ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

MINES
ParisTech

École doctorale n°84:
Sciences et technologies de l'information et de la communication

Doctorat européen ParisTech

THÈSE

pour obtenir le grade de docteur délivré par

l'École nationale supérieure des mines de Paris

Spécialité Informatique temps-réel, robotique et automatique

présentée et soutenue publiquement par

François CHUNG CHOONG LON

17 Janvier 2011

Regional appearance modeling for deformable model-based image segmentation

~ ~ ~

Modélisation de l'apparence de régions pour la segmentation d'images basée modèle

Directeur de thèse: **Hervé Delingette**
Co-encadrement de la thèse: **Nicholas Ayache**

Jury

Luc Soler, IRCAD - Institut de Recherche contre les Cancers de l'Appareil Digestif	Rapporteur
Andrew Todd-Pokropek, UCL - University College London	Rapporteur
Nadia Magnenat-Thalmann, UNIGE - Université de Genève	Examinateur
Daniel Thalmann, EPFL - École Polytechnique Fédérale de Lausanne	Examinateur
Hervé Delingette, INRIA - Institut National de Recherche en Informatique et Automatique	Directeur
Nicholas Ayache, INRIA - Institut National de Recherche en Informatique et Automatique	Co-directeur

THÈSE

MINES ParisTech
Centre de Mathématiques Appliquées
Rue Claude Daunesse B.P. 207, 06904 Sophia Antipolis Cedex, France

Regional appearance modeling
for deformable model-based image segmentation

Abstract

This thesis presents a novel appearance prior for model-based image segmentation. This appearance prior, denoted as *Multimodal Prior Appearance Model (MPAM)*, is built upon an EM clustering of intensity profiles with model order selection to automatically select the number of profile classes. Unlike classical PCA-based approaches, the clustering is considered as regional because intensity profiles are classified for each mesh and not for each vertex.

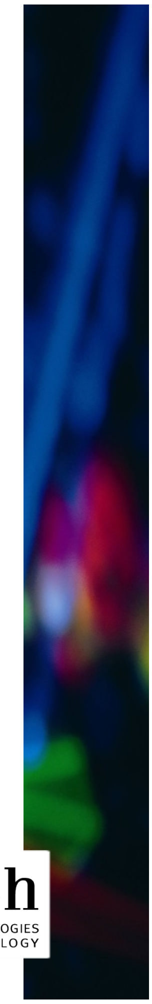
First, we explain how to build a MPAM from a training set of meshes and images. The clustering of intensity profiles and the determination of the number of appearance regions by a novel model order selection criterion are explained. A spatial regularization approach to spatially smooth the clustering of profiles is presented and the projection of the appearance information from each dataset on a reference mesh is described.

Second, we present a boosted clustering based on spectral clustering, which optimizes the clustering of profiles for segmentation purposes. The representation of the similarity between data points in the spectral space is explained. Comparative results on liver profiles from CT images show that our approach outperforms PCA-based appearance models.

Finally, we present methods for the analysis of lower limb structures from MR images. In a first part, our technique to create subject-specific models for kinematic simulations of lower limbs is described. In a second part, the performance of statistical models is compared in the context of lower limb bones segmentation when only a small number of datasets is available for training.

Keywords

Appearance modeling, unsupervised clustering, model-based image segmentation, medical imaging, liver, lower limbs.



Modélisation de l'apparence de régions pour la segmentation d'images basée modèle

Résumé

Cette thèse est consacrée à un nouveau modèle d'apparence pour la segmentation d'images basée modèle. Ce modèle, dénommé *Multimodal Prior Appearance Model (MPAM)*, est construit à partir d'une classification EM de profils d'intensité combinée avec une méthode automatique pour déterminer le nombre de classes. Contrairement aux approches classiques basées ACP, les profils d'intensité sont classifiés pour chaque maillage et non pour chaque sommet.

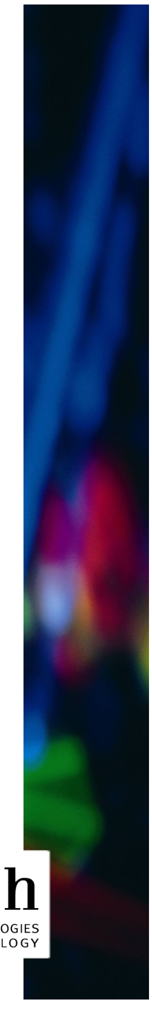
Tout d'abord, nous décrivons la construction du MPAM à partir d'un ensemble de maillages et d'images. La classification de profils d'intensité et la détermination du nombre de régions par un nouveau critère de sélection sont expliquées. Une régularisation spatiale pour lisser la classification est présentée et la projection de l'information d'apparence sur un maillage de référence est décrite.

Ensuite, nous présentons une classification de type spectrale dont le but est d'optimiser la classification des profils pour la segmentation. La représentation de la similitude entre points de données dans l'espace spectral est expliquée. Des résultats comparatifs sur des profils d'intensité du foie à partir d'images tomodensitométriques montrent que notre approche surpasse les modèles basés ACP.

Finalement, nous présentons des méthodes d'analyse pour les structures des membres inférieurs à partir d'images IRM. D'abord, notre technique pour créer des modèles spécifiques aux sujets pour des simulations cinématiques des membres inférieurs est décrite. Puis, la performance de modèles statistiques est comparée dans un contexte de segmentation des os lorsqu'un faible ensemble de données est disponible.

Mots clés

Modélisation de l'apparence, classification non supervisée, segmentation d'images basée modèle, imagerie médicale, foie, membres inférieurs.



MINES ParisTech

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

Acknowledgments

Talking about the life after the Ph.D., my little sister once told me, in a very philosophical way: "Anyway, after the Ph.D. thesis, what you'll really remember would be the people with whom you worked". She could not be more right. But something I know for sure, is that I'll also remember the many people I met and had the pleasure of getting to know all this time. Lots of people have helped me during my Ph.D. thesis, whether through their knowledge, availability, or (more importantly) friendship. I really believe those people are the most important contribution to this Ph.D. thesis, because they had (and will keep having) the biggest impact on my life.

Girona My Ph.D. journey began in Girona, Catalunya (Spain) where I started a Ph.D. in Computer Vision and Robotics in September 2005 at the Universitat de Girona. First of all, I would like to thank Olivier Delaunoy with whom I decided to start this Spanish experience and Rafael Garcia who welcomed me within the VICOROB group. From the Underwater Vision Lab, I am grateful to Armagan Elibol (el turco peligroso), Nuno Gracias, Jordi Ferrer (Linux lover), Marina Kudinova, Tudor Nicosevici, Ricard Prados (windows lover) and Josep Quintana (gràcies per la teva ajuda en el SAUCE'06). Special thanks to Carles Matabosch (el meu company de oficina). Thanks also to Birgit Moëller, Thierry Molinier, Jonathan Horgan and Emilien Bouley, who were visitors I had the pleasure to meet there. From the VICOROB group, thanks to Elisabet Batlle, Joan Batlle, Marc Carreras, Emili Hernández, Jordi Freixenet, Xavier Lladó, Lászlo Neumann, Radu Orghidan, Narcis Palomeras, David Ribas, Pere Ridao, Joaquim Salvi, Anna Bosch, Robert Martí and, last but not least, Arnau Oliver (lu secretari de porres).

My life at Girona was not only work-based. Apart from Ph.D. work, I learnt Catalan and Spanish. I am grateful to Raquel Rivera (la meva profesora de catala) and Susanna Dominguez (mi profesora de espanol) for teaching me these two languages. Special thanks to Elvina Chavtarayeva (del tàndem lingüístic). From the Spanish course, I'll remember the nice moments spent with some students, namely Silvia Fortuzzi (la chica guapa), Paul Kneppers (me casaré con mi bicicleta) and Xandra Pankratz (mirad mi video!) with whom we used to watch movies in Spanish (la película en la casa de Paul). During my stay in Girona, I met wonderful people I'd like to pay tribute, such as Albert Xuriguera (el meu company de pis), Claudia Dragoste, Patricia Gaspar, Silvana Aciar, Stefanie Heinz, Lidia Portal (y Ingrid). Some of them were Erasmus people, like Aurore Dupin and Johan Rutten. However, my lifestyle could not continue in calm waters because I had no grant to fund my Ph.D. With the help of Xevi Cufí, to whom I owe a great amount of gratitude (mai oblidaré la teva gran ajuda en la meva recerca de beques), I applied for Ph.D. fellowships in Europe.

Sophia-Antipolis In June 2007, my Ph.D. journey took a drastic change of direction. I was accepted for a Ph.D. fellowship within the Asclepios Team, at the INRIA Sophia-Antipolis, France. First of all, I would like to thank Nicholas Ayache for welcoming me within the team and Hervé Delingette, my supervisor, for guiding me all along my Ph.D. Special thanks to my Ph.D. jury who went all the way to the Côte d'Azur for attending my defense. I am grateful to Daniel Thalmann, president of the jury, and to Andrew Todd-Pokropek and Luc Soler, both reviewers, for having spent time to read, comment and correct my Ph.D. thesis.

From the Asclepios Team, I would like to thank Daniel Barbeau (Mr Muse), Florence Billet (je n'oublierai pas nos longues discussions sur la thèse), Benoît Bleuzé (Mr Geek), Marine Breuilly, Maxime Boucher, Aurélie Canale (apéroooo), Olivier Clatz, Olivier Commowick, Jean Cousty, Jimena Costa, Florence Dru, Guillaume Dugas-Phocion, Stanley Durrleman, Romain Fernandez (Mr Hamac), Pierre Fillard, Vincent Garcia (Mr thé), Ezequiel Geremia (merci de m'avoir hébergé en fin de thèse), Pascal Girard, Vladlena Gobunova, François Hebert, Tobias Heimann (Spanish-speaking 3DAH mate), Heike Hufnagel, Michael Knopke, Ender Konukoglu, Hans Lamecker, Damien Lepiller (merci de m'avoir fait découvrir Mozart), Hervé Lombaert (Brasilian-Canadian dude), Yonni Levy, Marco Lorenzi (Guitar Hero), Grégoire Malandain (XS bassist), Tommaso Mansi (compagnon d'infortune aux Mines), Kristin Mcleod (keep rocking on drums!), Bjoern Menze, Xavier Pennec, Erik Pernod (Mr Sofa), Jean-Marc Peyrat (Mr Surf), Adityo Prakosa, Liliane Ramus (merci pour ces longues discussions au labo), Jatin Relan, Stephan Schmitt, Maxime Sermesant, Alexander Schmidt-Richberg, Stefan Sommer, John Stark, Erin Stretton (ski mate), Andrew Sweet, Nicolas Toussaint (Mr CardioViz3D), Ken Wong (South Park rules!) and Thomas Yeo. Special thanks to Jean-Christophe Souplet, Stéphanie Marchesseau and Hugo Talbot, who had to endure me as an office mate, Isabelle Strobant, for her daily help with administrative stuff, and to Sébastien Novellas, for his advices as a Medical Doctor.

I had also the opportunity to meet very nice folks through INRIA (people). Warm greetings to Christophe Andre, Amélie Anglade, Elena Bukina, Xavier Cremaschi, Olivier Davidau, Alexandre Fournier, Rémi Hetmanski, Emmanuel Hutter, Stevens Le Blond, Geoffroy Le Prado, Barbara Mannino, Rasmus Paulsen, Radu Stefanescu (et Mimi!) and Sandrine Venet. From INRIA, I would like to thank Wassim Abou-jaoudé, Marc Barret, Luc Hogie, David Rey and Fabien Viale, with whom I enjoyed playing and/or talking music, Sapna Nundloll, my fellow-countrywoman, Minh Duc Hua, Mandar Harshe and Pierre Masci. From Mines ParisTech, I am thankful to Jean-Paul Marmorat and Dominique Micollier for their help with administrative stuff related to the Ph.D. program. Special thanks to Pierre Massip, with whom I used to attend Ph.D. courses. From Nice, warm greetings to Eléonore Gueit, Ana Laura Fabiano, Nicolas Legrand, Tsai Li and Stéphanie Rennesson.

3DAH My Ph.D. thesis has been funded by the 3DAH project, a Marie Curie Research Training Network through which I had the chance to meet great people from different countries. I am grateful to my ESR fellows: Mitchelle Chen, Ioana Ciuciu, Jose Iglesias, Margit Mikula, Niels Nijdam (stomach on feet), Caroline Öhman, Anders Sandholm and Jerome Schmid, the ER involved in the project: Daniel Espino, Jinman Kim, Aggelos Liapis (thanks for all the drinks offered!), Nicolas Pronost and Cédric Schwartz, as well as the scientists in charge: Massimiliano Baleani (please consider...), Andrea Giachetti and Mark de Zee (expert in lower limb shaving). Thanks for the nice moments shared during IMC, internal meetings, summer school and other workshops in Zermatt. Special thanks to the people I've met during my two stays in Geneva: Lazhari Assassi, Maher Ben Moussa, Sylvain Chagué, Caecilia Charbonnier, Jody Hausmann, Bart Kevelham and Yan, and during my many MRI acquisitions in London: Peter Del-Manso, Tryphon Lambrou and Freddy Odille.

Belgium Many times during my Ph.D., I came back to Fatherland Belgium. One reason was music. So first, I would like to thank musicians with whom I had the pleasure to play music: Raphaël Halbardier, Guillaume Etievant, Sébastien Bessard, Martin Favresse, David Detienne and Michel Lesens. Thanks also to my Belgian friends, who always made me feel like home: Jonathan Lousse, Jérémie Dewals, Pauline Humblet, Mica Hallin, Gaëlle Wastiaux, Patrick Janssens, Youssef Ben Achir, Angelo Tortolani and Delia Samyn. Special thanks to my little family: Lolo, Nath, Greg, and my parents, to whom I owe the greatest amount of gratitude for their love, patience, endless support and encouragement since the beginning of my studies. This Ph.D. thesis is dedicated to them.

Thanks to all! Merci à tous! Gràcies a tots! Gracias a todos!

François Chung

“Il n’y a pas d’homme cultivé, il n’y a que des hommes qui se cultivent.”
Ferdinand Foch

Contents

CHAPTER 1

Introduction

Contents

1.1 Objectives

Extracting anatomical structures of interest (*e.g.* organs, bones and tissues) from medical images is still an important topic in the medical imaging community. This process, known as *segmentation*, may be performed in several ways. Using a prior information about the shape, explicit model-based image segmentation has been extensively used in the literature. A mesh is deformed by means of both *internal* and *external* forces. Internal forces aim at ensuring a smooth deformation of the mesh during segmentation while external forces attract the mesh to image-based features. In the literature, such features have been widely based on image boundaries (*i.e.* local features) or image regions (*i.e.* global features). A review on appearance models is available in [Heimann 2009].

For instance, statistical models of appearance such as Active Shape Models (ASM) are widely used [Cootes & Taylor 1994]. Intensity profiles are trained and both mean profile and principal modes of variation for each landmark are extracted in a PCA fashion. Gabor wavelets have been successfully used in face recognition and tracking applications [Daugman 1988]. Other local feature-based approaches worth mentioning are kNN-classifier and boosting methods [Li *et al.* 2004]. The most popular global feature-based approach are the Active Appearance Models (AAM), where all intensities from the inner region of the mesh are used to create a large feature vector. Like in ASM, a PCA is built, though on textures this time, and both mean of textures and modes of variation are extracted. To overcome computational time or memory issues that may arise due to the feature vector size, only parts of the inner region may be used around specific landmarks, *e.g.* using bounding boxes [Cristinacce & Cootes 2008].

Appearance models based on PCA, *e.g.* those based on intensity profiles, suffer from several limitations. First, they assume that points found homologous based on the global shape should also be homologous by their appearance, which is not true in general. This is even less the case in the presence of pathologies (*e.g.* liver tumors). Second, they suffer from over-fitting since the number of samples (*i.e.* the number of training datasets) is usually way too small given the dimension of the feature (*i.e.* the profile length).

In this thesis, we propose an alternative to PCA-based appearance models that is based on a multimodal prior. This appearance prior, denoted as *Multimodal Prior Appearance Model (MPAM)*, is built upon an EM clustering of intensity profiles with model order selection to automatically select the number of profile classes, or *modes*. The clustering is considered as *regional* because intensity profiles are classified for each mesh, and not for each vertex (*i.e.* over a population of meshes). With the MPAM, each point may be associated with several profile modes and each profile mode is estimated on each subject and not across subjects. To make sure that the MPAM is suitable for segmentation purposes, the clustering of profiles is optimized in an iterative way using a boosting procedure. All mesh instances are then projected on a common reference in which similar modes are possibly merged in an attempt to be more efficient and to reduce the complexity of the MPAM.

The overall objective of this thesis is to propose, discuss and assess methodological approaches to improve the accuracy of model-based image segmentation, both in terms of qualitative and quantitative results. These methodological approaches could be integrated in a more complete set of segmentation tools, suitable for a clinical use. Although our proposed methods are assessed on liver and lower limb datasets throughout this thesis, our proposed methods are generic and could thus be used on any other anatomical structures.

1.2 Manuscript organization

Chapter 2 describes the medical context in which this thesis has been done. We start this chapter with some anatomical descriptions. We focus on the main anatomical features of both liver and lower limbs, with an emphasis on bones and muscles. Then, we present the field of medical imaging and describe both Computed Tomography and Magnetic Resonance Imaging, as we use these two imaging techniques in the framework of this thesis.

Chapter 3 describes model-based image segmentation. First, we review the main techniques used in medical imaging such as manual segmentation, thresholding and atlas-based approaches. Then, we describe explicit deformable models. In this thesis, we are particularly interested in simplex meshes, from which we describe the specific topology, geometry, evolution and quality improvement. Finally, we explain how to initialize a deformable model into an image and how to deform it by means of internal and external forces.

Chapter 4 presents our Multimodal Prior Appearance Model (MPAM) as a new regional appearance model based on a classification of intensity profiles. First, we explain the concepts of unsupervised clustering and describe Fuzzy C-Means (FCM), Expectation-Maximization (EM) and Neighborhood EM (NEM) algorithms. Then, we explain how to build a MPAM from a training set of meshes and images. We explain how to cluster intensity profiles using the EM algorithm. More precisely, we describe how to initialize the algorithm, how to cope with missing data and how to regularize covariance matrices. We present model order selection as a way to find the number of clusters, and explain how this number may be determined automatically by a novel model order selection criterion, denoted as *Overlap Separation Index (OSI)*. We present a spatial regularization approach to spatially smooth the clustering of profiles using the NEM algorithm. Finally, we describe how the appearance information from each dataset (*i.e.* clustering of profiles) is projected on a reference mesh where each vertex has a probability to belong to several intensity profile classes, denoted as *modes*. The Jaccard index is used to control the final number of modes associated with the MPAM.

Chapter 5 presents a boosted clustering based on spectral clustering, which aims at optimizing the clustering of profiles associated with the MPAM for segmentation purposes. First, we describe spectral clustering. We explain how to represent the similarity between data points in the spectral space through spectral graphs, spectral functions and affinity matrices. Then, we explain how to extract top eigenvectors from affinity matrices, how to select them in an automatic fashion, how to classify them using the EM algorithm and how to extract EM parameters from this clustering performed in the spectral space. Finally, we present the boosted clustering algorithm. We describe the similarity measures used during local search and propose different versions of the algorithm, namely the single pass boosted clustering, the cascading boosted clustering and the cascading boosted clustering with hierarchical approach. We conclude this chapter with a discussion about the algorithm.

Chapter 6 uses a database of thirty-five liver meshes with four resolutions to test the different methods presented in this thesis. Results are described and analyzed in a pipeline approach, from the search for an optimal intensity profile length to the comparison of our MPAM with PCA-based Appearance Prior (PCAP) in terms of segmentation. We are interested in model order selection and compare different criteria from information theory with the eigengap heuristic from spectral clustering. We test different thresholds on the Jaccard index and study how modes are merged when increasing the number of datasets into the MPAM. We also test the different versions of the boosted clustering. Finally, segmentation performance of both MPAM and PCAP are assessed by optimizing both internal and external forces separately, and then in a complete procedure including mesh initialization.

Chapter 7 describes two collaborative works that have been done within the framework of the 3D Anatomical Human project (3DAH). In a first part, we present our technique to create subject-specific models for kinematic simulations. We describe the data acquisitions and propose some pre-processing methods to reduce noise and artifacts present in MR images. Then, we describe the different steps to create anatomical and musculoskeletal models. In a second part, we compare the performance of statistical models in the context of lower limb bones segmentation using MR images when only a small number of datasets is available for training. To do so, we test different shape and appearance priors.

Chapter 8 concludes this thesis by listing our contributions and by giving some perspectives for the future.

Chapter 2

Medical Context

Contents

2.1 Introduction

In clinical environments, radiologists have to analyze a huge amount of data with increasing complexity. Indeed, medical imaging is still an active topic in which physicians are continuously trying to improve image quality and resolution. In the medical image analysis community, these improvements help researchers in their quest to accurately extract vital information from image data for both clinical purposes (*e.g.* detection of diseases, diagnosis and therapy planning) and medical science (*e.g.* study of anatomy and physiology).

We start this chapter with some anatomical descriptions of both liver (section 2.2.1) and lower limbs (section 2.2.2), with an emphasis on bones (section 2.2.2.1) and muscles (section 2.2.2.2). Liver meshes are used to test our proposed Multimodal Prior Appearance Model (MPAM) in chapter 6. Lower limbs, which are the main focus of the 3D Anatomical Human project (3DAH), are analyzed in chapter 7.

Then, we present some medical imaging modalities (section 2.3). After a brief description (section 2.3.1), we review both Computed Tomography (section 2.3.2) and Magnetic Resonance Imaging (section 2.3.3) as we use these two imaging techniques in the framework of this thesis.

2.2 Anatomy

2.2.1 Liver

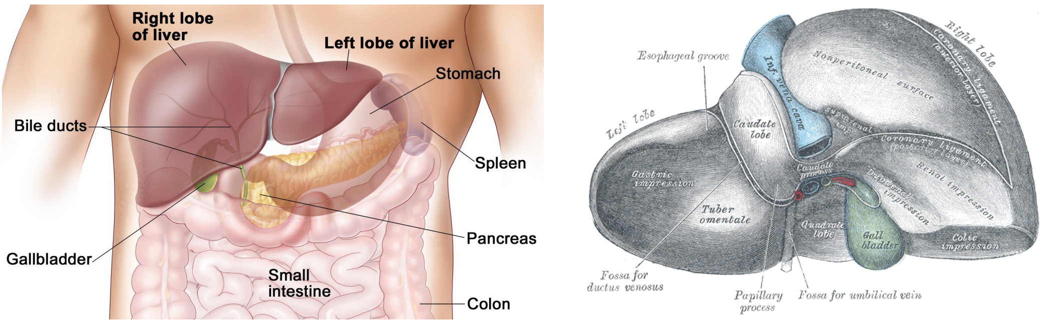
Since playing a major role in metabolism, the liver is considered as a vital organ. Constituting 2% of body weight (*i.e.* an average of 1.5 kg), the liver is the largest human organ and the organ that performs the largest number of chemical transformations. The medical discipline that incorporates the study of liver and its disorders is called *hepatology*.

The liver lies in the thoracic region of the abdomen, below the diaphragm and close to both the stomach and esophagus (see Figure 2.1, left). Main inner structure of the liver is the parenchyma. Main neighboring structures are the diaphragm, the stomach, the gallbladder and the thoracic cage.

The liver has several vital functions:

1. Nutritional function
    - Metabolism of carbohydrate (*e.g.* decomposition of insulin and other hormones).
    - Metabolism of lipids (*e.g.* cholesterol synthesis and degradation into bile acids).
    - Storage of glycogen and fat-soluble vitamins (A, D, E and K).
2. Blood function
    - Metabolism of proteins (*e.g.* production of coagulation factors).
    - Destruction of old red and white blood cells, as well as some bacteria in the blood.
    - Transformation of bilirubin (toxic) into conjugated bilirubin (not toxic).
3. Antitoxic function
    - Destruction of toxins and pharmaceutical drugs.
    - Conversion of ammonia into urea.

The liver consists of liver cells known as *hepatocytes* organized in bays around the sinusoids. The functional unit of the liver is the hepatic lobule. The exchanges with the rest of the body are mostly done through a dual blood supply coming from both the *hepatic portal vein* and the *hepatic artery* and ending up in a multitude of capillaries within the liver. The hepatic artery brings oxygenated blood while the hepatic portal vein brings nutrient-rich blood from the alimentary tract. Most of liver cells are hepatocytes (*i.e.* $\simeq 80\%$), but there are other types of cell such as bile duct cells, endothelial cells and hepatic stellate cells.



(a) (b)

Figure 2.1: Liver anatomy: (a) position of the liver and neighboring organs within the thoracic cage (Source: `http://www.sweetadditions.net/health/function-of-the-liver-in-digestion`), and (b) details of the inferior surface of the liver as depicted in *Gray's Anatomy* (Source: `http://www.wikimd.org/index.php/The_Liver`).

Because it is the only organ able to regenerate, the liver may be transplanted from a liver's portion. Most of the times, a damaged liver regenerates itself spontaneously. Otherwise, the liver may be surgically replaced. In case of transplantation, the liver may come either from a person in a state of brain death, or from a living donor. In the latter case, a portion of the donor's liver is extracted and transplanted to the donor. Each half of the liver regenerates towards a complete liver. However, the most common technique is *orthotopic transplantation*, in which the original liver is removed and replaced by the donor organ in the same anatomic location as the original liver.

Due to its strategic anatomical location and multiple functionalities, the liver is prone to many diseases:

- Primary liver cancer (*e.g.* hepatocellular carcinoma and angiosarcoma).
- Secondary neoplasms (*i.e.* secondary liver metastases coming from primary cancers such as carcinoma, endocrine cancer and melanoma).
- Benign liver tumors (*e.g.* hemangioma and adenoma).

Other common liver diseases are cirrhosis, haemochromatosis, both alcoholic and viral hepatitis, Wilson's disease and steatosis.

A blood test is the common technique to diagnose the liver function. This test may easily pinpoint the extent of liver damage. Should an infection be suspected, other serological tests must be done. Physical exam of the liver is not an accurate technique though, because it may only reveal the presence of tenderness or check the liver size. A convenient way to look at the liver consists in using medical imaging. In this case, a radiological study, such as an ultrasound or a CT scan (see section 2.3.2), is performed. Other typical medical imaging modalities for the liver are MRI and PET.

2.2.2 Lower limbs

In human anatomy, lower limbs, as opposed to upper limbs (*i.e.* arms and forearms), give humans the ability to stand up and move from one place to another. Human musculoskeletal system is mainly made up of both lower and upper limbs.

Lower limbs are composed of six parts:

1. The upper part known as *the gluteal region*, located between pelvis and thigh.
2. The femoral part known as *the thigh*, located between hip and knee.
3. The knee, located between thigh and leg.
4. The crural region known as *the leg*, located between knee and ankle.
5. The ankle, located between foot and leg.
6. The foot.

The musculoskeletal system rigidity is ensured by the skeleton, which provides structure as well as protection. Human skeleton is made up of bones (see section 2.2.2.1) attached to other bones with joints (*e.g.* knee and hip). Muscles (see section 2.2.2.2), which produce forces that allow motion, are attached to the skeleton by tendons. Bones are supported and completed by ligaments, tendons, muscles and cartilage. On top of the musculoskeletal system, the skin offers a protection to the external environment as well as other mechanisms (*e.g.* sensation, regulation and storage).

2.2.2.1 Bones

The main roles of bones are the support of body structures, the protection of internal organs and the generation of motion in conjunction with muscles (see main lower limb bones in Figure 2.2, left). Bones are also involved in the formation of blood cells, the calcium metabolism and the storage of minerals. They also play an important role in the detoxification of the organism by fixing and storing heavy metals (*e.g.* lead). Thanks to their structure, bones are light, flexible and robust.

Bones are mainly solid (*i.e.* 95% solid tissue and 5% water). The solid part is 65% mineral (*i.e.* mainly calcium ensuring bone rigidity) and 35% organic (*i.e.* mainly collagen fibers ensuring bone structure).

In anatomy, five types of bones are considered:

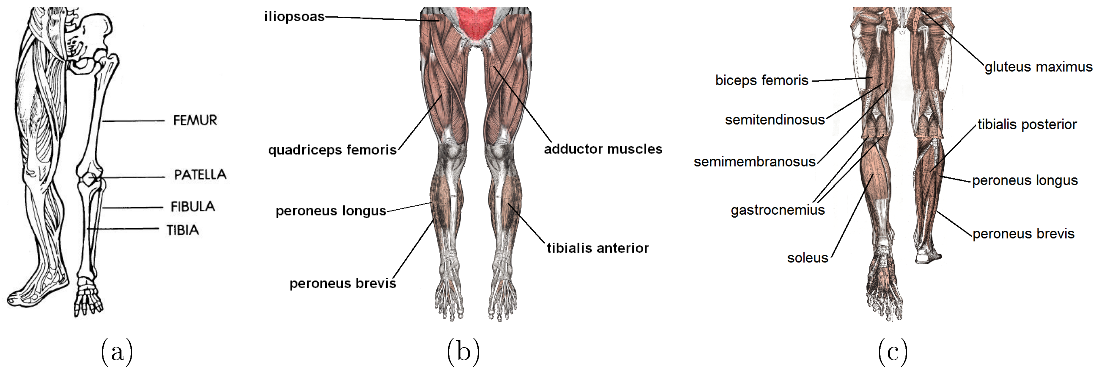


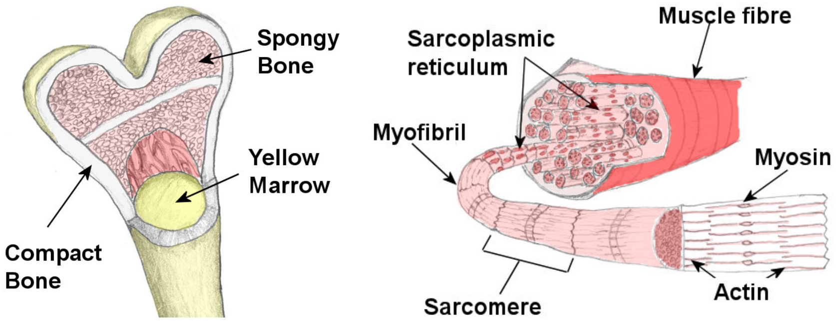
Figure 2.2: Lower limb anatomy: (a) focus on main bones (Source: http://www.tpub.com/content/armymedical/MD0010/MD00100027.htm), (b) anterior and (c) posterior views detailing some muscles (Source: http://en.wikipedia.org/wiki/Human_musculoskeletal_system).

1. Long bones: they are found in limbs and feature a curved and cylindrical shape with a central cavity filled with marrow (*e.g.* femur and tibia).
2. Short bones: they are combined to ensure strength and compactness to the skeleton (*e.g.* carpus and tarsus).
3. Flat bones: they provide large surfaces for muscle attachment (*e.g.* pelvis and skull bones).
4. Sesamoid bones: they are small bones surrounded by tendons and their role consists in transmitting forces and modifying pressure/friction (*e.g.* patella).
5. Irregular bones: they are non-categorized bones due to their unusual shape (*e.g.* coccyx and mandible).

Bones are not uniformly solid materials, they are composed of three different types of tissue (see an illustration for long bones in Figure 2.3, left):

1. Compact, or *cortical*, tissue: the harder and outer tissue of bones.
2. Cancellous, or *trabecular*, tissue (also called *the spongious bone*): the sponge-like and porous tissue inside bones.
3. Subchondral tissue: a smooth tissue covered with cartilage at the ends of bones.

Many disorders may affect the skeleton, but the most prominent is osteoporosis. This is a typical bone disease that leads to an increased risk of fracture. In osteoporosis, the bone mineral density is reduced, bone micro-architecture is disrupted and the amount and variety of non-collagenous proteins in bone is altered. The best ways of preventing osteoporosis consist in lifestyle advice and medication.



(a) (b)

Figure 2.3: Focus on main inner structures of (a) long bones and (b) muscles (Source: `http://www.teachpe.com/anatomy`).

Preventing people with known or suspected osteoporosis to fall remains an efficient way to prevent fractures. Osteoporosis may be treated with bisphosphonates and various other medical treatments. In medical imaging, CT (see section 2.3.2) and X-ray scans provide both an excellent bone contrast with respect to other tissues.

2.2.2.2 Muscles

Muscle is a contractile tissue whose contraction ensures the motion of both body parts (*i.e.* skeletal muscle) and substances within the body (*i.e.* smooth or cardiac muscle). Smooth and cardiac muscle contraction occurs without conscious thought and is necessary for survival (*e.g.* peristalsis that pushes food through the digestive system). Voluntary contraction of the skeletal muscles is used to move the body and can be finely controlled (*e.g.* movements of the eye). The medical discipline studying muscles is *myology*, though this field is primarily interested in skeletal muscle.

The human body is made up of more than 640 muscles whose size varies according to their function. All muscles represent from 40 to 45% of the dry weight of the body, which is the largest fraction of body mass in non-obese adults. Muscles are attached to tendons through which they can exert forces on bones (see an overview of lower limb muscles in Figure 2.2, middle and right). Tendon attachments may be very large and their site is correlated to bone shape. Skeletal muscles usually have two attachment sites known as the origin, which is the closest to the mass center, and the insertion site. Due to their hierarchical structure (see Figure 2.3, right), muscles can produce forces at a macroscopic scale due to microscopic events.

Muscles feature a large variety of shapes that may be reduced to three:

1. The long muscle spindles: they are muscles whose length is predominant. Their bodies are swollen and end up with strong and white tendons that attach them to bones. Some muscles have 2, 3 or 4 tendons at their extremities (*e.g.* biceps, triceps and quadriceps, respectively).

2. The broad and flat muscles: their shape is either flat, blade or ribbon. Spread like a fan but without any tendon, they attach to bones through a tendinous blade known as *aponeurosis*. They form the walls of large cavities within the body (*e.g.* pectoralis major muscle and diaphragm).
3. The short muscles: they are circular and define an opening. For instance, *orbicularis oculi* muscles are annular muscles surrounding a hollow organ. In this case, they are known as *sphincters* and open in response to a pressure.

Muscle diseases, which impair the functioning of muscles either directly via intrinsic muscle pathology or indirectly via nerve pathology, are among others:

- Myopathies, which consist of all diseases affecting the muscle itself, rather than its nervous control.
- Muscular dystrophies, which are a large group of diseases, mostly inherited, in which muscle integrity is compromised. This causes a gradual loss of strength, a high dependency and a shortened life.
- Musculoskeletal inflammatory diseases (*e.g.* polymyalgia rheumatica, polymyositis and dermatomyositis).
- Muscle tumors such as leiomyoma (or *liomyome*) on smooth muscles, habdomyoma and rhabdomyosarcoma on striated muscles, and metastases from another location (*e.g.* lung cancer).
- Other diseases such as tetanus, botulism, rhabdomyolysis, myasthenia gravis and Lambert-Eaton myasthenic syndrome.

Symptoms of muscle diseases may include weakness, spasticity, myoclonus and myalgia. Diagnostic procedures that may reveal muscular disorders include testing *creatine kinase* levels in the blood and electromyography, which is used to measure electrical activity in muscles (see an application in section 7.2.2). In some cases, a muscle biopsy may be done to identify a myopathy, as well as genetic testing to identify DNA abnormalities associated with specific myopathies and dystrophies. In medical imaging, MRI is the best technique to image muscles, as it is the reference for soft tissue acquisitions (see section 2.3.3). In addition, DT-MRI may be used to study the fiber direction of muscles.

2.3 Medical imaging

2.3.1 Brief description

Medical imaging consists in acquiring and restoring images from physical phenomena (*e.g.* magnetic resonance, ultrasonic waves reflection, radioactivity and x-ray absorption). This technology is used in radiology and enables the visualization of both the physiology and metabolism of the human body. However, such a technology is also used in other fields like industry, chemistry and archeology.

The purpose of medical imaging is to create a visual and comprehensible representation of the human body. The objective is to represent a large amount of data acquired in different ways in a relatively simple format (DICOM is the standard format used to deal with medical imaging data).

The reconstructed image may be processed to obtain:

- a 3D reconstruction of an organ.
- a movie showing the evolution of an organ over time (*e.g.* heart).
- a multimodal representation that registers several data within the same image (*e.g.* heart outlines and mobility of its walls).

Depending on the technique used, medical imaging provides the information about organ anatomy (*e.g.* size, volume, location and injury shape) or their function (*e.g.* physiology and metabolism). In the former case, we talk about *structural imaging* while in the latter case, we talk about *functional imaging*.

Here are some common medical imaging techniques:

1. Structural imaging
    - Tomographic methods, which are based on X-rays (*e.g.* Computed Tomography, see section 2.3.2) or on magnetic resonance (*e.g.* Magnetic Resonance Imaging, see section 2.3.3).
    - Ultrasound methods (*e.g.* diagnostic sonography).
    - Optical imaging methods, which use light beams (*e.g.* diffusive and ballistic imaging systems).
2. Functional imaging
    - Nuclear medicine methods, which are based on the emission of gamma rays from radioactive tracers (*e.g.* PET and SPECT).
    - Electrophysiological methods, which measure changes in the electrochemical condition of tissues (especially those linked with nerve activity).
    - Infrared thermography methods, which measure radiation in the infrared range of the electromagnetic spectrum.

This large number of techniques produce an important amount of data that is more and more difficult to analyze. This is why medical imaging analysis is a community becoming more and more important. Meshes segmented from medical images can help radiologists in their diagnosis. They can also have predictive capabilities.

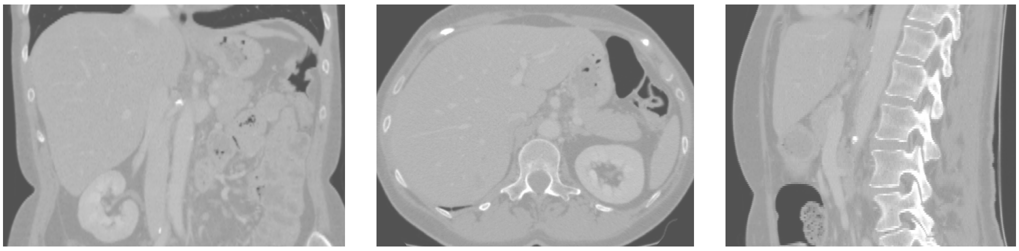


Figure 2.4: Three views of a CT image of the liver.

Another important advantage of medical imaging is that subject-specific models may be computed, which enables the analysis of intra and inter-subject variation. Unlike classical computer vision issues, medical imaging analysis has to deal with complex issues such as low resolution, anisotropic voxels, noise (*i.e.* computed using the signal-to-noise ratio (SNR)), partial voluming and may require a pre/post processing of images.

2.3.2 Computed Tomography (CT)

Computed Tomography (CT) is a medical imaging technique that measures the X-ray absorption of tissue. To acquire the images, the patient is subject to a scan of X-rays beams. This technique is used in radiology, archeology, biology, geophysics, oceanography, astrophysics and, in a high resolution version, materials science.

In radiology, the emitter turns around the patient at the same time than the receptor. Then, data is computed to reconstruct the image (see the CT image of a patient's liver in Figure 2.4). Image contrast may be enhanced using radioconstract agents (*e.g.* to highlight blood vessels).

Radiography was the first technique used to perform in-vivo anatomical imaging of internal structures. Coupling this technique with computer science to reconstruct the images, Godfrey Newbold Hounsfield and Allan MacLeod Cormack were awarded the Nobel Prize in Medicine in 1979. A detailed history of CT imaging is available in [Filler 2009].

In practice, intensity of X-ray beams going through the patient's body are measured by a sensor composed of phosphor plates. In computed radiography (CR) and digital radiography (DR), those sensors are digital. Since tissues absorb X-rays differently, anatomical structures may be highlighted with the resulting 2D projections. Combining consecutive projections, volumes with a size of typically 512x512x128 voxels are reconstructed with a 0.5x0.5x1.0 mm resolution.

As the image intensity is measured using the Hounsfield scale, a simple segmentation of tissues could be performed. However, different tissues may have the same intensity value and artifacts may disturb their identification in the image. With a particularly high bone-to-soft tissue contrast, CT is extensively used to study chest and bone. But this technique is also used to highlight blood vessels.

In a clinical environment, thresholding and simple post-processing techniques


(a) (b) (c)

Figure 2.5: Some MR images of the knee: (a) High resolution T_1 FLAIR Saggital, (b) T_1 FSPGR Saggital, and (c) its echo signal. FLAIR stands for Fluid Attenuated Inversion Recovery and FSPGR for Fast Spoiled Gradient-Recalled-Echo.

(*e.g.* morphological operations) are generally sufficient to segment bones or structures highlighted with suited radio-contrast agents (*e.g.* angiography). In orthopedics, where most of pathologies affect bones, CT is the privileged modality. However, this technique is usually considered as invasive because the body absorbs a high dose of X-ray that may damage cells and cause cancer. But the risk-benefit ratio linked to the radiation exposure is usually considered acceptable.

2.3.3 Magnetic Resonance Imaging (MRI)

Unlike CT imaging, which is mainly used to acquire structures like bones and blood vessels, Magnetic Resonance Imaging (MRI) is preferably used for soft tissue acquisitions (*e.g.* brain, heart and muscles). MRI scanners use non-ionizing radio frequencies (RF) that are suitable for non-calcified tissues. For this reason, they are extensively used in neurology (brain), cardiology (heart), biomechanics (musculoskeletal system) and oncology (see some MR images of the knee in Figure 2.5). A detailed history of MRI may be found in [Filler 2009].

Developed in 1973 by Paul Lauterbur and Peter Mansfield [Lauterbur 1973], MRI consists in using a magnetic field to align magnetic moments of water molecules (that compose 70% of the human body). More precisely, water hydrogen protons acquire a magnetization $\overrightarrow{m}$ (also called *magnetic moment*) when placed in a magnetic field $\overrightarrow{b}$. This magnetization $\overrightarrow{m}$ is proportional to $\overrightarrow{b}$ and subject to a precession motion around O_z (z-axis of $\overrightarrow{m}$).

The Larmor equation links the frequency of this rotation f_L with the magnitude of $\overrightarrow{m}$ (f_L is the resonance frequency from which radio waves are able to excite protons, *i.e.* $\sim$ 42 MHz/Tesla):

$$f_L = \frac{\Gamma}{2\pi}\|\overrightarrow{m}\| \tag{2.1}$$

where Γ is the gyromagnetic ratio.

The amplitude m is calculated by the Nuclear Magnetic Resonance (NMR) measure. Magnetic field $\overrightarrow{b}$ needs to be projected into O_{xy} plan (perpendicular to the direction of $\overrightarrow{b}$), in order to study what is known as the *relaxation phenomenon*. This is the macroscopic effect obtained when emitting a RF impulse at Larmor frequency f_L, perpendicularly to O_z, which is known as the *resonance phenomenon*. RF impulse is created by an antenna placed around the patient and its duration determines the angle from which $\overrightarrow{m}$ moves away from O_z.

When the RF impulse is done, $\overrightarrow{m}$ goes back to its original state by a precession motion at f_L frequency while emitting an electromagnetic field. This original state recovery (*i.e.* equilibrium) is called the *relaxation*, from which two components may be extracted: a longitudinal relaxation, corresponding to the recovery of the longitudinal component m_z towards $\overrightarrow{b}$, and a transversal relaxation for the recovery of the transversal component m_{xy}.

The kinetics of both components are defined as:

$$m_z(t) = m\ (1 - exp^{(-t/T_1)}) \tag{2.2}$$

$$m_{xy}(t) = m\ (exp^{(-t/T_2)}) \tag{2.3}$$

where T_1 and T_2 are the *longitudinal* and *transversal* relaxation times, respectively.

Values of T_1 and T_2 are chosen depending on the physicochemical properties of the tissue one wants to highlight (*e.g.* muscle, fat and white/gray matter). More precisely, T_1 depends on the mobility of hydrogen atoms, or the mobility of the molecules where they are present. The bigger the molecules, the shorter T_1. Similarly, T_2 also depends on the mobility of hydrogen atoms, but is always shorter than T_1.

In practice, the signal decreases faster than expected because of magnetic field inhomogeneities (known as *bias field*). These inhomogeneities come from both imperfections of the magnetic field inductors and magnetic susceptibility differences between tissues. This provokes a variation of rotation frequency inside the acquired volume and leads to an accelerated phase difference of microscopic magnetization. As a result, signal strength decreases faster than expected.

To measure relaxation times, several sequences may be used, such as *spin echo* and *gradient echo*. Spin echo sequence consists in the repetition of 90° and 180° RF pulses. The time between the 90° RF pulse and the MR signal sampling is called the echo time TE. Unlike spin echo, gradient echo sequence does not compensate the phase differences associated with bias field inhomogeneities and enables the measure of the signal exponential decrease in T_2^*:

$$m_{xy} = m \ (exp^{(-t/T_2^*)}), \ T_2^* < T_2 \tag{2.4}$$

Another important parameter is the repetition time TR, which measures the time between two consecutive sequences. The contrast of MR images depends on the values of TE and TR. Using T_1 and T_2, as well as $\|m\|$, enables the acquisition of T_1*-weighted*, T_2*-weighted* and *spin density weighted* MR images, respectively.

T_1-weighted MRI T_1-weighted MR images are obtained using a short repetition time TR ($400 \leq TR \leq 600$ ms) and a short echo time TE ($10 \leq TE \leq 20$ ms). This is the commonly used clinical modality and the short TR allows the fast acquisition of high resolution MR images (see some T_1 MR images in Figure 2.5). For the brain, white matter appears lighter than gray matter. Cerebrospinal fluid, situated between gray matter and bone, appears darker. T_1-weighted MR images are also used to highlight anomalies after the injection of a contrast agent.

T_2-weighted MRI T_2-weighted MR images are obtained using a long TR ($TR >$ 2000 ms) and a long TE ($TE >$ 80 ms) with a spin echo sequence. This modality is less sensitive to the bias field and is well suited to highlight water (in particularly edema, which is characterized by an increased water content).

T_2^*-weighted MRI T_2^*-weighted MR images are also obtained using a long TR and a long TE, but with a gradient echo sequence. This sequence does not have the extra refocusing pulse used in spin echo, so it is subject to additional losses above the normal T_2 signal decrease. As a consequence, this modality is more prone to susceptibility losses at air and tissue boundaries. However, contrast may be increased for certain types of tissue, such as venous blood.

Spin density weighted MRI Spin density weighted MR images are obtained using a long TR ($TR >$ 2000 ms) and a short TE ($10 \leq TE \leq 20$ ms). This modality tries to have no contrast from either T_2 or T_1 signal decrease, the only signal change coming from differences in the amount of available hydrogen nuclei in water. With $2000 \leq TR \leq 3000$ ms and $TE < 30$ ms, tissue (*e.g.* meniscus) is highlighted compared to water and fat. With $TR > 5000$ ms and $TE < 30$ ms, water is highlighted compared to tissue and fat.

In addition, more sophisticated modalities are worth mentioning such as *Diffusion-weighted imaging (DWI)*, *Magnetic resonance angiography (MRA)* and *Magnetic resonance spectroscopy (MRS)*.

Diffusion-weighted imaging (DWI) Using gradients of the magnetic field, this modality calculates the probability distribution of water molecule diffusion at each voxel of the image. Since the probability distribution is constrained by surrounding tissues, this modality provides the position, orientation and anisotropy of white matter fibers in the brain. Fiber tracking algorithms track fibers along their entire length (*e.g.* the corticospinal fiber, which transmits motor information, from the motor cortex to the spinal cord). Some common clinical applications are the localization of fibers affected by a lesion in the white matter (*e.g.* multiple sclerosis lesions) and the localization of main white matter fibers for surgical planning.

Magnetic resonance angiography (MRA) Used to visualize the arteries, this modality highlights abnormalities such as stenosis, occlusion, dissection, fistula, aneurysm and arteritis. For instance, the cerebral, cervical, renal and pulmonary arteries, as well as the thoracic and abdominal aorta, are extensively studied with this technique.

Magnetic resonance spectroscopy (MRS) Aimed at studying both the presence and concentration of certain metabolites, this modality not only requires high-field MRI but also a specific training for radiologists. The technique seems very promising though, especially in oncology where it provides a solution for the distinction between local recurrence and post-radiotherapy necrosis in an early stage with an accuracy that only a biopsy (invasive and sometimes risky) can equal.

2.4 Conclusion

In this chapter, we presented the medical context in which this thesis takes place. First, we described the main anatomical features of both liver and lower limbs because the methods proposed in this thesis are tested on these anatomical structures. Then, we described the main techniques used in medical imaging, with a deeper focus on Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). Both techniques are used in the framework of this thesis: CT for liver segmentation, which is the main application on which our methods are tested, and MRI for the segmentation of lower limb structures that we describe in chapter 7.

Chapter 3

Model-based Image Segmentation

Contents

3.1 Introduction

Although considered as a natural process for human beings, the task of extracting objects from images through segmentation still requires a tremendous attention in the computer vision community. Segmentation consists in extracting regions of interest out of images. With 2D images, this operation is done by grouping pixels (*i.e.* picture elements) having similar features (*e.g.* in terms of intensity), to form a specific region, or *partition*, of the image. A common usage is to extract objects out of the background for object recognition. If the number of partitions equals two, the segmentation is referred as *binarization* (see Figure 3.1).

In medical image analysis, the goal of segmentation is to extract anatomical structures of interest such as organs, bones and tissues. A thorough review on this matter may be found in [Fitzpatrick & Sonka 2000]. From model-based image segmentation, generated meshes may give more insight in terms of visualization than a set of binary stacks. These meshes may be used both for therapy planning [Kaus *et al.* 2007, Delingette *et al.* 2006] and quantification for diagnosis [Gérard *et al.* 2002]. Using statistics on shape, meshes also enable the interpretation of intra and inter-subject variability.

In this chapter, we first review the main techniques used in medical imaging to perform segmentation (section 3.2). After introducing manual segmentation (section 3.2.1), we briefly review approaches based on thresholding (section 3.2.2) and atlases (section 3.2.3).

In the framework of this thesis, we use explicit deformable models (section 3.3) as segmentation technique. We are particularly interested in simplex meshes (section 3.3.2), from which we describe the specific topology, geometry, evolution and quality improvement. Then, we explain how to initialize a deformable model into an image (section 3.3.3) and how to deform it by means of internal (section 3.3.4) and external (section 3.3.5) forces.

