Volume Medical Images for Collaborative Healthcare," *IEEE Trans Inform Tech Biomed*, vol. 4, no. 2, pp. 178–185, 2000.

[ZSN+04] R. A. Zoroofi, Y. Sato, T. Nishii, N. Sugano, H. Yoshikawa, and S. Tamura, "Automated segmentation of necrotic femoral head from 3D MR data," *Comput Med Imag Graph*, vol. 28, no. 5, pp. 267–278, 2004.

[ZT05] Z. Zhao and E. Teoh, "A novel framework for automated 3D PDM construction using deformable models," in *Proc. SPIE*, vol. 5747. SPIE, 2005, pp. 303–314.

[ZTK05] K. Zhang, M. Tang, and J. T. Kwok, "Applying neighborhood consistency for fast clustering and kernel density estimation," in *Proc. CVPR*, vol. 2. IEEE Computer Society, 2005, pp. 1001–1007.