3.2 Segmentation approaches

3.2.1 Manual segmentation

In order to model the various structures of interest in the human body with a high fidelity, manual segmentation approaches were first used on medical datasets. This task requires nonetheless a lot of work as the medical expert has to delineate 2D slices one by one (see an example in Figure 3.1). For this reason, more and more (semi-)automatic are proposed in the literature to overcome this issue. However, manual segmentation may still be helpful nowadays [Schmid *et al.* 2009].

First, manual segmentation may serve as a basis for automatic methods by providing some prior knowledge (*e.g.* internal forces in explicit deformable models that enforce the consistency with an *a priori* of a shape [Delingette 1999, Gilles *et al.* 2006]) or training material (*e.g.* methods based on Principal Component Analysis (PCA) [Cootes *et al.* 1993, Schmid & Magnenat-Thalmann 2008] or classification methods like K-means [Hartigan & Wong 1979] and Fuzzy C-means [Bezdek 1973]). In general, more than one mesh is needed to produce satisfactory results. Second, manual segmentation provides a ground truth for automatic and semi-automatic segmentation methods. Meshes generated by an automatic method may thus be compared with those obtained manually. Usually, manual segmentation is performed by a medical expert and is then used for validation purposes.

In any case, manual segmentation *per se* remains a difficult task, requiring a lot of time and prone to errors [Collier *et al.* 2003]. Indeed, the segmentation of the same structure always differs between different experts (so called *inter-expert* variability [Saarnak *et al.* 2000]). Even if the same expert performs the same segmentation, the result usually differs (so called *intra-expert* variability [Fiorino *et al.* 1998]).

3.2.2 Thresholding approaches

Thresholding approaches may be considered as the simplest semi-automatic methods. In the literature, thresholding is considered as belonging to *direct approaches*. In its simplest formulation, thresholding requires an intensity level (*i.e.* the threshold) from which intensities are considered either foreground (*i.e.* the object of in-

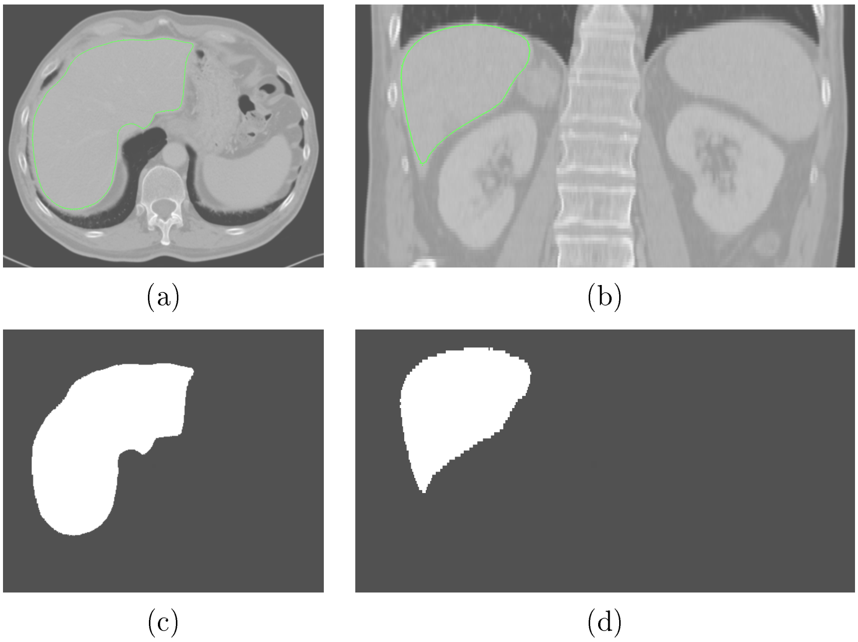


Figure 3.1: Manual segmentation of the liver in a CT image: (a)+(b) manual delineation of the liver in two views, and (c)+(d) corresponding binary images where white voxels correspond to the liver (*i.e.* foreground) and black voxels correspond to other structures of the image (*i.e.* background).

terest) or background. For multiple objects or regions of interest, several thresholds must be defined (see an example in Figure 3.2). The threshold may be given manually, either from empirical experiences or using heuristics [Sezgin & Sankur 2004].

Though easy to implement and use, thresholding approaches feature several drawbacks. First, they do not make any assumption about the shape of segmented regions. Also, they do not take into account any spatial information between intensities that may help the segmentation process. As they rely only on intensities, images are usually pre-processed to help the segmentation, since images may feature noise and artifacts.

In medical image analysis, these drawbacks are even worse because images are usually noisy or biased. This is especially the case with MR images. A pre-processing step consists in ensuring a good Signal-To-Noise ratio (SNR). Typical pre-processing techniques include, among others, bias field correction [Styner *et al.* 2000] and histogram equalization [Yang & Wu 2010]. As for manual segmentation algorithms, thresholding approaches may also need post-processing treatments. This is especially a requirement in medical image analysis since the objective is to accurately reconstruct structures of interest (*e.g.* organs, bones and tumors). This may be done by mesh extraction using first the Marching Cubes algorithm [Nielson 2003, Lorensen & Cline 1987] and then smoothing the generated mesh (see section 7.2.4 for more details).

In the literature, a popular approach based on thresholding is region growing

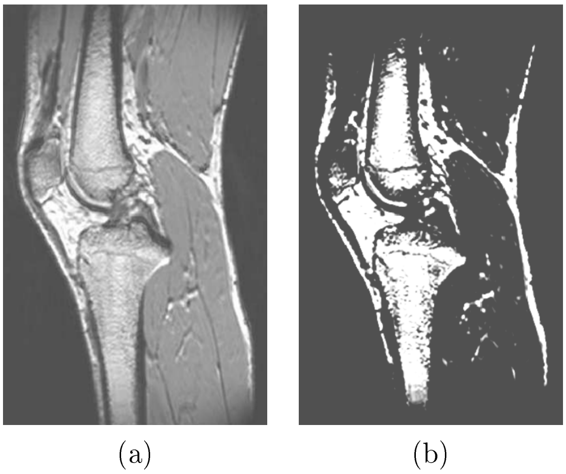

(a) (b)

Figure 3.2: MR image of a knee (a) before and (b) after bone segmentation using two thresholds. Note that fat tissue is also segmented, since its appearance in the MR image is close to spongious bone tissue.

[Petrou & Bosdogianni 1999]. From a starting point, usually chosen manually, the algorithm creates a region by enlarging the neighborhood until a stop criterion is reached. To consider if a neighborhood point belongs to the region, a criterion must be chosen, which is usually a threshold.

3.2.3 Atlas-guided approaches

In addition to an *a priori* on shape like deformable models (see next section), atlas-guided approaches use the position of the structures of interest as well. There are two major types of atlases, *mesh* and *image* atlases. A mesh atlas is composed of one or several meshes placed into an image at the exact location of structures (*e.g.* after manual segmentation or average construction) [Costa 2008]. An image atlas is a reference image where structures are labeled [Commowick 2007], which may be seen as a labeled representation of the anatomy.

In a segmentation framework, an atlas image, in which meshes are correctly located, is first registered to the image of a subject. Resulting deformation is then applied to the meshes. Depending on the registration quality, meshes are likely to be close to structures' boundaries of the subject. This makes atlas-guided approaches suitable for mesh initialization in a model-based segmentation [Pitiot *et al.* 2004].

In the literature, such approaches may be seen as *segmentation by registration* techniques. Atlases have been successfully used for the automatic segmentation of head and neck region [Commowick 2007], as well as brain [Souplet 2009, Bondiau *et al.* 2005, Dawant *et al.* 1999].

Using full *a priori* information on mesh shape and position is an advantage. For multiple structures, relative position between them is also kept by the atlas. A specific care needs to be taken for the registration between the atlas and the subject image [Vemuri *et al.* 2003, Lorenzo-Valdés *et al.* 2002, Maintz & Viergever 1998], whose

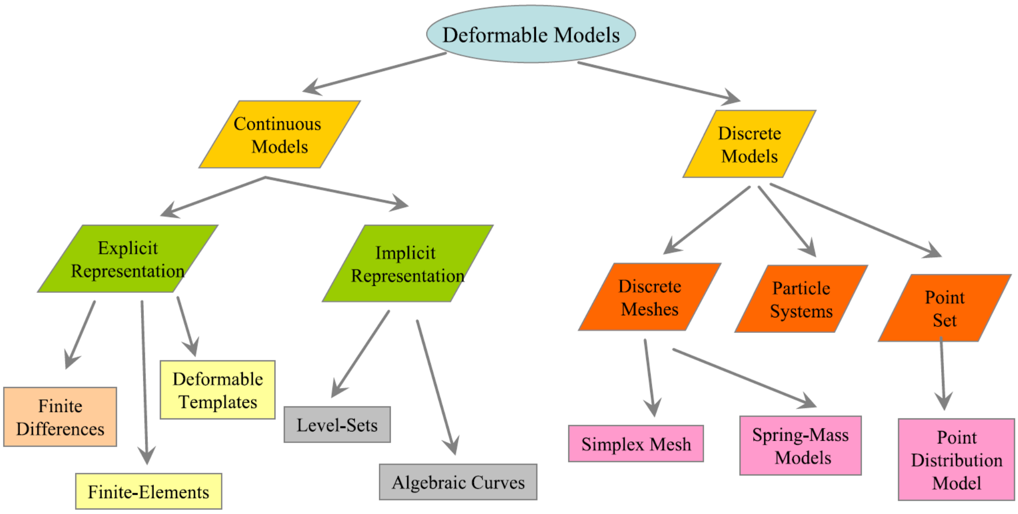


Figure 3.3: Classification of the different geometric representations used in model-based image segmentation (Source: [Heimann & Delingette 2010]).

quality is a key part of the segmentation result.

3.3 Deformable models

3.3.1 Geometric representations

Unlike thresholding approaches based solely on intensity, deformable models introduce an assumption about shape. This is why the geometric representation of deformable models is important in model-based segmentation. Each representation has pros and cons, since the choice of a representation is often problem-dependent. Common representations include level-sets, triangular meshes and spline curves or surfaces (see a classification of geometric representations in Figure 3.3).

In the literature, geometric representations are usually categorized into two main approaches, namely *implicit* and *explicit* methods.

Implicit methods They represent a surface S_f as the zero-crossing of a function f having values in $\mathbb{R}^3$:

$$S_f = \{\mathbf{x} \in \mathbb{R}^3 | f(\mathbf{x}) = 0\} \tag{3.1}$$

Algebraic surfaces are surfaces following equation 3.1 where f is a polynomial. They have been extensively used in graphics and modeling [Taubin 1994]. In medical imaging, two specific algebraic surfaces have been used: *superquadrics* [Barr 1981] and *hyperquadrics* [Han *et al.* 1993].

Level sets [Osher & Sethian 1988] are another implicit method-based technique that has been extensively used to segment medical images [Dindoyal *et al.* 2007, Malladi *et al.* 1995]. Extensions to the original energy formulation were proposed, such as an additional term, which deforms the model towards a previously learned shape [Leventon *et al.* 2000], and a region-based energy functional [Tsai *et al.* 2003]. A recent review on level sets may be found in [Cremers *et al.* 2007].

Explicit methods The first algorithm based on explicit methods was the snake algorithm [Kass *et al.* 1988], which has been used to extract contours on 2D images. The concept was then generalized [Terzopoulos *et al.* 1988] for the extraction of 3D shapes based on mesh representation (*e.g.* triangular [Ng-Thow-Hing 2000] and 2-simplex [Delingette 1999] meshes). Topology changes may also be implemented for deformable models [McInerney & Terzopoulos 1999]. Statistical models of shape and appearance are very common for 3D medical image segmentation, namely Active Shape Models (ASM) [Cootes & Taylor 1993] and Active Appearance Models (AAM) [Cootes & Taylor 2001].

In practice, a mesh is first initialized into the image and then deformed using forces that attract the mesh to fit regions of interest (see an example in Figure 3.4). Two main energies control the mesh deformation: *internal* and *external* energies. Internal energy constrains the mesh regularity (*i.e.* makes sure the mesh is consistent and smooth). For instance, local curvature at vertices level may be constrained (*e.g.* C_2 constraint in 2-simplex meshes [Delingette 1999]). External energy attracts the mesh to image-based features (*e.g.* intensity, gradient and edges) in order to fit a desired region of interest.

The mesh evolves as to minimize the sum of two energies:

$$E_V = \alpha E_{int} + \beta E_{ext} \tag{3.2}$$

where α and β are weights to control both the internal and external energies, respectively.

Other energies may be added to constraint the deformation. For instance, balloon energy [Cohen 1991] is used by internal forces to prevent the mesh from shrinking towards a point.

At a simple level, explicit deformable models may be initialized using basic circular or spherical shapes. Nonetheless, for regions of interest whose shape is known (*e.g.* livers and bones), using an *a priori* of the shape is more efficient. The objective is to use meshes whose shape is close to the region of interest [Montagnat 1999]. Starting the deformation with such an *a priori* is one of the main advantage of explicit deformable models. Also, they usually do not need post processing treatments as they are geometrically constrained by internal forces.

Explicit deformable models have been successfully used to segment the heart [Ecabert *et al.* 2008, Sermesant *et al.* 2006] and prostate [Costa *et al.* 2007]. They

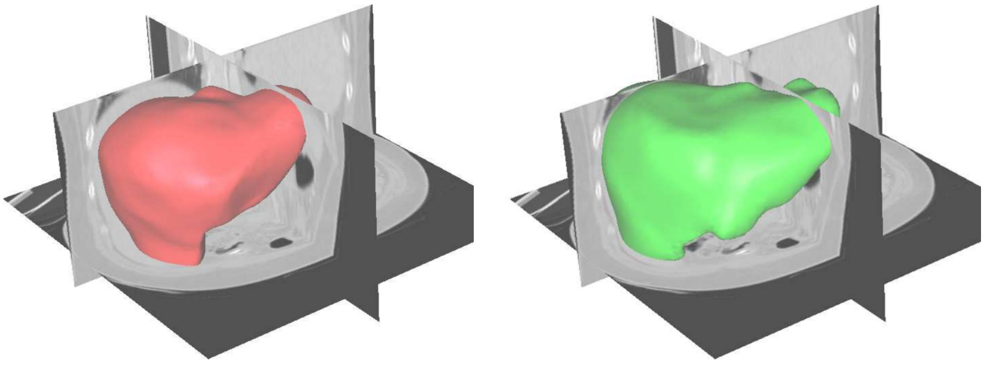

(a) (b)

Figure 3.4: Liver segmentation using a deformable model. As a first step, (a) the mesh is initialized into the image. Then, the mesh is deformed until reaching a steady state in which (b) the mesh fits the liver boundaries.

have also been used to improve hepatic surgical planning [Soler *et al.* 2001]. Reviews on deformable models may be found in [Montagnat *et al.* 2001, Jain *et al.* 1998, McInerney & Terzopoulos 1996].

3.3.2 Simplex meshes

Topologically dual of triangular meshes, simplex meshes offer a discrete representation of deformable models where surface shape is defined locally at each vertex. They have been introduced for 3D shape reconstruction and segmentation [Delingette 1999]. Then, they have been extended for 4D segmentation by adding global spatio-temporal constraints [Montagnat 1999]. Simplex meshes are particularly efficient in terms of flexibility and computational cost, since they feature a simple geometric description.

In this thesis, we use simplex meshes to represent structures of interest. In this section, we describe their topology, geometry and evolution. We also describe techniques to improve their quality. For more information, overviews and applications of simplex meshes may be found in [Heimann & Delingette 2010, Gilles 2007, Montagnat 1999].

3.3.2.1 Topology

A k-simplex mesh represents a manifold surface of dimension k and is defined as a $k+1$-cell (*e.g.* k=1 for 2D surfaces and k=2 for 3D surfaces).

A 0-cell is defined as a vertex $\mathbf{p}$ and a 1-cell as an edge (*i.e.* an unordered pair of distinct vertices). Recursively, a p-cell $\mathcal{C}$ ($p \geq 2$) is defined as the union of c $(p-1)$-cells ($c \in \mathbb{N}$) such that:

1. Each vertex $\mathbf{p}$ belonging to $\mathcal{C}$ also belongs to p distinct $(p-1)$-cells.

2. A $(p-2)$-cell belongs to two and only two $(p-1)$-cells.

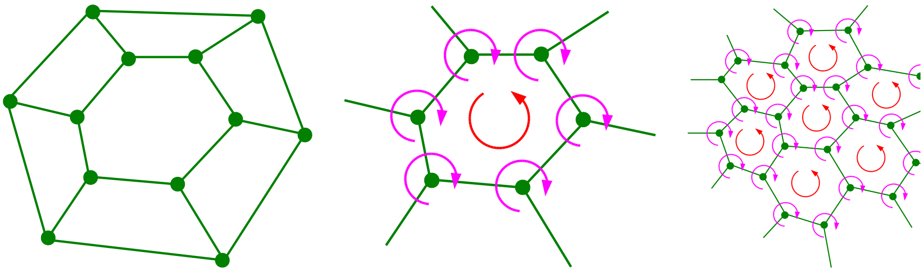

(a) (b) (c)

Figure 3.5: Specific topology of 2-simplex meshes: (a) each vertex has exactly three neighbors and one edge at most links two vertices, (b) every face and corresponding vertices feature a specific orientation, which leads to (c) an orientation on the whole mesh.

3. The intersection of two $(p-1)$-cells is empty or is a $(p-2)$-cell.
4. A p-cell is simply connected (*i.e.* given two vertices of that cell, there exists at least one set of edges that connect these two vertices).

Two topological properties result from this definition:

- A k-simplex mesh has a fixed vertex connectivity: every vertex is adjacent to $k+1$ distinct vertices.
- A k-simplex mesh is a topological dual of a k-triangulation. In this duality relationship, a triangle is associated with a vertex of a 2-simplex mesh, a triangulation edge with an edge and a vertex with a face (*i.e.* a 2-cell).

An illustration of topological properties of 2-simplex meshes may be found in Figure 3.5.

Orientation Neighboring rules applied to k-simplex meshes lead to an orientation of edges around vertices and therefore to an orientation of vertices. All vertices are oriented in a consistent manner, so that faces are oriented as well (see Figure 3.5, middle and right for an example on 2-simplex meshes).

Duality As previously mentioned, there is a topological equivalence, or *duality*, between 2-simplex meshes and triangular meshes (see Table 3.1 and Figure 3.6). In fact, this duality is valid for classical k-solid meshes (*e.g.* polylines and tetrahedral meshes) and is purely topological, as there is no homeomorphism allowing the transformation from 2-simplex mesh coordinates to triangular mesh coordinates.

Simplicial meshes feature a constant number of vertices per face, while simplex meshes feature a constant vertex connectivity. To create dual polygonal meshes, centers of adjacent simplex faces need to be linked (*i.e.* in case of 2-simplex meshes).

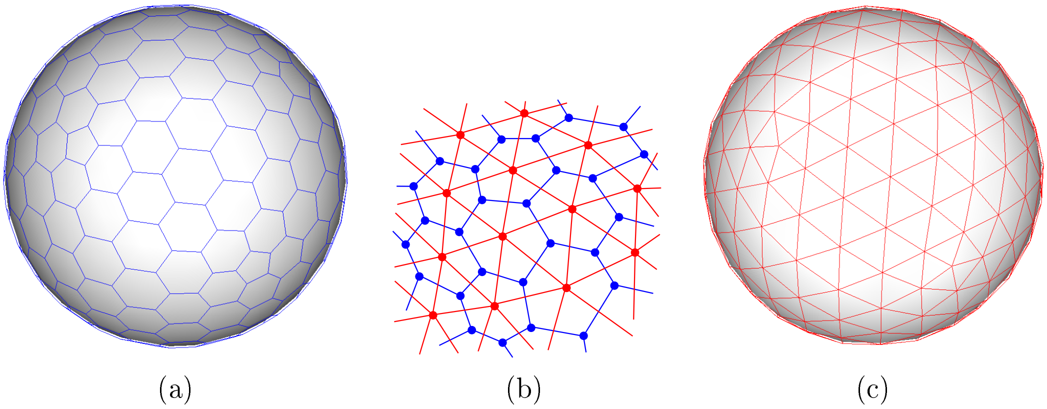


Figure 3.6: Duality shown on a sphere between (a) a 2-simplex mesh and (c) its dual triangular mesh. To create such a duality, (b) the centers of simplex faces (in blue), which become triangular vertices (in red), need to be linked.

2-simplex meshes	$\Longleftrightarrow$	triangular meshes
vertex	$\Longleftrightarrow$	face
edge	$\Longleftrightarrow$	edge
triangle	$\Longleftrightarrow$	vertex

Table 3.1: Duality between 2-simplex meshes and triangular meshes.

Converting simplex mesh to triangular mesh is important because triangular mesh is the standard format supported by graphics hardware.

Vertex neighborhood The set of vertices that are topologically separated from vertex $\mathbf{p}_i$ by a distance of at most η vertices constitutes the η-th order neighborhood of $\mathbf{p}_i$, noted $N_\eta(i)$ (see examples in Figure 3.7). This means there is a path of linking edges, whose length is at most η vertices, between each vertex from $N_\eta(i)$ and $\mathbf{p}_i$. The number of vertices belonging to neighborhood $N_\eta(i)$ is denoted by $|N_\eta(i)|$.

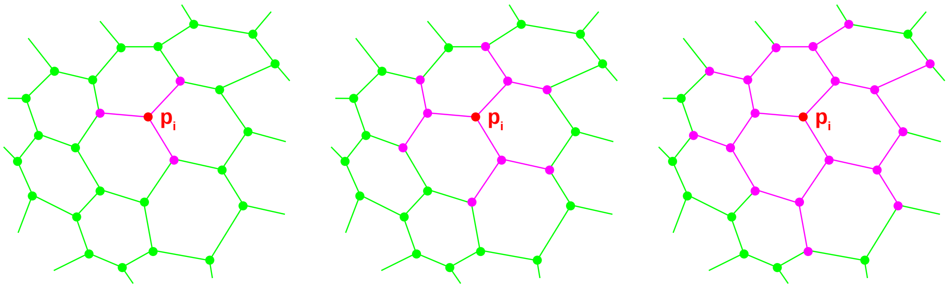
3.3.2.2 Geometry

In this section, 1-simplex meshes are first depicted to illustrate the geometrical concepts of simplex meshes. However, the geometry of 2-simplex meshes is more of interest since they are more commonly used for the segmentation of medical images.

1-simplex meshes In $\mathbb{R}^2$, a 1-simplex mesh is a polygonal line of the plane (see Figure 3.8, left).

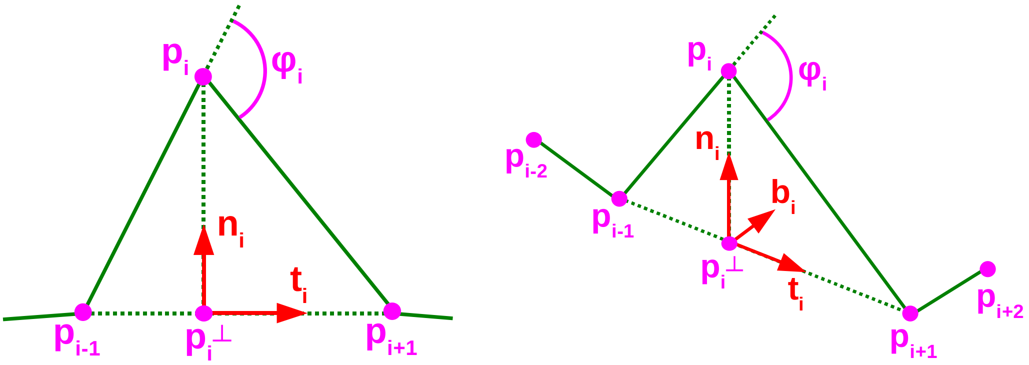
The discrete normal $\mathbf{n}_i$ and discrete tangent $\mathbf{t}_i$ at vertex $\mathbf{p}_i$ are defined as:

$$\mathbf{n}_i = \frac{\mathbf{p}_i^\perp \mathbf{p}_i}{\|\mathbf{p}_i^\perp \mathbf{p}_i\|} \quad \text{and} \quad \mathbf{t}_i = \frac{\mathbf{p}_{i-1}\mathbf{p}_{i+1}}{\|\mathbf{p}_{i-1}\mathbf{p}_{i+1}\|} \tag{3.3}$$



(a) (b) (c)

Figure 3.7: Different η-th order neighborhoods for vertex $\mathbf{p}_i$: (a) 1st order, (b) 2nd order, and (c) 3rd order.



(a) (b)

Figure 3.8: Specific geometry of 1-simplex meshes (a) in $\mathbb{R}^2$ and (b) in $\mathbb{R}^3$. In both cases, the local geometry at each vertex $\mathbf{p}_i$ is defined from its neighbors: φ_i is the simplex angle, $\mathbf{n}_i$ and $\mathbf{t}_i$ are the discrete normal and discrete tangent, respectively. In $\mathbb{R}^3$, $\mathbf{b}_i$ is the binormal vector used to build a local base at $\mathbf{p}_i$.

where $\mathbf{p}_i^{\perp}$ is the orthogonal projection of $\mathbf{p}_i$ on the line segment created by its two neighbors $\mathbf{p}_{i-1}$ and $\mathbf{p}_{i+1}$.

Metric parameters $(\varepsilon_i^1, \varepsilon_i^2)$ of $\mathbf{p}_i$ are defined as:

$$\varepsilon_i^1 = \frac{\|\mathbf{p}_i^{\perp}\mathbf{p}_{i+1}\|}{\|\mathbf{p}_{i-1}\mathbf{p}_{i+1}\|} \quad \text{and} \quad \varepsilon_i^2 = \frac{\|\mathbf{p}_{i-1}\mathbf{p}_i^{\perp}\|}{\|\mathbf{p}_{i-1}\mathbf{p}_{i+1}\|} = 1 - \varepsilon_i^1 \tag{3.4}$$

Simplex angle φ_i, which follows the classical notion of angle between two segments, is defined as:

$$\varphi_i = (\mathbf{p}_{i-1}, \widehat{\mathbf{p}_i}, \mathbf{p}_{i+1}) \tag{3.5}$$

From simplex angle φ_i, the discrete curvature of the mesh is defined as:

$$k_i = 2\frac{\sin(\varphi_i)}{\mathbf{p}_{i-1}\mathbf{p}_{i+1}} \tag{3.6}$$

In $\mathbb{R}^3$, the binormal vector $\mathbf{b}_i$ is defined for each vertex to build a local base at $\mathbf{p}_i$ (see Figure 3.8, right):

$$\mathbf{b}_i = \frac{\mathbf{p}_i\mathbf{p}_{i+1} \times \mathbf{p}_{i-1}\mathbf{p}_i}{\|\mathbf{p}_i\mathbf{p}_{i+1} \times \mathbf{p}_{i-1}\mathbf{p}_i\|} \tag{3.7}$$

from which discrete normal $\mathbf{n}_i$ may be computed as $\mathbf{n}_i = \mathbf{b}_i \times \mathbf{t}_i$.

2-simplex meshes The following parameters are illustrated in Figure 3.9.

Let C_i be the circumscribed circle to the three vertices $(\mathbf{p}_{V_1(i)}, \mathbf{p}_{V_2(i)}, \mathbf{p}_{V_3(i)})$ with center $\mathbf{c}_i$ and radius r_i.

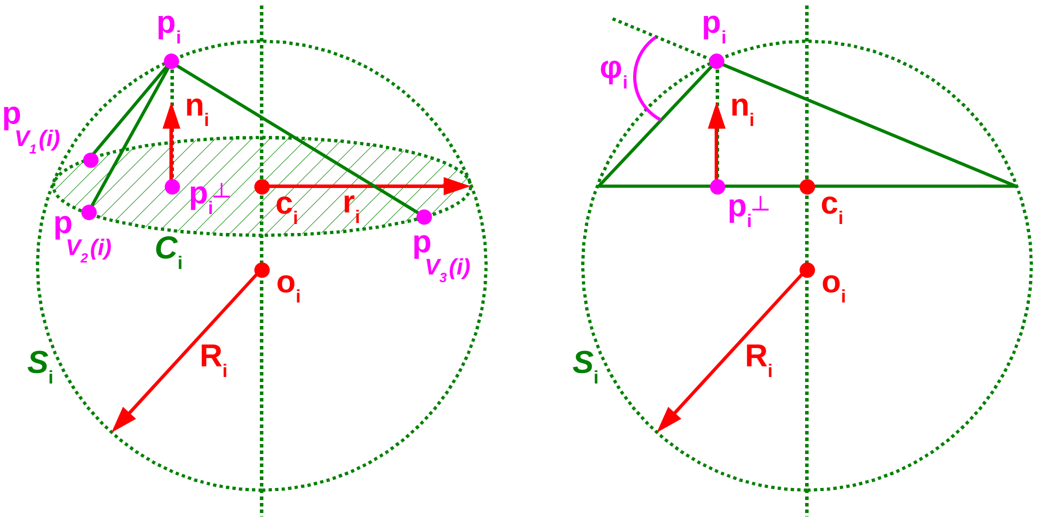
Let S_i be the circumscribed sphere to the tetrahedron $\{\mathbf{p}_i, \mathbf{p}_{V_1(i)}, \mathbf{p}_{V_2(i)}, \mathbf{p}_{V_3(i)}\}$ with center $\mathbf{o}_i$ and radius R_i.

The three neighbors of $\mathbf{p}_i$ create a plane normal to the vector $\mathbf{n}_i$:

$$\mathbf{n}_i = \frac{\mathbf{p}_{V_1(i)} \times \mathbf{p}_{V_2(i)} + \mathbf{p}_{V_2(i)} \times \mathbf{p}_{V_3(i)} + \mathbf{p}_{V_3(i)} \times \mathbf{p}_{V_1(i)}}{\|\mathbf{p}_{V_1(i)} \times \mathbf{p}_{V_2(i)} + \mathbf{p}_{V_2(i)} \times \mathbf{p}_{V_3(i)} + \mathbf{p}_{V_3(i)} \times \mathbf{p}_{V_1(i)}\|} \tag{3.8}$$

Metric parameters $(\varepsilon_i^1, \varepsilon_i^2, \varepsilon_i^3)$ are the barycentric coefficients of $\mathbf{p}_i^{\perp}$ (*i.e.* the projection of $\mathbf{p}_i$ on the plane created by its three neighbors).

Simplex angle φ_i, which is a geometric angle in the plane $(\mathbf{p}_i, \mathbf{c}_i, \mathbf{o}_i)$, is defined by the following relations:



(a) (b)

Figure 3.9: Specific geometry of 2-simplex meshes in $\mathbb{R}^3$: (a) position of a simplex vertex $\mathbf{p}_i$, its three neighbors $(\mathbf{p}_{V_1(i)}, \mathbf{p}_{V_2(i)}, \mathbf{p}_{V_3(i)})$ and its orthogonal projection $\mathbf{p}_i^\perp$ on the plane created by its three neighbors, and (b) focus on its simplex angle φ_i.

- $\varphi_i \in [-\pi, +\pi]$
- $\sin(\varphi_i) = \frac{r_i}{R_i} \operatorname{sign}(\mathbf{p}_i\mathbf{p}_{V_1(i)} \cdot \mathbf{n}_i)$
- $\cos(\varphi_i) = \frac{\|\mathbf{o}_i\mathbf{c}_i\|}{R_i} \operatorname{sign}(\mathbf{o}_i\mathbf{c}_i \cdot \mathbf{n}_i)$

3.3.2.3 Evolution

During segmentation, simplex meshes are deformed at each vertex $\mathbf{p}_i$ by means of *internal* and *external* forces, noted f_{int} and f_{ext}, respectively:

$$\mathbf{p}_i^{t+\Delta t} = \mathbf{p}_i^t + (1-\tau)(\mathbf{p}_i^t - \mathbf{p}_i^{t-\Delta t}) + \alpha f_{int}(\mathbf{p}_i^t) + \beta f_{ext}(\mathbf{p}_i^t) \tag{3.9}$$

where τ is the milieu viscosity, which is constant at each vertex, α and β are the weights for internal and external forces, respectively, and Δt is the time step integrated to parameters α, β and τ.

These parameters need to keep values small enough to ensure a stable explicit scheme.

Globally constrained deformations The amount of local deformation at each vertex may be controlled at a global level using Globally Constrained Deformations (GCD) [Montagnat & Delingette 1998]. Both deformable model and registration approaches are coupled in a general reconstruction framework based on the Iterative Closest Point (ICP) algorithm [Zhang 1994]. First, the model is globally registered. As the model converges, the number of degrees of freedom are then increased using the GCD scheme. The amount of deformation is gradually increased from affine transformations to Free Form Deformations (FFD). Using deformable models enables the use of both an *a priori* shape and relevant geometrical information throughout the deformation process.

3.3.2.4 Mesh quality improvement

Quality criteria After deformation, a mesh is expected to accurately adopt the shape of the structure of interest. Furthermore, the mesh needs to have good *topological* and *geometrical* attributes, so that an usable geometric representation of the structure of interest may be provided.

The topological quality of a mesh is defined as the regularity of the number of edges within each face. Ideally, a mesh should have regular faces with six vertices. For simplex meshes, six topological operations on k-simplex meshes, which are referred to as $\{TO_1^k, TO_2^k, \ldots, TO_6^k\}$, have been proposed to improve topological quality [Delingette 1999].

For instance, TO_1^k and TO_2^k do not modify mesh connectivity and are used to refine or decimate the mesh. Operations TO_1^k and TO_2^k consist in adding or removing a vertex (respectively, an edge) of a 1-simplex mesh (respectively, a 2-simplex mesh). Both operations are inverse of each other.

An example of application is the exchange operation, which consists in changing the repartition of edges across faces (see an illustration in Figure 3.10). This application combines TO_1^k and TO_2^k operations [Delingette 1999]. For dual triangular meshes, this corresponds to the swapping operator applied on two adjacent triangles.

The geometrical quality of a mesh depends on the vertex arrangement regularity. A uniform distribution of vertices where faces constitute regular polygons are preferred, rather than a distribution allowing lengthened or deformed faces. The geometrical quality preservation of meshes during deformation is a major problem in explicit model-based segmentation.

With simplex meshes, geometrical quality may be optimized by using metric parameters or by refining/decimating mesh surface [Montagnat 1999]. In the former case, metric parameters ε_i^j are adjusted at each vertex so that the area of all faces tend to be equal. In the latter case, geometrical quality is optimized by adding vertices in sparse zones while removing vertices in high concentration zones.

Global refinement The level of details that is used for a mesh usually depends on the specific level of details featured by the data to segment. Depending on the application, a precision criterion needs to be decided, knowing that a fine mesh leads to more computation time than a coarse mesh.

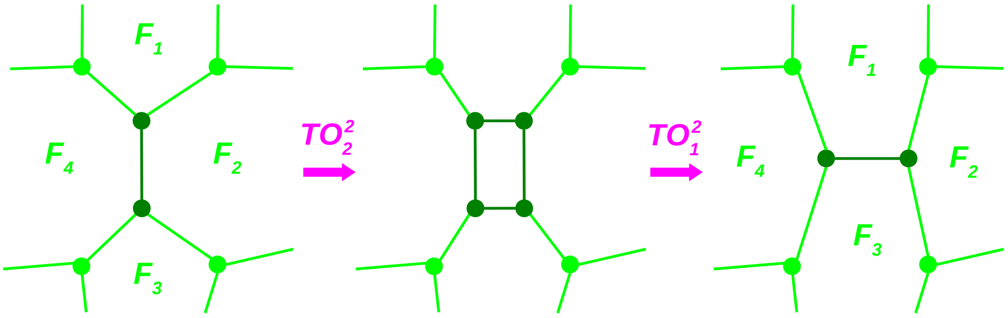


Figure 3.10: Exchange operation on 2-simplex meshes, which consists in changing the repartition of edges across faces using a combination of TO_1^k and TO_2^k operations.

A procedure consisting in using a coarse-to-fine algorithm is a standard approach [Schmid & Magnenat-Thalmann 2008, Lachaud & Montanverta 1999]. In this case, computation time is reduced by using first a coarse mesh for the important deformations and then a fine mesh for the last remaining deformations (*i.e.* when the mesh has almost converged).

With simplex meshes, a global refinement algorithm using a topologically dual operation on triangular meshes has been proposed [Montagnat 1999]. After one iteration, the number of faces is approximately tripled, should the surface be homeomorphic to a sphere. This refinement operation does not alter the topological and geometrical qualities of the mesh. In the refined mesh, a face is created at the center of each former face with the same number of edges. A face with exactly six edges is also created around each former vertex (see Figure 3.11 for an example on a liver mesh).

However, this method has two drawbacks. First, shape is not accurately preserved because low resolution vertices are not conserved in the higher resolutions (*i.e.* due to curvature averaging). Second, this method does not allow the constraining of vertices at specific locations, which is problematic when dealing with mesh attachments (*e.g.* tendon attachments on bones). To resolve this problem, a multi-resolution scheme based on the tessellation of dual triangular meshes was proposed [Gilles 2007].

Local refinement As previously mentioned, a less dense mesh is usually preferred to reduce computation time. However, mesh surface needs also to have a minimum level of details to correctly describe the shape of the structure. By finding a tradeoff between coarse and fine, geometrical criteria may be used to optimize mesh size. For instance, flat surfaces need only few vertices and big faces to be represented while high curvature zones require a dense mesh to take curvature details into account.

Therefore, a common criterion consists in gathering vertices in high curvature zones of the mesh (see Figure 3.12). However, deformations during segmentation have a tendency to stretch the mesh in some zones or concentrate it in other zones.

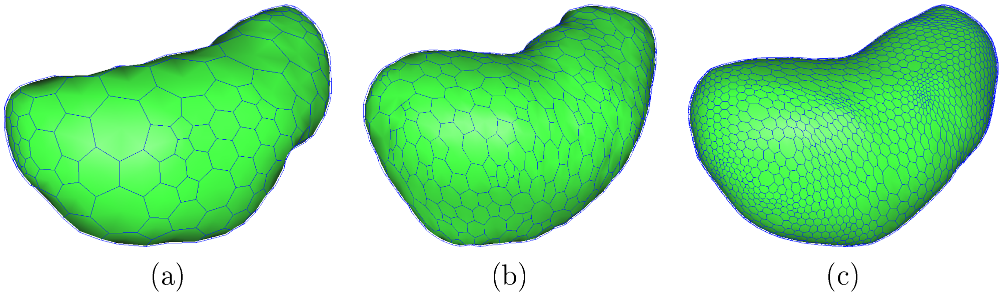

(a) (b) (c)

Figure 3.11: Decimation and refinement of a liver 2-simplex mesh: (b) original resolution, (a) resolution after two decimating iterations, and (c) resolution after two refining iterations.

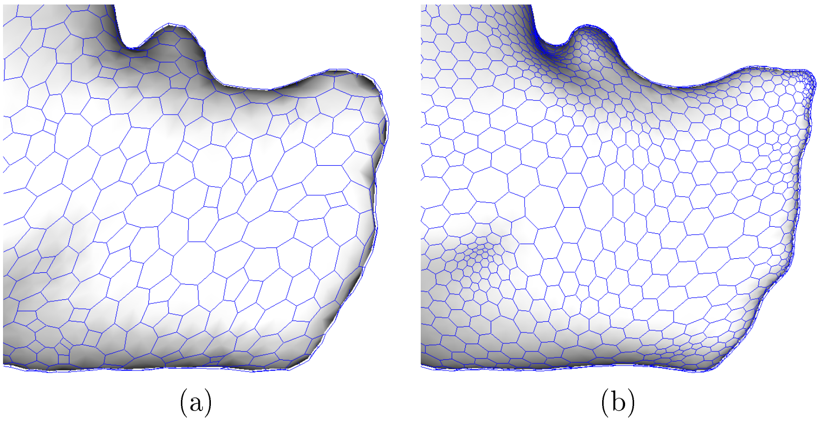

(a) (b)

Figure 3.12: Local refinement with 2-simplex meshes: (a) original mesh with a uniform distribution of vertices, and (b) mesh after gathering vertices in high curvature zones.

A uniform distribution of vertices leads thus to either less dense zones, or high concentration zones.

Another mesh quality optimization consists in distributing vertices in a more uniform way, either by making them slide along the surface, or by removing vertices in high concentration zones and creating them in sparse zones.

3.3.3 Model initialization

In medical image analysis, many structures of interest have a typical shape and location inside the body. Medical devices may also give information about the position and orientation of the patient. Gathering such information gives thus an idea of where the structure is located into the image, meaning the deformable model may be approximately initialized.

However, initial shape and location of the structure may not be known. A first solution consists in interactively placing the mesh into the image. For example, the mesh may be automatically displayed in the center of the image and then, the user is asked to align the mesh with the structure of interest, using translation, rotation and scale [Heimann 2009]. Though quite reliable, interactive methods may become a difficult task with a large number of datasets. In this case, automatic initialization procedures are to be used.

For example, structures may be roughly segmented to give an *a priori* of the position (*e.g.* using thresholding and morphological operations). As a result, a binary image containing the rough segmentation as foreground is generated. From this binary image, a simplex mesh may be generated and serve as initialization. Then, the mesh is normally deformed using external forces, which are supposed to compensate errors that may have occurred during the rough initialization step [Delingette 1998].

In the literature, several methods have been proposed for model initialization, usually depending on the application. With Statistical Shape Models (SSM), initialization may be done by performing a global search of the shape model in the image using methods based on the evolutionary algorithm [Heimann *et al.* 2007] and evolution strategies [Hamarneh *et al.* 2002, Schwefel 1993]. An initialization guided by rotation-invariant Gabor features has been proposed for prostate segmentation [Shen *et al.* 2003]. For the segmentation of patella cartilage, its strong spatial relationship to the underlying bone is used for the initialization [Fripp *et al.* 2005].

3.3.4 Internal forces

During mesh deformation, internal forces ensure mesh surface to be smooth and less sensible to noise present in images. Thus, they have a regularizing effect because they force the mesh to deform while respecting regularity criteria.

For continuous models, internal forces are mainly based on regularization terms using second order Tikhonov stabilizers [McInerney & Terzopoulos 1995a]. This is because these stabilizers express shape deformation through differential terms up to the second order. Differential terms are easy to compute because of the continuous nature of explicit models. However, this task is more difficult with discrete models since geometric approximations are to be made.

In this section, we describe some criteria used with internal forces, first on deformable contours, then on deformable surfaces. Detailed descriptions of internal forces may be found in [Gilles 2007, Montagnat 1999].

3.3.4.1 Deformable contours

Let $\mathcal{C}$ be an explicit deformable contour varying on interval Ω with r as parameter. Arc length s of $\mathcal{C}$ is linked to parameter r by metric g defined as:

$$s(r) = \int_0^r g(u)du \quad \text{with} \quad g = \left\| \frac{\partial \mathcal{C}}{\partial r} \right\| \tag{3.10}$$

where $u \in [0, r]$.

Many different regularizing internal forces have been proposed in the literature [Kimia *et al.* 1992, Cohen 1991, Kass *et al.* 1988]. They usually take contour curvature or higher order differential features into account. Here are the most common internal forces.

Mean curvature

$$E_{reg} = \int_{\Omega} \left\| \frac{\partial C}{\partial r} \right\| dr \tag{3.11}$$

This energy depends on curve's length. The Euler-Lagrange equation, when applied on a contour subject to this energy, leads to the expression of internal forces:

$$f_{int}(r) = \kappa(r)\mathbf{n}(r) \tag{3.12}$$

This force, which tends to minimize the curve's length, is directed towards the inside of a closed contour and commonly used with level sets [Ambrosio & Soner 1996].

Laplacian smoothing The elastic energy of the contour is defined as:

$$E_{reg} = \int_{\Omega} \left\| \frac{\partial C}{\partial r} \right\|^2 dr \tag{3.13}$$

The elastic energy minimization leads to the expression of internal forces:

$$f_{int}(r) = \frac{\partial^2 C}{\partial r^2} \tag{3.14}$$

The corresponding Lagrangian evolution equation is a diffusion equation:

$$\frac{\partial C}{\partial t} = \frac{\partial^2 C}{\partial r^2} \tag{3.15}$$

where t is the diffusion time.

At equilibrium, we have:

$$\frac{\partial^2 C}{\partial r^2} = 0 \tag{3.16}$$

Laplacian smoothing is a standard method in which vertices are attracted towards the barycenter of their neighbors. This method aims at minimizing the elastic energy of the model. Well-known drawback is that this method is likely to make the model shrink. Solutions consist in compensating Laplacian smoothing forces with balloon forces or removing their normal contributions (*i.e.* their tangential contributions preventing the model from stretching) [Montagnat *et al.* 2001].

Bending energy With bending energy, forces that rely on the averaging of the discrete curvature in the normal direction are generated.

For a thin rod, bending energy minimization defined as:

$$E_{reg} = \int_{\Omega} \left\| \frac{\partial^2 C}{\partial r^2} \right\|^2 dr \tag{3.17}$$

leads to the expression of internal forces:

$$f_{int}(r) = -\frac{\partial^4 C}{\partial r^4} \tag{3.18}$$

To estimate the curvature, an average of the normal contribution of Laplacian forces may be performed [Park *et al.* 2001]. Other solutions consist in smoothing the simplex angle [Montagnat *et al.* 2001] or the Gaussian curvature [Ghanei *et al.* 1998].

Balloon force This force uses a positive constant k, which corresponds to the minimization of the area delineated by a closed contour [Cohen 1991].

Balloon force is defined as:

$$f_{int}(r) = -k\ \mathbf{n}(r) \tag{3.19}$$

Internal force may be decomposed in tangential and normal components:

$$f_{int}(r) = f_{tan}(r)\ \mathbf{t}(r) + f_{nor}(r)\ \mathbf{n}(r) \tag{3.20}$$

Tangential component affects the parametrization of contour $\mathcal{C}$, but not its shape [Kimia *et al.* 1992]. This is why only the normal component is taken into account in level sets, in which no parametrization is required. However, this parametrization has a great impact for explicit discrete models.

Among the four aforementioned internal forces, mean curvature and balloon force have a tangential component that is null.

3.3.4.2 Deformable surfaces

As for muscles in section 7.2.4, main anatomical structures are considered as having smooth shapes in medical image analysis. When information about the structure of interest is not available, internal forces ensure a geometric continuity of the mesh. If the shape is known, this information is usually introduced into the deformation process. To this end, a shape constraint that forces the mesh to take back its initial shape may be used.

In triangular meshes, [Ghanei *et al.* 1998] define a deformable model with a discrete curvature notion at vertices from which they derive shape regularizing forces. Though deformable models implicitly handle deformation discontinuities at boundaries, this is not the case with registration based on Free Form Deformations (FFD), in which boundaries are embedded into the deformable volumetric lattice [Sederberg & Parry 1986].

Simplex meshes Metric parameters ε_i^j ($\sum_j \varepsilon_i^j = 1, \forall j, 0 < \varepsilon_i^j < 1$) ensure the control of the relative spacing between a vertex $\mathbf{p}_i$ and its neighbors (see Figure 3.13).

Let $\tilde{\mathbf{p}}_i$ be an ideal position defining the internal force at $\mathbf{p}_i$, and $\{\tilde{\varepsilon}_i^j, \tilde{\varphi}_i\}$ the metric parameters and simplex angle at vertex $\tilde{\mathbf{p}}_i$, respectively.

Let $\tilde{\mathbf{p}}_i^\perp$ be the projection of $\tilde{\mathbf{p}}_i$ on the plane tangent to $\mathbf{p}_i$.

Let $\mathbf{r}_i$ be an approximate binormal vector at $\mathbf{p}_i$ defined as:

$$\mathbf{r}_i = \frac{\mathbf{t}_i \times (\mathbf{p}_{i-2}\mathbf{p}_{i-1} \times \mathbf{p}_{i+1}\mathbf{p}_{i+2})}{\|\mathbf{t}_i \times (\mathbf{p}_{i-2}\mathbf{p}_{i-1} \times \mathbf{p}_{i+1}\mathbf{p}_{i+2})\|} \tag{3.21}$$

Internal force is defined by:

$$f_{int}(\mathbf{p}_i) = \alpha \mathbf{p}_i\tilde{\mathbf{p}}_i = \alpha(f_{tan}(\mathbf{p}_i) + f_{nor}(\mathbf{p}_i)) \tag{3.22}$$

where f_{tan} and f_{nor} are the *tangential* and *normal* components, respectively (see Figure 3.13):

$$f_{tan}(\mathbf{p}_i) = \tilde{\mathbf{p}}_i^\perp - \mathbf{p}_i^\perp = \sum_j (\varepsilon_i^j - \tilde{\varepsilon}_i^j)\mathbf{p}_{V_j(i)} \tag{3.23}$$

$$f_{nor}(\mathbf{p}_i) = (h(r_i, \varepsilon_i^1, \varepsilon_i^2, \varphi_i) - h(r_i, \tilde{\varepsilon}_i^1, \tilde{\varepsilon}_i^2, \tilde{\varphi}_i))\mathbf{n}_i \tag{3.24}$$

Tangential component f_{tan} is a restoring force from $\mathbf{p}_i^\perp$ towards $\tilde{\mathbf{p}}_i^\perp$, which ensures a control of the relative spacing between vertices. Normal component f_{nor} is a force that controls the shape of the mesh.

For a uniform distribution of vertices along mesh surface, metric parameters may be defined as $\tilde{\varepsilon}_i^j = \frac{1}{3}, \forall j$. When metric parameters are fixed, normal component f_{nor} ensures the control of surface regularity.

With simplex meshes, whose topology is dual to triangular meshes, *simplex angle regularity* and *shape memory* constraints may be used as internal forces.

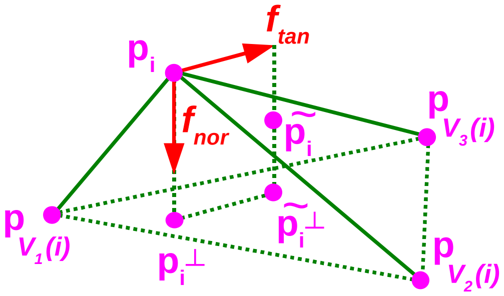


Figure 3.13: Focus on the tangential f_{tan} and normal f_{nor} components of the internal force applied on vertex $\mathbf{p}_i$ when using 2-simplex meshes.

Simplex angle regularity constraint Internal forces may be used for smoothing purposes on simplex meshes. A simplex angle regularity constraint that tends to force each vertex to a configuration defined by its neighbors has been proposed [Montagnat 1999]. This constraint, known as *rigidity*, acts with a scale parameter that specifies the range of this neighborhood and hence how spread is the smoothing. In practice, this constraint imposes a certain regularity of the simplex angle φ_i and thus of the surface's discrete curvature.

For each vertex $\mathbf{p}_i$, the new simplex angle $\tilde{\varphi}_i$ is a weighted sum of its neighboring values φ_j:

$$\tilde{\varphi}_i = \sum_{\mathbf{p}_j \in N_\eta(i)} \frac{\varphi_j}{|N_\eta(i)|} \tag{3.25}$$

where $N_\eta(i)$ is the η-th order neighborhood of vertex $\mathbf{p}_i$ and $|N_\eta(i)|$ is the number of vertices belonging to $N_\eta(i)$ (see section 3.3.2.1).

Resulting internal forces tend to impose a constant simplex angle continuity constraint to the mesh (*i.e.* a sphere for 2-simplex meshes and a circle for 1-simplex) [Delingette 1994]. Unlike the classical energy minimization of a thin plate where the area (respectively, the volume) of the contour (respectively, the surface) is minimized, the mesh does not shrink towards a point when there is not data attachment term (*i.e.* external forces).

The topological size parameter of the neighborhood η corresponds to the surface rigidity. Mesh deformations are more local with a small rigidity parameter η. Inversely, surface is affected in a larger range when η increases. The higher the rigidity, the smoother the surface.

Shape memory constraint A simplex mesh may converge towards its initial shape if both metric parameters and simplex angles (see section 3.3.2.3) are imposed at the beginning of the deformations. For instance, when there is no data attachment terms, an iterative scheme in which the mesh is converging towards its initial shape has been proposed [Montagnat 1999].

Only metric parameters and simplex angles, which are similarity transform invariant, are necessary for shape to be locally recovered. Although shape memory is a prior based on a single representative shape, the similarity invariance property confers more flexibility than encoding the shape as 3D points. This simple prior is similar to the notion of strain energy in mechanics.

PCA-based methods A set of P training shapes $\{\mathbf{S_1}, \ldots, \mathbf{S_P}\}$ with point correspondence (*i.e.* each landmark is located at the same position through all the P training shapes) are necessary with PCA-based methods using Point Distribution Models (PDM) [Cootes & Taylor 2004]. This point correspondence is usually ensured through a registration between training shapes. Then, the P training shapes are aligned in a common coordinate frame. The most popular method to solve this problem is the General Procrustes Analysis (GPA) [Goodall 1991, Gower 1975], which aligns the set of P training shapes to their unknown mean by minimizing the mean squared distance between two shapes in an iterative procedure.

Then, a Statistical Shape Model (SSM) [Cootes & Taylor 2001] is built by reducing the dimensionality of the training set with the aim of finding a small set of modes that best describes the observed variation. This is accomplished using PCA [Jolliffe 2002].

An arbitrary shape $\mathbf{S}$ is approximated from the computed statistics by:

$$\mathbf{S} \approx T(\overline{\mathbf{S}} + \Phi\ \mathbf{b}) \tag{3.26}$$

where vector $\overline{\mathbf{S}}$ is the mean shape, Φ is a matrix of L ($L \leq P$) principal modes (with respective variances λ_i), $\mathbf{b}$ is a vector of shape parameters and T denotes the alignment transform.

An iterative procedure may be used to estimate the unknown parameters $\mathbf{b}$ and T [Cootes & Taylor 1993]. To ensure the SSM specificity, two kinds of constraints may be considered: *hard* or *soft*.

Hard constraint is defined as:

$$-3\sqrt{\lambda_m} \leq b_m \leq 3\sqrt{\lambda_m}, \forall m \in [1, L] \tag{3.27}$$

Soft constraint scales b_m to lie inside a hyperellipsoid:

$$\sum b_m^2 / \lambda_m \leq C, \forall m \in [1, L] \tag{3.28}$$

where C is computed from the χ^2_L distribution [Cootes & Taylor 1993].

In addition, a multiresolution scheme may be used to produce various levels of detail (LOD) of the shapes [Gilles *et al.* 2006]. The LOD are then exploited in a coarse-to-fine fashion, improving the robustness and accuracy of the segmentation evolution. When shape priors are based on Statistical Shape Models (SSM), the mesh is constrained to only adopt variations expressed by the SSM [Heimann 2009]. In case of simplex meshes, both precomputed and current simplex parameters are used to derive new target point positions.

3.3.5 External forces

During segmentation, deformation is driven by external forces whose role is to attract the mesh to image-based features while ensuring the stability of the numerical resolution. This is why external forces are referred to *data attachment terms*. Different techniques have been presented in the literature, such as gradient forces, region-based approaches and intensity profiles. These techniques ensure the mesh to deal with various situations, depending on image features (*e.g.* intensity distribution and noise). This makes image segmentation from different modalities and multimodal registration possible.

Creating external forces requires a clear definition of boundaries between structures of interest. These boundaries usually appear where there are strong intensity differences, in between two different types of tissue. Various filters have been extensively used to highlight these boundaries. However, the mesh may be attracted to unexpected boundaries close to a structure of interest, which is likely to jeopardize the final segmentation result. To overcome this issue, boundaries need to be discriminant enough, *e.g.* by introducing more *a priori* knowledge about the structure and its neighbors.

3.3.5.1 Gradient forces

External forces may be computed as the gradient of the potential field $\|\nabla I\|^2$ [Kass *et al.* 1988]. Applying convolution to an image with Gaussian standard deviation σ, the response zone of the gradient operator around boundaries is controlled.

In this case, the external force is formulated as:

$$f_{ext} = \nabla(\|\nabla I \otimes G_\sigma\|^2) \tag{3.29}$$

This equation features several drawbacks. First, Gaussian operators (as well as Gaussian filters) are known to blur structure boundaries. This is not the case with anisotropic filters. The attraction potential around boundary points detected by the gradient operator has usually a small spatial range. As a consequence, the mesh may converge only if close to the data.

Second, the force as expressed in equation 3.29 is a function that does not depend on the distance to the data. This means that the force may not be zero when computed at boundaries, which leads to oscillations (*i.e.* mesh will not converge). A solution consists in progressively stabilizing the mesh by reducing the weighting coefficient of external forces β each time a vertex changes of direction [McInerney & Terzopoulos 1995b].

Finally, second derivatives are needed for the computation of this force. In case of noisy image, this computation is likely to be unstable.

As a solution, external forces may be normalized, not only to avoid important displacements but also to stay in the stability range of the discrete resolution:

$$f_{ext} = \frac{\nabla(\|\nabla I \otimes G_\sigma\|^2)}{\|\nabla(\|\nabla I \otimes G_\sigma\|^2)\|} \tag{3.30}$$

Using this new equation, stability is ensured and oscillations are limited to an amplitude smaller than voxel size. However, at a cost of a slower convergence speed.

Another issue with gradient information is that the mesh may converge to boundaries that are *local maxima*. These boundaries may be undesired noisy edges too. A solution consists in regularizing the gradient vector field with the Gradient Vector Flow [Xu & Prince 1998]. To make gradient information more discriminant, contour information may be added to local gradient information [Delingette 1999, McInerney & Terzopoulos 1993]. Distance maps, which may be extracted using the Euclidean distance or Chamfer masks, have been proposed to find the closest contour point from each voxel [Cohen 1992]. Contour force is thus defined as a function of the distance to the closest data points.

Another method consists in looking for a contour voxel along the normal at each vertex [Delingette 1999]. This method has the advantage to reduce the range of external forces. As a result, deformations are smoother without limiting model deformations. Searching for attracting points at a great scale, the contour force allows a fast convergence of the model. This force is homogeneous to the distance to data points and does not lead to oscillations.

In medical image analysis, gradient is the information commonly used to determine structure contours [Rosenfeld & Kak 1982]. A map containing main contours is usually generated using the extremum of the gradient norm in the direction of the gradient [Delingette 1999, McInerney & Terzopoulos 1993, Cohen 1992]. The whole gradient norm image may be used too [Montagnat 1999].

3.3.5.2 Region-based approaches

Should gradient-based information not be discriminant enough for extracting image contour points, a more robust information about structures to segment needs to be used. Though using contour extraction operators that are robust to noise may be a solution [Liu *et al.* 1995], other methods consist in looking for homogeneous regions in the image. Indeed, it has been shown in the literature that region information is more meaningful than gradient information [Zhu & Yuille 1996, Cocquerez & Philipp 1995, Ronfard 1994].

For instance, a method consisting in detecting regions of similar intensity has been proposed [Cohen & Cohen 1993]. This method features a region detector and an energy taking both these regions and regions outside the deformable contour into account.

Another method consists in using a force based on the detection of homogeneous intensity regions by means of both intensity and gradient profiles [Montagnat 1999]. In this case, region information is considered as a discriminative criterion ensuring the filtering of contour points. Gradient is extracted from the image as explained in previous section. Since gradient information is known, intensity profiles sampled along the normal may be filtered. The objective of using gradient intensity is to define zones not to be filtered, so that boundaries are not blurred (*i.e.* similarly to anisotropic filters). In such approach, only high gradient points closed to a region boundary are kept. The force is defined as a function of the distance between a vertex and its corresponding boundary point. In practice, only one region is necessary, corresponding either to the segmented structure (*i.e.* inside the mesh) or a neighboring structure (*i.e.* outside the mesh).

In the literature, other region-based approaches include the use of *histograms* and *texture*.

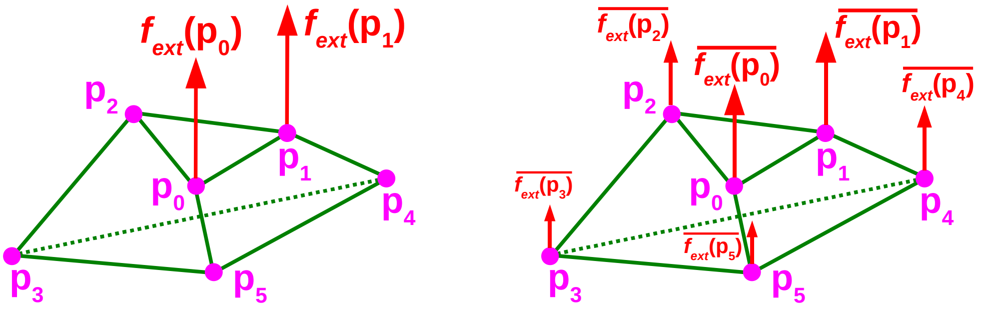
Histograms may be used to represent the different regions inside an image and their intensity properties [Paragios & Deriche 2000]. In this case, intensity regions are modeled by a mixture of Gaussian elements using a statistical analysis based on the Minimum Description Length criterion and the Maximum Likelihood Principle.

Textured images are characterized by local correlations between intensity values [Cremers *et al.* 2007]. The structure tensor, which is also called *second order moment matrix*, has been proposed to capture these local correlations [Bigün *et al.* 1991]. The structure tensor is defined by a matrix of partial derivatives smoothed by a Gaussian kernel. Despite its good properties for texture discrimination, the structure tensor is invariant to intensity changes. An extension of this work is the scale introduction via Total Variation (TV) flow, which combines texture features at different scales [Brox & Weickert 2004]. More sophisticated texture features have also been proposed [Leung & Malik 2001].

3.3.5.3 Simplex meshes

As previously described in section 3.3.5, external forces ensure the displacement of vertices towards image features during deformation. Image features include boundaries extracted by image gradient or voxel intensity.

In practice, external forces are expressed as a field of displacement vectors. This field is usually not smooth at all, but rather irregular, due to the presence of numerous false positives or false negatives detected during the boundary extraction



(a) (b)

Figure 3.14: External forces coupled with 2-simplex meshes are expressed as a field of displacement vectors. Due to the presence of numerous false positives or false negatives detected during the boundary extraction operation, (a) this field is usually irregular. With simplex meshes, (b) this field is regularized.

operation (see Figure 3.14, left).

With simplex meshes, the field of displacement vectors may be regularized using the average applied on the neighborhood of size η (see Figure 3.14, right):

$$\forall i, \overline{f_{ext}(\mathbf{p}_i)} = \frac{1}{|\nu_\eta(i)|} \sum_{\mathbf{p}_i \in \nu_\eta(i)} f_{ext}(\mathbf{p}_j) \tag{3.31}$$

The rigidity parameter η defines the neighborhood on which the local external force is smoothed.

3.3.5.4 Intensity profiles

Structures of interest, or their neighboring structures, may not be defined by homogeneous regions (*e.g.* presence of textures or organs whose intensity is variable along boundaries). By providing a finer information, intensity profiles have been proposed to make external forces more robust. They provide a local information about the intensity evolution along the normal at each vertex. Typically, for every sampling step, the intensity is sampled at one point along the normal using voxel interpolation, thus the profile does not feature any specific width. However, intensity profiles may also be constructed by averaging a number of pixels in the neighborhood of each sampling step [Brejl & Sonka 2000].

Intensity profiles were among the first image representation tools used to describe appearance for segmentation purposes. They may extend inward, outward or both sides, depending on the application (see Figure 3.15). In computer vision, intensity profiles act as feature vectors that could be replaced by any other local or global features (*e.g.* isophote curvature, texture descriptors and oriented filters).

For instance, intensity profiles are used to build Statistical Appearance Models (SAM) [Cootes & Taylor 1994, Cootes *et al.* 1993]. Intensity profiles are sampled

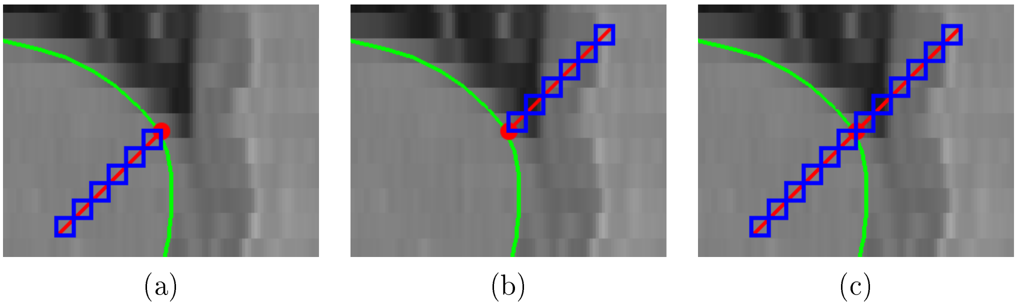


Figure 3.15: Three different ways to create intensity profiles. Using the normal (red line) at a vertex (red circle), intensities (blue squares) may be sampled: (a) inward, (b) outward, or (c) both sides.

in training images and both mean profile and its principal modes of variation are extracted for each landmark. During segmentation, a prior profile associated with each landmark is compared with current intensity profiles sampled at a certain search window around the current landmark. This comparison is done using similarity measures like the Mahalanobis distance or the correlation coefficient. In their basic form, profiles are made of plain pixel/voxel intensity values sampled directly from the image. However, the gradient of these profiles may be used too, as well as the normalized version of both profiles.

In another application, intensity and gradient profiles were used to optimize image forces of explicit deformable models [Gilles 2007]. The objective was to better discriminate organ contours in images by comparing intensity profiles, using two generic models and checking the similarity variation with the normalized cross-correlation. Normalized gradient profiles proved to be the best choice for face recognition [Cootes & Taylor 1993]. For bone segmentation in radiographs, normalized intensity profiles gave the best results [Behiels *et al.* 2002]. Combining these different type of profiles into a larger feature vector has also been implemented [Brejl & Sonka 2000]. A thorough study on intensity profiles for 3D medical image segmentation may be found in [Heimann 2009]. A study on optimal profile length and inward/outward ratio may be found in section 6.4.

When reference profiles are sampled, usually from the same structure in multiple images, they are used to create prior profiles on a reference mesh (*e.g.* using PCA [Cootes & Taylor 1994, Cootes *et al.* 1993]). During deformation, the objective is to deform the reference mesh so that profiles extracted from the image (*i.e.* current profiles) match to prior profiles.

At each vertex, the external force is computed as the translation that optimizes the match between the current and prior profiles (see an illustration in Figure 3.16). In practice, the prior profile is shifted along the normal with a certain sampling distance as displacement step. The current profile is sampled the same way than the prior profile (*i.e.* same length and sampling distance) using trilinear interpolation. Other interpolation methods like partial voluming may be used [Maes *et al.* 1997].

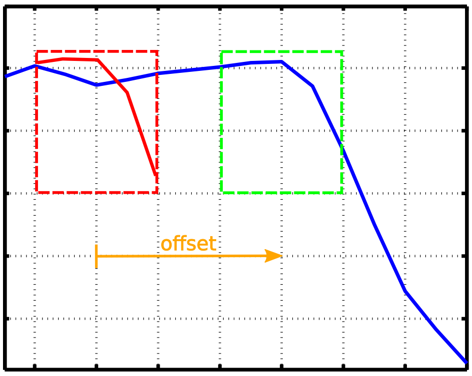


Figure 3.16: The external force is computed as the offset distance (orange arrow) that optimizes the match between the current intensity profile (blue line) and the prior intensity profile (red line, highlighted with the red dashed rectangle). The optimal position is highlighted with the green dashed rectangle.

The optimal match between current and prior profile uses a similarity measure, which evaluates how similar are the two profiles.

In the literature, several measures have been proposed [Buzug & Weese 1998, Penney *et al.* 1998, Roche *et al.* 1998b, Roche *et al.* 1998a]. Most common measures include sum of absolute differences, correlation coefficient and Mahalanobis distance. Below, we describe these similarity measures, as well as a linear criterion [Montagnat 1999].

Let $\{\mathbf{y}_i\}_{i\in[1,d]}$ be the set of intensities sampled from the image (*i.e.* current profile) and $\{\mathbf{x}_i\}_{i\in[1,d]}$ be the set of intensities from the reference profile, where d is the profile dimension that must be the same for both the current and reference profile.

Let $\bar{\mathbf{y}}$ and $\sigma_{\mathbf{y}}$ be the mean and standard deviation of the set $\{\mathbf{y}_i\}$, respectively.

Let $\{\boldsymbol{\mu}_i\}_{i\in[1,d]}$ and $\{\sigma_i\}_{i\in[1,d]}$ be the Gaussian parameters of prior profile (*i.e.* respectively, the mean and standard deviation of reference profile $\{\mathbf{x}_i\}$ over a population of meshes).

Let $\bar{\boldsymbol{\mu}}$ and $\sigma_{\boldsymbol{\mu}}$ be the mean and standard deviation of the set $\{\boldsymbol{\mu}_i\}$, respectively, defined as:

$$\bar{\boldsymbol{\mu}} = \frac{1}{d}\sum_{i=1}^{d}\{\boldsymbol{\mu}_i\} \tag{3.32}$$

$$\sigma_{\boldsymbol{\mu}} = \sqrt{\frac{1}{d}\sum_{i=1}^{d}(\{\boldsymbol{\mu}_i\} - \bar{\boldsymbol{\mu}})^2} \tag{3.33}$$

Sum of absolute differences This criterion, which assumes a linear relationship between the two sets of intensities, is a simple point-to-point comparison of intensity values:

$$C(\{\mathbf{y}_i\},\{\boldsymbol{\mu}_i\}) = \frac{1}{d}\sum_{i=1}^{d} |\{\mathbf{y}_i\} - \{\boldsymbol{\mu}_i\}| \tag{3.34}$$

Though simple and statically robust, comparison is only possible with intensities whose distribution and image modality are close (*i.e.* this criterion is not robust to noise nor able to deal with multimodal images).

Linear criterion This criterion, which also assumes a linear relationship between the two sets of intensities, is robust and normalized:

$$C(\{\mathbf{y}_i\},\{\boldsymbol{\mu}_i\}) = \frac{(\sum_{i=1}^{d}\{\mathbf{y}_i\}\{\boldsymbol{\mu}_i\})^2}{\sum_{i=1}^{d}\{\mathbf{y}_i\}^2\sum_{i=1}^{d}\{\boldsymbol{\mu}_i\}^2} \tag{3.35}$$

In the interval $[0, 1]$, linear criterion makes the hypothesis of a linear transformation between $\{\mathbf{y}_i\}$ and $\{\boldsymbol{\mu}_i\}$.

Correlation coefficient Though more computationally intensive, correlation coefficient is robust and normalized. Mean and standard deviation are used in addition to intensity values:

$$C(\{\mathbf{y}_i\},\{\boldsymbol{\mu}_i\}) = \frac{1}{d^2\sigma_{\mathbf{y}}^2\sigma_{\boldsymbol{\mu}}^2}\left(\sum_{i=1}^{d}(\{\mathbf{y}_i\} - \bar{\mathbf{y}})(\{\boldsymbol{\mu}_i\} - \bar{\boldsymbol{\mu}})\right)^2 \tag{3.36}$$

This criterion makes the hypothesis of an affine transformation between $\{\mathbf{y}_i\}$ and $\{\boldsymbol{\mu}_i\}$. This hypothesis improves robustness with respect to global intensity change across acquisition protocols. To eliminate the affine terms, cross correlation normalized by the standard deviation (*i.e.* Normalized Cross Correlation) may be used [Holden *et al.* 2000].

Mahalanobis distance In this case, the prior profile is computed from a statistical analysis of a set of meshes, *i.e.* the prior profile is related to a population of profiles. The Mahalanobis distance, which assumes that the current profile is sampled from a Gaussian distribution of intensity profiles, takes also the correlation of profiles into account:

$$C(\{\mathbf{y}_i\},\{\boldsymbol{\mu}_i\}) = \sqrt{\sum_{i=1}^{d}\frac{(\{\mathbf{y}_i\} - \{\boldsymbol{\mu}_i\})^2}{\{\sigma_i\}^2}} \tag{3.37}$$

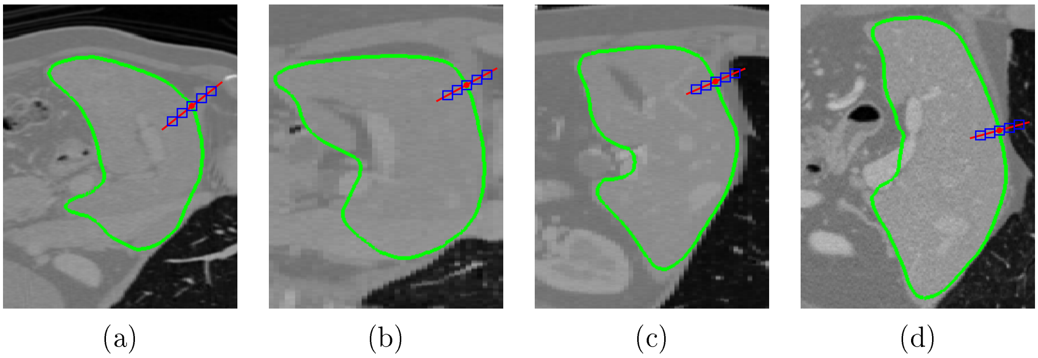

(a) (b) (c) (d)

Figure 3.17: Sampling of an intensity profile (blue squares) at the normal (red line) of the same vertex (red circle) from a reference mesh (green delineation) registered to four liver images {(a),(b),(c),(d)}. The reference mesh features $N = 11760$ vertices and has been registered to the four liver images to ensure point correspondence.

Each intensity value (*i.e.* each profile dimension) is weighted according to its standard deviation. In addition, Mahalanobis distance is scale-invariant (*i.e.* distance does not dependent on the scale of $\{\mathbf{y}_i\}$ and $\{\boldsymbol{\mu}_i\}$).

When using the covariance matrix $\mathbf{\Sigma}_i$ of the reference profile, Mahalanobis distance is defined as:

$$C(\{\mathbf{y}_i\}, \{\boldsymbol{\mu}_i\}) = \sqrt{(\mathbf{y} - \boldsymbol{\mu})^T \mathbf{\Sigma}_i^{-1} (\mathbf{y} - \boldsymbol{\mu})} \tag{3.38}$$

3.3.5.5 PCA-based methods

Typical PCA-based methods consist in computing the Gaussian distribution (*i.e.* mean $\boldsymbol{\mu}_i$ and covariance matrix $\mathbf{\Sigma}_i$) of profiles $\mathbf{x}_i$ at each vertex of a reference mesh registered to all datasets, *i.e.* point correspondence must be ensured (see an illustration in Figure 3.17).

As a similarity measure, the Mahalanobis distance d_M may be computed from equation 3.38 as:

$$d_M(\mathbf{y}) = \sqrt{(\mathbf{y} - \boldsymbol{\mu}_i)^T \mathbf{\Sigma}_i^{-1} (\mathbf{y} - \boldsymbol{\mu}_i)} \tag{3.39}$$

where $\mathbf{y}$ is the current profile.

However, a regularization of the covariance matrix $\mathbf{\Sigma}_i$ is usually necessary to avoid any singularities due to its inversion [Schäfer & Strimmer 2005]. A solution consists in using an alternative approach proposed by Cootes [Cootes & Taylor 1993], in which the Mahalanobis distance d_M is computed without inverting the covariance matrix.

Given a PCA performed on N profiles $\mathbf{x}_i$, expressed by the principal matrix Φ_i, the m eigenvalues λ_{il} and the mean intensity profile $\boldsymbol{\mu}_i$, the Mahalanobis distance d_M is defined as:

$$d_M(\mathbf{y}) = \sum_{l=1}^{m} \sqrt{\frac{b_{il}^2}{\lambda_{il}}} \tag{3.40}$$

where $b_i = (b_{i1}, \ldots, b_{im})^T$ is the model parameter vector of the best fit $\hat{\mathbf{y}}$ of $\mathbf{y}$ given the PCA model:

$$b_i = \Phi_i^T (\hat{\mathbf{y}} - \boldsymbol{\mu}_i) \tag{3.41}$$

Using PCA enables to reduce the number of profile dimensions by working on a subspace that still retains most of the variability observed with the profiles. In practice, only top eigenvectors from an eigen decomposition performed on covariance matrix $\boldsymbol{\Sigma}_i$ are selected to reduce its complexity. For that, eigenvectors are sorted in order of decreasing eigenvalues and a ratio on normalized eigenvalues $\mathcal{R}$ ($\mathcal{R} \in [0, 1]$) is specified. The number of top eigenvectors is selected so that their cumulative eigenvalues are above the ratio $\mathcal{R}$ (*e.g.* $\mathcal{R} = 0.9$).

During local search, other similarity measures may be used, such as those proposed in section 5.2. In the literature, external forces based on intensity profiles are usually computed using PCA with explicit deformable models [Gilles *et al.* 2006, Schmid & Magnenat-Thalmann 2008].

3.4 Conclusion

In this chapter, we presented different segmentation approaches that are used in medical image analysis to extract anatomical structures of interest such as organs, bones and tissues. We were particularly interested in explicit model-based segmentation, from which we described the main features, *i.e.* mesh initialization and both internal and external forces.

As explicit deformable models, we presented simplex meshes that are efficient in terms of flexibility and computational cost, since they feature a simple geometric description. We presented some topological and geometrical properties that are helpful in a segmentation approach. For instance, the rigidity defined from a certain vertex neighborhood in section 3.3.2.1 ensures the smoothing of meshes when building anatomical models in section 7.2.4. Another example is the global refinement described in section 3.3.2.4, which is used in chapter 6 to create meshes with different resolutions.

As the objective of this thesis is to provide a regional appearance modeling by means of new external forces, we first presented external forces coupled with simplex meshes, then external forces from intensity profiles (*i.e.* to regionally enhance model-based image segmentation), and finally external forces based on PCA (*i.e.* current state of the art to which we compare our new regional external forces).

Chapter 4

Multimodal Prior Appearance Models

Contents

4.1 Introduction

In order to model appearance around structures of interest, intensive research has been performed on Statistical Appearance Models (SAM), which are based on Principal Component Analysis (PCA) [Heimann 2009, Gilles 2007]. Active Shape Models (ASM) [Cootes & Taylor 1994] are a widely used technique. Intensity profiles are trained and both mean profile and principal modes of variation for each landmark are extracted in a PCA fashion. Active Appearance Models (AAM), which are a generalization of the ASM, use all the intensities from the inner region of the mesh to create a large feature vector. Like in ASM, a PCA is built, though on textures this time, and both mean of textures and modes of variation are extracted.

Several issues may be raised with such PCA-based methods. First, they require an accurate point-wise registration as the statistical analysis of shape and appearance is performed at each point. Defining homologous points for 3D structures is

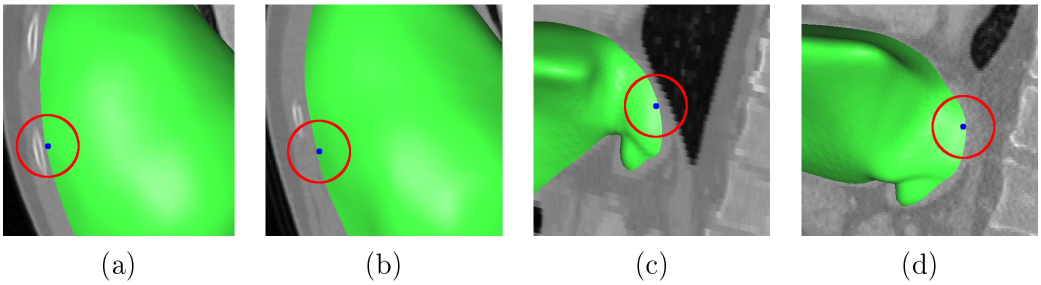


Figure 4.1: Two examples illustrating that shape is not necessarily correlated with appearance: {(a),(b)} at vertebrae level, and {(c),(d)} at lung level. The same reference mesh featuring $N = 11760$ vertices has been registered to the four liver images {(a),(b),(c),(d)} to ensure point correspondence. In each example, a corresponding vertex (blue point) is displayed to compare the neighboring appearance. From {(a),(b)}, note how the vertebrae (*i.e.* white ellipsoid-like region) close to the corresponding vertex in liver image (a) is shifted in liver image (b). From {(c),(d)}, note how the lungs (*i.e.* black region) close to the corresponding vertex in liver image (c) are shifted in liver image (d).

difficult and therefore registering those points accurately is still considered challenging [Dam *et al.* 2008, Josephson *et al.* 2005, Heitz *et al.* 2004, Guest *et al.* 2001].

A second limitation common to most appearance models such as AAM is that they are *monomodal*, *i.e.* they rely on the hypothesis that the probability density function is well described by a single Gaussian distribution. This hypothesis is often violated by the presence of pathologies but also by the fact that shape is not necessarily correlated with appearance. Indeed, a point correspondence between two meshes ensures a corresponding vertex to be at the same geometric location on these two meshes, but does not ensure the neighboring appearance around each corresponding vertex to be similar (see two examples from liver meshes in Figure 4.1).

These limitations lead to a Gaussian distribution featuring a large covariance at each point because the estimated profiles at each point may greatly vary around the average profile. Indeed, even if the point correspondence is accurate, the intensity profile at each point may greatly vary (*i.e.* shape is not necessarily correlated with appearance). Instead of having one Gaussian distribution with large covariance (*i.e.* due to a poor performance in evaluating the similarity between data points), it is preferable for image segmentation or image detection purposes to have several Gaussian distributions with lower covariance (*i.e.* as they are more representative and discriminant for a set of data points).

To cope with both the inaccuracy of point correspondence and the absence of correlation between shape and appearance, we propose Multimodal Prior Appearance Models (MPAM) as a novel way to model the appearance around structures of interest in medical images [Chung & Delingette 2009]. Unlike PCA, this method does not rely on an accurate pointwise registration and is based on a regional clustering of intensity profiles. The clustering is considered as *regional* because intensity profiles are classified for each mesh, and not for each vertex (*i.e.* over a population of meshes).

In Figure 4.2, we overview the consecutive steps our method requires to build the MPAM (section 4.3). The input is a set of P meshes $\mathcal{M}_p$ corresponding to the segmentation of the same structure whose appearance needs to be captured in different images. The meshes may have different number of vertices, or even different topologies. Unlike PCA-based approaches, our method does not require accurate pointwise correspondences.

At each vertex $\mathbf{p}_i$ of mesh $\mathcal{M}_p$, d regularly sampled intensities are extracted to build a d-dimensional intensity profile along the normal direction, noted $\mathbf{x}_i^p$, that may extend inward, outward or both sides. This part may be seen as the first step of the training, where the method is capturing raw appearance information throughout the datasets.

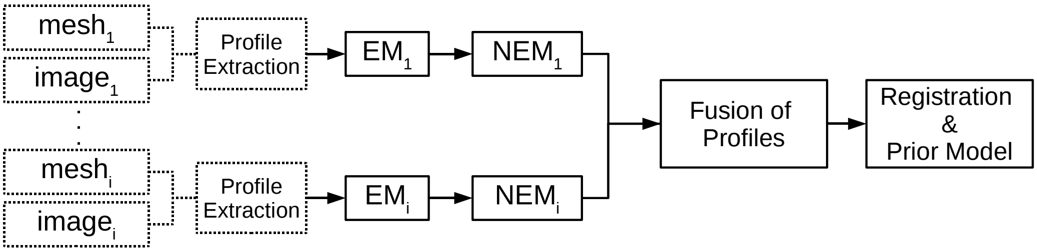


Figure 4.2: Pipeline for Multimodal Prior Appearance Model (MPAM) construction, which includes the Expectation-Maximization (EM) and Neighborhood EM (NEM) algorithms.

Then, we propose to automatically cluster intensity profiles from each mesh using an EM classification (section 4.2.3), whose centers are initialized using the Fuzzy C-Means (section 4.2.2). EM algorithm and Fuzzy C-Means are both unsupervised clustering techniques (section 4.2), which are used as we do not want to make any assumption about the profile classification. We detail how to cope with missing data and present new regularization strategies for covariance matrices (section 4.3.1).

The number of classes (*i.e.* the number of appearance regions), a hyperparameter, is selected in an automatic fashion through a model order selection (section 4.3.2.1) based on a new criterion (section 4.3.2.2). To have a spatially smooth classification of profiles, we present a spatial regularization approach (section 4.3.3) based on the Neighborhood EM (section 4.2.4).

Building a MPAM out of P datasets requires the projection of their associated EM classifications to a common framework. Two steps are needed for such a task. First, all intensity profile classes from the P datasets are compared and classes corresponding to the same tissue are possibly merged, depending on a threshold on the Jaccard index $\mathcal{J}$ (section 4.3.4). The objective is to control the final number of classes associated with the MPAM. Finally, all P meshes are registered to the same

Figure 4.3: Two typical images used in computer vision to test segmentation algorithms: Baboon (on the left) and Lena (on the right). From the original images $\{(a),(c)\}$, the K-Means algorithm was launched to classify image pixels using: (b) $K = 6$ clusters, and (d) $K = 8$ clusters.

reference mesh $\mathcal{M}^\star$. The objective is to project the classifications from P meshes into a reference mesh and to estimate the posterior probabilities associated with the MPAM (section 4.3.5). Each vertex $\mathbf{p}_i$ of reference mesh $\mathcal{M}^\star$ is given a probability $\tilde{\gamma}_i^m$ to belong to several intensity profile classes.

The MPAM construction is illustrated on datasets of both livers and tibia bones (section 4.3.6).

4.2 Unsupervised clustering

4.2.1 Principles

Clustering [Kotsiantis & Pintelas 2004] is a statistical method for data analysis whose aim is to separate a set of data points, or *observations*, into K clusters, or *subsets* (see an example in Figure 4.3). The objective is to group data points into clusters so that data points from a same cluster share similar features. The notion of similarity is usually defined with a proximity criterion using a distance measure (*e.g.* Euclidean, Mahalanobis, Manhattan and Hamming distances).

Most popular unsupervised clustering methods are Fuzzy C-Means [Bezdek 1981] and Expectation-Maximization [Dempster *et al.* 1977]. Other methods include hierarchical algorithms [Hastie *et al.* 2009], either agglomerative (*i.e.* bottom-up) or divisive (*i.e.* top-down), density-based clustering algorithms [Wan *et al.* 2009] and subspace clustering methods [Kriegel *et al.* 2009]. A review on unsupervised clustering methods may be found in [Jain *et al.* 1999].

4.2.2 Fuzzy C-Means

The Fuzzy C-Means (FCM) [Bezdek 1981, Dunn 1973] is a clustering algorithm inspired from the fuzzy logic [Zadeh 1968, Zadeh 1965]. The objective of FCM is to classify N data points $\mathbf{x}_i$ to a specified number of clusters K (see Figure 4.4 for an example with 2D data points).

The method consists in an iterative procedure during which an objective function J_m is minimized:

$$J_m = \sum_{i=1}^{N} \sum_{k=1}^{K} (\gamma_i^k)^m \parallel \mathbf{x}_i - \boldsymbol{\mu}_k \parallel^2 \tag{4.1}$$

where N is the number of data points, K is the number of clusters, m is a fuzziness coefficient ($m \leqslant 1 \leqslant \infty$), γ_i^k is the degree of membership of $\mathbf{x}_i$ in cluster k, $\mathbf{x}_i$ is the d-dimensional i-th data point, $\boldsymbol{\mu}_k$ is the d-dimensional center of cluster k and $\parallel \mathbf{x}_i - \boldsymbol{\mu}_k \parallel$ is a norm expressing the similarity between $\mathbf{x}_i$ and $\boldsymbol{\mu}_k$.

During the iterative optimization of J_m, degrees of membership γ_i^k and cluster centers $\boldsymbol{\mu}_k$ are updated as:

$$\gamma_i^k = \frac{1}{\sum_{k=1}^{K} \left(\frac{\|\mathbf{x}_i - \boldsymbol{\mu}_k\|}{\|\mathbf{x}_i - \boldsymbol{\mu}_k\|} \right)^{\frac{2}{m-1}}} \tag{4.2}$$

$$\boldsymbol{\mu}_k = \frac{\sum_{i=1}^{N} (\gamma_i^k)^m \mathbf{x}_i}{\sum_{i=1}^{N} (\gamma_i^k)^m} \tag{4.3}$$

This iterative procedure stops either after a certain number of iterations, or when a steady state is reached:

$$\max_{ik}\{\parallel (\gamma_i^k)^{(t+1)} - (\gamma_i^k)^{(t)} \parallel\} < \varepsilon \tag{4.4}$$

where ε is a termination criterion and t an iteration number.

The minimum reached by FCM is not guaranteed to be global (*i.e.* it may be local). The local minimum depends on the algorithm initialization, which is often done using random centers.

Unlike K-Means [MacQueen 1967] where each $\mathbf{x}_i$ is sharply assigned to only one cluster ($\gamma_i^k = 0|1$), FCM allow data points to be given a degree of membership (*i.e.* a probability: $0 < \gamma_i^k < 1$, $\sum_{k=1}^{K} \gamma_i^k = 1$) to all clusters using fuzziness coefficient m. This means that data points at the edge of a cluster k are assigned less degree of membership than those close to its center $\boldsymbol{\mu}_k$. FCM bring thus more information into the classification compared to K-Means, which only specify whether a data point belongs to a cluster or not.

The fuzziness coefficient m plays an important role in the classification result. When m is close to 1, data points close to the cluster centers are given much more

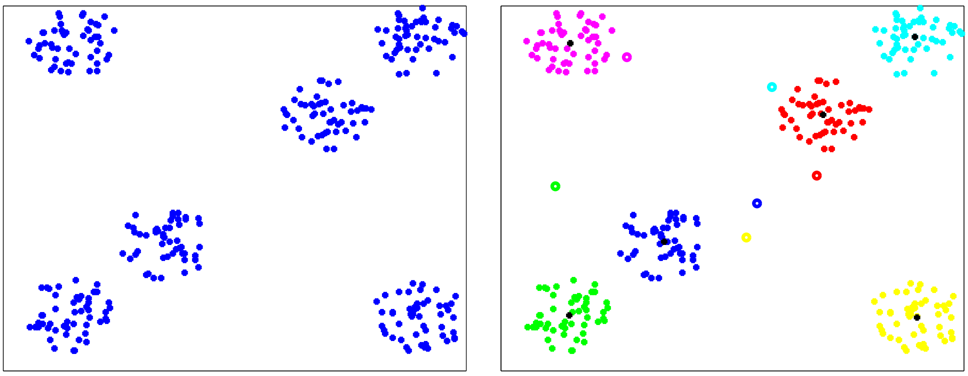

(a) (b)

Figure 4.4: FCM algorithm applied on a set of 300 2D data points: (a) original data points to be clustered, and (b) result of the FCM algorithm with $K = 6$ clusters. Each cluster has a different color and its initial center is depicted as a circle with the same color (original centers are generated from random values). Final cluster centers are depicted as black crosses.

weight than those far away and the algorithm is similar to K-Means. The higher m, the fuzzier the degrees of membership γ_i^k. In most applications, m is set to 2. In the literature, FCM-based algorithms have been used, among others, for tissue classification [Parveen *et al.* 2006], and both image segmentation [Forsyth & Ponce 2002] and compression [Karayiannis & Pai 1995].

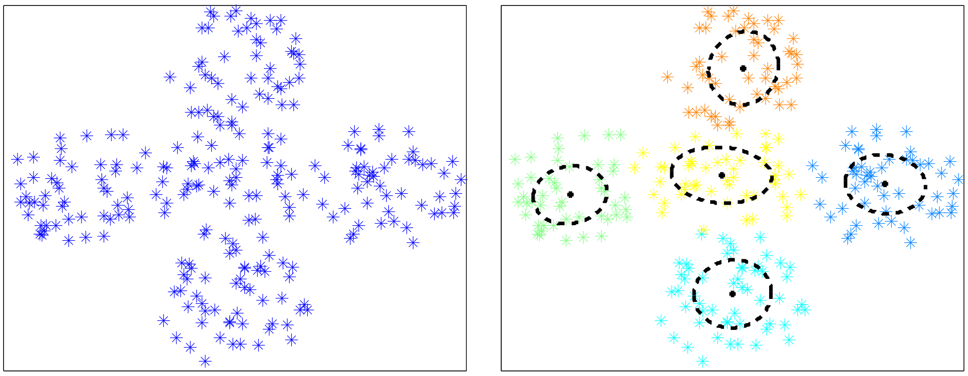
4.2.3 Expectation-Maximization

Compared to FCM, Expectation-Maximization algorithm (EM) is a more statistically formalized method with broad applicability, which has better convergence properties and is easier to derive [Dempster *et al.* 1977, McLachlan & Krishnan 1996]. Furthermore, EM provides an additional information about the variance within each class thanks to the covariance matrix $\mathbf{\Sigma}_k$ (see Figure 4.5 for an example with 2D data points). EM is an iterative procedure whose objective is to compute the Maximum Likelihood (ML) estimate in the presence of a hidden variable, which is the belonging of each data point $\mathbf{x}_i$ to a particular class k.

Several data distributions may be modeled by the EM algorithm, such as:

- Poisson distribution [Nuyts *et al.* 2001, Hebert & Leahy 1989]
- Weibull distribution [Jiang & Kececioglu 1992, Jiang 1991]
- Bernouilli distribution [McLachlan & Peel 2000, Lazarsfeld & Henry 1968]
- Gaussian distribution [Chung & Delingette 2009, Xu & Jordan 1996]

However, the Gaussian distribution is the most often used due to its simplicity and well-known representation. In this case, the clustering needs to be performed



(a) (b)

Figure 4.5: EM algorithm applied on a set of 250 2D data points: (a) original data points to be clustered, and (b) result of the EM algorithm with $K = 5$ clusters. Each cluster has its own color and both its center and covariance highlighted.

using Gaussian Mixture Models (GMM). Each cluster is assumed to be well defined by a Gaussian distribution whose parameters (*i.e.* center and variance) are to be computed. These parameters are considered latent variables that need to be inferred from the observations through a mathematical model [Borsboom *et al.* 2003, Celeux & Govaert 1995]. With GMM, data points are considered to be identically independently distributed (i.i.d) from the linear superposition of K Gaussian distributions (*i.e.* K *classes*, known as *clusters* in FCM literature).

The unknown parameters, or *latent variables*, to be estimated through the EM iterations are K, which is the number of classes, and Θ:

$$\Theta = \{\pi_1, \ldots, \pi_K, \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K, \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_K\} \tag{4.5}$$

where π_k is the mixing coefficient, $\boldsymbol{\mu}_k$ the center and $\boldsymbol{\Sigma}_k$ the covariance matrix of class k.

GMM are aimed at providing a richer class of density models than a single Gaussian distribution:

$$G_{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k}(\mathbf{x}_i) = \frac{\exp\Big(-0.5\ (\mathbf{x}_i - \boldsymbol{\mu}_k)^T\ \boldsymbol{\Sigma}_k^{-1}\ (\mathbf{x}_i - \boldsymbol{\mu}_k)\Big)}{(2\pi)^{\frac{d}{2}}\ |\boldsymbol{\Sigma}_k|^{\frac{1}{2}}} \tag{4.6}$$

EM algorithm consists in computing the Maximum Likelihood Estimate (MLE) of the unknown parameters using the likelihood of the observed data:

$$p(\mathbf{x}_i|\Theta) = \sum_{k=1}^{K} \pi_k\ p(\mathbf{x}_i|k, \theta_k) \tag{4.7}$$

where π_k is the mixing coefficient ($0 < \pi_k < 1$, $\sum_{k=1}^{K} \pi_k = 1$), which indicates the contribution of class k into the GMM, and $p(\mathbf{x}_i|k, \theta_k)$ denotes the density of a Gaussian distribution:

$$p(\mathbf{x}_i|k, \theta_k) = G_{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k}(\mathbf{x}_i) \tag{4.8}$$

We assume in equation 4.6 that the inverse of $\boldsymbol{\Sigma}_k$ is non singular. We shall see further down that this assumption may not be met.

The objective of EM is to estimate the model parameters for which these observed data are the most likely to maximize the log-likelihood:

$$L(\Theta) = \sum_{i=1}^{N} \log(p(\mathbf{x}_i|k, \theta_k)) \tag{4.9}$$

For each EM iteration, these parameters are estimated through two steps: the *E-step* and the *M-step*.

During the E-step, or *expectation*, the hidden variable (*i.e.* the belonging of each data point to a particular class) is estimated with posterior probabilities γ_i^k (*i.e.* corresponding to FCM's *degrees of membership*).

Given the observed data and current estimate of the model parameters, posterior probabilities γ_i^k are calculated using Bayes' rule:

$$\gamma_i^k = \frac{\pi_k \ G_{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k}(\mathbf{x}_i)}{\sum_{l=1}^{K} \pi_l \ G_{\boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l}(\mathbf{x}_i)} \tag{4.10}$$

where $G_{\boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l}(\mathbf{x}_i)$ is calculated from equation 4.6.

During the M-step, or *maximization*, the likelihood function is maximized under the assumption that the hidden variable is known.

The estimated posterior probabilities from the E-step are used to calculate model parameters π_k, $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$:

$$\pi_k = \frac{1}{N} \sum_{i=1}^{N} \gamma_i^k \tag{4.11}$$

$$\boldsymbol{\mu}_k = \frac{\sum_{i=1}^{N} \gamma_i^k \ \mathbf{x}_i}{\sum_{i=1}^{N} \gamma_i^k} \tag{4.12}$$

$$\boldsymbol{\Sigma}_k = \frac{\sum_{i=1}^{N} \gamma_i^k (\mathbf{x}_i - \boldsymbol{\mu}_k)(\mathbf{x}_i - \boldsymbol{\mu}_k)^T}{\sum_{i=1}^{N} \gamma_i^k} \tag{4.13}$$

The computation of π_k, $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ does not provide a closed-form solution because the posterior probabilities γ_i^k depend on these parameters (see equation 4.10). This is why an iterative scheme, which consists in updating the model parameters to estimate the posterior probabilities in the next E-step, is used.

Convergence is assured since EM is guaranteed to increase the likelihood at each iteration [Neal & Hinton 1998, Hathaway 1986, Csiszár & Tusnády 1984]. In practice, convergence is considered to be reached when computed change in the log-likelihood, or in the estimated parameters, is less than a given threshold.

In the literature, a relationship between the EM algorithm and the Fuzzy C-Means has been demonstrated [Flandin 2004]. More precisely, Fuzzy C-Means may be seen as a special case of the EM algorithm in which classes feature identical spherical covariance matrices $\boldsymbol{\Sigma}_k$ and equal mixing coefficients π_k. Also, it has been proven that EM algorithm with GMM is formally equivalent to an alternate optimization of a fuzzy function [Hathaway 1986].

This fuzzy function is an alternate optimization of the L functional described in equation 4.9:

$$L(\Theta) = \sum_{k=1}^{K}\sum_{i=1}^{N} \gamma_i^k \ \log(\pi_k \ G(\mathbf{x}_i|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)) - \sum_{k=1}^{K}\sum_{i=1}^{N} \gamma_i^k \log(\gamma_i^k) \tag{4.14}$$

where $G(\mathbf{x}_i|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ is the Gaussian probability density function.

As for FCM, EM initialization plays an important part of the algorithm. That may be done either by using known initial values for the means, covariance matrices and mixing coefficients, or by using complete random values. However, EM is usually initialized using K-Means [MacQueen 1967]. In this case, both means $\boldsymbol{\mu}_k$ and covariance matrices $\boldsymbol{\Sigma}_k$ are initialized using the mean and covariance of data points belonging to each K-Means cluster. Concerning mixing coefficients π_k, they may be initialized using the ratio of data points belonging to each K-Means cluster.

Applying the maximum likelihood framework to GMM may be an issue due to the presence of singularities. For instance, if all data points $\mathbf{x}_i$ are equal to the center $\boldsymbol{\mu}_k$ of the k-th GMM component, elements of covariance matrix $\boldsymbol{\Sigma}_k$ will be zeros (see equation 4.13). In fact, the rank of covariance matrix $\boldsymbol{\Sigma}_k$ may be lower than d if there is a correlation between the dimensions of the feature vector (*i.e.* data point). In this case, covariance matrix $\boldsymbol{\Sigma}_k$ from equation 4.6 cannot be inverted and log-likelihood from equation 4.9 cannot be computed. This proves that the log-likelihood is not a well posed problem and that caution needs to be taken.

Several heuristics have been proposed in the literature to overcome singularities. In homogeneous processes, all GMM components are constrained to have the same covariance matrices. Another heuristic consists in using a Singular Value Decomposition (SVD) at each EM iteration and checking for small values on one of the covariance matrix dimensions. In case of small values, the component is reinitialized with a higher value (*e.g.* randomly chosen value or initial value). To avoid such singularities, we propose in section 4.3.1.3 three different methods to regularize the covariance matrix.

4.2.4 Neighborhood EM

Taking spatial information about data into account may be of a great help in a classification procedure, as it may enhance similar data points to be neighbors. The simplest method consists in adding spatial coordinates of data points as new dimensions, and thus increasing the feature space. Another possibility is using Markov Random Fields (MRF), a well-known method to take spatial constraints into account [Zhang *et al.* 2001, Leemput *et al.* 2001]. Although some procedures for FCM have been proposed [Chuang *et al.* 2006, Udupa & Saha 2003], EM is usually preferred because its derivation is easier to perform.

The Neighborhood EM (NEM) [Ambroise *et al.* 1997] is a clustering algorithm for spatial data using the alternate optimization of the $L(\Theta)$ functional described in equation 4.14.

The energy function is penalized with a parameter β_n taking spatial homogeneity into account:

$$\mathcal{L}(\Theta) = L(\Theta) + \beta_n \sum_{k=1}^{K} \sum_{i=1}^{N} \sum_{j=1}^{N} (\gamma_i^k \; \gamma_j^k \; v_{ij}) \tag{4.15}$$

where v_{ij} is a similarity value between neighboring data points $\mathbf{x}_i$ and $\mathbf{x}_j$.

The former term $L(\Theta)$ leads to the classical EM while the latter is a spatial regularization term controlled by β_n. The more correlation between neighboring data points, the greater this term is.

The similarity value v_{ij} comes from a neighborhood matrix V that sets the amount of correlation between neighboring data points:

$$\begin{aligned} v_{ij} &= \mathit{\Omega} > 0 && \text{if } \mathbf{x}_i \text{ and } \mathbf{x}_j \text{ are neighbors} \\ &= 0 && \text{if } \mathbf{x}_i \text{ and } \mathbf{x}_j \text{ are not neighbors} \end{aligned}$$

To compute matrix V, a function computing the spatial distance between data points must be used. For instance, graph structures may be used as neighboring concept [Ambroise *et al.* 1997]. Indeed, defining neighborhood relationships is equivalent to building a graph where each data point $\mathbf{x}_i$ is represented by a node and each neighborhood relationship by an edge.

Two steps are required to cluster data points with such a spatial constraint. First, the neighborhood graph must be defined using standard algorithms like Delaunay triangulation [Green & Sibson 1978] or Gabriel Graph [Gabriel & Sokal 1969]. Then, EM must be run while taking the spatial constraints into account.

NEM minimizes the $\mathcal{L}(\Theta)$ functional with an alternate optimization, leading to a modified E-step where the posterior probabilities are iteratively estimated until a fixed value is reached.

From equation 4.10, new posterior probability $\{\gamma_i^{k,t}\}_{NEM}$ at NEM iteration t is computed as:

$$\{\gamma_i^{k,t}\}_{NEM} = \frac{\pi_k \; G_{\boldsymbol{\mu}_k,\boldsymbol{\Sigma}_k}(\mathbf{x}_i) \; \exp\left(\beta_n \; \sum_{j=1}^{N}(\gamma_j^{k,t} \; v_{ij})\right)}{\sum_{l=1}^{K}\left[\pi_l \; G_{\boldsymbol{\mu}_l,\boldsymbol{\Sigma}_l}(\mathbf{x}_i) \; \exp\left(\beta_n \; \sum_{j=1}^{N}(\gamma_j^{l,t} \; v_{ij})\right)\right]} \tag{4.16}$$

where β_n is the parameter weighting the spatial homogeneity, π_l the mixing coefficient of class l, and $\gamma_j^{l,t}$ the posterior probability of neighbor $\mathbf{x}_j$ to belong to class l at NEM iteration t (at first NEM iteration, γ_j^k from the previous EM iteration is used).

The E-step becomes iterative because the posterior probability at each data point $\mathbf{x}_i$ is computed with respect to its neighbors, whose posterior probability is updated during the same NEM iteration. This iterative procedure is thus needed to reach a steady state in terms of posterior probabilities. At each NEM iteration, the spatial information modifies the classification according to the importance of β_n. This iterative procedure stops when there are no significant differences between two consecutive NEM iterations.

Unlike E-step, the M-step stays unchanged and EM parameters $\{\pi_k,\ \boldsymbol{\mu}_k,\ \boldsymbol{\Sigma}_k\}$ are thus estimated using the new posterior probabilities $\{\gamma_i^k\}_{NEM}$.

4.3 Building Multimodal Prior Appearance Models

4.3.1 EM classification of intensity profiles

The classification associated with a given mesh $\mathcal{M}_p$ is formulated in the context of a probability density estimation using GMM, as described in equation 4.7. The approach is semi-parametric because the number of classes K_p associated to each mesh $\mathcal{M}_p$ is an unknown parameter. As formulated in equation 4.5, other unknown parameters to be computed by the EM classification include Gaussian parameters (*i.e.* mean and variance), as well as the weight of each Gaussian distribution into the mixture (*i.e.* mixing coefficients).

In the sequel, the intensity profile class is denoted by *mode* and we write $\boldsymbol{\mu}_k^p$, $\boldsymbol{\Sigma}_k^p$, π_k^p and ${}^p\gamma_i^k$ for, respectively, the mean, covariance matrix, mixing coefficient and posterior probabilities of mode k associated to mesh $\mathcal{M}_p$ $(1 \leq p \leq P)$.

4.3.1.1 Initialization

As explained in section 4.2.3, initialization is a very important part of the method because EM is known to be very sensitive to its initialization. A good initialization is supposed to yield robust solutions that are not prone to errors nor local minima.

Though EM is usually initialized with K-Means, we initialize it with FCM, since tests on synthetic data proved FCM to be more stable. As a first step, profiles are thus classified using FCM.

After convergence, FCM data membership values $\{{}^p\gamma_i^k\}_{\text{fcm}}$ are directly used to initialize EM posterior probabilities $\{{}^p\gamma_i^k\}_{\text{em}}$:

$$\{{}^p\gamma_i^k\}_{\text{em}} = \{{}^p\gamma_i^k\}_{\text{fcm}} \tag{4.17}$$

Both FCM data membership values $\{{}^p\gamma_i^k\}_{\text{fcm}}$ and cluster centers $\{\boldsymbol{\mu}_k^p\}_{\text{fcm}}$ are used to initialize EM mode means $\{\boldsymbol{\mu}_k^p\}_{\text{em}}$, covariance matrices $\boldsymbol{\Sigma}_k^p$ and mixing coefficients π_k^p in one M-step:

$$\{\boldsymbol{\mu}_k^p\}_{\text{em}} = \frac{\sum_{i=1}^N \{{}^p\gamma_i^k\}_{\text{fcm}}\, \mathbf{x}_i}{\sum_{i=1}^N \{{}^p\gamma_i^k\}_{\text{fcm}}} \tag{4.18}$$

$$\boldsymbol{\Sigma}_k^p = \frac{\sum_{i=1}^N \{{}^p\gamma_i^k\}_{\text{fcm}} (\mathbf{x}_i - \{\boldsymbol{\mu}_k^p\}_{\text{fcm}})(\mathbf{x}_i - \{\boldsymbol{\mu}_k^p\}_{\text{fcm}})^T}{\sum_{i=1}^N \{{}^p\gamma_i^k\}_{\text{fcm}}} \tag{4.19}$$

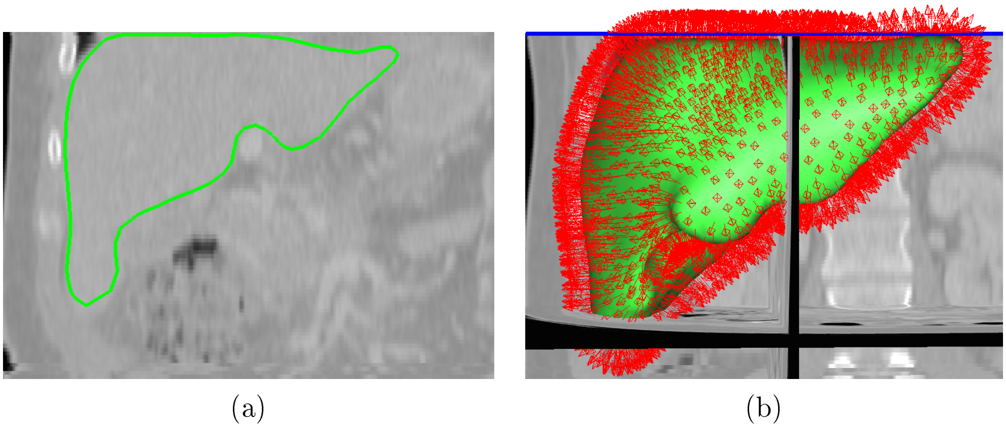
$$\pi_k^p = \frac{1}{N} \sum_{i=1}^N \{{}^p\gamma_i^k\}_{\text{fcm}} \tag{4.20}$$

FCM are themselves initialized with random cluster centers.

4.3.1.2 Coping with missing data

Since segmented structures may be close to the image boundaries, profiles are likely to be incomplete (see an illustration in Figure 4.6). The original EM algorithm cannot deal with this kind of issue, as data points are supposed to have all the same dimension (*i.e.* they should be represented by feature vectors having the same size).

If a profile partly lies outside the image, a solution could be to associate a default intensity value to this part (*e.g.* 0.0 or any other background value). An intensity



(a) (b)

Figure 4.6: Liver mesh close to image boundaries: (a) delineation in a 2D slice, and (b) 3D view of the mesh inside the image. At the top of image, intensity profiles that extend outward (red arrows) lay outside image boundaries (blue line) and are thus incomplete (*i.e.* they have missing data information).

interpolation, which would guess the missing intensities based on the boundary's neighborhood, could also be used. In both cases, the problem remains the same: data is missing and any substitute would only be based on hypothesis. Furthermore, any hypothesis is likely to misguide the EM classification, depending on the influence of incomplete profiles with respect to complete profiles.

To avoid any hypothesis on missing intensities, we prefer to rely only on intensities we have into the image and protect the classification from missing data influence. To use incomplete profiles in the classification, a valid dimension range $d_i \leq d$ is determined for each profile $\mathbf{x}_i$, where d is the expected number of dimensions featured by the profiles.

In the E-step, the term $\exp((\mathbf{x}_i^p)^T(\mathbf{\Sigma}_k^p)^{-1}(\mathbf{x}_i^p))$ is only computed for the valid dimension range and scaled by d/d_i. In the M-step, the mode means and covariance matrices are normalized by the total number of valid values for each index u of the profile $(1 \leq u \leq d)$. FCM initialization has also been extended in similar fashion to cope with incomplete profiles.

4.3.1.3 Covariance matrix regularization

The EM log-likelihood maximization, as described in equation 4.9, may lead to local maxima or degeneracy. In section 4.2.3, we explained why singularities may arise and how they may be dealt with in the literature. In our case, degeneracy of covariance matrices $\mathbf{\Sigma}_k^p$ for d-dimensional profiles (typically $d \approx 10$) may occur due to a coarse sampling (typically $N \approx 4000$ vertices) of this high dimensional space. This phenomenon is known as *curse of dimensionality* [Bishop 2007], which is due to the number of data points that is not enough to sample properly data featuring such a high dimensional space.

Regarding our EM classification, we propose three distinct methods to regularize

$\mathbf{\Sigma}_k^p$ based on a regularization parameter h $(0 \leq h \leq 1)$ [Chung & Delingette 2009].

The first method is based on *Spectral Regularization*. The covariance matrix is diagonalized $\mathbf{\Sigma}_k^p = \mathbf{P\Lambda P}^T$ and the lowest eigenvalues (*i.e.* the normalized eigenvalues $< h$) are set to 1% of the highest eigenvalue, thus leading to a new diagonal matrix $\hat{\mathbf{\Lambda}}_h$.

The inverse is then computed as:

$$^h(\mathbf{\Sigma}_k^p)^{-1} = \mathbf{P}\hat{\mathbf{\Lambda}}_h^{-1}\mathbf{P}^T \tag{4.21}$$

This may be seen as performing PCA and filtering the covariance matrix by discarding high frequencies.

In a second approach, *Diagonal Regularization*, the covariance matrix is regularized towards a diagonal matrix controlled by parameter h.

The u, v element of matrix $^h\mathbf{\Sigma}_k^p$ is computed as:

$$^h(\mathbf{\Sigma}_k^p)_{u,v} = (1-h)(\mathbf{\Sigma}_k^p)_{u,v} + h\ \delta_{u,v}\ (\mathbf{\Sigma}_k^p)_{u,v} \tag{4.22}$$

where $\delta_{u,v}$ is the Kronecker symbol.

The higher h, the more diagonally dominant the covariance matrix. This approach has been used in climate modeling to cope with missing values [Schneider 2001].

The last approach, *Constant Regularization*, regularizes the covariance matrix towards a constant matrix $\mathbf{I}d\ tr(\mathbf{\Sigma}_k^p)/M$ controlled by parameter h:

$$^h(\mathbf{\Sigma}_k^p) = (1-h)\ (\mathbf{\Sigma}_k^p) + h\ \mathbf{I}d\ tr(\mathbf{\Sigma}_k^p)/M \tag{4.23}$$

With a high value of h, the covariance matrix converges towards a diagonal matrix with the same variance.

The choice of a covariance matrix regularization method depends on the nature of the data. We have tested the three methods on profiles and we found that *Diagonal* and *Constant Regularization* methods with $h = 0.9$ lead to the most intuitive classification results.

4.3.2 Model order selection

4.3.2.1 State of the art

The objective of model order selection is to find the number of clusters, or *modes*, that best represents the data points using a specific criterion. The result is a partition composed by the number of clusters found. This criterion must optimize the goodness of fit (*i.e.* ensures that the clusters fit at best the data). Such a task is usually performed by launching the clustering algorithm with an increasing number of clusters (*i.e.* $K = 2,\ 3,\ \ldots,\ k_{max}$) and by searching an optimal tradeoff between under and over-fitting through the obtained models.

Under-fitting occurs when the number of clusters is not enough to correctly represent the data points. As a result, non similar data points that should be in two different clusters may be represented by the same cluster. In other words, the model does not have enough degrees of freedom to fit the data points.

In over-fitting, data points are represented by too many clusters, which leads to similar data points represented by more than one cluster. In this case, the model has too many degrees of freedom and features an unnecessary sensitivity to data variation. Most criteria have a tendency to over-fit, their value decreasing, or increasing, in an asymptotic way when the number of clusters increases [Kwon 1998, Pal & Bezdek 1995].

When related to FCM, criteria are known as *cluster validity indices* while they are referred to as *model order selection criteria* when used with EM. A detailed study may be found in [Gheissari & Bab-Hadiashar 2008].

Many cluster validity indices have been proposed [Kim *et al.* 2001, Boudraa 1999, Zahid *et al.* 1999, Kwon 1998, Rezaee *et al.* 1998, Fukuyama & Sugeno 1989]. They typically involve two optimizations: *compactness* and *separation*. Clusters are encouraged to be as compact as possible (*i.e.* minimizing intra-cluster variation) and as separate as possible (*i.e.* maximizing inter-cluster separation) [Kim *et al.* 2004].

Both Bezdek's partition coefficient υ_{PC} and partition entropy υ_{PE}, as well as Xie-Beni's index υ_{XB}, have been initially used. More recently, Kim's υ_{OS} and Saha's Fuzzy Vector Quantization υ_{FVQ} have been proposed.

Bezdek's partition indices Partition coefficient υ_{PC} and partition entropy υ_{PE} proposed by Bezdek [Bezdek 1974b, Bezdek 1974a] are both considered as the first cluster validity indices:

$$\upsilon_{PC} = \frac{\sum_{i=1}^{N} \sum_{k=1}^{K} (\gamma_i^k)^2}{N} \tag{4.24}$$

$$\upsilon_{PE} = -\frac{1}{N} \sum_{i=1}^{N} \sum_{k=1}^{K} [\gamma_i^k \log_a(\gamma_i^k)] \tag{4.25}$$

The best model is obtained by maximizing υ_{PC}, or minimizing υ_{PE}, since high values of γ_i^k correspond to compact clusters. The main issue is that only posterior probabilities γ_i^k are used for the computation of both indices. Therefore, they do not take into account the data structure of clusters and usually feature weak performances.

Xie-Beni's index In addition to compactness, Xie and Beni proposed an index υ_{XB} [Xie & Beni 1991] that encourages the separation between clusters:

$$\upsilon_{XB} = \frac{\sum_{i=1}^{N} \sum_{k=1}^{K} (\gamma_i^k)^2 \, \|\mathbf{x}_i - \boldsymbol{\mu}_k\|^2}{N(\min_{k \neq l} \, \|\boldsymbol{\mu}_k - \boldsymbol{\mu}_l\|^2)} \tag{4.26}$$

where $\boldsymbol{\mu}_k$ and $\boldsymbol{\mu}_l$ are the centers of clusters k and l, respectively.

The numerator accounts for the cluster *compactness* while the denominator accounts for the *separation* between clusters. Compactness means that the average distance between data points and their clusters should be as small as possible. Separation means that the distance between cluster centers should be as large as possible. Combining both compactness and separation in the equation shows that the criterion must be minimized.

Kim's overlap/separation index Claiming that most indices are limited in their ability to compute both compactness and separation, Kim proposed υ_{OS} [Kim *et al.* 2004] to focus on a ratio between inter-cluster overlap and separation:

$$Overlap = \frac{2}{K(K-1)} \sum_{p=1}^{K-1} \sum_{q=p+1}^{K} \times \left[\sum_{\gamma} \sum_{i=1}^{N} \delta(\mathbf{x}_i, \gamma : C_p, C_q) \; \omega(\mathbf{x}_i) \right] \tag{4.27}$$

where (C_p, C_q) are two clusters and $\omega(\mathbf{x}_i)$ $(0 < \omega(\mathbf{x}_i) < 1)$ is a weight factor for each data point, which is determined by the degree of overlap of $\mathbf{x}_i$ between clusters C_p and C_q.

For all pairs of clusters, the separation measures the minimum distance between two clusters. A large separation indicates a partition with well-separated clusters:

$$Separation = 1 - \min_{p \neq q} \left[\max_{x \in X} \, \min(\gamma_{Cp}(x), \gamma_{Cq}(x)) \right] \tag{4.28}$$

Because the overlap and separation measures may have different scales, they must be normalized. After normalization, the criterion is defined as a simple ratio between overlap and separation:

$$\upsilon_{OS} = \frac{Overlap}{Separation} \tag{4.29}$$

Since inter-cluster overlap needs to be minimized in the numerator and inter-cluster separation maximized in the denominator, Kim's υ_{OS} must be minimized. Thus, a small value of υ_{OS} indicates a partition in which the clusters are overlapped to a lesser degree and are more separated from each other.

Saha's Fuzzy Vector Quantization index Unlike most of the existing cluster validity indices, whose objective is to optimize cluster compactness, the Fuzzy Vector Quantization index υ_{FVQ} proposed by Saha [Saha & Bandyopadhyay 2007] aims at optimizing data reconstruction.

Based on a fuzzy quantization-dequantization criterion and the fuzzy vector quantization theory [Pedrycz & Hirota 2007], this index identifies the clusters' ability to properly reconstruct data points using membership function γ (*i.e.* posterior probabilities):

$$\upsilon_{FVQ} = \frac{\sum_{i=1}^{N} \parallel \mathbf{x}_i - \mathbf{x}'_i \parallel^2}{N \ d_{min}(C_p, C_q)} \tag{4.30}$$

where $d_{min}(C_p, C_q)$ represents the minimum distance between any pair of cluster centers and $\mathbf{x}'_i$ is an approximation, or *reconstruction*, of $\mathbf{x}_i$ using cluster centers and membership function:

$$\mathbf{x}' = \frac{\sum_{k=1}^{K} (\gamma_i^k)^m \boldsymbol{\mu}_k}{\sum_{k=1}^{K} (\gamma_i^k)^m} \tag{4.31}$$

where m is the fuzziness coefficient from the FCM algorithm (see equation 4.1).

The interesting point of this index is that it considers FCM as a compression algorithm in which the clusters, referred as *the prototypes*, are a compressed version of data points. The probabilities, referred as *the code book*, are used to reconstruct data points from the cluster centers.

The method used is similar to the Minimum Message Length (MML) theory [Wallace & Freeman 1987], which is based on the idea that statistical inference may be viewed as data compression. If we may build a short code, or *compressed version*, for the available data, we will have a good data generation model [Rissanen 1989].

Regarding EM model order selection, several criteria based on information theory [Bishop 1995] have been proposed but the most commonly used are AIC, $AICc$ and BIC.

AIC The Akaike Information Criterion (AIC) [Akaike 1974] is a parsimonious approach to the estimation of relative expected Kullback-Leibler divergence:

$$AIC = -2 \ \log(p(X|\Theta_K)) + 2\upsilon_K \tag{4.32}$$

where Θ_K is the maximum likelihood estimator and υ_K is the number of degrees of freedom of the model defined as:

$$\upsilon_K = K - 1 + Kd + K(d(d+1)/2) \tag{4.33}$$

for a Gaussian mixture model with full covariance matrices. K is the number of clusters and d is the number of dimensions featured by data points.

This criterion is grounded in the concept of entropy and describes the tradeoff between bias and variance. However, it has a tendency to choose too complex models (*i.e.* over-fitting).

AICc A second-order AIC [Burnham & Anderson 1998] has been derived with an additional bias correction term to better handle cases when there are too many parameters with respect to the sample size:

$$AICc = -2\,\log(p(X|\Theta_K)) + 2\,\upsilon_K \left(\frac{N}{N - \upsilon_K - 1}\right) \tag{4.34}$$

This criterion may be seen as a corrected version of the AIC for samples featuring a small size.

BIC The Bayesian Information Criterion (BIC) [Schwarz 1978], or *Schwarz Criterion* (also *SBC* and *SBIC*), is often preferred:

$$BIC = -2\,\log(p(X|\Theta_K)) + \upsilon_K\,\log(N) \tag{4.35}$$

This criterion is inspired by Bayes but is equivalent to the Minimum Description Length (MDL) criterion [Rissanen 2007, Grünwald 2007, Rissanen 1978], which provides a model featuring the best data compression. Indeed, BIC penalizes the over-fitting by introducing a penalty term for the number of parameters in the model, which is stronger than AIC's. As for AIC and AICc, BIC is an asymptotic result derived under the assumption that data distribution is in the exponential family.

To overcome the issue of choosing the number of clusters K, an interesting method called Mean Shift was proposed [Comaniciu & Meer 2002]. Though claimed as a non-parametric technique, a scale parameter needs to be specified and thus, the search for the optimal K is replaced by the search for an optimal scale parameter. An approach was proposed to automatically select this scale parameter [Bugeau & Pérez 2007], but its application appears complex for our method especially if needed to be performed for every dataset.

Another method worth mentioning proposes a model order selection starting from a maximum number of clusters until a minimum [Figueiredo & Jain 2002]. The method is claimed to integrate both estimation and model selection in a single algorithm, not requiring a careful initialization. Although the proposed EM improvements seem very attractive, some drawbacks must be taken into account. Indeed, this method may fail to find the global optimal solution (*i.e.* the right number of clusters) when there is great disparity between the components' mixing

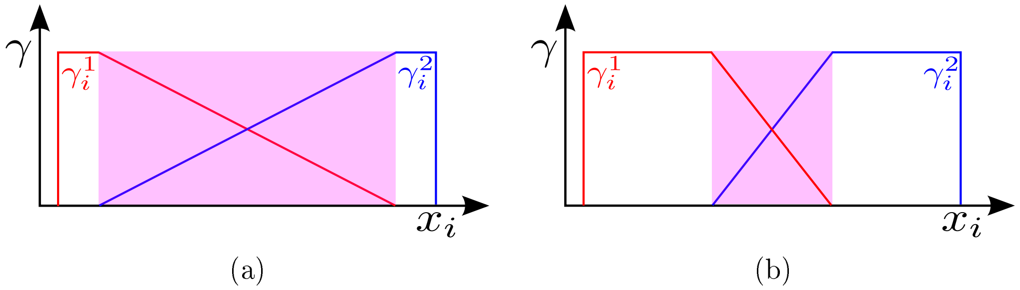


Figure 4.7: Overlap between two clusters. For simplicity, each cluster is assumed to be represented by a triangular fuzzy set. Zone in violet depicts the overlap between cluster #1 (red) and cluster #2 (blue), which is greater in (a) than in (b). Overlap is computed as $[2\ \gamma_i^2/(\gamma_i^1+\gamma_i^2)]$, where cluster #1 and cluster #2 are the clusters data point $\mathbf{x}_i$ has the highest probability γ_i to belong to. The overlap measure penalizes vague data points and encourages them to sharply belong to one specific cluster.

proportions [Wang *et al.* 2004]. Also, as the algorithm starts by a maximum number of clusters towards a minimum number, the required computation time may be quite important, knowing that the maximum number of clusters must be chosen with caution as we may miss the global optimal solution otherwise.

4.3.2.2 *OSI* criterion

In addition, we worked on a criterion that would measure the quality of the EM classification for a given number of modes K while being suitable for our needs. The proposed new non parametric model order selection criterion is called *Overlap Separation Index* (OSI) and is inspired by Kim *et al.* [Kim *et al.* 2004]. This criterion is solely based on the posterior probability ${}^p\gamma_i^k$, which is the probability for profile $\mathbf{x}_i$ from mesh $\mathcal{M}_p$ to belong to mode k, and penalizes the overlap between modes while encouraging their separation.

More precisely, the criterion is computed as $[OSI = \frac{C_1}{C_2}]$. The first term C_1 sums the amount of overlap $[2\ {}^p\gamma_i^s/({}^p\gamma_i^r +^p \gamma_i^s)]$ between the best two modes r and s for profile $\mathbf{x}_i$, *i.e.* modes with the highest ${}^p\gamma_i^k$ (see Figure 4.7). The second term C_2 is the minimum separation between any pair of modes. The separation between pair of modes r and s is computed as the sum of $[2\ {}^p\gamma_i^s/({}^p\gamma_i^r +^p \gamma_i^s)]$ for all profile $\mathbf{x}_i$ being classified to mode r and $[2\ {}^p\gamma_i^r/({}^p\gamma_i^r +^p \gamma_i^s)]$ for all profile $\mathbf{x}_i$ being classified to mode s (see Figure 4.8).

As a first step, we tested the performance of *OSI* criterion on standard datasets in the literature. Experiments on 2D data points [Bandyopadhyay & Maulik 2002a, Maulik & Bandyopadhyay 2000, Bandyopadhyay *et al.* 1998] and 3D data points [Bandyopadhyay & Pal 2007, Bandyopadhyay & Maulik 2002b] showed the robustness of *OSI* criterion with respect to classical model order selection criteria.

Then, we tested *OSI* criterion on intensity profiles from liver meshes. Table 4.1

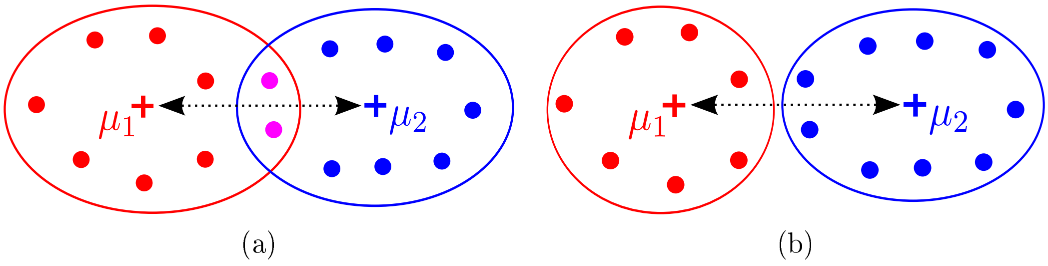


Figure 4.8: Separation between two clusters. Unlike conventional separation indices that base their separation measure on $\|\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\|$, which gives the same value for both (a) and (b), *OSI* criterion computes the separation as the sum of $[2\ \gamma_i^2/(\gamma_i^1 + \gamma_i^2)]$ for all data points $\mathbf{x}_i$ being classified to cluster #1 (red) and $[2\ \gamma_i^1/(\gamma_i^1 + \gamma_i^2)]$ for all data points $\mathbf{x}_i$ being classified to cluster #2 (blue). Using probabilities γ_i^k of data points $\mathbf{x}_i$ to belong to cluster $\#k$ ensures to take the overall geometric structure between clusters into account [Kim *et al.* 2004].

		Outward profiles					Inward profiles				
Regularization	h	*OSI*	*FVQ*	*AIC*	*AICc*	*BIC*	*OSI*	*FVQ*	*AIC*	*AICc*	*BIC*
Diagonal	0.9	4	2	3	3	2	2	2	4	4	2
Diagonal	1.0	3	2	3	3	2	2	2	4	4	2
Constant	0.9	4	2	5	5	2	2	2	6	6	2
Constant	1.0	3	2	3	3	2	2	2	5	5	2

Table 4.1: Selection of the optimal number of EM modes for outward and inward profiles of a liver mesh when using different model order selection criteria {*OSI*,*FVQ*,*AIC*,*AICc*,*BIC*} with different combinations for the regularization of the covariance matrix {Diagonal,Constant,h}, where h is the regularization parameter presented in section 4.3.1.3.

shows the performance of *OSI* criterion with respect to different model order selection criteria on a liver mesh with varying regularization methods. The number of modes being tested varies between 2 and 10. With outward profiles, the expected number of modes is at least three (for air, bones and soft tissue) while two modes are expected with inward profiles (for parenchyma and non-parenchyma tissue).

From these experiments, we found that classical model order selection criteria have limited performances when dealing with intensity profiles. *OSI* criterion gives the most consistent results with a limited sensitivity to regularization methods and h parameter. More comparative results between model order selection criteria may be found in section 6.3.

4.3.3 Spatial regularization

The EM algorithm does not take the neighborhood information of profiles into account, since every data point is given an independent posterior probability ${}^p\gamma_i^k$ to belong to clusters. This leads to non smooth posterior probability maps, which may impair the fusion of appearance regions.

Since this fusion is an important feature in our MPAM construction, we propose accounting for the connectivity between profiles. For that, we use the Neighborhood EM algorithm (NEM) [Ambroise *et al.* 1997], as presented in section 4.2.4, since it nicely extends EM and leads to efficient computation (compared to Markov Random Field).

In our case, as we extract profiles from 2-simplex meshes [Delingette 1999, Montagnat & Delingette 1998], each vertex has exactly three neighbors. Thus, v_{ij} has only three non-zero neighboring values to take into account in equation 4.15, which substantially speeds-up the computation. In practice, less than 5 iterations are necessary to obtain stable posterior probabilities.

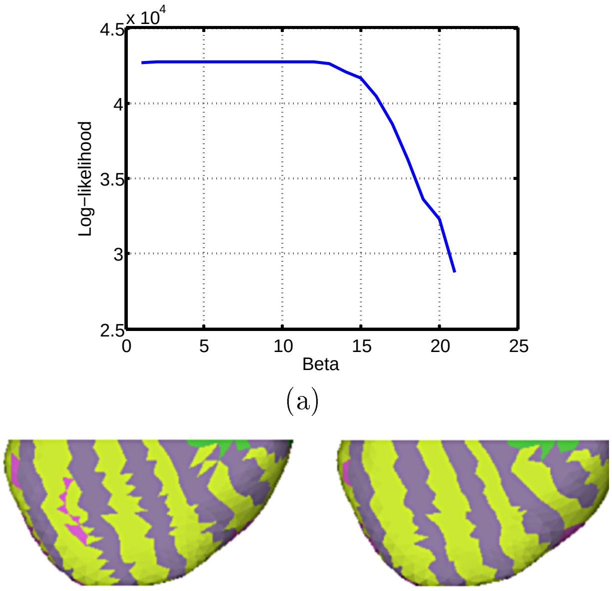
To compute the neighborhood parameter v_{ij} itself, we must define a similarity measure between profiles to define whether they are neighbors or not. Euclidean, median, or Mahalanobis distances have been commonly used in the literature for such a task, as well as the correlation, cross-correlation and normalized cross-correlation.

In our case, we choose to use the correlation coefficient between neighboring profiles $\mathbf{x}_i$ and $\mathbf{x}_j$ because of its simple and fast computation. With this choice, the spatial regularization of posterior probabilities is stronger between similar neighboring profiles, similarly to anisotropic diffusion in image processing. This prevents the blurring of tissue modes that would have occurred with a constant v_{ij} value.

Last but not least, the choice of parameter β_n is an important issue as it will define the spatial regularization importance. For a too small β_n, the spatial regularization would be insufficient and useless. Conversely, an artificially high value of β_n may damage the EM classification result.

This is why we propose to set this parameter automatically by using an heuristic proposed by Dang [Dang & Govaert 1998]. This heuristic consists in using NEM with increasing values of β_n, and then detecting the β_n value above which the log-likelihood $\mathcal{L}(\Theta)$ described in equation 4.15 sharply decreases (see an example in Figure 4.9, top). This sharp decrease is due to the fact that too much spatial regularization leads to a significantly worse profile classification captured by the $\mathcal{L}(\Theta)$ functional.

In practice, NEM is launched in an iterative way starting with a small value of β_n. At the end of the first iteration, log-likelihood $\{\mathcal{L}(\Theta)\}^{t=0}$ is stored. After each iteration, the new $\{\mathcal{L}(\Theta)\}^{t+1}$ is compared with the previous $\{\mathcal{L}(\Theta)\}^{t}$. Before the sharp decrease, the new $\{\mathcal{L}(\Theta)\}^{t+1}$ is slowly increasing, *i.e.* $(\{\mathcal{L}(\Theta)\}^{t+1} - \{\mathcal{L}(\Theta)\}^{t}) = \varepsilon$. The beginning of the sharp decrease is considered as reached when the difference between both values is negative, *i.e.* $(\{\mathcal{L}(\Theta)\}^{t+1} - \{\mathcal{L}(\Theta)\}^{t}) < 0$. The classification associated with the last iteration is then saved. The proposed approach is a fully automatic way to spatially regularize posterior probabilities (see an example in Figure 4.9, bottom).



(a)

(b) (c)

Figure 4.9: Spatial regularization using Neighborhood EM algorithm applied on the EM classification of outward profiles from a liver mesh: (a) parameter β_n is chosen as the value above which the log-likelihood $\mathcal{L}(\Theta)$ described in equation 4.15 sharply decreases, liver mesh with associated profile classification (b) before and (c) after NEM regularization (see section 4.3.6 for more results).

4.3.4 Fusion of modes

After classifying the profiles from P meshes, we propose to merge the generated modes that may be similar between meshes. These modes, which may be seen as *appearance regions*, are extracted on the same structure for P different subjects. Profiles of each mesh $\mathcal{M}_p$ have been classified and lead to K_p modes (*i.e.* K_p may vary among meshes, *e.g.* due to the occurrence of pathologies). This fusion step is not a requirement *a priori*, as we could just keep all the modes from the P meshes and add them directly to the MPAM. However, we expect modes from different meshes to be similar (*i.e.* tissue around the same structure in different images is expected to be mainly the same). The objective is thus to reduce the complexity of the MPAM, by making sure that every mode is as unique as possible.

In order to have a meaningful comparison between modes, an intensity normalization may be required (*e.g.* to cope with images from different imaging systems that are likely to feature different intensity distributions). At image level, this may be done by histogram normalization [Gonzalez & Woods 2008, Acharya & Ray 2005], among others [Vovk *et al.* 2007, Likar *et al.* 2004, Weisenfeld & Warfteld 2004].

With CT images, we found it best not to perform any normalization as the intensity is naturally normalized with the Hounsfield scale. For MR images, which are usually subject to intensity inhomogeneities (*e.g.* artifacts, phantoms and bias field), we propose a method to normalize all modes before fusion that does not require the normalization of the whole images' intensity. To do so, a normalization factor $\mathcal{N}^p_{\boldsymbol{\mu},\sigma}$ is computed through all modes and through all dimensions.

To normalize modes in the interval [0-100], normalization factor $\mathcal{N}^p_{\boldsymbol{\mu},\sigma}$ is defined as:

$$\mathcal{N}^p_{\boldsymbol{\mu},\sigma} = \frac{100}{(\max^p_{\boldsymbol{\mu},\sigma} - \min^p_{\boldsymbol{\mu},\sigma})} \tag{4.36}$$

where $\max^p_{\boldsymbol{\mu},\sigma}$ and $\min^p_{\boldsymbol{\mu},\sigma}$ are, respectively, the maximum and minimum value of the curves created by the mode means with their corresponding standard deviations, *i.e.* $\boldsymbol{\mu}^p_k \pm \sqrt{\sigma^p_k}$, where σ^p_k is the diagonal of the covariance matrix $\boldsymbol{\Sigma}^p_k$ (see some examples of these curves in Figure 4.10).

Then, mean $\boldsymbol{\mu}^p_k$ and covariance matrix $\boldsymbol{\Sigma}^p_k$ are normalized as $\tilde{\boldsymbol{\mu}^p_k}$ and $\tilde{\boldsymbol{\Sigma}^p_k}$, respectively:

$$\tilde{\boldsymbol{\mu}^p_k} = \mathcal{N}^p_{\boldsymbol{\mu},\sigma} \times (\boldsymbol{\mu}^p_k - \min^p_{\boldsymbol{\mu},\sigma}) \tag{4.37}$$

$$\tilde{\boldsymbol{\Sigma}^p_k} = (\mathcal{N}^p_{\boldsymbol{\mu},\sigma})^2 \times \boldsymbol{\Sigma}^p_k \tag{4.38}$$

After this proposed normalization procedure, similar modes may be merged. To do so, we measure the similarity between any pair of modes M_k^p and M_l^q ($M_k^p = \{\boldsymbol{\mu}_k^p, \boldsymbol{\Sigma}_k^p\}$ and $M_l^q = \{\boldsymbol{\mu}_l^q, \boldsymbol{\Sigma}_l^q\}$) by using the Jaccard index (*i.e.* ratio of the intersection of two sets over their union).

Jaccard index $\mathcal{J}$ between M_k^p and M_l^q is defined as:

$$\mathcal{J} = \frac{\sum_{i=1}^{i=d} \left[\boldsymbol{\mu}_k^p(i) - \sqrt{\sigma_k^p(i)}, \boldsymbol{\mu}_k^p(i) + \sqrt{\sigma_k^p(i)}\right] \cap \left[\boldsymbol{\mu}_l^q(i) - \sqrt{\sigma_l^q(i)}, \boldsymbol{\mu}_l^q(i) + \sqrt{\sigma_l^q(i)}\right]}{\sum_{i=1}^{i=d} \left[\boldsymbol{\mu}_k^p(i) - \sqrt{\sigma_k^p(i)}, \boldsymbol{\mu}_k^p(i) + \sqrt{\sigma_k^p(i)}\right] \cup \left[\boldsymbol{\mu}_l^q(i) - \sqrt{\sigma_l^q(i)}, \boldsymbol{\mu}_l^q(i) + \sqrt{\sigma_l^q(i)}\right]}$$

where $[a, b] \cap$ (respectively, $\cup$) $[c, d]$ is the intersection (respectively, the union) of the two modes, and $\boldsymbol{\mu}_k^p \pm \sqrt{\sigma_k^p}$ is the region spanned by both mean $\boldsymbol{\mu}_k^p$ and standard deviation $\sqrt{\sigma_k^p}$ (σ_k^p is the diagonal of covariance matrix $\boldsymbol{\Sigma}_k^p$, see Figure 4.10).

This index is equal to 1 when the 2 modes have the same mean and variance and 0 when they have no values in common. A threshold $\mathcal{J}(M_k^p, M_l^q)$ in the interval $[0, 1]$ is used to decide whether the two modes M_k^p and M_l^q are equivalent. By defining two modes as equivalent depending on a threshold on the Jaccard index $\mathcal{J}$, we have a convenient way to control the global number of modes. Results on a liver database may be found in section 6.7.

In fact, using the Jaccard index is equivalent to creating a graph where nodes represent the modes and arcs link the modes found to be equivalent. The number of connected components of this graph is the number of independent modes K. For connected components with only one node (*i.e.* without equivalence), modes are directly included into the MPAM with a new index m.

For connected components having r equivalent nodes, mean $\tilde{\boldsymbol{\mu}}_m$ of the new mode m is computed as the weighted sum of the means $\boldsymbol{\mu}_r$ of these r equivalent nodes:

$$\tilde{\boldsymbol{\mu}}_m = \frac{\sum_{j=1}^{r} \sum_{i=1}^{N} \gamma_i^j \boldsymbol{\mu}_j}{\sum_{j=1}^{r} \sum_{i=1}^{N} \gamma_i^j} \tag{4.39}$$

Similarly, the covariance matrix $\tilde{\boldsymbol{\Sigma}}_m$ of the new mode m is computed as the weighted sum of the covariance matrices $\boldsymbol{\Sigma}_r$ of these r equivalent nodes:

$$\tilde{\boldsymbol{\Sigma}}_m = \frac{\sum_{j=1}^{r} \left(\sum_{i=1}^{N} \gamma_i^j (\mathbf{x}_i^j - \boldsymbol{\mu}_j)(\mathbf{x}_i^j - \boldsymbol{\mu}_j)^T\right)}{\sum_{j=1}^{r} \left(\sum_{i=1}^{N} \gamma_i^j\right)} \tag{4.40}$$

The computation of both mean $\tilde{\boldsymbol{\mu}}_m$ and covariance matrix $\tilde{\boldsymbol{\Sigma}}_m$ leads to K independent modes $\{\tilde{\boldsymbol{\mu}}_m, \tilde{\boldsymbol{\Sigma}}_m\}$ and an equivalence table $\eta(p, m)$ establishing the new index m of mode k associated with mesh $\mathcal{M}_p$.

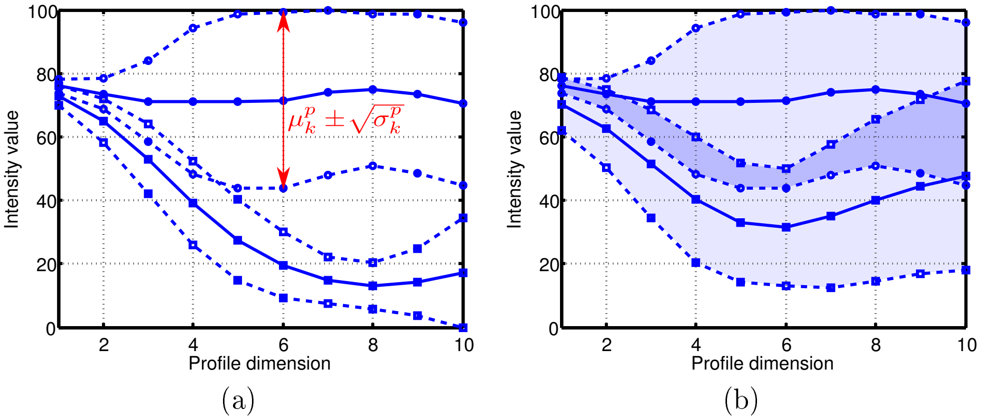


Figure 4.10: Fusion of two modes using the Jaccard index: (a) Mean $\boldsymbol{\mu}_k^p$ (solid lines) and standard deviation $\sqrt{\sigma_k^p}$ (dashed lines) of two modes (circle and square) after EM classification of profiles (σ_k^p is the diagonal of $\boldsymbol{\Sigma}_k^p$), and (b) similarity between pairs of modes defined as the ratio of the intersection (dark blue) over the union (light blue) of their variance surface.

An alternative to this first step could be to perform an EM classification of all profiles for all P subjects with model order selection to find the optimal number of modes. The advantage would be to launch only once the EM classification procedure, instead of classifying each dataset's profiles separately. However, such an approach would lead to a more time consuming task, which would need to be performed each time a new dataset is added.

Instead, we prefer to achieve a separate clustering of each dataset followed by a merging of all modes. This allows us to possibly retrieve one dataset from the global framework without recomputing the others, should this dataset be weak in terms of appearance information. Another advantage of our approach is that it is not biased by the variation of mesh resolution between datasets. In case of EM classification of all datasets' profiles, the mesh resolution would need to be exactly the same throughout the datasets, so that no dataset gets more profiles and thus no more impact than the others in the classification.

4.3.5 Projection into a reference framework

To build the MPAM, a reference mesh $\mathcal{M}^\star$ is required to store the appearance information. The objective is to provide a geometric embedding for the independent modes, in which the P EM classifications need to be *registered*. More precisely, posterior probabilities ${}^p\gamma_i^k$ from the P datasets should be projected into this common reference framework. Because all profile modes are computed independently to any registration, and because the MPAM is multimodal (*i.e.* less dependent to the point correspondence), our approach is not as sensitive to registration errors as the classical PCA-based methods [Heimann 2009]. However, there is still a need to build a mapping between each mesh $\mathcal{M}_p$ and the reference mesh $\mathcal{M}^\star$.

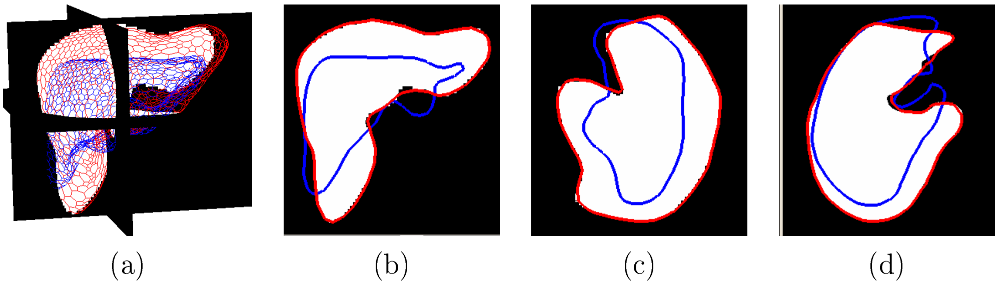

(a) (b) (c) (d)

Figure 4.11: Liver mesh registered towards a binary image with globally-constrained deformations: (a) 3D view of original mesh (blue) and registered mesh (red), and {(b),(c),(d)} corresponding 2D delineations in the three reference planes.

Reference mesh $\mathcal{M}^\star$ construction Our first idea consists in non-rigidly registering all P meshes $\mathcal{M}_p$ towards the same reference mesh with a coarse-to-fine deformable surface approach. More precisely, each mesh is registered towards a binary image with the globally constrained deformation framework described in section 3.3.2.3 (see an illustration in Figure 4.11). Though results are usually quite satisfactory, our experiments showed that some meshes may diverge from the expected result (*i.e.* boundaries of the binary image) during mesh deformation. To overcome this problem, we change several times the mesh rigidity (see section 3.3.4.2) during mesh deformation in a coarse-to-fine fashion. However, this solution requires a user intervention to optimize this coarse-to-fine procedure for each mesh. Also, the registration using a binary image as target (*i.e.* mesh-to-image registration) is not strictly accurate due to the rasterization (*i.e.* the accuracy depending on voxel size).

This is why we prefer to use a mesh-to-mesh registration. The method consists in first randomly initializing the reference mesh as one of the P meshes. As a first rigid transformation, we use a robust point set registration using Gaussian mixture models to perform a rigid registration without point correspondence [Jian & Vemuri 2005]. Then, we register the current reference mesh on all meshes before finally recomputing the reference mesh as the mean of the deformed instances (see an illustration in Figure 4.12, left). This process is iterated several times until the mesh reaches a steady state, *i.e.* typically 3 iterations in our case (see an example in Figure 4.12, right).

Registration between $\mathcal{M}^\star$ and $\mathcal{M}_p$ After defining a reference mesh $\mathcal{M}^\star$, a registration between the reference mesh $\mathcal{M}^\star$ and the P instance meshes needs to be performed. This registration may be done in both ways: deforming each instance towards the reference mesh (see an illustration in Figure 4.13, left), or deforming the reference mesh towards each mesh (see an illustration in Figure 4.13, right).

In both cases, we propose to use a mesh-to-mesh registration based on currents [Glaunes 2005, Vaillant & Glaunes 2005], which have been successfully used on brain [Durrleman *et al.* 2009] and heart [Mansi 2010] meshes. The mesh defor-

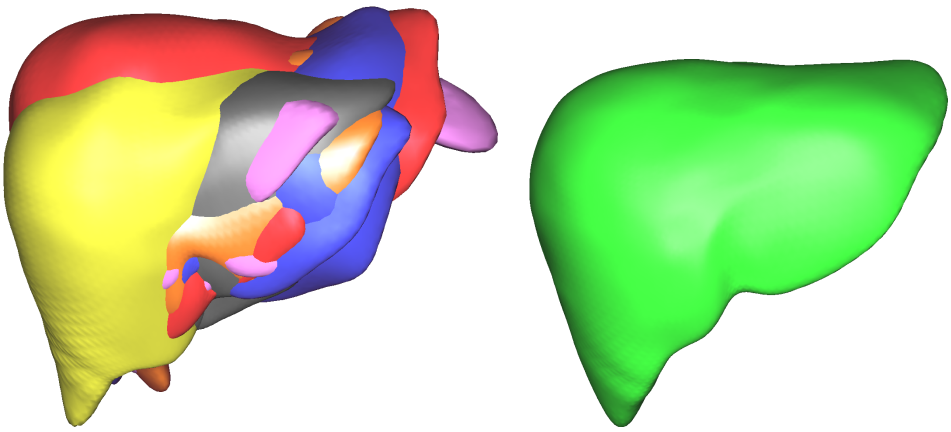

(a) (b)

Figure 4.12: Creation of the reference mesh $\mathcal{M}^\star$ from 16 liver meshes using a mesh-to-mesh registration based on currents: (a) registration of all instance meshes towards the current reference mesh, and (b) creation of the final reference mesh $\mathcal{M}^\star$ as the mean of all registered instance meshes.

mation during registration is smooth and a coarse-to-fine framework may be used by changing the width of the Gaussian kernel used to represent both currents and diffeomorphisms. The former parameter is a fit parameter that attracts $\mathcal{M}_p$ to $\mathcal{M}^\star$ while the latter is a rigidity parameter that tries to keep $\mathcal{M}_p$ as smooth as possible.

Transfer of appearance information Last step consists in an interpolation method required to transfer appearance information from $\mathcal{M}_p$ to $\mathcal{M}^\star$ (*i.e.* the clustering associated with $\mathcal{M}_p$ needs to be projected into $\mathcal{M}^\star$). For this purpose, we search for each vertex $\mathbf{p}_i$ of the reference mesh its closest points $Cl(\mathbf{p}_i)$ on mesh $\mathcal{M}_p$, and also consider its neighbors. We then assign in $\mathbf{p}_i$ the modes of $Cl(\mathbf{p}_i)$ and its neighbors and interpolate the posterior probabilities, the interpolation being inversely proportional to the distance. This interpolation not only compensates any misregistration that may have occurred but also smooths the clustering projection.

To ensure that the clustering information from all P meshes $\mathcal{M}_p$ is captured by reference mesh $\mathcal{M}^\star$, the latter should be finer in terms of resolution (*i.e.* number of vertices) than the finest mesh $\mathcal{M}_p$. This way, there is always at least one vertex from $\mathcal{M}^\star$ to capture the clustering information from the associated $\mathcal{M}_p$ vertices (see an illustration in Figure 4.14).

Finally, for each vertex $\mathbf{p}_i$ of $\mathcal{M}^\star$, we compute the posterior probability $\tilde{\gamma}_i^m$ by summing and normalizing the posterior probabilities associated to each mode:

$$\tilde{\gamma}_i^m = \frac{\sum_{p=1}^{P} {}^{p}\gamma_i^{\eta(p,m)}}{\sum_{m=1}^{K}\sum_{p=1}^{P} {}^{p}\gamma_i^{\eta(p,m)}} \tag{4.41}$$

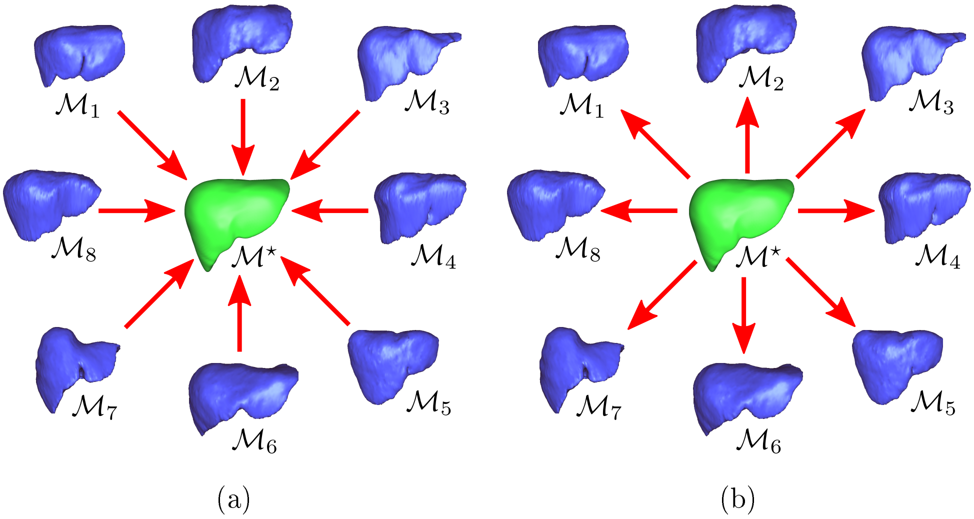


Figure 4.13: The registration between the reference mesh $\mathcal{M}^\star$ and 8 instance meshes $\{\mathcal{M}_1, \ldots, \mathcal{M}_8\}$ may be done in both ways: (a) deforming each instance mesh towards the reference mesh, or (b) deforming the reference mesh towards each instance mesh.

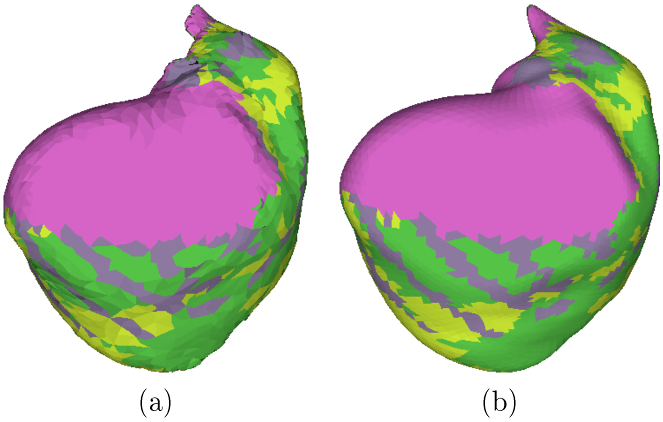


Figure 4.14: Transfer of appearance information from (a) instance mesh $\mathcal{M}_p$ ($N =$ 4530 vertices) to (b) reference mesh $\mathcal{M}^\star$ ($N = 11760$ vertices) using a closest point approach. Note how similar is the clustering despite the change of mesh resolution.

Liver	Resolution(mm)	Image size
#1	0.58x0.58x1.6	512x512x90
#2	0.70x0.70x2.0	512x512x81
#3	0.67x0.67x4.0	512x512x48
#4	0.67x0.67x3.0	512x512x62
#5	0.61x0.61x5.0	512x512x34
#6	0.56x0.56x2.0	512x512x79
#7	0.54x0.54x2.0	512x512x71

Table 4.2: Details of the seven CT images of the liver.

Tibia	Resolution(mm)	Image size
#1	0.50x0.50x1.0	512x512x128
#2	0.29x0.29x0.6	512x512x128
#3	0.50x0.50x1.0	512x512x120
#4	0.50x0.50x1.0	512x512x120

Table 4.3: Details of the four MR images of the tibia.

where P is the number of meshes $\mathcal{M}_p$, K is the number of modes associated with $\mathcal{M}^\star$ and $\eta(p, m)$ is the equivalence table described in section 4.3.4 that establishes the new index m of mode k from mesh $\mathcal{M}_p$.

In practice, this approach leads to sparse posterior probabilities where only a few modes have non-zero posterior probabilities (*i.e.* non-negligible values). By filtering negligible posterior probabilities, we reduce the number of modes associated to each vertex during the local search performed by both the regional external forces presented in section 5.2 for segmentation and the boosted clustering presented in section 5.4, which greatly improves the computational time.

4.3.6 An example on livers and tibias

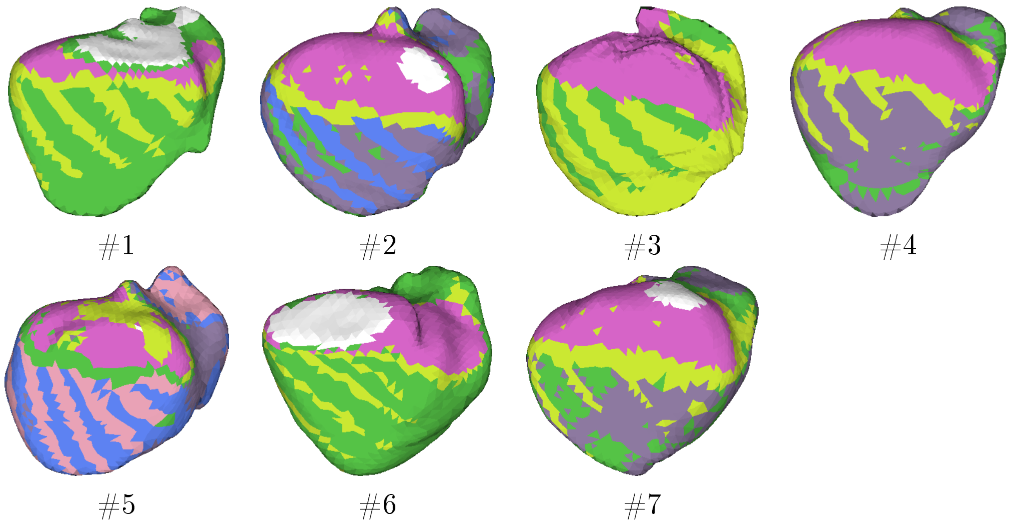
To illustrate the MPAM construction, we have tested our method on seven livers segmented from CT images (see Figure 4.15) and four tibias cropped at knee level segmented from MR images (see Figure 4.18). The CT images of the liver come from 3Dircadb1[1], which is a database provided by IRCAD[2], the French Research Institute against Digestive Cancer. The MR images of the knee are acquired at the University College London Hospital (UCLH) using a T1 FLAIR sequence on a 1.5T MRI device. Details of both CT and MR images may be found in Table 4.2 and Table 4.3, respectively.

For both structures, outward profiles composed of 10 samples extracted every mm are generated from meshes with ≈ 4000 vertices. The optimal number of modes

---

[1] http://www.ircad.fr/softwares/3Dircadb/3Dircadb1

[2] http://www.ircad.fr



#1 #2 #3 #4

#5 #6 #7

Figure 4.15: EM classification of outward profiles performed on 7 liver meshes $\{\#1, \ldots, \#7\}$. For every mesh, the number of modes is: $K_1 = 3$, $K_2 = 5$, $K_3 = 3$, $K_4 = 4$, $K_5 = 6$, $K_6 = 3$ and $K_7 = 4$, respectively. White zones depict vertices associated to incomplete profiles (see section 4.3.1.2).

is estimated in an iterative fashion with the OSI criterion, starting from 2 modes up to 10 modes. At each iteration, EM classification is initialized with Fuzzy C-Means that are themselves initialized with the same number of random cluster centers. For both structures, an average of 4 modes is found. *Constant Regularization* with $h = 0.9$ is used to regularize the EM covariance matrices. Results of the EM classification may be found in Figure 4.15 and Figure 4.18 for liver and tibia meshes, respectively. A plot of mode means and standard deviations (SD) for both liver #4 and tibia #4 may be found in Figure 4.21.

The EM classification is then spatially smoothed with the NEM algorithm. To find the optimal parameter β_n, the heuristic described in section 4.3.3 is used. For that, NEM is launched in an iterative way starting with $\beta_n = 1$ and increasing β_n with a step of 1 at each new iteration. The sharp decrease is considered as reached when the difference between the log-likelihood of two consecutive iterations is negative (*i.e.* $(\{\mathcal{L}(\Theta)\}^{t+1} - \{\mathcal{L}(\Theta)\}^{t}) < 0$). For both structures, $\overline{\beta_n} = 12$ is necessary to reach this sharp decrease (*i.e.* $\overline{\beta_n} = 13$ for liver meshes and $\overline{\beta_n} = 11$ for tibia meshes). Results of the spatial regularization using NEM may be found in Figure 4.16 and Figure 4.19 for liver and tibia meshes, respectively. Probability for each vertex of both liver #4 and tibia #4 meshes (*i.e.* each corresponding intensity profile) to belong to one of their four associated modes may be found in Figure 4.17 and Figure 4.20, respectively.

Both liver mesh #4 (see Figure 4.15) and tibia mesh #4 (see Figure 4.18) are chosen as reference meshes $\mathcal{M}^\star$ for liver and tibia meshes, respectively. To ensure clustering information to be entirely captured by $\mathcal{M}^\star$, both liver #4 and tibia #4 meshes are refined (see section 3.3.2.4). Then, the meshes of both structures are

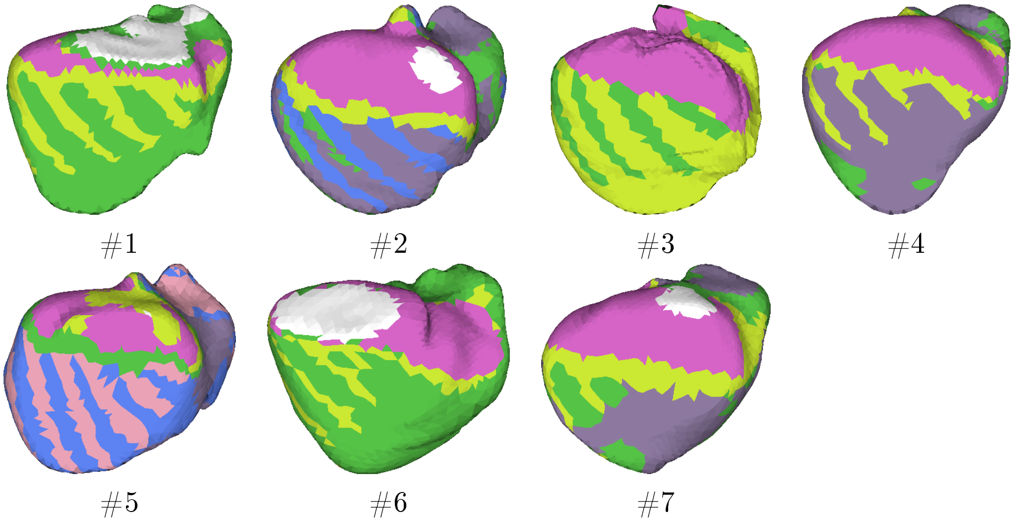

#1 #2 #3 #4
#5 #6 #7

Figure 4.16: Spatial regularization of the EM classification of outward profiles performed on 7 liver meshes $\{\#1, \ldots, \#7\}$ (see Figure 4.15) using the NEM algorithm with parameter $\overline{\beta_n} = 13$.

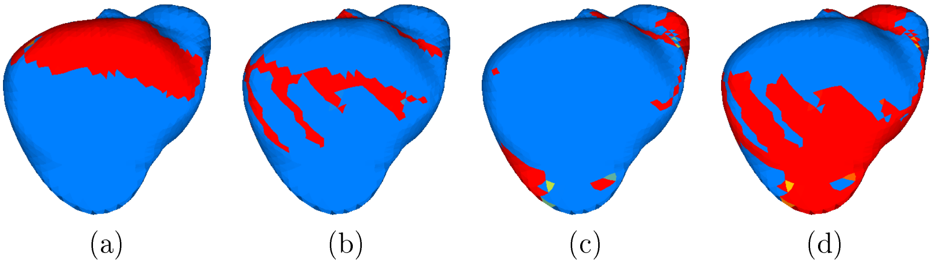

(a) (b) (c) (d)

Figure 4.17: Probability for each vertex of liver mesh #4 (*i.e.* each corresponding intensity profile) to belong to one of its four associated modes depicted in Figure 4.16: (a) pink mode, (b) light green mode, (c) dark green mode, and (d) violet mode. The color map is going from blue (low EM posterior probability) to red (high EM posterior probability).

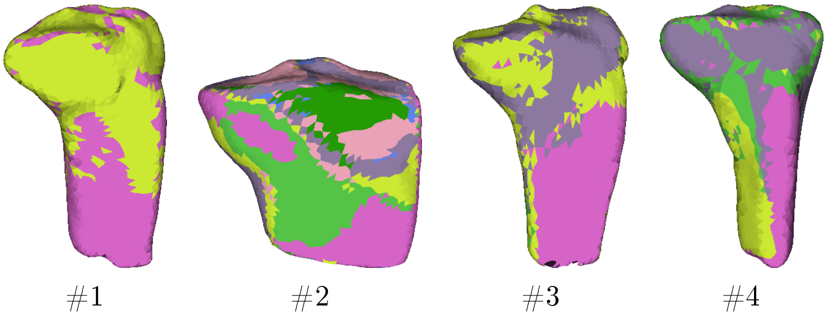


Figure 4.18: EM classification of outward profiles performed on 4 tibia meshes cropped at knee level {#1,#2,#3,#4}. For every mesh, the number of modes is: $K_1 = 2$, $K_2 = 7$, $K_3 = 4$ and $K_4 = 4$, respectively. Tibia mesh #2 has a smaller mesh size compared to the others (*i.e.* only the tibial head is segmented), which is due to a reduced FOV featured by its corresponding High Resolution MR image.

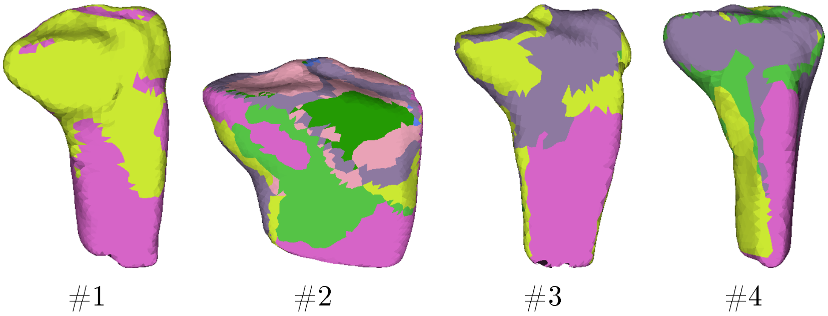


Figure 4.19: Spatial regularization of the EM classification of outward profiles performed on 4 tibia meshes cropped at knee level {#1,#2,#3,#4} (see Figure 4.18) using the NEM algorithm with parameter $\overline{\beta_n} = 11$.

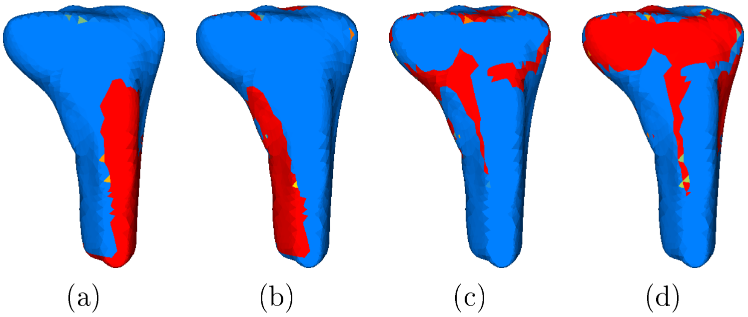
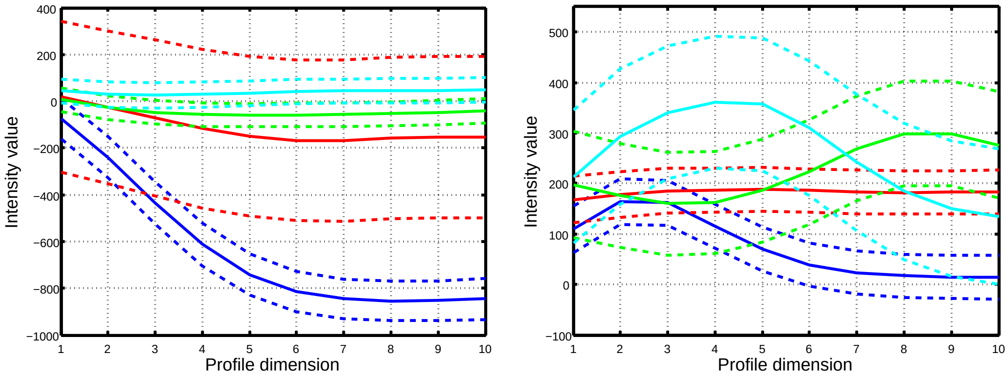

(a) (b) (c) (d)

Figure 4.20: Probability for each vertex of tibia #4 (*i.e.* each corresponding intensity profile) to belong to one of its four associated modes depicted in Figure 4.19: (a) pink mode, (b) light green mode, (c) dark green mode, and (d) violet mode. The color map is going from blue (low EM posterior probability) to red (high EM posterior probability).



(a) (b)

Figure 4.21: Mean (solid lines) ± SD (dashed lines) of the four modes (red, green, blue, cyan) from (a) liver #4 and (b) tibia #4 meshes.

MPAM	Threshold on Jaccard index $\mathcal{J}$									
	1.0	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
Livers	28	27	23	16	11	6	3	1	1	1
Tibias	10	10	10	8	7	5	3	3	1	1

Table 4.4: Evolution of the final number of modes K for both liver and tibia outward profiles when varying the threshold on Jaccard index $\mathcal{J}$.

registered to their respective reference mesh $\mathcal{M}^\star$ using the mesh-to-mesh registration based on currents (see section 4.3.5). Their associated EM classification is projected on $\mathcal{M}^\star$ and modes are possibly merged with Jaccard index. For tibia meshes, the normalization proposed in section 4.3.4 has been performed to cope with intensity inhomogeneities present in MR images. Also, we decide not to include tibia #2 (see Figure 4.15) into the MPAM due to its smaller mesh size compared to the others.

For liver meshes, an initial total number of 28 modes leads to 11 new modes after the merging of profiles with $\mathcal{J} = 0.6$. For tibia meshes, an initial total number of 10 modes leads to 5 new modes with $\mathcal{J} = 0.5$ (see more results in Table 4.4 and the plot of mode means in Figure 4.23). Jaccard index provides thus a simple way to taylor the complexity of the MPAM. Two interesting trends may be highlighted: there is no significant difference between $\mathcal{J} = 1.0$ and $\mathcal{J} = 0.9$ and there is a significant decrease of the final number of modes between $\mathcal{J} = 0.7$ and $\mathcal{J} = 0.6$. These trends are confirmed in section 6.7.

The Multimodal Prior Appearance Models (MPAM) built from both liver and tibia meshes are depicted in Figure 4.22.

4.4 Conclusion

In this chapter, we proposed the construction of Multimodal Prior Appearance Models (MPAM) from the regional clustering of intensity profiles as a novel way to model

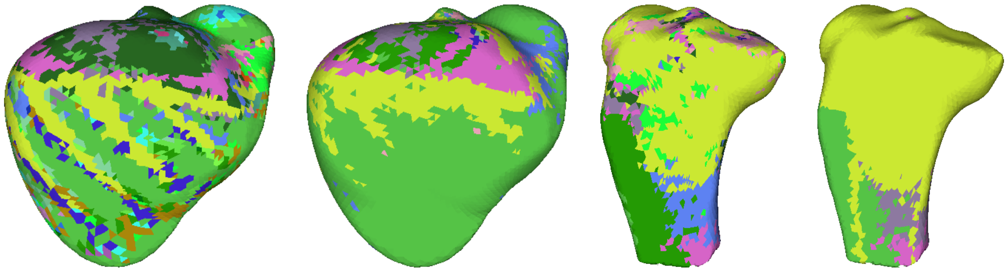

(a) (b) (c) (d)

Figure 4.22: Multimodal Prior Appearance Model (MPAM) built from both {(a)+(b)} liver and {(c)+(d)} tibia meshes when using different thresholds on Jaccard index $\mathcal{J}$: (a) $\mathcal{J}$ = 1.0 (K = 28 modes), (b) $\mathcal{J}$ = 0.6 (K = 11 modes), (c) $\mathcal{J}$ = 1.0 (K = 10 modes), and (d) $\mathcal{J}$ = 0.5 (K = 5 modes). Mode means are plotted in Figure 4.23.

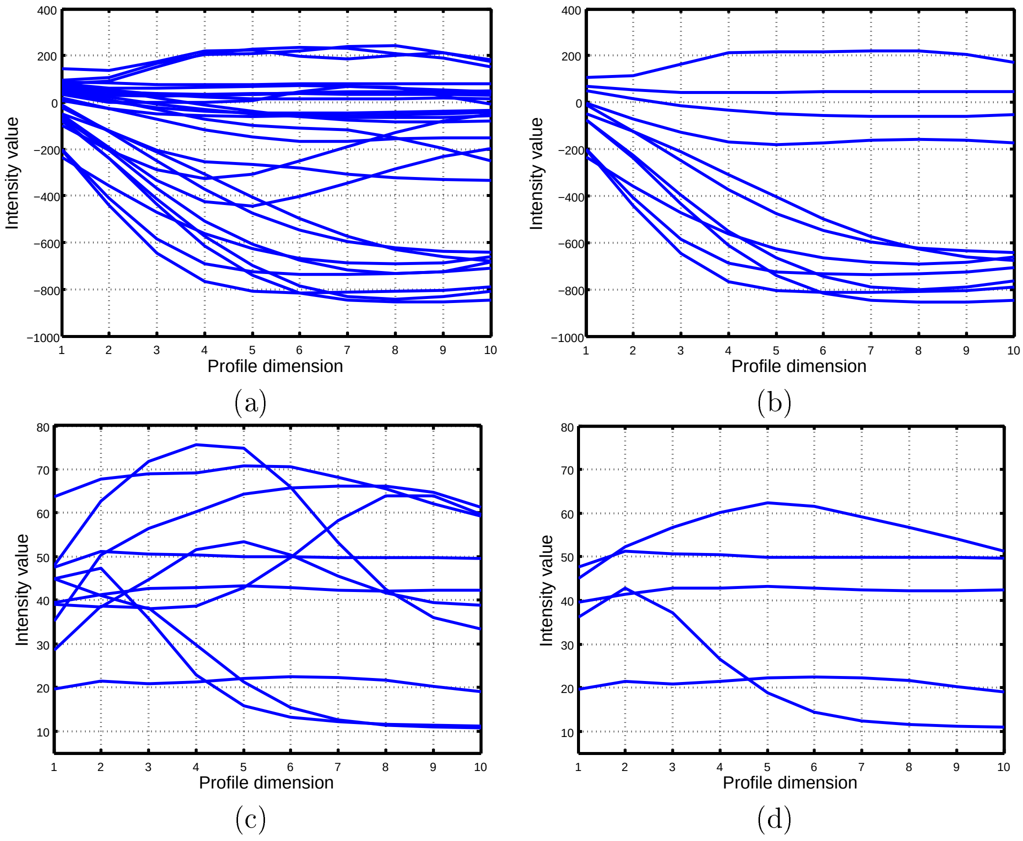


Figure 4.23: Mode means from the MPAM built from both {(a)+(b)} liver and {(c)+(d)} tibia meshes when using different thresholds on Jaccard index $\mathcal{J}$: (a) $\mathcal{J}$ = 1.0 (K = 28 modes), (b) $\mathcal{J}$ = 0.6 (K = 11 modes), (c) $\mathcal{J}$ = 1.0 (K = 10 modes), and (d) $\mathcal{J}$ = 0.5 (K = 5 modes).

the appearance around structures of interest in medical images. The clustering is considered as *regional* because intensity profiles are classified for each mesh, and not for each vertex (*i.e.* over a population of meshes). Unlike PCA-based methods that need an accurate point correspondence, our approach creates modes without requiring an accurate pointwise registration. This is because the statistical analysis of appearance is performed at each mesh, and not at each vertex.

One advantage of this feature is that a meaningful MPAM may be built with very few datasets (in fact one dataset suffices), which makes it well suited for a bootstrapping strategy. Since multimodal, the MPAM is able to cope with large variations of appearance. This is not the case with PCA-based methods, which rely on the hypothesis that the probability density function is well described by one single Gaussian distribution. As we explained, this hypothesis is often violated by the presence of pathologies but also by the fact that shape is not necessarily correlated with appearance.

Intensity profiles are classified using the EM algorithm, and we explained how to cope with missing data and how to regularize the covariance matrix to avoid any singularities due to its inversion. After introducing the principles of unsupervised clustering and model order selection, we presented the *OSI* index as a novel way to automatically determine the number of clusters. To have a spatial anisotropic regularization of EM classification, we presented a spatial regularization approach based on the Neighborhood EM algorithm. Finally, we explained how to project the classification from several datasets into a reference mesh where each vertex has a probability to belong to several intensity profile classes. We also presented the Jaccard index as a way to control the final number of modes associated with the reference mesh, in an attempt to reduce the MPAM complexity.

Chapter 5

Regional External Forces

Contents

5.1 Introduction

In chapter 4, we presented Multimodal Prior Appearance Models (MPAM) for the modeling of appearance around structures of interest in medical images. These models may be seen as appearance priors because appearance is learned through datasets using an EM clustering of intensity profiles. Somehow, MPAM could be just used for statistical studies on tissue appearance (*e.g.* in a population-based study). In this case, the EM clustering would have the classification of intensity profiles as sole and final objective.

However, in our case, our final goal is to efficiently use MPAM for segmentation purposes. In practice, our aim is first to estimate the similarity of profile modes contained in the MPAM with profiles sampled during a local search and then, to compute external forces out of this comparison in a deformable model-based segmentation framework (section 5.2). To do so, we propose to change the construction of the MPAM as proposed in Figure 4.2 and replace both the EM and NEM steps by two other steps, namely *spectral clustering* and *boosted clustering* (see Figure 5.1). We explain the motivation of these changes in the remainder of this introduction.

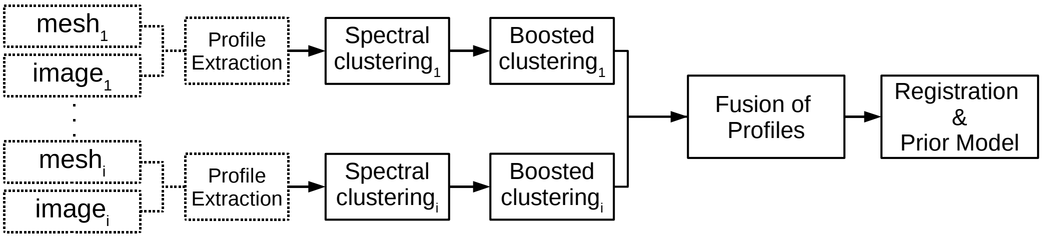
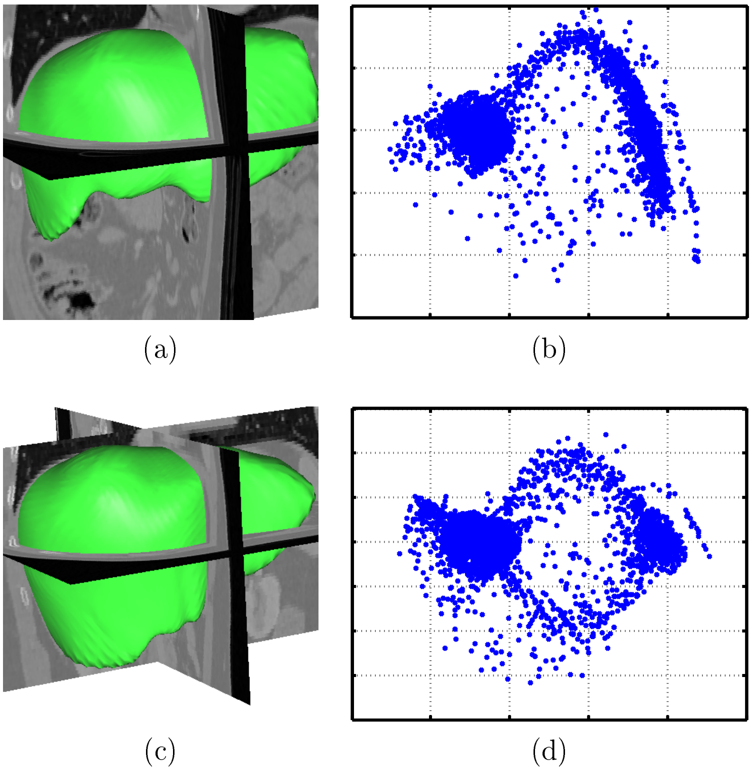


Figure 5.1: New pipeline for Multimodal Prior Appearance Model (MPAM) construction, which includes spectral clustering and our proposed boosted clustering.

Spectral clustering In section 4.3.1, we proposed to classify intensity profiles with the EM algorithm, for which we assumed profiles to be represented by Gaussian Mixture Models (GMM), implying that each mode is defined by its mean profile and associated covariance matrix. Though quite efficient to represent appearance around structures at a given resolution, tests have shown that EM algorithm lacks robustness when changing mesh resolution. Indeed, the well-known curse of dimensionality seems to affect the results for different resolutions (*i.e.* number of profiles changes while profile length stays the same). In fact, the main issue with generative models, such as K-Means, Fuzzy C-Means and especially EM algorithm, is that they usually assume data to have a Gaussian distribution. In medical image analysis, this hypothesis may not always hold [Heimann 2009]. This seems to be the case with intensity profiles too, whose distribution appears not to be Gaussian (this is exemplified in Figure 5.2). As a consequence, EM algorithm is known to be sensitive to its initialization and is thus likely to converge to local maxima. In this case, several restarts are usually needed to converge to global maxima, which makes EM algorithm unstable.

To overcome these drawbacks, we present in this chapter an alternative clustering method, the *spectral clustering* [Donath & Hoffman 1973], which is used in computer vision, machine learning, pattern recognition and VLSI design [Luxburg 2007, Malik *et al.* 2001, Ng *et al.* 2001, Alpert & Yao 1994]. The idea is to perform clustering on a lower dimensional space built from the spectral analysis of similarity matrices. Experiments in the literature have shown that graph vertices have the advantage to reinforce the Gaussian hypothesis on input data points [Luxburg 2007, Ng *et al.* 2001] (see Figure 5.3). Another advantage is to ease the initialization issue, since data points are meant to be better represented with the similarity graph. Last but not least, curse of dimensionality is reduced. Indeed, data dimension is not related to profile length anymore, but rather to spectral data featuring a reduced number of dimensions.


(a) (b)

(c) (d)

Figure 5.2: Plot of intensity profiles (extending from 10 mm inward to 10 mm outward with samples every 1 mm) from two liver meshes {(a),(c)} when projected on a 2D subspace using PCA {(b),(d)}. This 2D subspace corresponds to the two largest eigenvalues. For both meshes, $N = 7586$ profiles are extracted. Note that the distribution of intensity profiles on their respective 2D subspace is not Gaussian.

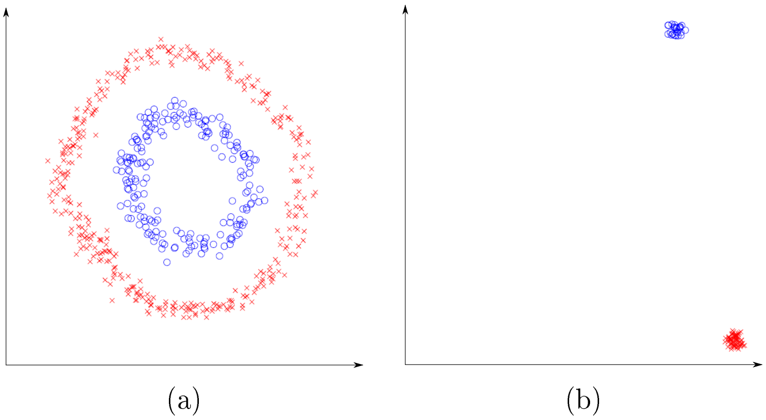

(a) (b)

Figure 5.3: An example of spectral clustering applied on a set of 2D data points: representation of the dataset (a) before and (b) after spectral clustering (Source: [Ng *et al.* 2001]).

The consecutive steps for spectral clustering are depicted in Algorithm 1 and Figure 5.4. First, a similarity function needs to be defined between intensity profiles (section 5.3.1). This similarity function ensures the creation of a spectral graph (section 5.3.2), whose nodes represent the data points (*i.e.* intensity profiles). Two nodes are connected if corresponding data points are similar, *i.e.* if their similarity value is positive or larger than a threshold, and the connecting edge is weighted with the value returned by the similarity function. Spectral clustering consists in finding partitions of the graph such that edges between cluster of vertices have low weights (*i.e.* data points in different cluster are dissimilar from each other) and edges within a cluster have high weights (*i.e.* data points within the same cluster are similar from each other). Then, top eigenvectors are extracted from the affinity matrix associated with the spectral graph (section 5.3.3). The number of top eigenvectors may be computed in an automatic fashion with a model order selection based on a heuristic (section 5.3.4). These top eigenvectors are the new representation of data points, *i.e.* spectral data, which are classified with EM clustering (section 5.3.5).

Algorithm 1 Spectral clustering algorithm (inspired from [Ng *et al.* 2001]).

```
Given a set of intensity profiles $S = \{\mathbf{x}_i, \dots, \mathbf{x}_N\}$ in $\mathbb{R}^d$:
  Define a similarity function and a similarity graph.
  From the similarity graph, form affinity matrix $A \in \mathbb{R}^{N \times N}$
    defined by $A_{ij} = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/(2\sigma^2))$ if $i \neq j$ and $A_{ii} = 0$.
  Define diagonal matrix $D$ whose $(i,i)$-element is the sum of $A$'s $i$-th row.
  Construct Laplacian matrix $L = D^{-1/2}AD^{-1/2}$.
  Find $\{v_1, v_2, \dots, v_K\}$, the $K$ top eigenvectors of $L$.
  Form matrix $X = [v_1 v_2 \dots v_K] \in \mathbb{R}^{N \times K}$ by stacking the eigenvectors in columns.
  Form matrix $Y$ from $X$ by re-normalizing each of $X$'s rows
    to have unit length, i.e. $Y_{ij} = X_{ij}/(\sum_{j=1}^{N} X_{ij}^2)^{1/2}$.
  Cluster the rows of $Y$ (i.e. as points in $\mathbb{R}^K$) into $K$ clusters using EM algorithm.
  Finally, compute EM parameters $\{\boldsymbol{\mu}_k^p, \boldsymbol{\Sigma}_k^p, \pi_k^p\}$ using posterior probabilities ${}^p\hat{\gamma}_i^k$
    from the EM clustering performed in the spectral space.
```

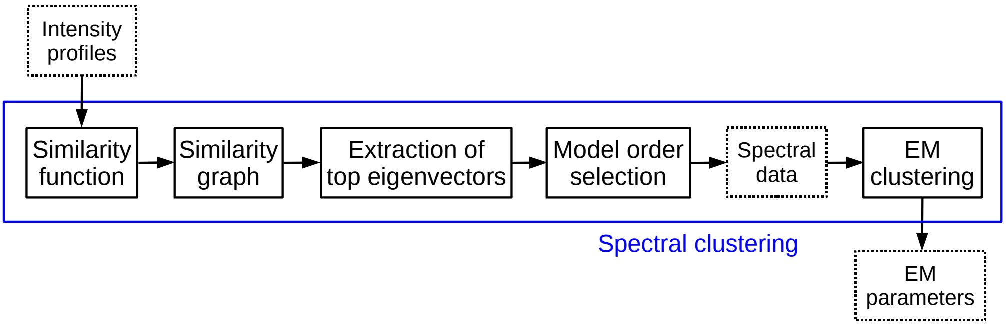


Figure 5.4: Pipeline for spectral clustering.

Boosted clustering To efficiently use the clustering of intensity profiles associated with the MPAM as an appearance prior for segmentation purposes, one question is worth raising:

> How well the clustering (set of centers and covariance matrices for each mode) is suited for image segmentation?

To answer this question, our objective is to perform a clustering under a localization constraint, so that a profile mode should be able to describe in a discriminative way the boundary between the inside and the outside of the structure.

To do so, we propose a boosted clustering that relies on a classifier of intensity profiles (section 5.4). The objective is to tune an EM classification of profiles including a condition on offset α_i (see Figure 5.5), which states whether a profile is well localized by its associated modes. The condition related to offset α_i is based on the hypothesis that at least one of the modes to which the profile belongs to must be able to localize well the boundary along that profile, similarly to what is done with regional external forces (see section 5.2.2). Otherwise, no boundary information would be found at this particular vertex because no suitable mode would be able to find the right boundary.

Similarly to section 4.1, the objective of boosted clustering is to encourage the creation of multiple modes with low variance, instead of few modes with large variance. It provides more discriminant modes to the MPAM (*i.e.* modes that are more representative for a set of profiles), so that these modes are robust when searching for the right boundaries during segmentation.

First, we describe a single pass version of the algorithm (section 5.4.1). Then, to improve the clustering results, we present a cascading boosted clustering (section 5.4.2) and improve it with a model order selection based on a hierarchical approach (section 5.4.3). Finally, we discuss about the proposed algorithm (section 5.4.4).

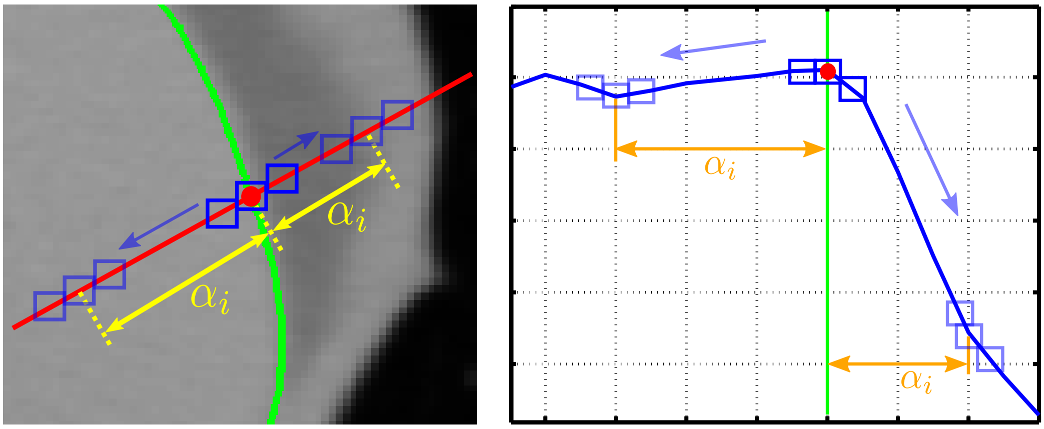
5.2 Localization criterion and MPAM external forces

5.2.1 Localization criterion

The objective of this localization criterion is to test whether a given profile mode M_k (*i.e.* mean $\boldsymbol{\mu}_k$ and covariance matrix $\boldsymbol{\Sigma}_k$) is able to unambiguously determine the boundary at a given vertex. More precisely, the offset which maximizes a similarity criterion (or minimizes a distance) between the current profile and the associated profile mode is computed and must be small enough. This localization criterion is used later on in this chapter by the boosted clustering.

First step consists in choosing an intensity similarity measure. Several measures may be used (see section 3.3.5.4). In our case, we decide to use *sum of absolute differences*, *linear criterion* and *Mahalanobis distance.*

The objective is to compute the offset α_i that maximizes a similarity value (for linear criterion) or minimizes a distance (for sum of absolute differences and



(a) (b)

Figure 5.5: In boosted clustering, the search for the optimal offset α_i along the normal (red line) at a given vertex (red circle) consists in shifting a profile (blue squares) inward and outward (blue arrows) mesh surface (green line): (a) 2D slice of the CT image from which profiles are sampled, and (b) 2D plot of the corresponding sampled intensity values.

Mahalanobis distance). In practice, the profile is regularly sampled along the normal (*e.g.* every millimeter) and profile offsets are obtained by shifting the profile inward and outward (see an illustration in Figure 5.5).

Let $\mathbf{x}(\mathbf{p}_i, \mathbf{n}_i)$ be the intensity profile at vertex $\mathbf{p}_i$ along normal $\mathbf{n}_i$.

Let α_i be the offset of that profile along the normal as $\mathbf{x}(\mathbf{p}_i + \alpha_i \mathbf{n}_i, \mathbf{n}_i)$.

Let mean $\boldsymbol{\mu}_k$ and covariance matrix $\boldsymbol{\Sigma}_k$ be the EM parameters of profile mode M_k the profile $\mathbf{x}(\mathbf{p}_i, \mathbf{n}_i)$ belongs to.

Offset $\{\alpha_i\}_{\text{line}}$ needs to maximize the linear criterion:

$$\{\alpha_i\}_{\text{line}} = \arg\max_{\alpha_i, k} \left(\frac{(\mathbf{x}(\mathbf{p}_i + \alpha_i \mathbf{n}_i, \mathbf{n}_i) \cdot \boldsymbol{\mu}_k)^2}{(\mathbf{x}(\mathbf{p}_i + \alpha_i \mathbf{n}_i, \mathbf{n}_i)^2) \cdot (\boldsymbol{\mu}_k^2)} \right) \tag{5.1}$$

Both offsets $\{\alpha_i\}_{\text{diff}}$ and $\{\alpha_i\}_{\text{maha}}$ need to minimize the sum of absolute differences and the Mahalanobis distance, respectively:

$$\{\alpha_i\}_{\text{diff}} = \arg\min_{\alpha_i, k} \left(|\mathbf{x}(\mathbf{p}_i + \alpha_i \mathbf{n}_i, \mathbf{n}_i) - \boldsymbol{\mu}_k| \right) \tag{5.2}$$

$$\{\alpha_i\}_{\text{maha}} = \arg\min_{\alpha_i, k} \left((\mathbf{x}(\mathbf{p}_i + \alpha_i \mathbf{n}_i, \mathbf{n}_i) - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1} (\mathbf{x}(\mathbf{p}_i + \alpha_i \mathbf{n}_i, \mathbf{n}_i) - \boldsymbol{\mu}_k) \right) \tag{5.3}$$

A small α_i means there exists one mode that can well localize the boundary at that point. Otherwise, it means either that the current profile has no suitable feature to localize the structure boundary, or that profile mode M_k (*i.e.* mean $\boldsymbol{\mu}_k$ and covariance matrix $\boldsymbol{\Sigma}_k$) is not suitable to represent it. Should α_i be greater than a defined threshold (*e.g.* 2mm), current point is considered as not being well represented by the current classification (*i.e.* considered as an outlier). Using offset α_i as a weak classifier, we thus enforce the classification associated with the MPAM to produce profile modes that are suitable for image segmentation.

5.2.2 External forces based on the MPAM

As previously introduced in section 3.3.5, external forces attract the mesh to image-based features in order to fit a desired region of interest. These forces, which are expressed as a displacement field, compute the displacement of each vertex using information coming from the MPAM. For intensity profiles, a local search is performed (see an illustration in Figure 5.5).

The new vertex position $\mathbf{p}(\mathbf{x}_i)'$, or *target position*, is computed as:

$$\mathbf{p}(\mathbf{x}_i)' = \mathbf{p}(\mathbf{x}_i) + (\alpha_i \ \mathbf{n}_i) \tag{5.4}$$

where $\mathbf{n}_i$ is the normal at vertex $\mathbf{p}(\mathbf{x}_i)$ and α_i is the offset distance from $\mathbf{p}(\mathbf{x}_i)$ and along $\mathbf{n}_i$ for which the profile maximizes its similarity measure with one of the modes M_k profile $\mathbf{x}_i$ belongs to.

In this thesis, we propose regional external forces based on the MPAM (see chapter 4). Though only one offset per vertex is considered, the comparison is done with multiple modes during local search. Indeed, since the MPAM is multimodal, more than one mode may be assigned to $\mathbf{x}_i$. More precisely, every current profile sampled during local search is compared with every mode M_k, hence the multimodality of the MPAM.

In practice, only modes whose posterior probability is higher than a threshold (*e.g.* 10^{-3}) are considered. This may be done because EM classification leads to unfuzzy posterior probabilities whose values are either very high or very low (*i.e.* $\lll 10^{-3}$).

To compare profile $\mathbf{x}_i$ with multiple modes, we use the three similarity measures presented in previous section. For the Mahalanobis distance, only the diagonal term σ_k of covariance matrix $\boldsymbol{\Sigma}_k$ may be considered. This has the advantage of reducing computational time, which may be considerable since the comparison is multimodal.

Finally, the regional external force ${}^{R}f_{ext}(\mathbf{p}(\mathbf{x}_i))$ drives the vertex $\mathbf{p}(\mathbf{x}_i)$ in the direction of the target position $\mathbf{p}(\mathbf{x}_i)'$:

$${}^{R}f_{ext}(\mathbf{p}(\mathbf{x}_i)) = \beta(\mathbf{p}(\mathbf{x}_i)' - \mathbf{p}(\mathbf{x}_i)) \tag{5.5}$$

where β is the global weight controlling the strength of the external forces (see equation 3.2).

To improve segmentation results, a confidence Ψ_i based on the value returned by the intensity similarity measure may be used to locally weight the regional external force ${}^{R}f_{ext}(\mathbf{p}(\mathbf{x}_i))$ at each vertex $\mathbf{p}(\mathbf{x}_i)$ (unlike β, which is a global weight equally applied on all external forces). Indeed, the external force applied to a vertex whose associated modes feature a high similarity measure during the local search is expected to give good segmentation results.

For the Mahalanobis distance, confidence Ψ_i is defined as:

$$\Psi_i = \frac{d}{\sqrt{\{\alpha_i\}_{\text{maha}}}} \tag{5.6}$$

where d is the number of dimensions featured by profile $\mathbf{x}_i$ and $\{\alpha_i\}_{\text{maha}}$ is computed from equation 5.3.

Using confidence Ψ_i, the regional external force ${}^{R}f_{ext}(\mathbf{p}(\mathbf{x}_i))$ from equation 5.5 is defined as:

$${}^{R}f_{ext}(\mathbf{p}(\mathbf{x}_i)) = \beta\Psi_i(\mathbf{p}(\mathbf{x}_i)' - \mathbf{p}(\mathbf{x}_i)) \tag{5.7}$$

In practice, this external force expression needs a normalization step because the Mahalanobis distance computed at each vertex may greatly vary. However, such a normalization may be done only after computing the Mahalanobis distance at each vertex. This requires a second pass of computation at each deformation iteration, which may substantially increase the computation time needed for the segmentation. This is why we decide not to use this expression of the external force in the remainder of this manuscript. Nevertheless, this expression could be an improvement in the future.

5.3 Spectral clustering

5.3.1 Similarity function

First step in spectral clustering consists in choosing the similarity function to determine how close two feature vectors are (*i.e.* how similar two data points are). This function determines the weight of the edge between two vertices from the similarity graph. This is why this function must be chosen carefully, so that local neighborhoods induced by this function are meaningful [Luxburg 2007]. In general, this function depends on the domain the data comes from.

In the common case where feature vectors belong to the Euclidean space $\mathbb{R}^d$, the Gaussian kernel function is to be used. This is also our case because we want data points to be represented by a Gaussian Mixture Model. Other kernel functions may be found in the literature [Balcan *et al.* 2007, Blum 2007, Claussen 1985].

The Gaussian kernel function $S(\mathbf{x}_i, \mathbf{x}_j)$ is defined as:

$$S(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/(2\sigma^2)) \tag{5.8}$$

where σ is the variance controlling the scale of the neighborhoods.

Variance σ plays an important role in the graph construction, and thus on the spectral clustering. As there seems to be no theory on how to determine σ, empirical experiments need to be done to optimize its value. For instance, one solution consists in iteratively trying different values of σ and choosing the value that gives the tightest clusters after clustering, so that distortion is minimized [Ng *et al.* 2001]. Some rules of thumb are usually used in the literature to compute σ [Luxburg 2007]. When using a k-nearest neighbor graph, σ may be chosen in the order of the mean distance of a vertex to its k-nearest neighbor. When using a ε-neighborhood graph, σ may be chosen as $\sigma = \varepsilon$ after determining ε using the minimal spanning tree heuristic [Feremans *et al.* 2004]. Other methods may be found in the literature [Meila & Shi 2000].

With intensity profiles, we decide to compute σ as the intensity standard deviation (SD) of the tissue that features at best the structure of interest, multiplied by the number of dimensions. For instance, we choose parenchyma as the tissue that features at best the liver and estimate the standard deviation of its intensity in CT images at $SD = 60$. For a profile of length 16, this means we have $\sigma_{liver} = 16 \times 60 = 960$. Our experiments showed that this value gives satisfactory results.

In addition, we added a neighboring weight ω_{nei} specific to simplex meshes to spatially regularize the clustering. Similarly to section 4.3.3, the objective is to account for the connectivity between profiles by giving a higher similarity to neighboring profiles.

Should profile $\mathbf{x}_j$ be among the three neighbors of profile $\mathbf{x}_i$, we decrease their pairwise distance by ω_{nei} such that:

$$S(\mathbf{x}_i, \mathbf{x}_j)' = \omega_{nei}\ S(\mathbf{x}_i, \mathbf{x}_j) \tag{5.9}$$

where $\omega_{nei} \in [0, 1]$.

Tests showed that the SVD computation, which is used afterwards for the extraction of top eigenvectors, features convergence issues for $\omega_{nei} < 0.8$. However, $\omega_{nei} = 0.8$ gives good results, leading to an increased local neighborhood (see an illustration in Figure 5.6).

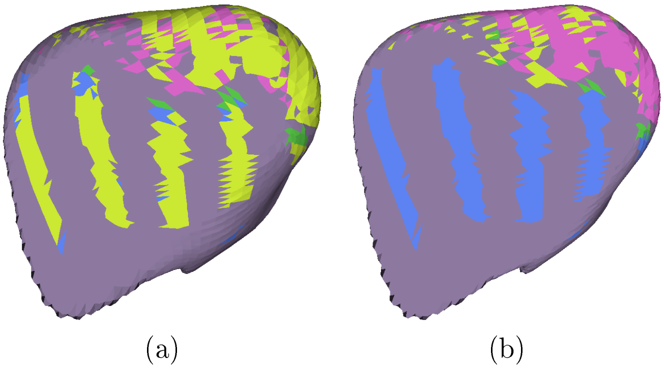


Figure 5.6: Spectral clustering of liver profiles spanning both sides with respect to neighboring weight ω_{nei}: (a) $\omega_{nei} = 1.0$ (no spatial regularization) and (b) $\omega_{nei} = 0.8$. See how the thoracic cage appearance becomes spatially smoother.

5.3.2 Similarity graph

After choosing the appropriate similarity function, next step consists in transforming data points into vertices of a similarity graph.

In the literature, several types of graph have been proposed:

- Fully connected graph: all vertices are connected with each other.
- ε-neighborhood graph: two vertices are connected only if their pairwise distance is smaller than a threshold ε.
- k-nearest neighbor graph: two vertices are connected only if one vertex is among the k-nearest neighbors of the other.

Fully connected graph is to be used only if all points are known to have a similarity with each other. Also, depending on the number of vertices N, this graph may require large computation resources since a Singular Value Decomposition (SVD) needs to be performed on the associated $N \times N$ matrix to extract eigenvectors. This is especially the case with high resolution meshes, where the number of vertices N, and thus of profiles, may be greater than 20 000.

To overcome this potential issue, *ε-neighborhood graph* features a threshold ε that is use to decide whether two vertices are connected or not (*i.e.* vertices are connected only if their are similar).

Finally, *k-nearest neighbor graph* consists in connecting only the k-nearest neighbors of each vertex. This implementation leads to a directed graph though, since neighborhood relationships are not symmetric (*i.e.* two vertices are not necessarily among the first k-neighbors of each other). To make the graph undirected, like for both *fully connected and ε-neighborhood graphs*, a solution consists in allowing the connection of two vertices only if they are among the k-nearest neighbors of each other.

Should either *ε-neighborhood* or *k-nearest neighbor graph* be chosen, the SVD algorithm must be able to deal with sparse matrices, as some entries of the similarity matrix are going to be empty.

We tested the three methods for the clustering of intensity profiles. Depending on mesh resolution, fully connected graph requires lots of memory, especially with high resolution meshes. We thus tested the ε-neighborhood graph, with varying values of ε to overcome this problem. Using a threshold offers a fast and straightforward solution to define a neighborhood compared to a k-nearest neighborhood where a sort needs to be performed among all possible neighbors. This implementation is not stable enough though, because some profiles are dissimilar to many others. Depending on the value of ε, some vertices may have many neighbors while others may only have a few neighbors (*i.e.* corresponding rows of the affinity matrix are almost full or empty). As a consequence, SVD computation is sometimes not able to find a solution.

This is why we use a k-nearest neighbor graph, as suggested in [Luxburg 2007], so that each vertex, and thus each row of the associated affinity matrix, has exactly the same number of neighbors. Empirical tests on liver and bone profiles showed that $k = 14\% \times N$ gives the optimal results. Below this value, SVD computation does not converge (*i.e.* there is not enough information to compute eigenvectors). Conversely, increasing k does not seem to bring more information in SVD computation (*i.e.* eigenvectors are not significantly changing).

When both the similarity function and the similarity graph have been chosen, affinity matrix A must be computed so that eigenvectors may be extracted. Affinity matrix A is a full matrix with fully connected graphs and becomes sparse with ε-neighborhood and k-nearest neighbor graphs. As previously said, caution needs to be taken to avoid convergence issues, depending on how sparse is the matrix.

Affinity matrix A is defined as:

$$A_{ij} = S(\mathbf{x}_i, \mathbf{x}_j) \quad \text{if } i \neq j \tag{5.10}$$

where $A \in \mathbb{R}^{N \times N}$, $S(\mathbf{x}_i, \mathbf{x}_j)$ is the Gaussian kernel function (see equation 5.8), $i \in k$-nearest neighbors of j (and *vice versa*) and $A_{ii} = 0$ (see Algorithm 1, line 2).

5.3.3 Top eigenvectors extraction

To compute top eigenvectors, the Laplacian matrix L needs first to be computed (see Algorithm 1, line 4). Laplacian matrices are the main concept of spectral clustering and lots of research has been conducted to study these matrices [Chung 1997].

However, three types of Laplacian matrix are usually used in most common spectral clustering algorithms:

- Unnormalized Laplacian matrix L [Mohar 1997]:

$$L = D - A \tag{5.11}$$

- Normalized Laplacian matrix L_{rw} [Shi & Malik 2000]:

$$L_{rw} = D^{-1}A \tag{5.12}$$

- Normalized Laplacian matrix L_{sym} [Ng *et al.* 2001]:

$$L_{sym} = D^{\frac{-1}{2}}AD^{\frac{-1}{2}} \tag{5.13}$$

where A is the affinity matrix (see equation 5.10) and D is a diagonal matrix whose (i, i)-element is the sum of A's i-th row (see Algorithm 1, line 3).

The normalized Laplacian matrix from equation 5.12 is denoted L_{rw} because this matrix is closely related to a random walk [Lovász 1993]. The normalized Laplacian matrix from equation 5.13 is denoted L_{sym} as it is a symmetric matrix. To cluster intensity profiles, we decide to use L_{sym} because the associated spectral clustering algorithm [Ng *et al.* 2001] features an additional normalization step (*i.e.* by creating matrix Y) that is suitable for our proposed initialization of the EM algorithm using Gray code (see section 5.3.5).

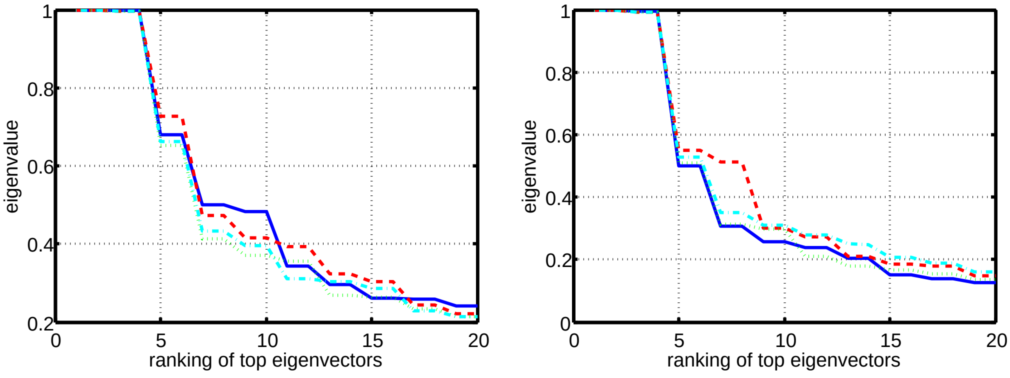
Next step consists in finding the K top eigenvectors of Laplacian matrix L (see Algorithm 1, line 5). As there are as many eigenvectors as rows in L, computation requirement to extract eigenvectors may be considerable when using fully connected graphs, especially for high resolution meshes. This is why ε-neighborhood and k-nearest neighbor are usually preferred, since they lead to sparse matrices that have the advantage to reduce the computation resources. Another point is that computing all eigenvectors is not necessary. Only top eigenvectors are needed for spectral clustering. The most popular methods to extract top eigenvectors are the Krylov methods [Golub & Loan 1996].

5.3.4 Model order selection

Instead of performing a model order selection on original data points, spectral clustering is aimed at using top eigenvectors for this operation. As previously explained, the advantage is that spectral data points feature a much stronger Gaussian distribution [Luxburg 2007, Ng *et al.* 2001]. This feature results in more efficient GMM-based model order selection criteria, and all methods described in section 4.3.2 should be reinforced thanks to this new data representation.

An interesting property that is particularly suited for spectral clustering, and which is known as the *eigengap heuristic*, consists in combining a model order selection with the selected number of top eigenvectors. The objective is to select the number of modes K as the number of eigenvalues $\lambda_1, \ldots, \lambda_k$ so that there is a significant gap δ between λ_k and λ_{k+1}. This heuristic has been justified in both perturbation and spectral graph theories [Mohar 1997, Chung 1997].

Advantage of such an heuristic is its invariance to eigenvalues, since the selection is not based on the eigenvalue *per se*, but rather on the difference between consecutive eigenvalues. This feature is particularly interesting in our case because



(a) (b)

Figure 5.7: Plot of eigenvalues from 20 top eigenvectors, which are extracted from the profiles of 2 livers {(a)-(b)} using the same reference mesh in four resolutions: $N = 3856$ (blue solid line), $N = 7586$ (red dashed line), $N = 11760$ (green dotted line), and $N = 35280$ (cyan dash-dot line). Note how eigenvalues feature the same steep decreases in the four resolutions. More results of eigengap heuristic robustness with respect to resolution may be found in Figure 6.4.

the number of intensity profiles may greatly vary depending on the mesh resolution. Our experiments showed that such a difference may have an impact on eigenvalues. They also showed that a gap $\delta = 0.15$ gives good results in selecting the number of modes for liver profiles (see a plot of eigenvalues in Figure 5.7). Results showing the performance of the eigengap heuristic compared to the OSI criterion may be found in section 6.4. Results showing the eigengap heuristic robustness with respect to the resolution may be found in section 6.5.

5.3.5 Clustering of spectral data

At this stage, we have found the K top eigenvectors of Laplacian matrix L (see Algorithm 1, line 5). These K top eigenvectors have been stacked in columns to form matrix X (see Algorithm 1, line 6). Since we are using the normalized Laplacian matrix L_{sym} (see section 5.3.3), the last transformation consists in normalizing matrix X (see Algorithm 1, line 7).

For that, matrix Y is formed by normalizing each row of matrix X to have unit length [Ng *et al.* 2001]:

$$Y_{ij} = \frac{X_{ij}}{\sqrt{\sum_{j=1}^{N} X_{ij}^2}} \tag{5.14}$$

Matrix Y contains the new data points that are to be clustered. In the literature, this clustering is usually performed using K-Means [Ng *et al.* 2001]. In our case, we

0	0
0	1
1	1
1	0

(a)

0	0	0
0	0	1
0	1	1
0	1	0
1	1	0
1	1	1
1	0	1
1	0	0

(b)

0	0	0	0
0	0	0	1
0	0	1	1
0	0	1	0
0	1	1	0
0	1	1	1
0	1	0	1
0	1	0	0
1	1	0	0
1	1	0	1
1	1	1	1
1	1	1	0
1	0	1	0
1	0	1	1
1	0	0	1
1	0	0	0

(c)

Figure 5.8: Examples of Gray codes when using (a) 2 bits, (b) 3 bits, and (c) 4 bits. Gray code, which is a binary numerical system in which two successive values differ in only one bit, is used to initialize the EM mode centers from spectral data.

chose the EM algorithm as it is the most suited method to represent GMM (see Algorithm 1, line 8).

A specific care needs to be taken when initializing the EM algorithm with spectral data points. Indeed, some eigenvalues may be similar. This is particularly the case with intensity profiles. Our experiments showed that corresponding eigenvectors are likely to lead to similar mode centers trapped in the same local minimum (see an illustration in Figure 5.9).

Since spectral data points are normalized when using Laplacian matrix L_{sym} (*i.e.* $X_{ij} \in \{0, 1\}$), we decide to initialize mode centers using the Gray code [Savage 1997] (see some Gray code examples in Figure 5.8). Gray code is a binary numerical system in which two successive values differ in only one bit.

Doing so, we prevent mode centers to end up in similar minima, since they are initialized orthogonal one from each other (*i.e.* unlike a classical random initialization where two mode centers may be close to each other). Caution needs to be taken though, as we have to ensure there are enough Gray codes for the number of mode centers K (*i.e.* $K < 2^d$, where d is the number of dimensions featured by the mode centers, which corresponds to the number of top eigenvectors in spectral clustering).

After the EM clustering, EM parameters from spectral clustering, which are referred to as $\hat{\pi}_k^p$, $\hat{\boldsymbol{\mu}}_k^p$ and $\hat{\boldsymbol{\Sigma}}_k^p$, may not be directly used with original data points because these parameters are computed in the spectral space (*i.e.* they do not have the same number of dimensions). However, posterior probabilities ${}^p\hat{\gamma}_i^k$ do not

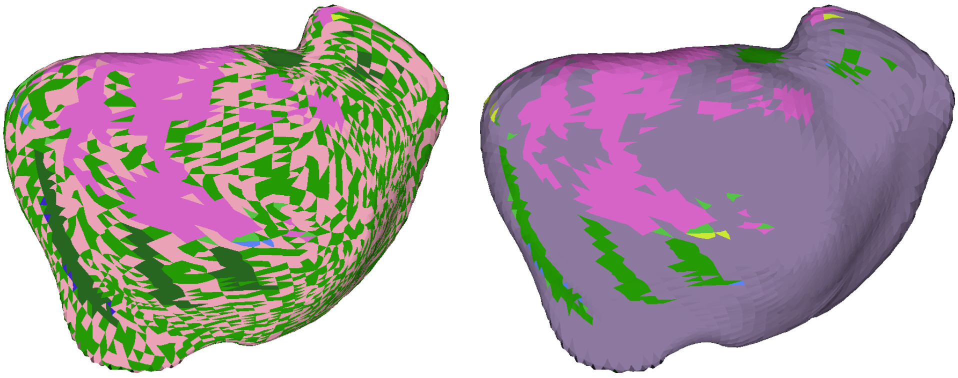

(a) (b)

Figure 5.9: Spectral clustering of liver profiles spanning both sides and whose mode centers ($K = 9$ modes) are initialized: (a) using a random EM initialization, and (b) using the Gray code. Note how the two main modes on the left (in green and light pink), whose centers are similar and trapped in the same local minimum, are separated on the right (main mode is in purple now) thanks to the Gray code initialization.

have this dimensional issue and may thus be used to recompute $\{\pi_k^p, \boldsymbol{\mu}_k^p, \boldsymbol{\Sigma}_k^p\}$ in one M-step on original data points $\mathbf{x}_i$ (see Algorithm 1, line 9).

From equations 4.11, 4.12 and 4.13, mixing coefficients π_k^p, means $\boldsymbol{\mu}_k^p$ and covariance matrices $\boldsymbol{\Sigma}_k^p$ are computed as:

$$\pi_k^p = \frac{1}{N}\sum_{i=1}^{N} {}^p\hat{\gamma}_i^k \tag{5.15}$$

$$\boldsymbol{\mu}_k^p = \frac{\sum_{i=1}^N {}^p\hat{\gamma}_i^k \; \mathbf{x}_i}{\sum_{i=1}^N {}^p\hat{\gamma}_i^k} \tag{5.16}$$

$$\boldsymbol{\Sigma}_k^p = \frac{\sum_{i=1}^N {}^p\hat{\gamma}_i^k (\mathbf{x}_i - \boldsymbol{\mu}_k^p)(\mathbf{x}_i - \boldsymbol{\mu}_k^p)^T}{\sum_{i=1}^N {}^p\hat{\gamma}_i^k} \tag{5.17}$$

Results illustrating the spectral clustering performed on liver profiles may be found in Figure 6.3.

5.4 Boosted clustering

5.4.1 Single pass boosted clustering

As depicted in Figure 5.10 and Algorithm 2, the boosted clustering consists in applying a weak classifier under the constraint that offset α_i is smaller than a threshold

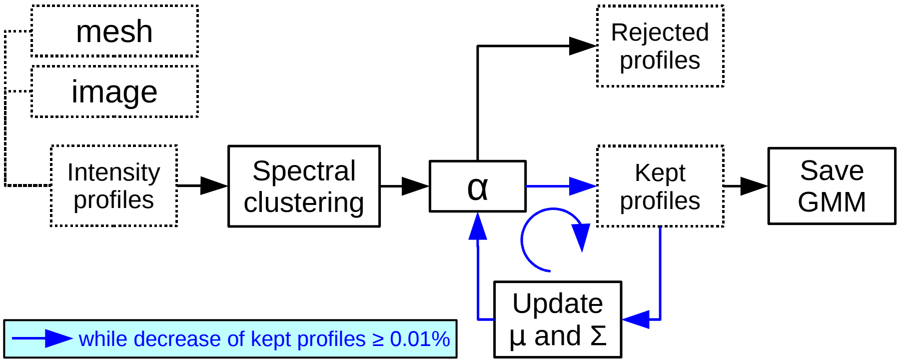


Figure 5.10: Pipeline for single pass boosted clustering.

for all profiles. Profiles that are rejected (*i.e.* whose offset α_i is greater than the threshold) are not considered anymore in the remainder of the algorithm. Profiles that are kept may not be saved as such though, because their associated EM parameters (*i.e.* $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$) are computed from profiles that have been discarded (*i.e.* rejected profiles).

To overcome this issue, EM parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ must be updated so that they are only computed from kept profiles. But then, the weak classifier must be launched once again to check whether kept profiles are still featuring an offset α_i lesser or equal to the threshold with their updated EM parameters. If so, profiles are kept; otherwise, they are rejected. This procedure is iterated until a steady state is reached, *i.e.* when the number of rejected profiles becomes negligible (see Figure 5.10, blue arrows).

However, our experiments showed that the boosted clustering only keeps between 30% and 60% of the total number of profiles, which is not enough for a further segmentation procedure. This is explained by the important number of profiles that are rejected, not only because their associated classification is not discriminant enough with respect to offset α_i at the first iteration, but also because their updated classification does not improve during the iterative procedure. In next section, we propose to iteratively re-estimate rejected profiles as new GMM in an attempt to increase the number of kept profiles.

5.4.2 Cascading boosted clustering

In this new version of the boosted clustering, depicted in Figure 5.11 and Algorithm 3, the profiles that have been rejected are now considered as input to another iteration of the algorithm. At each iteration, rejected profiles are represented by a new GMM (*i.e.* new $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$). By cascading the weak classifier, the number of profiles that are classified while satisfying offset α_i is constantly increasing after each iteration (see Figure 5.11, red arrows).

In theory, the boosted clustering should stop when offset α_i is met by all profiles. In practice, the algorithm stops when a steady state is reached, *i.e.* when the

Algorithm 2 Single pass boosted clustering algorithm.

```
apply spectral clustering on $N$ profiles with the number of modes selected by
  the eigengap heuristic on 20 top eigenvectors.
while decrease of kept profiles $\geq 0.01\%$ do
  for $i = 1$ to $N$ do
    if profile $j$ matches one of its modes at a distance $\leq 2.0$ then
      keep profile $j$.
    else
      reject profile $j$.
    end if
  end for
  compute one M-step to update centers and covariance matrices.
end while
save probabilities, means and covariance matrices as a GMM and consider re-
  jected profiles as outliers.
```

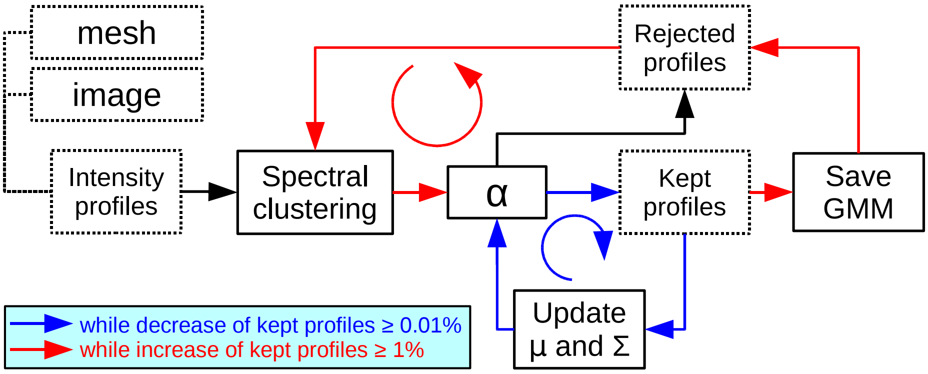


Figure 5.11: Pipeline for cascading boosted clustering.

additional number of kept profiles is below a defined threshold. Each iteration of the weak classifier increases the number of profile modes that are found to be adequate to represent the appearance of the structure in the image.

Our experiments showed that cascading boosted clustering keeps around 60% to 70% of the total number of profiles. Although these results are better than those described in the previous section, they are still unsatisfactory and could be improved. These unsatisfactory results are likely to be due to the spectral clustering performance on a number of profiles that may decrease dramatically after each iteration. Also, there is no guarantee that the eigengap heuristic used for model order selection performs the same way with this decreased number of profiles. As a solution, we propose in the next section to change the model order selection based on the eigengap heuristic by a hierarchical approach.

Algorithm 3 Cascading boosted clustering algorithm.

```
apply spectral clustering on $N$ profiles with the number of modes selected by
the eigengap heuristic on 20 top eigenvectors.
while increase of kept profiles $\geq 1\%$ do
  while decrease of kept profiles $\geq 0.01\%$ do
    for $i = 1$ to $N$ do
      if profile $j$ matches one of its modes at a distance $\leq 2.0$ then
        keep profile $j$.
      else
        reject profile $j$.
      end if
    end for
    compute one M-step to update centers and covariance matrices.
  end while
  save probabilities, means and covariance matrices as a GMM.
  apply spectral clustering on rejected profiles with the number of modes se-
  lected by the eigengap heuristic on 20 top eigenvectors.
end while
consider rejected profiles as outliers.
```

5.4.3 Cascading boosted clustering with hierarchical approach

In Figure 5.12 and Algorithm 4, we depict the cascading boosted clustering with hierarchical approach. Though the initialization is still performed using the eigengap heuristic, the model order selection performed at each iteration of the cascading boosted clustering is now replaced by a hierarchical approach [Bishop 2007]. Instead of systematically saving a new GMM after each iteration (see Figure 5.11, red arrows), the consistency of the GMM is now tested. To do so, the ratio of kept profiles (*i.e.* number of kept profiles with respect to the total number of profiles N) is computed.

If this percentage is higher than a threshold ρ (*e.g.* 3 % of the total number of profiles N), the GMM is saved (*i.e.* considered as consistent) and a new iteration of the cascading boosted clustering is launched (see Figure 5.12, red arrows). A low threshold ρ enables the creation of GMM composed of modes with a few number of profiles, which may be useful to capture the appearance of very small or very specific regions around the structure.

Otherwise, kept profiles are rejected and model order selection based on the hierarchical approach is launched (see Figure 5.12, green arrows). The objective is to classify rejected profiles using spectral clustering, but with a minimum number of modes K_{min} this time. At the next iteration, if kept profiles still do not reach threshold ρ, they are discarded again and re-classified using K_{min+1} modes. This procedure is performed in an iterative way until K_{max} modes are reached.

The number of modes is increased at each iteration so that the GMM receives

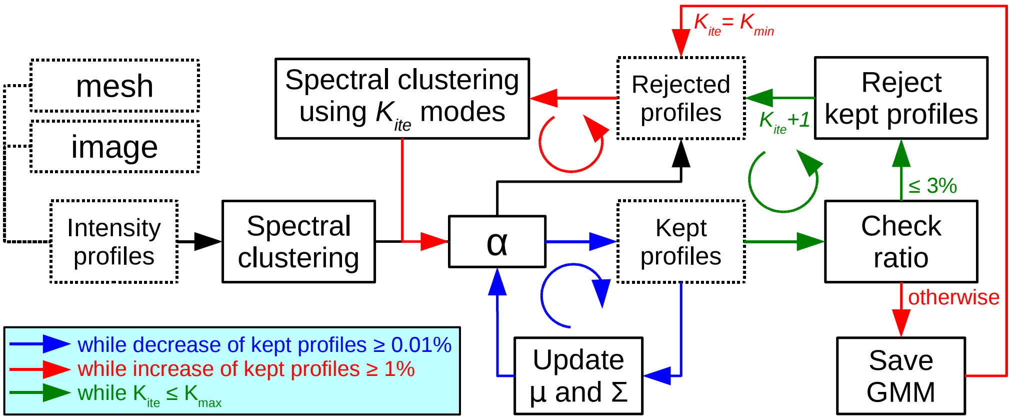


Figure 5.12: Pipeline for cascading boosted clustering with hierarchical approach.

more and more degrees of freedom to represent the data points (*i.e.* kept profiles). This is an important principle of unsupervised clustering: the more clusters, the better data are represented, but at a cost of a more complex model, which usually leads to over-fitting. An appropriate K_{max} must thus be carefully chosen to avoid this undesirable effect.

The cascading boosted clustering with hierarchical approach may end in two different ways. Either K_{max+1} modes have been reached at the end of the model order selection based on the hierarchical approach (*i.e.* meaning that K_{max+1} modes are not enough to ensure the classification of at least ρ% of kept profiles), or a steady state is reached (*i.e.* the additional number of kept profiles is below a defined threshold, similarly to cascaded boosted clustering).

Cascading boosted clustering with hierarchical approach leads in average to 75% to 99% of kept profiles (see an example in Figure 5.13), which we consider as a satisfying number of kept profiles for a further segmentation procedure. More results of cascading boosted clustering may be found in section 6.6.

5.4.4 Discussion

The advantage of using a hierarchical approach is to test the GMM consistency at each iteration. Should one GMM not be consistent (*i.e.* when number of associated kept profiles is below threshold ρ), the classification is optimized by increasing the number of modes K_{ite} (*i.e.* by giving the GMM more degrees of freedom to represent kept profiles). Our experiments showed that a $\rho = 3\%$ gives satisfactory results. Another advantage of this hierarchical approach is to enforce a minimum mode size using threshold ρ, which prevents the creation of tiny modes.

The profiles that are not classified by the cascading boosting clustering correspond to small regions with a specific appearance but with not enough profiles to robustly estimate a mean and covariance matrix. These profiles could also be associated with regions with no salient features to describe the boundaries. At the end

Algorithm 4 Cascading boosted clustering algorithm with hierarchical approach.

```
apply spectral clustering on $N$ profiles with the number of modes selected by
the eigengap heuristic on 20 top eigenvectors.
while increase of kept profiles $\geq 1\%$ AND $K_{ite} \leq K_{max}$ do
  initialize $K_{ite}$ with $K_{min}$.
  break = false.
  repeat
    while decrease of kept profiles $\geq 0.01\%$ do
      for $i = 1$ to $N$ do
        if profile $j$ matches one of its modes at a distance $\leq 2.0$ then
          keep profile $j$.
        else
          reject profile $j$.
        end if
      end for
      compute one M-step to update centers and covariance matrices.
    end while
    if number of kept profiles $\leq 3\%$ then
      reject kept profiles associated with current GMM.
      increase the number of modes $K_{ite}$.
      apply spectral clustering on rejected profiles using $K_{ite}$ modes.
    else
      save probabilities, means and covariance matrices as a GMM.
      re-initialize $K_{ite}$ with $K_{min}$.
      break = true.
    end if
  until $K_{ite} > K_{max}$ OR break = true
end while
consider rejected profiles as outliers.
```

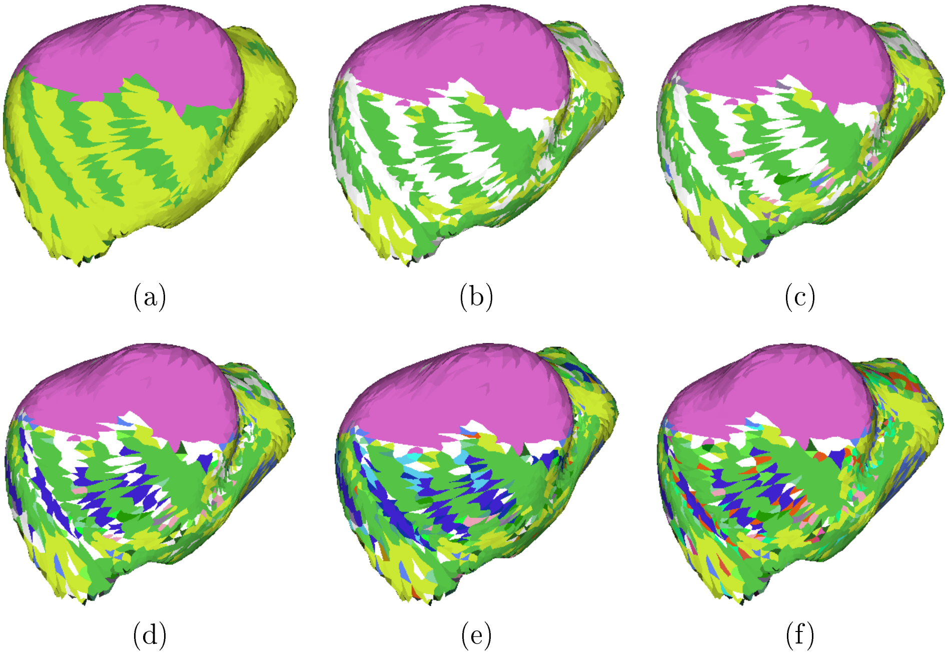


Figure 5.13: Cascaded boosted clustering with hierarchical approach applied on profiles sampled both sides of a liver mesh. The algorithm needed 5 global iterations to converge. First, (a) spectral clustering is applied as initialization step. Note that the thoracic cage is represented by only one mode (dark green stripes). After the first global iteration, (b) only 58% of profiles are considered as well classified by the offset α_i. See how profiles representing the thoracic cage have been discarded (discarded profiles are highlighted by white vertices). Then, the cascading boosted clustering optimizes the classification to reach (c) 67%, (d) 77%, (e) 85%, and (f) 92% of kept profiles. Note that the thoracic cage is now represented by multiple modes.

of the cascading boosted clustering on mesh $\mathcal{M}_p$, we obtain K_p modes M_k^p and a set of posterior probabilities ${}^p\gamma_i^k$ indicating the probability that a vertex $\mathbf{p}_i$ belongs to mode M_k^p.

5.5 Conclusion

In this chapter, we presented regional external forces based on the boosted clustering with the aim of improving our Multimodal Prior Appearance Model (MPAM) for segmentation purposes. In addition to learning appearance through datasets, the boosted clustering tunes the clustering of profiles during an iterative local search, which provides some guaranty about the boundary localization by the profile modes.

We used the spectral clustering to classify intensity profiles because this algorithm is robust when dealing with different resolutions. Also, spectral clustering is less affected by classical issues featured by the EM algorithm, such as local minima, initialization and curse of dimensionality. We explained how to transform original data points into vertices of a similarity graph using a similarity function, which leads to an affinity matrix. In fact, the new representation of data points used by spectral clustering comes from top eigenvectors extracted from this affinity matrix and we explained how to use sparse matrices to deal with high resolution meshes. To automatically determine the number of top eigenvectors, we described the eigengap heuristic. Finally, the classical EM algorithm is used with these top eigenvectors as input data and we proposed to initialize centers with the Gray code to deal with similar eigenvalues. From our experiments, we noticed that spectral clustering is much more faster than original EM algorithm. This is because model order selection is not launched for every possible K, but performed once using the eigengap heuristic. Also, the use of sparse matrices makes the spectral clustering faster.

During an iterative procedure, the spectral clustering and a hierarchical approach are used by the cascading boosted clustering to optimize the classification of profiles. We presented a localization criterion to decide whether the clustering associated to a vertex is able to localize well the boundary along the profile. If not, the profile associated to the vertex is rejected and possibly clustered at a next iteration. At the end of the iterative procedure, if the cascading boosted clustering is still not able to localize well the boundary along corresponding profiles, associated vertices are considered as outliers. The objective is to avoid their influence during the segmentation.

Chapter 6

Results from Liver Database

Contents

6.1 Introduction

In this chapter, we use a database of 35 liver meshes with four resolutions to test the different methods presented in this thesis. First, we present our database and explain how we create reference meshes at different resolutions (section 6.2). Then, we compare the performance of our proposed OSI criterion with classical model order selection criteria (section 6.3) and search for an optimal profile length and inward/outward ratio (section 6.4). The results of spectral clustering applied on intensity profiles are then displayed. We test the robustness of the eigengap heuristic with respect to the resolution (section 6.5) and the different versions of the boosted clustering (section 6.6). We also evaluate different thresholds on the Jaccard index and study how modes are merged when increasing the number of datasets into the MPAM (section 6.7).

In a second part, the segmentation performances of both MPAM and PCAP are assessed (section 6.8). First, we optimize the internal force (section 6.8.1) and the tradeoff between internal and external forces (section 6.8.2). Then, we compare the external forces generated from both MPAM and PCAP (section 6.8.3). Finally, both MPAM and PCAP are compared in a complete segmentation procedure including mesh initialization (section 6.8.4).

6.2 Data

To compute significant statistical results, we use a database of 35 liver meshes, coming from both 3Dircadb1[1] and SLIVER07[2] databases. 3Dircadb1 is a database of liver meshes provided by IRCAD[3], the French Research Institute against Digestive Cancer. SLIVER07 is a database of segmented liver images provided by the Workshop on 3D Segmentation in the Clinic that was held in conjunction with MICCAI 2007[4] conference.

3Dircadb1 contains meshes with different number of vertices (*i.e.* no point correspondence). SLIVER07 contains binary images as segmentation and the Marching Cube algorithm [Lorensen & Cline 1987] has thus been used to generate meshes from these binary images. Since mesh generation from Marching Cube algorithm depends on mesh volume and shape, generated meshes from SLIVER07 database also feature different number of vertices.

To have a meaningful comparison between datasets, a reference mesh $\mathcal{M}^\star$ is created using the method presented in section 4.3.5. More precisely, one liver mesh from 3Dircadb1 is randomly selected as the reference mesh $\mathcal{M}^\star$ and registered to 16 other liver meshes from 3Dircadb1 using the mesh-to-mesh registration based on currents [Glaunes 2005, Vaillant & Glaunes 2005]. Then, an average mesh is computed from the 16 registered meshes (using the average position at each vertex) and constitutes the new version of the reference mesh $\mathcal{M}^\star$. The whole registration procedure is performed twice to reach a steady state (*i.e.* a reference mesh $\mathcal{M}^\star$ that is stable between two iterations).

In addition, reference mesh $\mathcal{M}^\star$ is sampled in four resolutions $\{\mathcal{M}^\star_0, \mathcal{M}^\star_1, \mathcal{M}^\star_2, \mathcal{M}^\star_3\}$ using decimation/refinement operations (see section 3.3.2.4) to study the influence of resolution on results. Reference mesh $\mathcal{M}^\star_1$ features the original resolution with $N = 11760$ vertices. Reference mesh $\mathcal{M}^\star_1$ is refined once to create $\mathcal{M}^\star_0$ with $N = 35280$ vertices and decimated twice to create $\mathcal{M}^\star_2$ and $\mathcal{M}^\star_3$ with $N = 7586$ and $N = 3856$ vertices, respectively (see Figure 6.1).

Finally, the four resolutions of reference mesh $\{\mathcal{M}^\star_0, \mathcal{M}^\star_1, \mathcal{M}^\star_2, \mathcal{M}^\star_3\}$ are registered to the 35 liver meshes of our database using the mesh-to-mesh registration based on currents. These registrations serve as ground truth for our experiments and are denoted as *ground truth liver meshes* in the remainder of this chapter.

6.3 Model order selection

As presented in section 4.3.2, model order selection aims at finding the number of clusters, or *modes*, that best represents the data points using a specific criterion (*i.e.* the objective is to optimize the goodness of fit). Such a task is usually performed by launching the clustering algorithm with an increasing number of clusters (*i.e.*

---

[1] http://www.ircad.fr/softwares/3Dircadb/3Dircadb1

[2] http://sliver07.isi.uu.nl

[3] http://www.ircad.fr

[4] http://www.miccai2007.org

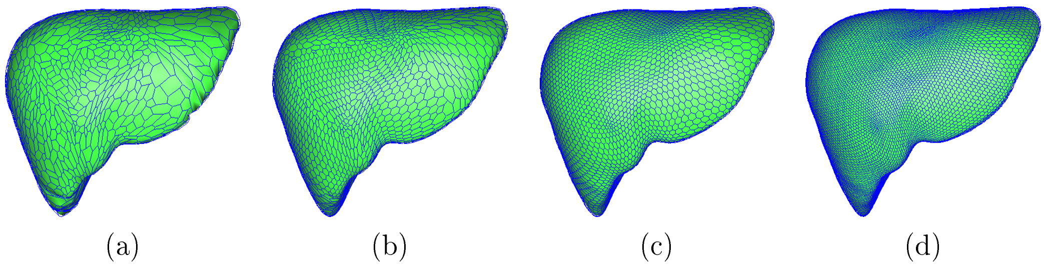


Figure 6.1: Four resolutions of reference mesh $\mathcal{M}^{\star}$: (a) $\mathcal{M}_3^{\star}$ ($N = 3856$ vertices), (b) $\mathcal{M}_2^{\star}$ ($N = 7586$ vertices), (c) $\mathcal{M}_1^{\star}$ ($N = 11760$ vertices), and (d) $\mathcal{M}_0^{\star}$ ($N = 35280$ vertices).

$K = 2, 3, \ldots, k_{max}$) and by searching an optimal tradeoff between under and over-fitting through the obtained models.

We described several criteria based on information theory for model order selection, namely *AIC*, *AICc*, *BIC*, *FVQ* and our proposed *Overlap Separation Index* (*OSI*). Though a certain number of modes is expected to be related to the number of inner and neighboring structures of the liver, a good criterion must also be stable for different datasets. In fact, a good criterion should feature two qualities. First, it should find a reasonable number of modes, avoiding under and over-fitting. Second, it should be robust and stable for different datasets (*i.e.* the standard deviation on the number of modes must be minimized).

To test the quality of these criteria on our liver database, we sample profiles on reference mesh $\mathcal{M}_3^{\star}$ registered to the 35 liver meshes of our database, from [02,04,06,08,10,12] mm inward to [02,04,06,08,10,12] mm outward (*i.e.* 36 profile lengths) with intensities sampled every mm. We launch the EM clustering on these profiles with a number of modes K going from $K_{min} = 2$ to $K_{max} = 20$. For each K, EM classification is initialized with Fuzzy C-Means that are themselves initialized with the same number K of random cluster centers. Also, for each K, the EM classification is launched three times to change the random cluster centers used in the initialization of the Fuzzy C-Means (in an attempt to get better results with the same K). *AIC*, *AICc*, *BIC*, *FVQ* and *OSI* criteria are used to select the best classification (*i.e.* the optimal number of modes).

Similarly to the experiments performed in section 4.3.2.2, we expect the model order selection criteria to find two modes inside the liver mesh (*i.e.* corresponding to the parenchyma and non-parenchyma tissue) and three modes outside liver mesh (*i.e.* corresponding to air, bones and soft tissue). Unlike these experiments, we sample profiles both sides of the liver mesh this time. Thus, we expect model order selection criteria to find six modes (*i.e.* combining inner and outer structures). However, since inner structures (parenchyma and non-parenchyma) feature less discriminative features in terms of intensity than outer structures (air, bones and soft tissue), the inner structures may not be discriminant enough and thus considered as only one mode. In this case, number of modes would be around three (*i.e.* only taking into account the outer structures).

Mean (SD) of the number of modes over the 35 liver meshes and for the different profile lengths may be found in Table 6.1 and Table 6.2, respectively. Results show that AIC is clearly affected by over-fitting, with an average of more than 19 modes for all profile lengths. The worst case being for profiles extending from 2 mm inward to 10 mm outward with the maximum number of clusters ($K_{max} = 20$) found for every dataset (*i.e.* SD is null). $AICc$ and BIC behave the same way with K always greater than 11 ($AICc$) and 14 (BIC), and low SD only from $K > 11$ ($AICc$) and $K > 19$ (BIC). This means that these three classical EM-based criteria are not suited for intensity profiles, probably because intensity profiles do not feature a Gaussian distribution (see section 5.1), as it is assumed with these criteria. On the other hand, FVQ is clearly affected by under-fitting with a majority of profile lengths featuring an average of $K = 2$ ($=K_{min}$). Besides, this trend is confirmed by a SD that appears to be minimum for $K = 2$. Though slightly encouraging under-fitting, OSI seems to be a rather good tradeoff between under and over-fitting, with K always greater than 2 but never greater than 8. Profiles extending 8 mm both sides give the best results with a mean K between 3 and 4 and a SD = 1.2, which is an expected result from previous paragraph.

6.4 Intensity profile

As previously said in section 3.3.5.4, intensity profiles feature both a length and an inward/outward ratio. These parameters depend on the structures whose appearance needs to be represented. If inner structures are studied, inward profiles should be used. Conversely, outward profiles should be used for outer structures. In section 4.3.6, we chose outward profiles composed of 10 samples extracted every mm to study the appearance of structures neighboring both the liver and the tibia. In the literature, profiles extending from 12.5 mm inward to 2.5 mm outward sampled every 0.5 mm were used for the segmentation of lower limb bones [Gilles 2007]. For the segmentation of livers, profiles extending from 9 mm inward to 6 mm outward sampled every mm were successfully used [Heimann 2009].

In our case, intensity profiles are used for the creation of the MPAM (see section 4.3) and the generation of external forces from boosted clustering (see section 5.4). To select an optimal profile length and inward/outward ratio, the objective is to minimize the standard deviation on the number of modes, similarly to previous section. For that, we use the same profiles sampled on reference mesh $\mathcal{M}_3^\star$ (*i.e.* extending from [02,04,06,08,10,12] mm inward to [02,04,06,08,10,12] mm outward with a sampling step of 1 mm). Two criteria are used to select the best profile length and inward/outward ratio, namely the OSI criterion (as already done in previous section) and eigengap heuristic (see section 5.3.4). Results may be seen in Table 6.2 and Table 6.3, respectively.

When using the OSI criterion, profiles extending 8 mm both sides give the best results (SD = 1.2), followed by profiles extending from 4 mm inward to 12 mm

		Outward profile length					
AIC		2	4	6	8	10	12
Inward profile length	2	19.9 (0.1)	19.9 (0.2)	19.9 (0.1)	19.9 (0.1)	20.0 (0.0)	19.9 (0.2)
	4	19.9 (0.2)	19.9 (0.2)	19.9 (0.2)	19.8 (0.3)	19.9 (0.2)	19.9 (0.2)
	6	19.9 (0.2)	19.8 (0.3)	19.9 (0.3)	19.9 (0.2)	19.9 (0.2)	19.9 (0.3)
	8	19.9 (0.3)	19.9 (0.1)	19.8 (0.4)	19.8 (0.3)	19.9 (0.3)	19.9 (0.2)
	10	19.9 (0.1)	19.8 (0.3)	19.8 (0.3)	19.9 (0.2)	19.8 (0.3)	19.9 (0.1)
	12	19.8 (0.3)	19.8 (0.3)	19.8 (0.3)	19.8 (0.3)	19.9 (0.2)	19.8 (0.3)

		Outward profile length					
AICc		2	4	6	8	10	12
Inward profile length	2	19.9 (0.2)	19.9 (0.2)	19.9 (0.2)	19.5 (0.7)	17.2 (1.7)	14.1 (1.5)
	4	19.9 (0.2)	19.8 (0.4)	19.7 (0.5)	18.9 (0.9)	15.4 (1.4)	13.0 (0.9)
	6	19.8 (0.3)	19.8 (0.3)	19.0 (1.1)	16.0 (0.8)	13.4 (1.0)	19.0 (0.3)
	8	19.8 (0.4)	19.1 (0.7)	16.2 (0.9)	13.5 (0.7)	19.0 (0.0)	16.0 (0.2)
	10	19.2 (0.8)	16.2 (1.2)	13.6 (0.7)	19.0 (0.0)	16.0 (0.0)	13.0 (0.1)
	12	16.1 (1.2)	13.5 (0.8)	19.0 (0.0)	16.0 (0.0)	13.0 (0.0)	11.0 (0.3)

		Outward profile length					
BIC		2	4	6	8	10	12
Inward profile length	2	19.4 (1.1)	19.0 (1.2)	18.4 (2.2)	17.4 (2.6)	15.9 (3.0)	14.3 (2.7)
	4	19.8 (0.4)	19.4 (0.8)	19.0 (1.0)	18.2 (1.8)	17.1 (2.4)	16.6 (2.6)
	6	19.6 (0.5)	19.4 (0.8)	18.9 (1.2)	18.4 (1.5)	17.0 (2.1)	16.2 (2.2)
	8	19.3 (0.9)	19.1 (0.9)	18.6 (1.3)	17.8 (1.6)	16.8 (2.2)	15.8 (2.3)
	10	19.3 (0.9)	19.1 (1.0)	18.4 (1.4)	17.5 (1.7)	16.1 (2.3)	14.9 (2.3)
	12	19.0 (1.0)	18.2 (1.5)	17.4 (1.9)	16.5 (2.2)	15.2 (2.0)	14.1 (2.1)

		Outward profile length					
FVQ		2	4	6	8	10	12
Inward profile length	2	5.4 (5.4)	4.2 (3.8)	3.5 (3.1)	2.8 (1.9)	2.5 (1.1)	2.7 (1.8)
	4	2.5 (2.1)	2.6 (1.8)	2.5 (1.5)	2.5 (1.5)	2.4 (1.3)	2.4 (1.5)
	6	2.1 (0.5)	2.0 (0.1)	2.0 (0.1)	2.0 (1.5)	2.0 (0.2)	2.2 (0.8)
	8	2.0 (0.1)	2.0 (0.2)	2.0 (0.2)	2.0 (0.2)	2.0 (0.2)	2.0 (0.1)
	10	2.0 (0.2)	2.0 (0.1)	2.0 (0.1)	2.0 (0.2)	2.0 (0.2)	2.0 (0.2)
	12	2.0 (0.2)	2.0 (0.2)	2.0 (0.3)	2.0 (0.2)	2.0 (0.1)	2.0 (0.2)

Table 6.1: Mean (SD) of the number of modes K over the 35 liver meshes and for different profile lengths (2 to 12 mm) and inward/outward ratios when using AIC, $AICc$, BIC, FVQ criteria and reference mesh $\mathcal{M}_3^\star$.

		Outward profile length					
OSI		2	4	6	8	10	12
Inward profile length	2	3.7 (4.2)	4.5 (3.4)	4.1 (2.3)	4.4 (3.7)	3.6 (1.4)	4.4 (3.1)
	4	3.5 (2.9)	5.3 (4.7)	4.7 (4.1)	3.8 (2.6)	4.0 (1.9)	3.7 (1.3)
	6	2.7 (1.4)	5.3 (4.9)	4.1 (2.9)	3.9 (2.4)	4.8 (3.7)	4.2 (2.1)
	8	4.1 (3.3)	6.7 (6.3)	6.2 (5.7)	3.4 (1.2)	3.6 (1.9)	4.3 (3.1)
	10	4.2 (3.9)	7.9 (6.3)	5.4 (4.7)	4.5 (3.7)	3.7 (1.4)	3.7 (1.4)
	12	3.8 (3.4)	7.2 (6.2)	5.3 (4.7)	4.3 (3.3)	4.4 (3.1)	4.0 (2.5)

Table 6.2: Mean (SD) of the number of modes K over the 35 liver meshes and for different profile lengths (2 to 12 mm) and inward/outward ratios when using OSI criterion and reference mesh $\mathcal{M}_3^\star$.

		Outward profile length					
		2	4	6	8	10	12
Inward profile length	2	5.8 (5.7)	3.7 (1.6)	4.6 (1.5)	5.1 (1.2)	5.3 (1.0)	5.1 (0.8)
	4	5.6 (4.5)	4.4 (2.8)	4.5 (1.4)	5.1 (1.0)	5.3 (1.0)	5.1 (0.9)
	6	8.6 (6.9)	4.4 (3.2)	4.5 (1.5)	5.0 (1.1)	5.2 (1.2)	5.1 (0.9)
	8	8.4 (6.7)	6.5 (4.8)	4.7 (1.6)	5.1 (1.3)	5.1 (1.1)	5.0 (0.9)
	10	8.1 (5.9)	7.8 (6.3)	5.3 (2.1)	5.3 (1.5)	5.1 (1.1)	5.1 (1.0)
	12	9.9 (5.8)	9.4 (6.1)	6.3 (3.6)	5.5 (1.6)	5.3 (1.4)	5.2 (1.1)

Table 6.3: Mean (SD) of the number of modes K over the 35 liver meshes and for different profile lengths (2 to 12 mm) and inward/outward ratios when using eigengap heuristic and reference mesh $\mathcal{M}_3^\star$.

outward (SD = 1.3). Corresponding number of modes varies between 3 and 4. When using eigengap heuristic, profiles extending from 2 mm inward to 12 mm outward give the best results (SD = 0.8), followed by profiles extending from [04,06,08] mm inward to 12 mm outward (SD = 0.9). Corresponding number of modes is around 5. From both OSI criterion and eigengap heuristic, a trend seems to indicate that profiles extending more outward than inward give better results in terms of model order selection robustness, which is probably because there are more discriminative features outward (*i.e.* several neighboring structures) than inward (*i.e.* only liver parenchyma).

When comparing OSI criterion with eigengap heuristic, the latter method seems to be more robust. First, results in terms of SD are better (*e.g.* best result is 0.8 for eigengap heuristic compared to 1.2 for OSI criterion). Also, eigengap heuristic features a smoother behavior with respect to both mean and SD (see 3D surface plots in Figure 6.2 for an illustration). From 3D surface plots of eigengap heuristic results, we notice that, regardless of the inward length chosen, a profile extending

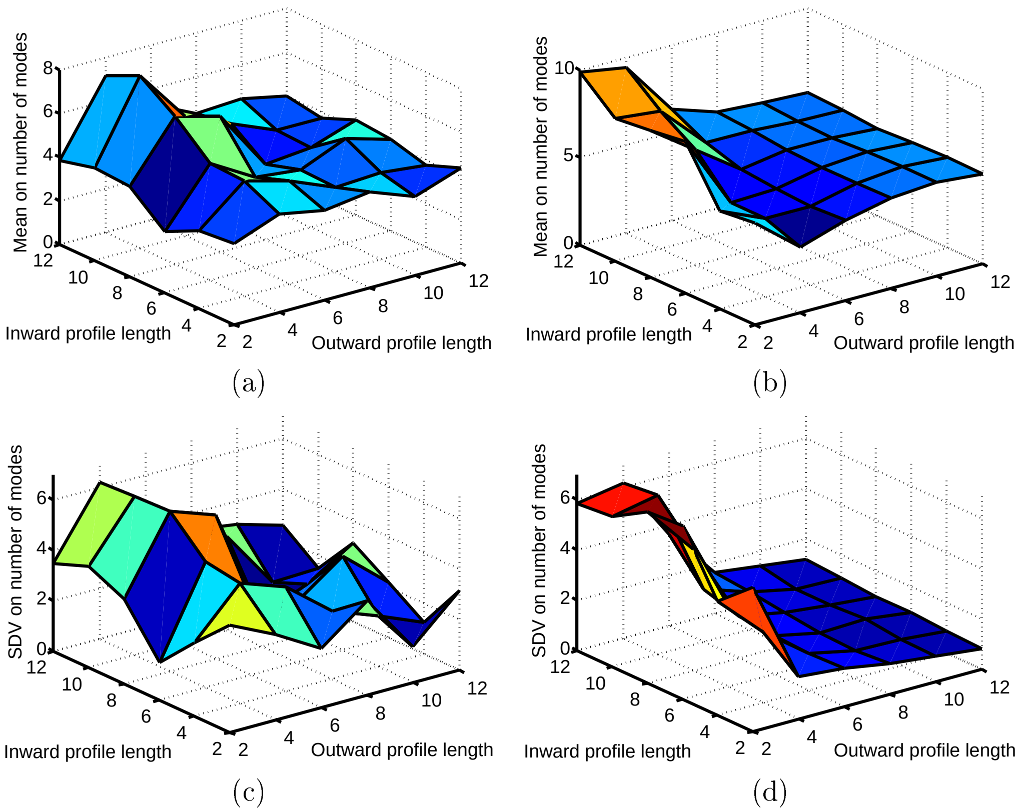


Figure 6.2: 3D surface plot of the mean (top) and standard deviation (bottom) on the number of modes (z-axis) with respect to both inward (y-axis) and outward (x-axis) profile length from reference mesh $\mathcal{M}_3^\star$ using {(a),(c)} *OSI* criterion, and {(b),(d)} eigengap heuristic.

between 8 and 12 mm outward gives stable and smooth results.

6.5 Spectral clustering

In previous section, results showed that profiles extending from 2 mm inward to 12 mm outward feature the best results with spectral clustering, in terms of profile length and inward/outward ratio. Illustrations of these results on the whole database of 35 livers may be found in Figure 6.3.

The number of modes has been found using the model order selection based on the eigengap heuristic. From profiles of reference meshes $\{\mathcal{M}_3^\star, \mathcal{M}_2^\star, \mathcal{M}_1^\star, \mathcal{M}_0^\star\}$, 20 top eigenvectors have been extracted and their eigenvalues have been used by the eigengap heuristic. A plot of these eigenvalues may be found in Figure 6.4. Our experiments show that the four resolutions feature the same steep decreases (*i.e.* at the 5th top eigenvector), which proves that the eigengap heuristic is robust to resolution change.

Though profiles extending from 2 mm inward to 12 mm outward feature the

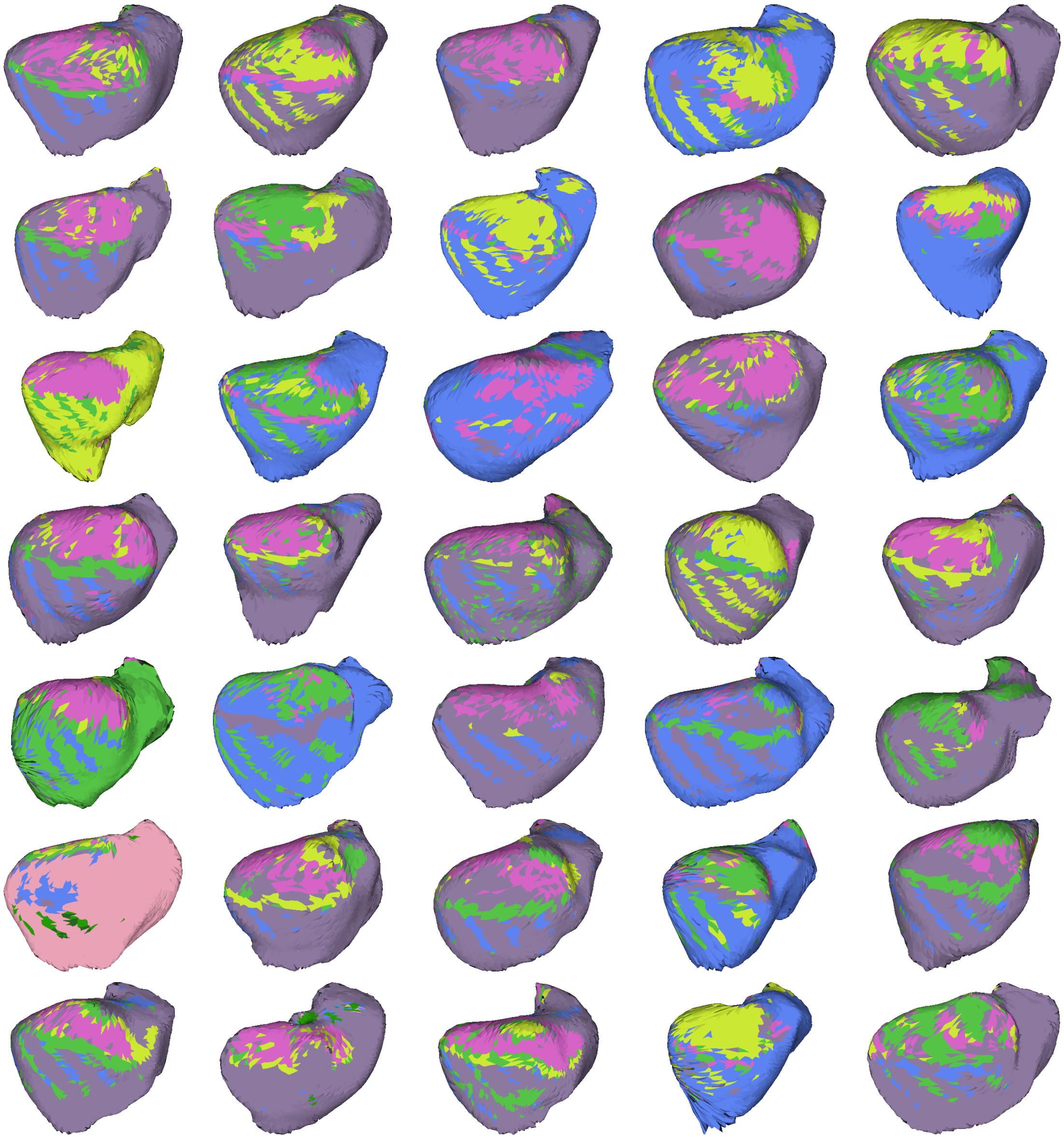
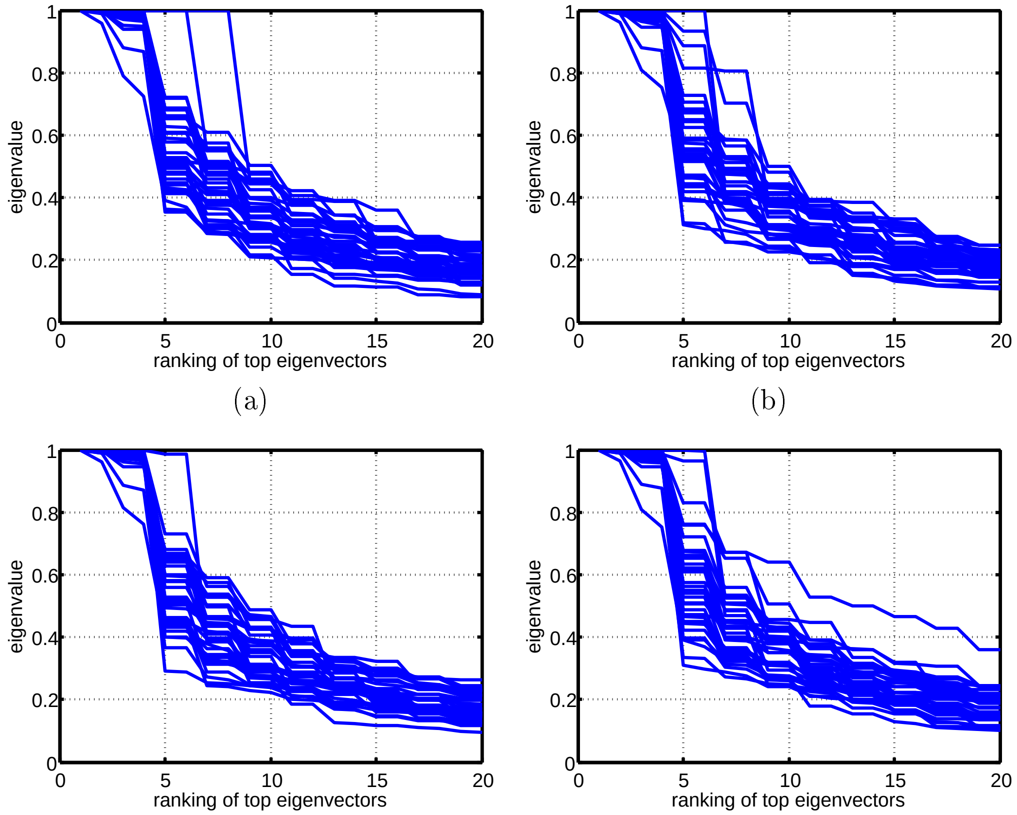


Figure 6.3: Spectral clustering of profiles extending from 2 mm inward to 12 mm outward (*i.e.* optimal profile length and inward/outward ratio from section 6.4). Intensity profiles are sampled from reference mesh $\mathcal{M}_3^\star$ registered to the 35 liver meshes of our database. Note that the color for each mode is randomly selected (*i.e.* the color for each mode does not match from one liver mesh to another).



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

Figure 6.4: Plot of eigenvalues from 20 top eigenvectors, which are extracted from profiles of reference meshes $\{\mathcal{M}^\star_3,\mathcal{M}^\star_2,\mathcal{M}^\star_1,\mathcal{M}^\star_0\}$. The four resolutions of reference mesh $\mathcal{M}^\star$ are registered to the 35 liver meshes of our database: (a) $\mathcal{M}^\star_3$ ($N = 3856$), (b) $\mathcal{M}^\star_2$ ($N = 7586$), (c) $\mathcal{M}^\star_1$ ($N = 11760$), and (d) $\mathcal{M}^\star_0$ ($N = 35280$). Note how eigenvalues feature the same steep decreases for the four resolutions (*i.e.* at the 5th top eigenvector). This feature is used for model order selection, as explained in section 5.3.4.

best results in terms of spectral clustering, this combination may not be optimal during the local search performed by both boosted clustering and external forces. Indeed, using only 2 mm inward comes to consider very few inward information, which may weaken the boundary localization during local search. From Table 6.3 and Figure 6.2, profiles extending between 2 mm and 10 mm inward, and between 6 mm and 12 mm outward, feature quite stable and satisfactory results in terms of mean and standard deviation. This coincides with the profile parameters used for the segmentation of livers in [Heimann 2009], *i.e.* profiles extending from 9 mm inward to 6 mm outward sampled every mm. In order to be robust during local search, and to use profile parameters that have been successfully used for liver segmentation, we decide to use this combination in the remainder of this chapter.

6.6 Boosted clustering

When testing the boosted clustering, we expect two main features from the algorithm. First, a reasonable percentage of kept profiles for every dataset to ensure a good representation of all appearance regions around the structure. Second, a similar behavior through the different resolutions in terms of both percentage of kept profiles and number of modes.

To test these features, we launch the boosted clustering with the four resolutions of reference mesh $\{\mathcal{M}_3^\star, \mathcal{M}_2^\star, \mathcal{M}_1^\star, \mathcal{M}_0^\star\}$. We compute the average percentage of kept profiles per dataset ($\overline{\%}$) and the average number of modes per dataset after boosted clustering ($\overline{K}$), *i.e.* $\overline{K} = (\sum_{p=1}^{P} K_p)/P$. We also test the three versions of the boosted clustering presented in section 5.4, namely the single pass boosted clustering [SingPassBoost], the cascading boosted clustering [CascBoost], and the cascading boosted clustering with a model order selection based on a hierarchical approach [CascBoostHier]. Results are displayed in Table 6.4.

Regarding the average number of modes per dataset $\overline{K}$, [SingPassBoost] features the best results with $\overline{K} = 5$ for all resolutions. [CascBoostHier] is also quite stable with $\overline{K} = 40$ for almost all resolutions. [CascBoost], which does not use a hierarchical approach, features an unstable $\overline{K}$ between resolutions with $11 \leq \overline{K} \leq 45$. Regarding the average of the percentage of kept profiles per dataset $\overline{\%}$, [CascBoostHier] has the best results with $\overline{\%} = 91$ for all resolutions, followed by [CascBoost] and [SingPassBoost] with $\overline{\%} \simeq 55$ and $\overline{\%} \simeq 42$, respectively.

Good results of [SingPassBoost] in terms of $\overline{K}$ stability are explained by the fact that only one iteration of the boosted clustering is performed. In this case, the final number of modes associated with the MPAM is directly coming from the eigengap heuristic, which has proven in section 6.4 to be stable when changing the resolution. However, the percentage of kept profiles per dataset $\overline{\%}$ featured by [SingPassBoost] is the worst. Using the cascading boosted clustering [CascBoost] leads to a slight increase of $\overline{\%}$, but this value remains insufficient. Also, $\overline{K}$ is unstable through the different resolutions. Overall, [CascBoostHier] features the best results, in terms of both stability of $\overline{K}$ and percentage of kept profiles per dataset $\overline{\%}$. This confirms that the cascading boosted clustering gives better results when combined with the hierarchical approach.

In addition, we evaluate the computation time needed to perform the three versions of the boosted clustering over four liver meshes when using the four resolutions of reference mesh $\mathcal{M}^\star$. Results are displayed in Table 6.5. As expected, the higher the resolution, the larger the computation time, as the local search needs to be performed on a larger number of vertices. Also, the computation time dramatically increases from [SingPassBoost] to [CascBoostHier], which is due to the larger number of iterative loops.

	$\mathcal{M}_3^\star$		$\mathcal{M}_2^\star$		$\mathcal{M}_1^\star$		$\mathcal{M}_0^\star$	
	$\overline{\%}$	$\overline{K}$	$\overline{\%}$	$\overline{K}$	$\overline{\%}$	$\overline{K}$	$\overline{\%}$	$\overline{K}$
[SingPassBoost]	41	5	41	5	43	5	43	5
[CascBoost]	55	14	56	45	56	12	53	11
[CascBoostHier]	91	40	91	40	91	40	91	41

Table 6.4: Results of the boosted clustering for the four resolutions of reference mesh $\{\mathcal{M}_3^\star, \mathcal{M}_2^\star, \mathcal{M}_1^\star, \mathcal{M}_0^\star\}$: average of the percentage of kept profiles ($\overline{\%}$) and average number of modes per dataset ($\overline{K}$). The three versions of the boosted clustering are tested: single pass boosted clustering [SingPassBoost], cascading boosted clustering [CascBoost], and cascading boosted clustering with a model order selection based on a hierarchical approach [CascBoostHier].

	computation time [min]			
	$\mathcal{M}_3^\star$	$\mathcal{M}_2^\star$	$\mathcal{M}_1^\star$	$\mathcal{M}_0^\star$
[SingPassBoost]	0.12 (0.03)	0.17 (0.05)	0.35 (0.10)	1.25 (0.50)
[CascBoost]	0.40 (0.16)	1.37 (0.47)	3.50 (0.57)	17.3 (4.72)
[CascBoostHier]	1.50 (0.57)	5.50 (1.73)	13.2 (1.73)	38.5 (2.46)

Table 6.5: Mean (SD) of the computation time needed to perform the boosted clustering over four liver meshes when using the four resolutions of reference mesh $\{\mathcal{M}_3^\star, \mathcal{M}_2^\star, \mathcal{M}_1^\star, \mathcal{M}_0^\star\}$. The three versions of the boosted clustering are tested: single pass boosted clustering [SingPassBoost], cascading boosted clustering [CascBoost], and cascading boosted clustering with a model order selection based on a hierarchical approach [CascBoostHier].

6.7 Fusion of modes

In section 4.3.4, we presented the Jaccard index $\mathcal{J}(M_k^p, M_l^q)$ to measure the similarity between any pair of modes and a threshold in the interval $[0, 1]$ to decide whether two modes M_k^p and M_l^q are equivalent. The objective is to reduce the complexity of the MPAM. As we capture the appearance of the same structure through all the datasets, we expect the final number of modes to asymptotically reach a constant value.

In this section, we test the ability of the MPAM to reach this constant value after the fusion of a minimum number of datasets. To this end, we add datasets to the MPAM and plot the final number of modes for different thresholds on $\mathcal{J}$ after both spectral and boosted clustering. Doing so, we aim at finding what may be the optimal number of datasets necessary to have a meaningful appearance prior (*i.e.* under the hypothesis that reaching this asymptotic value leads to a prior that has fully captured all possible appearance). We also test the Jaccard index with five different thresholds {1.0,0.8,0.7,0.6,0.5}, with $\mathcal{J} = 1.0$ meaning that no modes are merged (*i.e.* all modes from all liver meshes are added to the MPAM). The objective is to consider the optimal threshold as the minimum $\mathcal{J}$ (*i.e.* minimum level of fusion) necessary to have a stable MPAM. Finally, we also test the four resolutions and study their influence on results.

Results after spectral clustering (see Figure 6.6) show that the four resolutions feature the same behavior for every threshold and that a constant value is reached when using both threshold $\mathcal{J} = 0.5$ and threshold $\mathcal{J} = 0.6$. This constant value seems also to be reached with threshold $\mathcal{J} = 0.7$, but with a better emphasis for high resolutions (especially with resolution $\mathcal{M}_0^\star$).

Results after boosted clustering (see Figure 6.7) also show that the four resolutions feature a similar behavior for every threshold $\mathcal{J}$, but the constant value seems to be only reached when using threshold $\mathcal{J} = 0.5$. Also, there is almost no difference between $\mathcal{J} = 1.0$ and $\mathcal{J} = 0.8$ this time, and a very small difference with $\mathcal{J} = 0.7$, indicating that modes are rather dissimilar. This comes from the boosted clustering that encourages the creation of numerous modes with lower variance (*i.e.* rather than few modes with large variance) and thus, encourages their dissimilarity.

With $\mathcal{J} = 0.5$, a rather constant value seems to be reached when adding between 5 and 15 datasets into the MPAM. This constant value does not exactly feature an asymptotic behavior though, as we notice the presence of bumps. This means that more than one mode previously added to the MPAM (for instance blue and red modes in Figure 6.5, right) was merged with one mode from a new dataset (for instance green mode in Figure 6.5, right), thus reducing the final number of modes. This is possible because $\mathcal{J} = 0.5$ easily allows the fusion of modes, *i.e.* only 50% of similarity between modes is required for merging them (see an illustration in Figure 6.5).

In both cases (*i.e.* after spectral and boosted clusterings), the final number of modes is slightly higher for low resolutions than high resolutions, regardless of the threshold on Jaccard index used. Indeed, after spectral clustering and without mode fusion, *i.e.* $\mathcal{J} = 1.0$ (see Figure 6.6), the final number of modes is $K = 183$ ($\mathcal{M}_3^\star$ with $N = 3856$ vertices), $K = 168$ ($\mathcal{M}_2^\star$ with $N = 7586$ vertices), $K = 165$ ($\mathcal{M}_1^\star$ with $N = 11760$ vertices) and $K = 159$ ($\mathcal{M}_0^\star$ with $N = 35280$ vertices). After boosted clustering and also without mode fusion (see Figure 6.7), the final number of modes is $K = 1291$ ($\mathcal{M}_3^\star$), $K = 1167$ ($\mathcal{M}_2^\star$), $K = 1151$ ($\mathcal{M}_1^\star$) and $K = 1087$ ($\mathcal{M}_0^\star$). This coincides with clustering theory, which states that a higher number of data points (*i.e.* profiles from high resolution mesh) is better to represent clusters. This feature helps the fusion of modes when using a high resolution mesh and leads thus to a smaller final number of modes.

As a conclusion, a threshold $\mathcal{J} = 0.5$ seems to be a reasonable choice after boosted clustering, since results showed that, with this threshold, a rather constant value is reached when adding between 5 and 15 datasets into the MPAM. However, this threshold is likely to create merged modes with large variance that may not be discriminant enough for segmentation. The different thresholds are thus tested during segmentation in the next section.

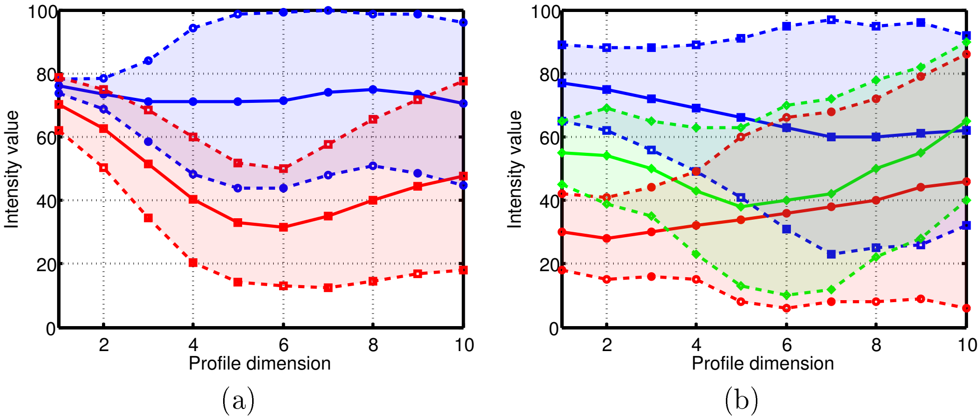


Figure 6.5: Fusion of multiple modes using Jaccard index: (a) two modes (blue and red), and (b) three modes (blue, green and red). Even for a non-restrictive threshold on Jaccard index (*e.g.* $\mathcal{J} = 0.5$), the two modes in (a) are not merged because their intersection is quite small. However, the three modes in (b) are merged because of the large intersection between them, which is due to the green mode that overlaps both blue and red modes. In this case, the three modes are merged into one mode.

6.8 Segmentation

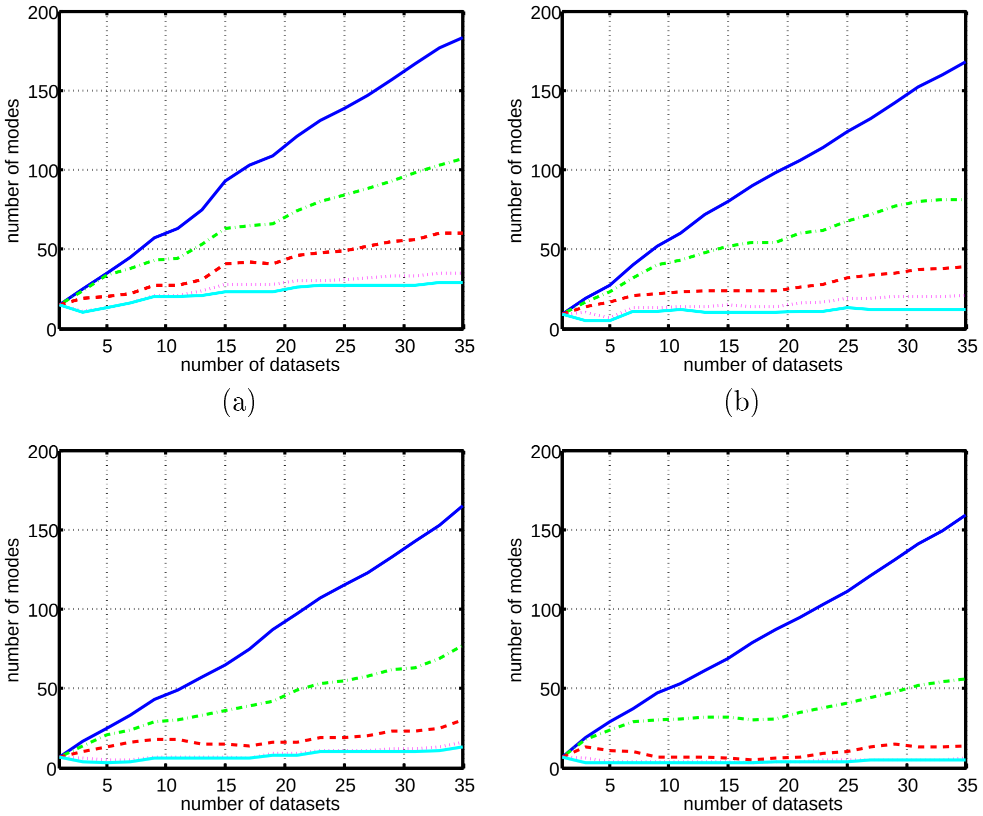
When using explicit deformable models for segmentation (see section 3.3), we have to deal with both internal and external forces. Since this thesis is based on the appearance around structures, we are mainly interested on external forces, which are the main focus of this section.

First, we study the sole influence of the internal forces on mesh deformation. Then, we choose a reasonable tradeoff between internal and external forces. From this tradeoff, we test our regional external forces using MPAM and compare them with external forces from PCA-based methods. Finally, we add an initialization stage to study how our MPAM behaves in a complete segmentation procedure.

6.8.1 Internal forces

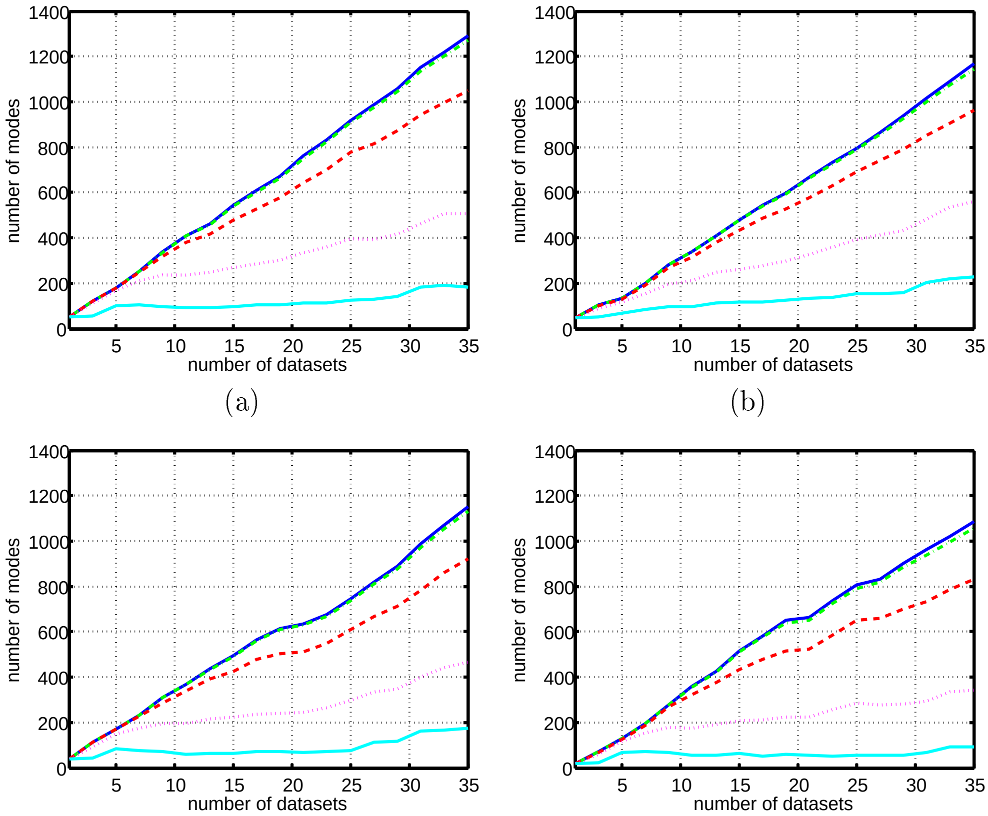
As described in section 3.3.4, internal forces ensure mesh surface to be smooth and less sensible to noise present in the image during mesh deformation. Thus, they have a regularizing effect because they force the mesh to deform while respecting regularity criteria. For our tests, we decide to use internal forces associated with simplex meshes coupled with rigidity parameter η (see section 3.3.4.2).

As a first step before segmentation, we study the influence of the only internal force ($\alpha = 1$, $\beta = 0$) on the deformation of reference mesh $\mathcal{M}_2^\star$ ($N = 7586$) registered to four liver meshes, *i.e.* four ground truth liver meshes at $\mathcal{M}_2^\star$ resolution (there is no initialization stage here because we are only studying the internal forces). Results may be seen in Figure 6.8. Deformations are applied during 100 iterations with rigidity parameter $\eta = 1$.



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

Figure 6.6: Plot of the final number of modes when adding datasets into the MPAM after spectral clustering. The fusion of modes described in section 4.3.4 is used with four resolutions: (a) $\mathcal{M}_3^\star$ ($N = 3856$), (b) $\mathcal{M}_2^\star$ ($N = 7586$), (c) $\mathcal{M}_1^\star$ ($N = 11760$), and (d) $\mathcal{M}_0^\star$ ($N = 35280$). For each resolution, five thresholds on Jaccard index $\mathcal{J}(M_k^p, M_m^l)$ are used to decide whether two modes M_k^p and M_m^l are equivalent: $\mathcal{J} = 1.0$ (blue solid line), $\mathcal{J} = 0.8$ (green dash-dot line), $\mathcal{J} = 0.7$ (red dashed line), $\mathcal{J} = 0.6$ (magenta dotted line), and $\mathcal{J} = 0.5$ (cyan solid line). Note that the four resolutions feature the same behavior for every threshold $\mathcal{J}$ and that a constant value is reached when using a threshold $\mathcal{J} = 0.5$.



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

Figure 6.7: Plot of the final number of modes when adding datasets into the MPAM after boosted clustering. The fusion of modes described in section 4.3.4 is used with four resolutions: (a) $\mathcal{M}_3^\star$ ($N = 3856$), (b) $\mathcal{M}_2^\star$ ($N = 7586$), (c) $\mathcal{M}_1^\star$ ($N = 11760$), and (d) $\mathcal{M}_0^\star$ ($N = 35280$). For each resolution, five thresholds on Jaccard index $\mathcal{J}(M_k^p, M_m^l)$ are used to decide whether two modes M_k^p and M_m^l are equivalent: $\mathcal{J} = 1.0$ (blue solid line), $\mathcal{J} = 0.8$ (green dash-dot line), $\mathcal{J} = 0.7$ (red dashed line), $\mathcal{J} = 0.6$ (magenta dotted line), and $\mathcal{J} = 0.5$ (cyan solid line). Note that the four resolutions feature the same behavior for every threshold $\mathcal{J}$ and that a rather constant value seems to be reached when using a threshold $\mathcal{J} = 0.5$.

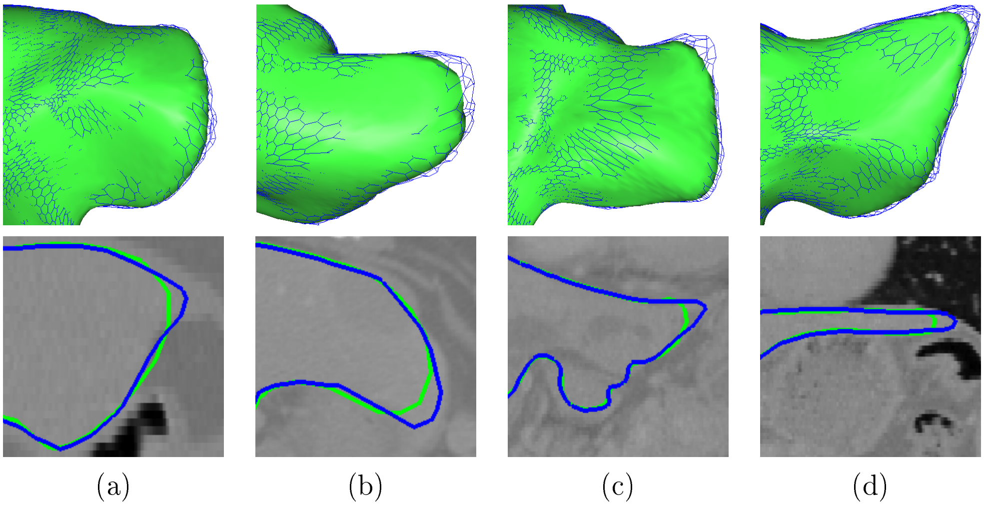


Figure 6.8: Influence of the only internal force ($\alpha = 1$, $\beta = 0$) on the deformation of four liver meshes {(a),(b),(c),(d)} with reference mesh $\mathcal{M}_2^\star$ ($N = 7586$): resulting meshes (top row) and corresponding delineations (bottom row) with original mesh in blue and deformed mesh in green. Deformations are applied during 100 iterations with a rigidity parameter $\eta = 1$. Note how high curvature zones have been smoothed by internal forces.

When comparing deformed meshes with original meshes, we notice that most of the surface is preserved except in high curvature zones where the mesh is subject to a shrinking effect. This is because the internal forces associated with simplex meshes tend to deform the mesh towards a constant curvature shape (*i.e.* a sphere for 2-simplex meshes). Since high curvature zones do not feature a constant curvature, they are the first regions to undergo this smoothing effect. To reduce the global impact of the internal force in high curvature zones, rigidity parameter η must be locally increased in these zones.

6.8.2 Tradeoff on forces

After studying the effect of the only internal force on mesh deformation, next step consists in choosing a reasonable tradeoff between internal and external forces (*i.e.* finding an optimal couple $\{\alpha,\beta\}$ with rigidity parameter η). Too large internal forces lead to a mesh whose excessive shape constraints prevent it to be attracted to image-based features. Conversely, too large external forces lead to a mesh that is excessively deformed at salient image-based features.

For this study, we use external forces generated from [CascBoostHier]. The cascading boosted clustering is applied on profiles of $\mathcal{M}_2^\star$ registered to all liver meshes of our database. The K_p modes generated from every dataset are added to the MPAM without any fusion (*i.e.* $\mathcal{J} = 1.0$), to avoid any bias on results (*i.e.* here we are just interested in defining a tradeoff between α and β, not in studying the influence of $\mathcal{J}$ on the results). We segment four liver images using their associated ground truth liver mesh at $\mathcal{M}_2^\star$ resolution, and we include their classification into the MPAM. The objective is also to test the robustness of the MPAM, similarly to what we have done with the localization criterion from the boosted clustering (see section 5.4) but with the MPAM this time. The classification of every dataset is included into the MPAM, even the classification of the four segmented datasets, so a good segmentation result is expected. This way, the only unknowns are parameters α and β. Deformations are applied during 100 iterations. Results on four liver meshes may be seen in Figure 6.9.

Our first test consists in using equal weights, *i.e.* a segmentation with $\{\alpha = 0.1, \beta = 0.1, \eta = 1\}$ (see Figure 6.9, first row). We notice that resulting meshes feature a quite irregular shape. Some of them are even irregular because of holes on their surface. To overcome that, we first increase the mesh rigidity using $\eta = 5$ (see Figure 6.9, second row). Though still irregular, the surfaces are now smoother and regular. To get smoother deformed meshes, we then increase the weight on internal forces $\{\alpha = 0.3\}$ and finally rigidity parameter $\{\eta = 10\}$ (see Figure 6.9, third and fourth row, respectively), which gives reasonable results in terms of mesh regularity after 100 deformation iterations. This is why we decide to use these parameters (*i.e.* $\{\alpha = 0.3, \beta = 0.1, \eta = 10\}$) for the segmentations performed in the remainder of this section.

6.8.3 External forces

Segmentation quality indices Before testing the external forces, the segmentation quality must be assessed. This can be done by means of *segmentation quality indices*. In our case, we decide to assess the segmentation results with Volumetric Overlap Error (VOE) [%], Signed Relative Volume Difference (SRVD) [%], Average Symmetric Surface Distance (ASSD) [mm] and Maximum Surface Distance (MSD) [mm], which were used to compare segmentation results during the MICCAI 2007 Grand Challenge workshop [Heimann *et al.* 2009].

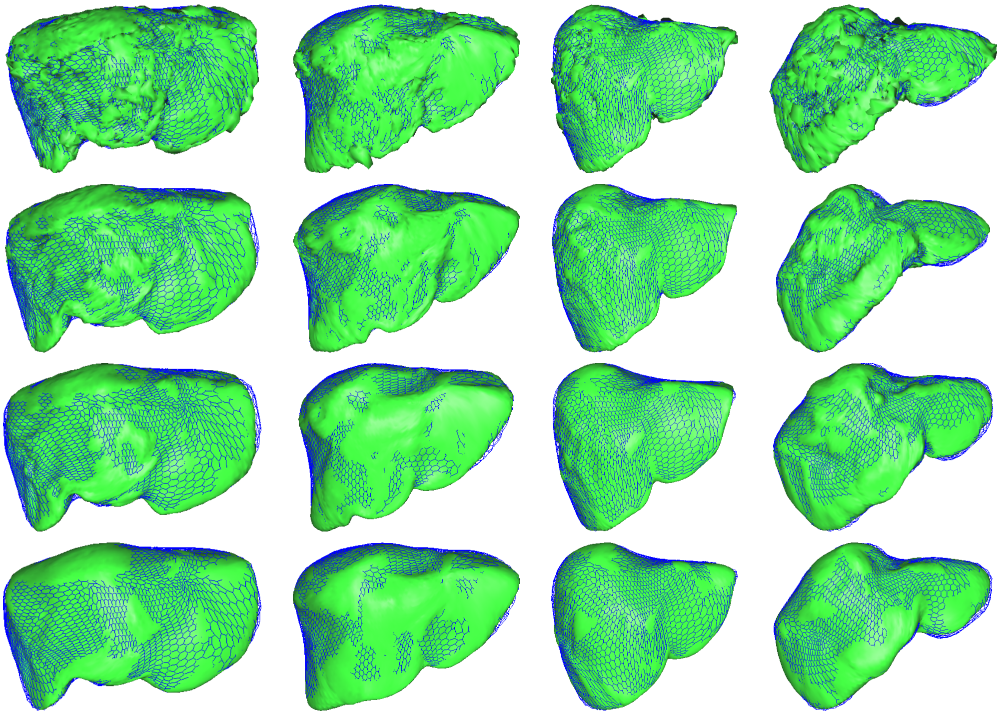
Let A and B be two sets of voxels from a binary image representing segmentation result mesh and ground truth mesh, respectively, *i.e.* for both, the binary image is composed of foreground voxels (inner part of the mesh) and background voxels (rest of the image).

Let $S(A)$ be the set of surface voxels from A, and $d(\mathcal{V}, S(A))$ be the shortest distance of an arbitrary voxel $\mathcal{V}$ to $S(A)$ defined as:

$$d(\mathcal{V}, S(A)) = \min_{s_A \in S(A)} \parallel \mathcal{V} - s_A \parallel \tag{6.1}$$

VOE is one of the most popular methods to evaluate segmentation accuracy:

$$\mathrm{VOE} = 100 \left[1 - \left(\frac{|A \cap B|}{|A \cup B|} \right) \right] \tag{6.2}$$



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

Figure 6.9: Influence of internal forces coupled with external forces on the deformation of four liver meshes $\{(a),(b),(c),(d)\}$ with reference mesh $\mathcal{M}_2^\star$ ($N = 7586$): $\{\alpha = 0.1,\ \beta = 0.1,\ \eta = 1\}$ (first row), $\{\alpha = 0.1,\ \beta = 0.1,\ \eta = 5\}$ (second row), $\{\alpha = 0.3,\ \beta = 0.1,\ \eta = 5\}$ (third row), and $\{\alpha = 0.3,\ \beta = 0.1,\ \eta = 10\}$ (fourth row). Original mesh is in blue wireframe and deformed mesh in green solid surface. Deformations are applied during 100 iterations.

Given in percentage, VOE is 0 for a perfect segmentation and 100 if segmentation result and ground truth do not overlap at all.

SRVD is also given in percentage and is 0 if segmentation result and ground truth are identical:

$$\text{SRVD} = 100\left(\frac{|A|-|B|}{|B|}\right) \tag{6.3}$$

However, this does not imply that they are identical, or that they overlap with each other. For this reason, SRVD should never be used as the only measure of segmentation quality. In combination with other measures though, it reveals if a method tends to over or under-segment. For this purpose, results of SRVD are given as signed numbers.

Given in millimeters, ASSD is based on the surface voxels of both the segmentation result and ground truth, and is 0 for a perfect segmentation:

$$\text{ASSD} = \frac{1}{|S(A)|+|S(B)|}\left(\sum_{s_A\in S(A)} d(s_A,S(B)) + \sum_{s_B\in S(B)} d(s_B,S(A))\right) \tag{6.4}$$

MSSD, which is also known as *Hausdorff distance* [Huttenlocher *et al.* 1993], is given in millimeters and is 0 for a perfect segmentation:

$$\text{MSD} = \max\{\max_{s_A\in S(A)} d(s_A,S(B)), \max_{s_B\in S(B)} d(s_B,S(A))\} \tag{6.5}$$

Multimodal Prior Appearance Model Since we have found a reasonable trade-off between internal and external forces (*i.e.* $\{\alpha=0.3,\ \beta=0.1,\ \eta=10\}$), the external forces of our MPAM may now be tested. For that, we use the cascading boosted clustering [CascBoostHier] applied on profiles of $\mathcal{M}_2^\star$ registered to 31 liver meshes of our database and we segment the 4 remaining liver images using their associated ground truth liver meshes at $\mathcal{M}_2^\star$ resolution (there is still no initialization stage here because we are focusing on the external forces).

This time, final modes K_p from every dataset are possibly merged into the MPAM with different thresholds on the Jaccard index $\mathcal{J}$. The objective is to study the effect of Jaccard index on segmentation results. Also, we do not include the classification of segmented datasets into the MPAM this time, so that external forces are fully studied without bias (*i.e.* without using appearance information from the datasets that are segmented). Final number of modes K for the MPAM when using different thresholds on the Jaccard index $\mathcal{J}$ may be found in Table 6.6.

Table 6.7, Table 6.8 and Table 6.9 depict these segmentation results when using, respectively, Mahalanobis distance, sum of absolute differences and linear criterion

	K			
$\mathcal{J}$	L01	L02	L03	L04
1.0	1240	1254	1255	1264
0.7	1013	1027	1017	1037
0.6	503	503	501	511
0.5	184	179	182	187
0.4	60	58	52	63
0.3	37	36	38	38
0.2	30	29	31	31
0.1	28	27	29	29

Table 6.6: Final number of modes K for the MPAM used for the segmentation of four liver images {L01,L02,L03,L04} when using different thresholds on the Jaccard index $\mathcal{J}$.

MPAM with Mahalanobis distance								
	ASSD [mm]				VOE [%]			
$\mathcal{J}$	L01	L02	L03	L04	L01	L02	L03	L04
1.0	1.92	1.24	1.16	1.02	11.61	7.66	7.51	6.83
0.7	5.46	1.91	1.28	3.93	29.56	11.28	8.30	21.75
0.6	10.59	4.95	2.43	9.29	52.22	23.43	12.98	45.34
0.5	9.37	4.11	2.35	7.89	46.69	19.67	12.92	39.01
0.4	8.26	3.74	2.28	7.29	42.17	17.90	12.65	36.11
0.3	8.25	3.67	2.28	7.05	41.59	17.61	12.61	35.12
0.2	8.01	3.77	2.29	7.11	40.89	17.98	12.64	35.20
0.1	7.86	3.77	2.29	7.29	40.16	18.00	12.62	36.08

Table 6.7: Results of the segmentation performed on four liver images {L01,L02,L03,L04} when using our MPAM and Mahalanobis distance. MPAM is based on [CascBoostHier] with different thresholds on the Jaccard index $\mathcal{J}$. Results are assessed with both ASSD and VOE. Corresponding number of modes K are in Table 6.6. Some corresponding delineations may be found in Figure 6.10.

during local search. Segmentation delineations when using Mahalanobis distance may be found in Figure 6.10.

Best results are obtained when using the Mahalanobis distance (see Table 6.7). These results are clearly not linear since they get first worst when decreasing threshold on the Jaccard index from $\mathcal{J} = 1.0$ to $\mathcal{J} = 0.6$, then they get slightly better until reaching a rather constant value at $\mathcal{J} = 0.1$. Also, the best segmentation accuracy is usually obtained when using MPAM with threshold on Jaccard index $\mathcal{J} = 1.0$. This is the case with Mahalanobis distance and sum of absolute differences. For linear criterion, best results are obtained for threshold on Jaccard index $\mathcal{J} = 0.7$. In general, segmentation accuracy decreases with $\mathcal{J}$, which is due to the increased mode variance from mode fusion that leads to less discriminant modes during local search. Overall, results show that the best similarity measure is the Mahalanobis distance combined with threshold on Jaccard index $\mathcal{J} = 1.0$.

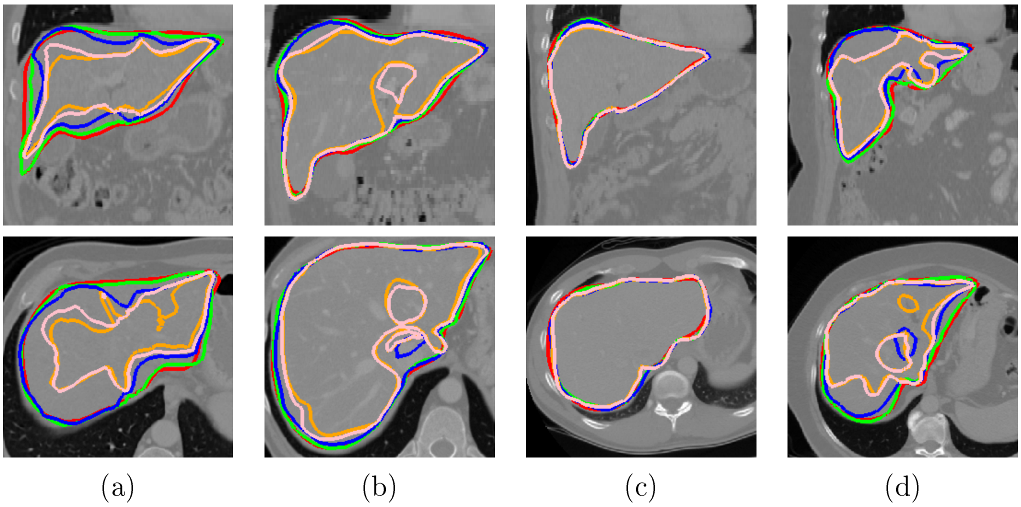

(a) (b) (c) (d)

Figure 6.10: Segmentation of four liver meshes {(a),(b),(c),(d)} with reference mesh $\mathcal{M}_2^\star$ ($N = 7586$) and the Mahalanobis distance. External forces are generated from our MPAM based on [CascBoostHier] and from different thresholds on the Jaccard index: $\mathcal{J} = 1.0$ (green), $\mathcal{J} = 0.7$ (blue), $\mathcal{J} = 0.6$ (orange), and $\mathcal{J} = 0.5$ (pink). We do not include $\mathcal{J} = 0.8$ because plots in Figure 6.7 do not reveal a significant difference between $\mathcal{J} = 1.0$ and $\mathcal{J} = 0.7$. Original mesh is depicted in red. Deformations are applied during 100 iterations. Corresponding statistics may be found in Table 6.7.

MPAM with sum of absolute differences								
	ASSD [mm]				VOE [%]			
$\mathcal{J}$	L01	L02	L03	L04	L01	L02	L03	L04
1.0	2.63	1.86	1.49	1.59	15.59	11.23	9.32	10.04
0.7	2.52	2.02	1.73	1.75	14.43	11.40	10.11	10.75
0.6	2.61	2.87	2.28	2.11	15.04	14.53	13.39	12.45
0.5	2.44	2.93	2.78	2.45	14.29	15.15	16.06	13.97
0.4	2.46	2.96	2.86	2.45	14.40	15.43	16.77	14.01
0.3	2.39	2.92	2.85	2.39	13.96	15.52	16.75	13.94
0.2	2.41	2.95	2.94	2.43	14.04	15.38	16.98	14.13
0.1	2.41	2.95	2.94	2.40	14.08	15.38	17.01	13.98

Table 6.8: Results of the segmentation performed on four liver images {L01,L02,L03,L04} when using our MPAM and sum of absolute differences. MPAM is based on [CascBoostHier] with different thresholds on the Jaccard index $\mathcal{J}$. Results are assessed with both ASSD and VOE. Corresponding number of modes K are in Table 6.6.

PCA-based Appearance Prior At this stage, our MPAM may also be compared with PCA-based Appearance Priors (PCAP) described in section 3.3.5.5. To have a fair comparison between both methods, we use the same framework, *i.e.* PCA applied on profiles of $\mathcal{M}_2^\star$ registered to 31 liver meshes of our database, external forces generated from the PCAP and the segmentation of the four remaining liver images using their associated ground truth liver meshes at $\mathcal{M}_2^\star$ resolution (there is still no initialization stage here because we are focusing on the external forces, the objective is to compare the impact of external forces on both MPAM and PCAP). We also use the same internal forces, the same tradeoff between internal and external forces (*i.e.* $\{\alpha = 0.3,\ \beta = 0.1,\ \eta = 10\}$), the same number of iterations (*i.e.* 100) and the same ratios of complexity reduction, *i.e.* ratio on eigenvalues $\mathcal{R}$ from 1.0 to 0.1 with $\mathcal{R} = 0.98$ and $\mathcal{R} = 95$, which are two typical values chosen in PCA-based methods.

Figure 6.11 and Table 6.10 depict segmentation results when using Mahalanobis distance. Table 6.11 depicts segmentation results when using sum of absolute differences and linear criterion, which both use only mean profile at each vertex (*i.e.* covariance matrix is not used).

Results show that both sum of absolute differences ($\overline{ASSD} = 2.82$) and linear criterion ($\overline{ASSD} = 3.14$) feature the best results compared to the Mahalanobis distance ($\overline{ASSD} = 3.41$ for both $\mathcal{R} = 0.95$ and $\mathcal{R} = 0.90$). If we take the standard deviation into account, $\mathcal{R} = 0.90$ (SD = 0.4) is more stable than $\mathcal{R} = 0.95$ (SD = 0.9). When decreasing values of $\mathcal{R}$ from $\mathcal{R} = 1.0$, results get better around $\mathcal{R} = 0.90$ for the Mahalanobis distance, which is a typical value at which low modes of variation considered as noise are removed. Then, they feature a rather constant value until $\mathcal{R} = 0.1$ since only the main modes of variation remains (*i.e.* no other modes of variation are removed from the PCA).

MPAM with linear criterion								
	ASSD [mm]				VOE [%]			
$\mathcal{J}$	L01	L02	L03	L04	L01	L02	L03	L04
1.0	4.03	1.78	1.93	2.25	22.54	11.60	11.63	14.09
0.7	2.41	1.76	1.70	1.92	13.77	10.68	10.33	11.87
0.6	2.67	2.59	3.77	2.55	15.64	15.04	21.89	15.01
0.5	2.91	2.87	4.26	2.79	16.88	16.92	24.38	16.37
0.4	3.08	2.89	4.29	2.86	17.73	17.14	24.60	16.70
0.3	3.09	2.88	4.29	2.84	17.77	17.09	24.61	16.64
0.2	3.04	2.89	4.30	2.86	17.56	17.08	24.63	16.71
0.1	3.06	2.89	4.30	2.85	17.66	17.09	24.62	16.67

Table 6.9: Results of the segmentation performed on four liver images {L01,L02,L03,L04} when using our MPAM and linear criterion. MPAM is based on [CascBoostHier] with different thresholds on the Jaccard index $\mathcal{J}$. Results are assessed with both ASSD and VOE. Corresponding number of modes K are in Table 6.6.

Comparison between MPAM and PCAP When comparing the PCAP with our MPAM using, respectively, ratio on eigenvalues $\mathcal{R}$ and Jaccard index $\mathcal{J}$, we notice they have similar features. At maximum value (*i.e.* 1.0), every variation mode (PCA) and all clustering modes (MPAM) are taken into account. When reducing this value, complexity of the system is reduced as modes with low eigenvalues are removed (PCAP) and as similar clustering modes are merged (MPAM). Though Jaccard index $\mathcal{J}$ associated with our MPAM may be used with every similarity measure presented in section 5.2, ratio on eigenvalues $\mathcal{R}$ may only be used with the Mahalanobis distance. This is because both sum of absolute differences and linear criterion do not make use of the covariance matrix on which $\mathcal{R}$ is applied.

In Figure 6.12, we compare the segmentation results of MPAM and PCAP using both Mahalanobis distance, *i.e.* with Jaccard index $\mathcal{J} = 1.0$ (MPAM) and ratio on eigenvalues $\mathcal{R} = 0.9$ (PCAP). In Figure 6.13, we compare the best segmentation results of MPAM and PCAP, *i.e.* with Jaccard index $\mathcal{J} = 1.0$ (MPAM) and sum of absolute differences (PCAP).

With Mahalanobis distance, the PCAP gives better results when reducing the number of variation modes since this reduction removes modes of variation that are not significant. This is not the case with our MPAM since reducing the number of clustering modes creates modes with increased variance and makes thus the segmentation more sensitive to outliers. From the three similarity measures tested, our MPAM outperforms the PCAP despite the fact that less profile modes are used. Indeed, our MPAM has a maximum of ~1250 modes (see Table 6.6) while the PCAP has a constant number of 7586 modes (*i.e.* the number of vertices of $\mathcal{M}_2^\star$).

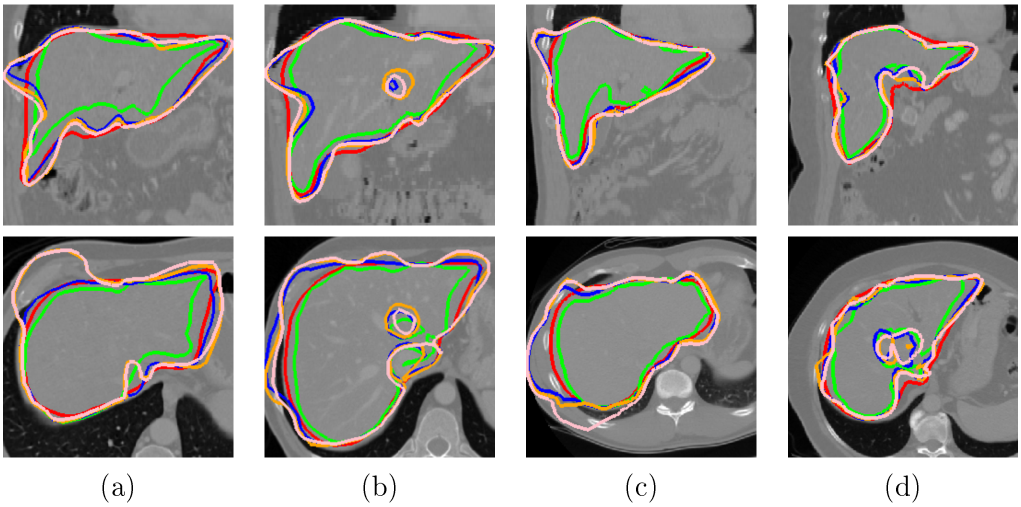

(a) (b) (c) (d)

Figure 6.11: Segmentation of four liver meshes {(a),(b),(c),(d)} with reference mesh $\mathcal{M}_2^\star$ ($N = 7586$) and the Mahalanobis distance. External forces are generated from PCAP using different ratios on eigenvalues: $\mathcal{R} = 1.0$ (green), $\mathcal{R} = 0.9$ (blue), $\mathcal{R} = 0.8$ (orange), and $\mathcal{R} = 0.6$ (pink). Original mesh is depicted in red. Deformations are applied during 100 iterations. Corresponding statistics may be found in Table 6.10.

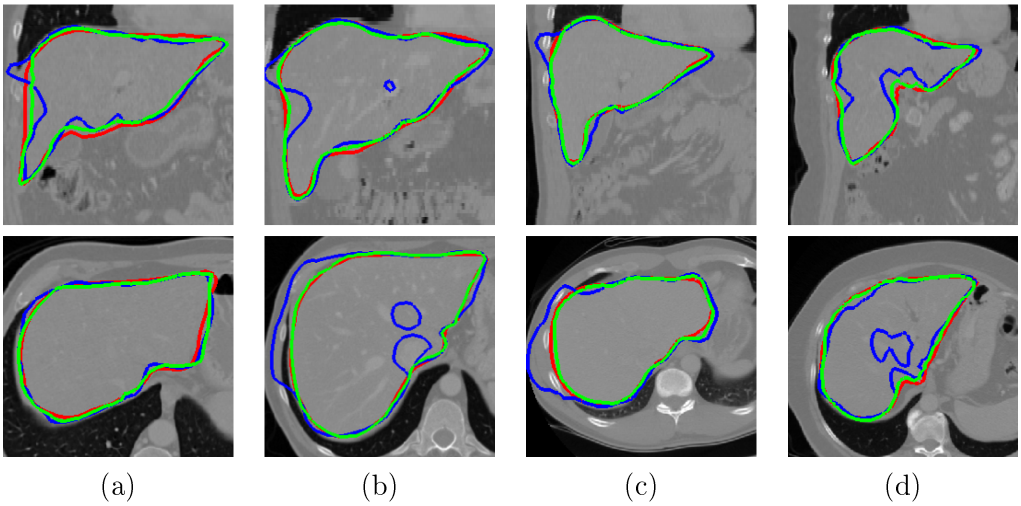

(a) (b) (c) (d)

Figure 6.12: Comparison between MPAM and PCAP using both Mahalanobis distance for the segmentation of four liver meshes {(a),(b),(c),(d)} with reference mesh $\mathcal{M}_2^\star$ ($N = 7586$): MPAM using threshold on the Jaccard index $\mathcal{J} = 1.0$ (green) and PCAP using ratio on eigenvalues $\mathcal{R} = 0.9$ (blue). Original mesh is depicted in red. Deformations are applied during 100 iterations.

PCAP with Mahalanobis distance								
	ASSD [mm]				VOE [%]			
$\mathcal{R}$	L01	L02	L03	L04	L01	L02	L03	L04
1.00	7.93	4.82	4.90	5.84	39.76	26.52	26.03	32.13
0.98	6.11	3.47	2.14	3.76	33.05	20.44	12.27	22.28
0.95	4.62	3.34	2.39	3.31	25.21	19.16	13.18	19.63
0.90	4.09	3.31	3.11	3.13	22.32	19.75	16.61	18.61
0.80	4.09	3.47	3.75	2.85	22.16	20.40	19.58	17.19
0.60	3.97	3.17	5.03	3.12	21.38	18.54	25.27	17.64
0.40	3.96	3.23	5.48	3.17	21.33	19.04	27.10	17.79
0.20	3.97	3.22	5.50	3.17	21.44	19.16	27.19	17.86
0.10	3.97	3.22	5.50	3.17	21.44	19.16	27.19	17.86

Table 6.10: Results of the segmentation performed on four liver images {L01,L02,L03,L04} when using PCAP with Mahalanobis distance and different ratios on eigenvalues $\mathcal{R}$. Results are assessed with both ASSD and VOE. Some corresponding delineations may be found in Figure 6.11.

PCAP with sum of absolute differences and linear criterion								
	ASSD [mm]				VOE [%]			
Similarity measure	L01	L02	L03	L04	L01	L02	L03	L04
Sum of absolute differences	3.57	2.93	1.83	2.97	20.04	16.61	10.24	17.74
Linear criterion	3.52	3.02	3.22	2.81	19.54	18.14	16.44	17.38

Table 6.11: Results of the segmentation performed on four liver images {L01,L02,L03,L04} when using PCAP with both sum of absolute differences and linear criterion, and with different ratios on eigenvalues $\mathcal{R}$. Results are assessed with both ASSD and VOE.

6.8.4 Comparison after initialization based on affine registration

The initialization is the first step of any explicit model-based segmentation procedure. The mesh may be approximately initialized into the image because many structures of interest have a shape and location into the body that are well known. In the literature, different initialization techniques have been proposed, depending on the model used (see section 3.3.3).

In our case, this initialization constitutes the last step to study how our MPAM behaves in a complete segmentation procedure. Also, this initialization ensures a more thorough performance comparison with respect to the PCAP. In previous section, reference mesh $\mathcal{M}_2^\star$ was registered to four liver meshes to study the external forces (*i.e.* using four ground truth liver meshes). In this section, the initialization is performed by registering reference mesh $\mathcal{M}_2^\star$ to liver meshes, but with a similarity transformation this time. The objective is to deform the original shape of reference mesh (see Figure 6.1, right), as it is usually the case in explicit model-based segmentation (see an illustration in Figure 6.14).

For our tests, we first register reference mesh $\mathcal{M}_2^\star$ on 12 liver meshes from our

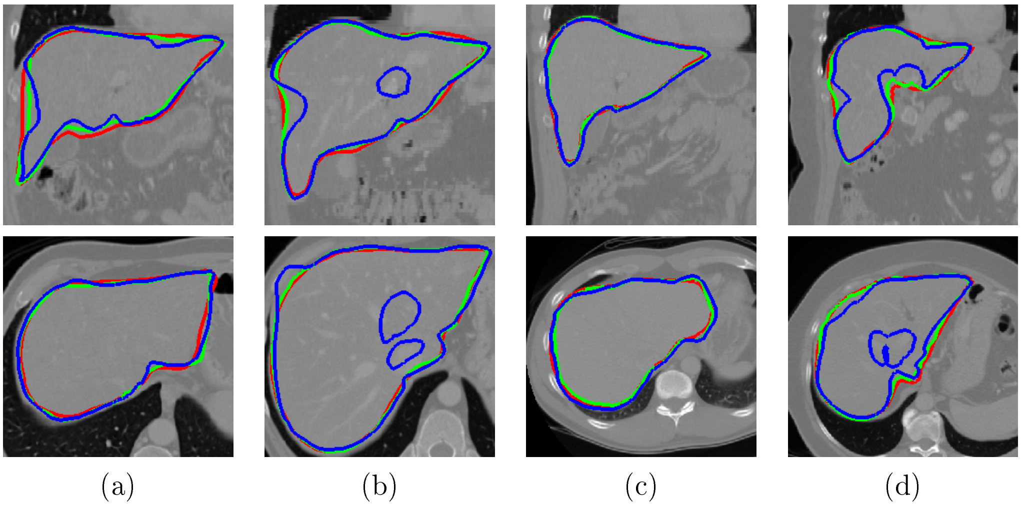

(a) (b) (c) (d)

Figure 6.13: Comparison between MPAM using Mahalanobis distance and PCAP using sum of absolute differences for the segmentation of four liver meshes $\{(a),(b),(c),(d)\}$ with reference mesh $\mathcal{M}_2^{\star}$ ($N = 7586$): MPAM using threshold on the Jaccard index $\mathcal{J} = 1.0$ (green) and PCAP using sum of absolute differences (blue). Original mesh is depicted in red. Deformations are applied during 100 iterations.

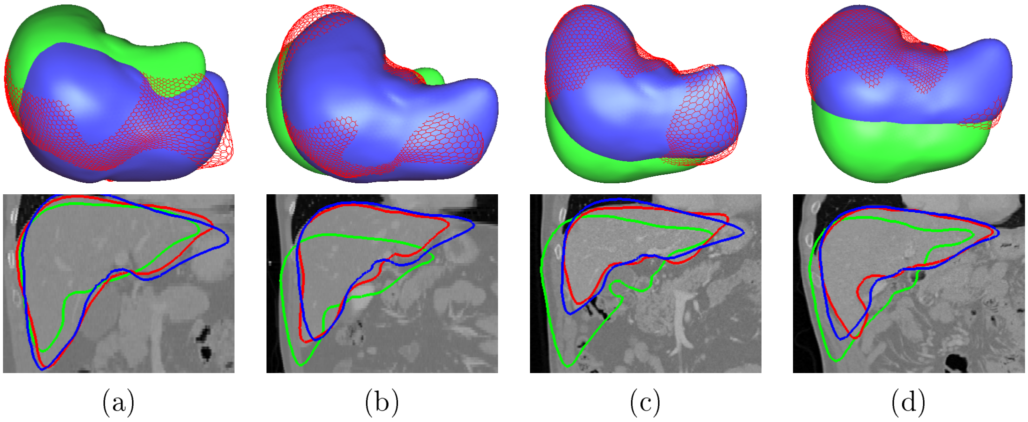

(a) (b) (c) (d)

Figure 6.14: Registration of reference mesh $\mathcal{M}_2^{\star}$ ($N = 7586$) towards four liver meshes $\{(a),(b),(c),(d)\}$ using a similarity transformation. Results are depicted at mesh (top) and image (bottom) level. Original reference mesh is in green, target liver mesh in red and registered reference mesh in blue.

database using the similarity transformation. The classification associated with each of these 12 liver meshes is not included into both the MPAM and PCAP (*i.e.* using a Leave-One-Out cross validation), so that external forces coupled with this initialization stage are fully studied without bias (*i.e.* without using appearance information from the datasets that are segmented).

Then, we segment the 12 corresponding liver images using the different combinations of appearance priors and similarity measures that gave the best results in previous section, namely MPAM with Mahalanobis distance and Jaccard index $\mathcal{J} = 1.0$ (MPAM+maha), PCAP with sum of absolute differences (PCAP+diff), PCAP with linear criterion (PCAP+line) and PCAP with ratio on eigenvalues $\mathcal{R} = 0.9$ (PCAP+maha). Results are assessed with ASSD, VOE, MSD and SRVD (see Table 6.12). Mean and standard deviation of these segmentation results may be found in Table 6.13. Mean and standard deviation of the computation time needed to perform the 12 segmentations may be found in Table 6.14.

Results assessed by ASSD, VOE and MSD show that our MPAM outperforms the PCAP, even with this initialization stage. SRVD seems to indicate that all methods slightly under-segment, which is probably due to the influence of the larger inner part of the profile compared to its outer part, *i.e.* profiles are extending from 9 mm inward to 6 mm outward. Also, the inner part of the liver (*i.e.* the parenchyma) features a lesser variance, which makes it more discriminant. MPAM needs slightly more computation time for the segmentation than PCAP though, which is due to the multimodality of the method (*i.e.* more than one mode may be compared to current profiles during local search).

6.9 Conclusion

In this chapter, we used a database of 35 liver meshes with four resolutions to test the different methods presented in this thesis. Results were described and analyzed in a pipeline approach.

First, we compared the performance of our proposed OSI criterion with classical model order selection criteria. Results showed that the former is a rather good tradeoff between under and over-fitting (though slightly encouraging under-fitting), while the latter is affected by both under and over-fitting. Then, we searched for an optimal profile length and inward/outward ratio using both the OSI criterion and the eigengap heuristic. Our experiments showed that, regardless of the inward length chosen, a profile extending between 8 and 12 mm outward gives stable and smooth results.

Regarding spectral clustering, our experiments showed that, for the four resolutions of the reference mesh, the eigenvalues associated with the 20 top eigenvectors feature the same steep decreases, which proves that the eigengap heuristic is robust to resolution change. Regarding boosted clustering, we tested the different versions of the algorithm. Results showed that the cascading boosted clustering with hierarchical approach gives the best results in terms of stability of both the final number of modes and the percentage of kept profiles per dataset through the different resolutions.

We also tested the ability of the MPAM to reach a constant value after the fusion of a minimum number of datasets. To this end, we added datasets to the MPAM and plotted the final number of modes for different thresholds on Jaccard index $\mathcal{J}$ and after both spectral and boosted clustering. For the four resolutions of the reference mesh, results featured a similar behavior for every threshold and after both spectral and boosted clustering. In both cases, a threshold $\mathcal{J} = 0.5$ showed to be a reasonable choice since, with this threshold, a rather constant value seems to be reached when adding between 5 and 15 datasets into the MPAM.

In a second part, we assessed the segmentation performance of both MPAM and PCAP. First, we optimized the only internal force by studying its effect on mesh deformation. Without external force, we noticed that most of the mesh surface is preserved except in high curvature zones where the mesh is subject to a shrinking effect. As a solution, rigidity parameter η associated with simplex meshes must be locally increased in these zones. Then, we searched for an optimal tradeoff between internal and external forces. Our experiments showed that parameters $\{\alpha = 0.3, \beta = 0.1, \eta = 10\}$ give reasonable results in terms of mesh regularity after 100 deformation iterations.

Finally, we compared the external forces generated from both MPAM and PCAP using, respectively, different thresholds on Jaccard index $\mathcal{J}$ and different ratios on eigenvalues $\mathcal{R}$. We noticed that $\mathcal{J}$ and $\mathcal{R}$ have similar features. At maximum value (*i.e.* 1.0), every variation mode (PCAP) and all clustering modes (MPAM) are taken into account. When reducing this value, complexity of the system is reduced as modes with low eigenvalues are removed (PCAP) and as similar clustering modes are merged (MPAM). With MPAM, segmentation accuracy decreases with $\mathcal{J}$, which is due to the increased mode variance from mode fusion that leads to less discriminant modes during local search. Best segmentation results are obtained when using Mahalanobis distance with $\mathcal{J} = 1.0$ (MPAM) and sum of absolute differences (PCAP).

Overall, results showed that our MPAM outperforms the PCAP, even when including an initialization stage using a similarity transformation, and despite the fact that less profile modes are used with our MPAM. In both cases, segmentation results showed a slight under-segmentation, which is probably due to the influence of the larger inner part of the profile compared to its outer part (*i.e.* profiles are extending from 9 mm inward to 6 mm outward).

ASSD [mm]	L01	L02	L03	L04	L05	L06	L07	L08	L09	L10	L11	L12
[MPAM+maha]	1.84	3.22	1.24	4.87	1.42	3.34	1.99	1.36	2.43	2.25	1.67	1.34
[PCAP+diff]	3.38	2.57	2.21	5.13	2.83	7.69	2.99	2.78	4.71	4.55	3.56	2.54
[PCAP+line]	3.87	4.55	3.36	4.59	3.23	8.24	4.15	3.17	3.10	3.57	4.77	3.11
[PCAP+maha]	3.64	2.99	3.17	5.23	3.62	6.74	3.48	3.69	4.89	4.61	4.18	3.46

VOE [%]	L01	L02	L03	L04	L05	L06	L07	L08	L09	L10	L11	L12
[MPAM+maha]	10.75	18.93	9.29	24.26	9.33	18.13	11.32	8.35	14.88	12.48	9.63	8.59
[PCAP+diff]	18.88	16.52	15.32	27.58	17.68	38.15	16.68	16.39	24.96	24.08	18.76	11.36
[PCAP+line]	21.16	25.63	22.31	24.38	20.22	38.05	20.85	18.25	16.71	18.57	23.91	16.86
[PCAP+maha]	20.45	19.20	21.31	28.00	21.64	33.36	19.39	21.67	26.02	23.85	21.67	16.59

MSD [mm]	L01	L02	L03	L04	L05	L06	L07	L08	L09	L10	L11	L12
[MPAM+maha]	14.29	43.75	18.86	33.23	16.25	32.73	34.76	21.33	22.05	28.79	22.88	20.05
[PCAP+diff]	28.96	21.29	19.66	39.80	23.03	35.21	33.02	35.01	34.05	37.61	29.49	40.00
[PCAP+line]	28.38	37.50	21.84	35.19	27.39	47.74	39.86	28.83	29.16	29.43	39.96	33.84
[PCAP+maha]	26.97	18.12	20.54	34.48	23.59	29.86	33.77	33.60	33.09	37.45	28.77	39.09

SRVD [%]	L01	L02	L03	L04	L05	L06	L07	L08	L09	L10	L11	L12
[MPAM+maha]	-7.87	-3.94	+1.91	-16.2	-7.43	+5.01	-4.16	-4.10	-12.0	-6.46	-7.96	-4.74
[PCAP+diff]	-9.25	-11.0	-2.12	-20.8	-14.7	-18.7	-13.0	-12.1	+18.3	+7.81	+2.38	-1.23
[PCAP+line]	+4.69	+2.39	+4.08	-13.0	-12.4	+1.34	-6.61	-7.30	-3.73	+2.65	+5.49	+4.14
[PCAP+maha]	-1.24	-8.00	+6.90	-18.4	-11.2	-13.4	-11.1	-9.76	+20.9	+6.59	+6.61	+8.03

Table 6.12: Results of the segmentation performed on 12 liver images $\{$L01,. . .,L12$\}$ with reference mesh $\mathcal{M}_2^\star$ ($N = 7586$) when using: MPAM with mahalanobis distance and Jaccard index $\mathcal{J} = 1.0$ (MPAM+maha), PCAP with sum of absolute differences (PCAP+diff), PCAP with linear criterion (PCA+line), and PCAP with ratio on eigenvalues $\mathcal{R} = 0.9$ (PCAP+maha). Results are assessed with ASSD, VOE, MSD and SRVD.

	ASSD [mm]	VOE [%]	MSD [mm]	SRVD [%]
[MPAM+maha]	2.24 (1.08)	12.99 (5.04)	25.74 (8.85)	-5.66 (5.59)
[PCAP+diff]	3.74 (1.56)	20.53 (7.17)	31.74 (7.00)	-6.20 (11.5)
[PCAP+line]	4.14 (1.43)	22.24 (5.76)	33.26 (7.13)	-1.52 (6.78)
[PCAP+maha]	4.14 (1.07)	22.76 (4.52)	29.90 (6.58)	-2.00 (11.7)

Table 6.13: Mean (SD) of the segmentation results performed on 12 liver images and assessed with ASSD, VOE, MSD and SRVD when using MPAM and PCAP with different similarity measures: MPAM with Mahalanobis distance and Jaccard index $\mathcal{J} = 1.0$ (MPAM+maha), PCAP with sum of absolute differences (PCAP+diff), PCAP with linear criterion (PCA+line), and PCAP with ratio on eigenvalues $\mathcal{R} = 0.9$ (PCAP+maha).

	computation time [min]
[MPAM+maha]	7.75 (0.45)
[PCAP+diff]	5.33 (0.49)
[PCAP+line]	5.50 (0.52)
[PCAP+maha]	6.66 (0.49)

Table 6.14: Mean (SD) of the computation time needed to perform the segmentation of 12 liver images when using MPAM and PCAP with different similarity measures: MPAM with Mahalanobis distance and Jaccard index $\mathcal{J} = 1.0$ (MPAM+maha), PCAP with sum of absolute differences (PCAP+diff), PCAP with linear criterion (PCA+line), and PCAP with ratio on eigenvalues $\mathcal{R} = 0.9$ (PCAP+maha).

Chapter 7

Analysis of Lower Limb Structures

Contents

7.1 Introduction

In this chapter, we describe two collaborative works that have been done within the framework of the 3D Anatomical Human project (3DAH).

The first work, which has been done with Jérôme Schmid (University of Geneva, Switzerland) and Anders Sandholm (EPFL, Switzerland) [Schmid *et al.* 2009], consists in creating subject-specific models for kinematic simulations (section 7.2). We explain our motivation (section 7.2.1) and the way we acquire data (section 7.2.2). Since medical images are known to be noisy and prone to artifacts, we also talk about pre-processing (section 7.2.3) with the aim of reducing these undesirable effects. Then, we describe the different steps to create anatomical (section 7.2.4) and musculoskeletal (section 7.2.5) models.

The second work, which has been done with Jérôme Schmid (University of Geneva, Switzerland) [Chung *et al.* 2011], compares the performance of different statistical models in the context of lower limb bone segmentation using MR images when only a small number of datasets is available for training (section 7.3). We explain our motivation (section 7.3.1) and present the MRI data used for this study (section 7.3.2). Then, we describe the shape and appearance priors from which statistical models are created (section 7.3.3), and both the segmentation framework (section 7.3.4) and experimental setup (section 7.3.5) used for this study. Finally, we present the results (section 7.3.6) and conclude by a discussion (section 7.3.7).

7.2 Subject-specific models for kinematic simulations

7.2.1 Motivation

Although eventually non-lethal, musculoskeletal disorders such as joint degeneration and hip osteoarthritis are the cause of severe and long term pain for patients. This is why predicting musculoskeletal behavior with computer models and simulations may be of a great help for medical diagnosis (*e.g.* orthopedics, movement optimization, rehabilitation and ergonomics).

In the medical imaging community, medical scanners, which are more and more precise, available, standardized and less and less invasive, are extensively used as diagnosis tools. In the biomechanics community, computer simulations of the musculoskeletal system are widely used to study the mechanisms behind human gait and its disorders. Musculoskeletal models are commonly generic, *i.e.* based on data derived from anatomical and biomechanical studies of cadaverous specimens. Within the 3DAH project, one objective of this thesis is to replace those *generic musculoskeletal* models by *subject-specific anatomical* models of the lower limb structures [Schmid *et al.* 2009].

In biomechanics, replacing generic models by subject-specific models consists in scaling the generic models so that they match at best specific features of the subject. However, such a scaling is not accurate. Indeed, studies have shown that some subject-specific anatomical differences may not be taken into account. As a result, errors are introduced in the models. Other methods were proposed to overcome this scaling issue but a major limitation of these approaches is to heavily rely on user interactions. Based on subject-specific medical data (*i.e.* using MR images), an accurate modeling of lower limb anatomical structures may be produced while avoiding this scaling operation.

In [Schmid *et al.* 2009], we introduced a method to create subject-specific musculoskeletal models from MRI datasets and motion capture data. Anatomical models are reconstructed using a model-based semi-automatic segmentation approach, where muscles, tendons, bones and corresponding attachments are identified (see section 7.2.4). These anatomical models coupled with motion capture data, joint kinematics information and muscle-tendons actuators are then used to create subject-specific musculoskeletal models (see section 7.2.5).

7.2.2 Data acquisitions

For musculoskeletal models in biomechanics, data acquisition consists mainly in a motion capture. In [Schmid *et al.* 2009], eight video cameras were used to capture motion data of reflective skin markers placed on subjects at specific lower limb locations. Subjects were also equipped with electromyography (EMG) sensors to capture electrical activity in eight large lower limb muscles. During motion acquisition, the subjects walked on two force platforms that captured ground reaction forces.

For anatomical models, CT and MRI acquisitions are required since these modalities are the references in bone and soft tissue acquisitions, respectively. A protocol needs to be defined, usually with radiologists. A medical-ethical committee must approve the acquisition and subjects need to give their written consent. For lower limb structures, this protocol needs to highlight both soft and bony tissues.

In [Schmid *et al.* 2009], we acquired MR images with the following protocol: Axial 2D T1 Turbo Spin Echo, TR = 578 ms, TE = 18 ms, FOV = 40 cm, FA = 90°, matrix = 512x512, resolution = 0.78x0.78 mm. To acquire the full lower limbs, six acquisitions were performed at the University Hospital of Geneva (Switzerland) on a 1.5T MRI device with varying slice thickness, *e.g.* a higher thickness was used for long bones to speed up the acquisition process. All acquisitions were finally registered and merged in a unique volume thanks to an appropriate overlap between consecutive slabs.

7.2.3 Pre-processing

As previously said in section 2.3.1, medical images are usually corrupted by anisotropic voxels, noise, partial voluming and artifacts. To overcome these issues, pre-processing may be required on images before any further processing. This is because most medical image analysis algorithms rely on image intensities.

When applying low pass filters on images, noise is usually reduced. Anisotropic filters are particularly suited since they smooth the image without altering edge sharpness (see an example in Figure 7.1). Intensity normalization may be applied between subjects, sessions or even slabs. Consecutive slabs are acquired when one acquisition is not sufficient to cover an entire region (*e.g.* whole lower limbs) due to the small field of view (FOV) of the scanner. Bias field, as introduced in section 2.3.3, may also be removed, or at least reduced, as a pre-processing step.

MR images are usually corrupted with noise and known to be prone to artifacts. This is because MR images are acquired in a complex form and only the magnitude of the signal is kept (*i.e.* as image intensity) while the phase information is discarded. To remove noise, anisotropic filters are commonly used, though other methods exist.

A first method consists in correlating two images acquired successively so that only the signal remains [Sijbers *et al.* 1998]. Another method called non-local means (NL-means) [Buades *et al.* 2005] has proven to be very efficient for the denoising of MR images [Coupé *et al.* 2006]. In this algorithm, the filtered value of a voxel is computed as the average of all voxel intensities using weights that are defined by a

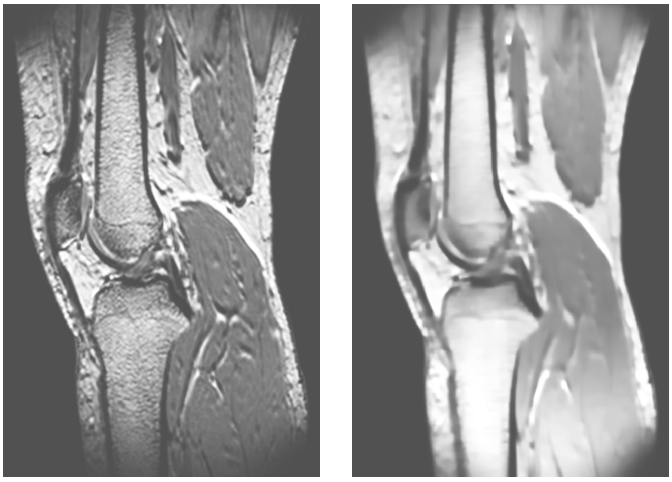

(a) (b)

Figure 7.1: An MR image of the knee (a) before and (b) after anisotropic filtering.

similarity function.

When several images are acquired from different patients or sessions, intensity normalization may be used so that image contrast and brightness match between images. Such normalization may be done using a piecewise affine transform of the intensities. Finally, intensity inhomogeneities (*e.g.* bias field) are likely to be important in MR images and several methods to correct it have been proposed in the literature [Vovk *et al.* 2007, Likar *et al.* 2004, Weisenfeld & Warfteld 2004, Styner *et al.* 2000].

7.2.4 Anatomical models construction

To model lower limb anatomical structures such as bones, muscles and tendons in a (semi-)automatic segmentation procedure, shape information needs to be gathered. After the MRI acquisition during which the lower limbs of a subject are scanned, a manual segmentation of anatomical structures is performed by a medical expert. After generating meshes from these segmentations, they appear not to be smooth. Furthermore, meshes are not anatomically attached whereas they should be. In this section, we describe the various steps to force meshes to be anatomically correct. The objective of this method is to reuse these meshes in (semi-)automatic segmentation methods. We presented the following method and results in [Chung *et al.* 2009, Schmid *et al.* 2009].

Manual delineation The manual delineation of bones and muscles is performed by a medical expert (see Figure 7.2). For each structure, a binary image is produced and then processed by the Marching Cubes algorithm [Lorensen & Cline 1987] to generate 2-simplex meshes [Delingette 1999] representing the structures of interest.

However, the generated meshes appear not to be smooth. This is mostly ex-

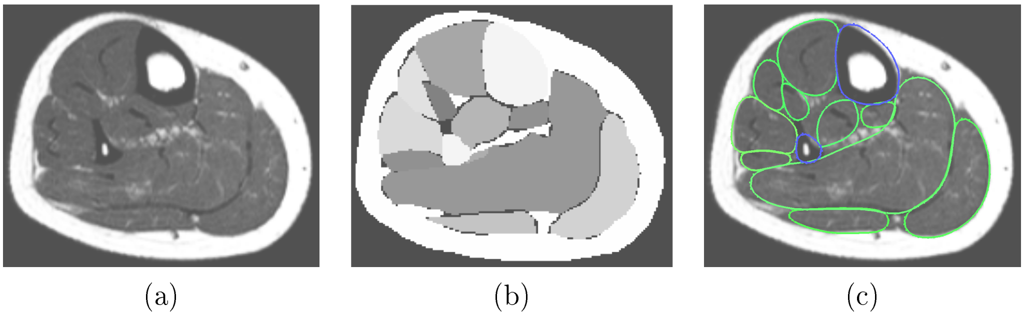

(a) (b) (c)

Figure 7.2: Consecutive steps related to manual delineation. First, (a) an MR image is analyzed by a medical expert (note how difficult it is to see where are muscle boundaries). Then, (b) a manual delineation defines structures of interest by means of labels (a gray value range is used here). Finally, (c) meshes are generated from these labels (contours represent the intersection between MRI slice and generated meshes). Blue contour delineates the tibia bone and green contours delineate the muscles.

plained by the dataset resolution and especially the large inter-slice distance. The ubiquitous error made during a manual segmentation needs also to be taken into account, since it remains a subjective task. 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 and acquisition time restrictions). Furthermore, a lower resolution is preferred to reduce the memory footprint, which results in a speed-up automatic algorithm.

Since muscles are known to be anatomically smooth, the objective is to find a solution that makes them appear so. The technique we use consists first in refining the meshes and then in applying internal constraints on them. More precisely, we apply forces on the meshes so that their rigidity is maximal. For that, we use the simplex angle regularity constraint used for shape recovery (see section 3.3.4.2).

Tendons Although muscles are difficult to delineate manually, tendons are easier to detect in MR images. Indeed, tendons have a uniform appearance in images since they appear as dark structures with respect to muscle tissue. An automatic method based on maximum intensity ridge tracing [Aylward & Bullitt 2002, Pock *et al.* 2005] is first used. This method relies on the assumption that tendons are roughly tubular structures, which is especially the case for leg tendons. These tracing methods have proved to be robust to noise. They provide also a simple way to model a structure as a series of centerline points (*i.e.* a position and a radius). Furthermore, such representation could be reused in another automatic method.

However, this method may not always segment correctly or entirely tendons because in some regions their intensity is corrupted by strong artifacts (*e.g.* noise, adjacent structures and partial volume effect). As a result, tendons are difficult to follow through slices. Moreover, some tendons are so close that distinguishing them

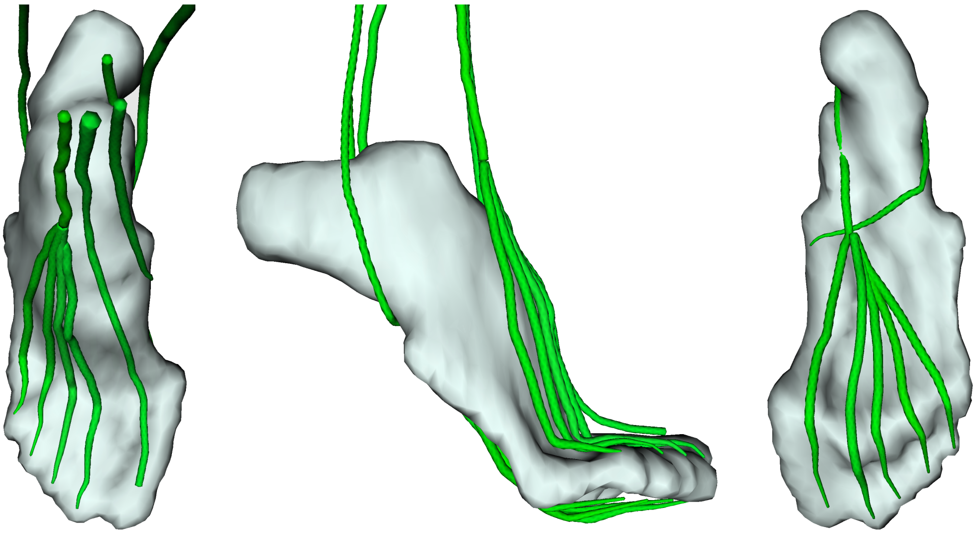


Figure 7.3: Three views showing the modeling of tendons in the foot region.

without some prior knowledge is difficult. To cope with these problems, missing centerline points are manually placed on high curvature points of tendons. Eventually, when the centerlines are complete using either an automatic or manual modeling, simplex meshes representing the tendons are automatically created (see Figure 7.3).

Attachments An attachment is defined as the linking region between two anatomical structures. For instance, tendons are usually attached to bones. In this case, tendon attachments are defined as tendon tissues sharing a common region with bone tissues. Comparatively, 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). Our solution consists in manually placing splines to define attachment regions as described in [Gilles 2007]. Meshes are then deformed until they are stuck. This provides a simple but efficient way to model attachment regions.

Interpenetration removal Smoothing and attachment procedures previously described may create non realistic interpenetrations between meshes. To solve this problem, we decide not to use implicit surfaces [Alexa & Adamson 2004] because the hypothesis of vicinity to surfaces may be invalid (*i.e.* in case of big interpenetrations). Instead, we propose an iterative geometrical post-processing method (see [Schmid *et al.* 2009] for details). Nevertheless, a perfect contact between meshes is not ensured but the results are satisfactory when mesh resolution is not too low (see some results in Figure 7.4). To speed up the process, precomputed signed distance maps may be used as well as the efficient golden section search technique [Press *et al.* 1992].

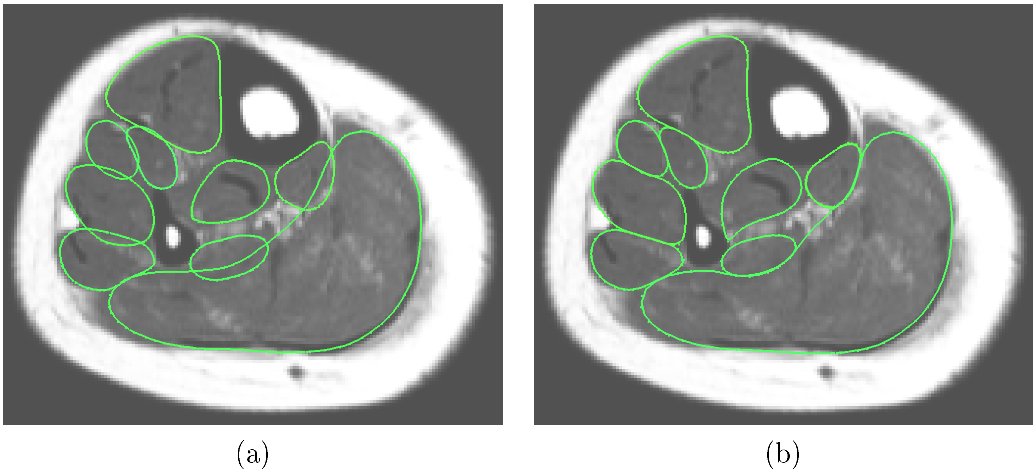

(a) (b)

Figure 7.4: Axial view of a lower limb MR image showing muscle delineation (a) before and (b) after interpenetration removal.

Semi-automatic segmentation As a result, we have a collection of anatomical meshes representing various soft and bony structures (see Figure 7.5). The objective is to re-exploit these meshes as priors in a (semi-)automatic segmentation. This method could be then applied on any similar dataset coming from other subjects. Our semi-automatic segmentation procedure is mostly based on deformable models [Gilles *et al.* 2006], which are explained more into details in section 3.3.

7.2.5 Musculoskeletal model generation

To generate subject-specific musculoskeletal models, anatomical models need first to be aligned with the motion capture data. For each bone (*i.e.* pelvis, femur and tibia), a rigid transformation is computed using corresponding skin markers. Doing so, each bone may be correctly aligned with the motion capture data. Similarly, this rigid transformation ensures skin markers to be expressed in the MR image coordinate space.

In a second step, information about muscles and joints needs to be included. Each joint center, computed from the anatomical models, is used as center point for the moment force generated by the muscles. Then, a function is defined to describe the kinematic behavior of each joint. By using this function and a given joint angle, the relative attachment point of joints may be determined. To calculate the force that each muscle and its tendons may generate, each muscle is expressed into a set of one or more muscle-tendon actuators. These muscle-tendon actuators stretch between the muscles insertion points (*i.e.* attachment regions) either as a straight line or via a wrapping point.

When a subject-specific simulation is carried out, generated musculoskeletal models are loaded along with motion capture data. Models match the motion capture data, so that no initial scaling needs to be done.

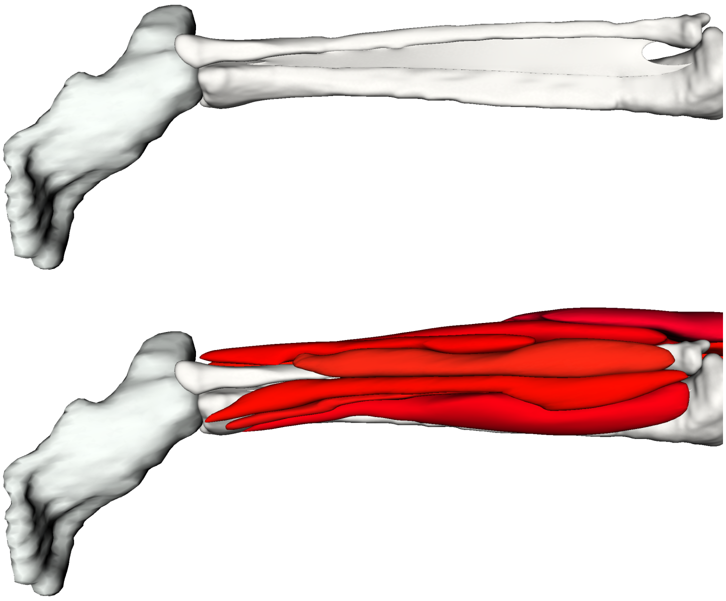


Figure 7.5: Some anatomical models of the lower limbs after 3D reconstruction, including bones (top) and muscles (down).

7.3 Bone segmentation using few training datasets

7.3.1 Motivation

Building meaningful priors for model-based image segmentation purposes is an important topic in medical imaging. To account for a large variety in both shape and appearance, numerous datasets are usually required in a training stage. Atlas-based methods are designed in the hypothesis that a significant number of training datasets may yield a reasonable mean model (*i.e.* in the sense of a population mean) as well as meaningful modes of variation.

Those methods usually define a reasonable estimation of the prior using the principal modes of variation (*i.e.* variations may be controlled with few parameters). They have been successfully used in coarse-to-fine approaches where the algorithm starts with a rough approximation of the prior (*i.e.* with few degrees of freedom) and evolves adding more variation until reaching a steady state [Heimann *et al.* 2007, Schmid & Magnenat-Thalmann 2008].

In the literature, shape variations are often described using Principal Component Analysis (PCA) [Seim *et al.* 2008, Schmid & Magnenat-Thalmann 2008] of Point Distribution Models (PDM) [Cootes & Taylor 2004] (see section 3.3.4.2). Regarding appearance, methods such as Active Shape Models (ASM) [Cootes & Taylor 1993] and Active Appearance Models (AAM) [Cootes & Taylor 2001] were proposed to account for the main intensity variation around and within structures of interest, respectively.

However, the number of datasets required to account for both the shape and appearance of a structure may be an issue. First, medical imaging acquisitions require time and resources, and this may limit the number of datasets available for training. Second, the large number of acquisition protocols and hardware characteristics (especially in case of versatile modalities such as MRI) tend to produce images with a large variety of intensity distribution for the same structure of interest. Additionally, noise and artifacts (*e.g.* patient movement and partial volume effect) are likely to corrupt the image intensity.

These factors strongly affect the construction of priors, as they bring meaningless intensity information into the appearance priors. Also, the manual segmentation of images by an expert, which is required for the extraction of shapes, is a tedious task and is a limiting factor for the availability of training datasets. Finally, the large natural variability of shape and appearance cannot be well represented by a Gaussian distribution assumed by PCA and thus capturing all the shape variations is still very challenging. As a result, the number of required datasets often seems insufficient to fully capture variations in both shape and appearance, especially in case of 3D modeling [Heimann 2009]. Various works such as FEM vibrational modes [Cootes & Taylor 1996] have been proposed to artificially produce additional modes of variation, but it seems that this approach is mainly restricted to cope with intra-subject variability.

To study the influence of few training datasets on segmentation results, we propose to test the performance of various shape and appearance priors in the context of lower limb bone segmentation using MR images [Chung *et al.* 2011]. In practice, two training sets are tested: one with only three datasets featuring a rather homogenous intensity distribution and the other with three more datasets featuring MRI artifacts. For shape modeling, both PCA and shape memory strategies are tested. PCA-based methods are known to need several datasets to be meaningful, while shape memory method requires in practice only one dataset. For appearance modeling, methods based on intensity profiles are tested, namely mean intensity profiles, multivariate Gaussian distributions of profiles and multimodal profiles from EM clustering.

Our objective is to find the most efficient strategy, *i.e.* the strategy that is robust against the low number of datasets and their intensity inhomogeneities. This strategy would have the advantage to be more easily integrated in a clinical environment where the need of quick results is vital, regardless of the number of datasets.

7.3.2 MRI data

7.3.2.1 Acquisitions

For this study, six acquisitions were performed on six different subjects (4 females and 2 males, aged between 25 and 35). Protocols used for each acquisition are detailed in Table 7.1. The acquisitions took place in two different locations. Three subjects were scanned at the St Mary's Hospital, London, UK on a GE Medical

Subject	TR/TE(ms)	FOV(cm)/Matrix	Resolution(mm)
#1	4.15/1.69	35/256x256	1.37x1.37x5
#2	4.15/1.69	35/256x256	1.37x1.37x5
#3	4.15/1.69	35/256x256	1.37x1.37x5
#4	5.06/2.23	43/256x256	0.84x0.84x2
#5	4.34/1.56	40/256x256	0.78x0.78x2
#6	5.09/2.22	43/256x256	0.84x0.84x2

Table 7.1: MRI protocols used to scan the six subjects. TR, TE and FOV stand for, respectively, repetition time, echo time and field of view.

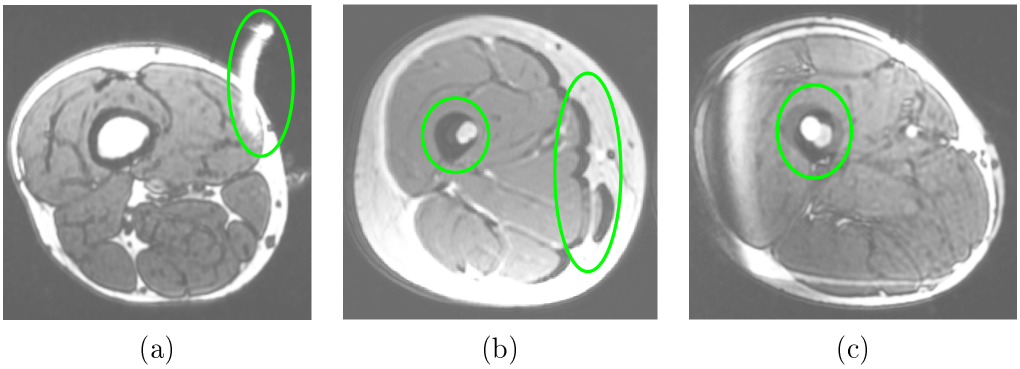

(a) (b) (c)

Figure 7.6: Some common MRI artifacts: (a) surface coil artifact on one distinctive spot, (b) tissue shift artifact on two distinctive spots (bone and fat tissue shifts), and (c) subject motion artifact at the overlap between two consecutive slabs where the femur is clearly shifted.

Systems 1.5T MRI device (subjects #4, #5 and #6) and the three other subjects at the University Hospital of Geneva, Switzerland on a Philips Medical Systems 1.5T MRI device (subjects #1, #2 and #3). Both acquisitions feature 256x256 matrices but with a different slice thickness (*i.e.* 2 mm at London and 5 mm at Geneva). An institutional medical-ethical committee approved this study and subjects gave their written consent.

Some MR images present strong artifacts like subject motion, surface coil, slice-to-slice interference and bias field (see Figure 7.6). Subject motion is an artifact created by the displacement of a structure, which arises when the subject slightly moves between two consecutive acquisitions. Surface coil is characterized by a very strong signal due to the close proximity of the subject with the surface coil. Slice-to-slice interference artifact is due to the cross-excitation of adjacent slices with contrast loss in reconstructed images. Finally, bias field is a very common artifact in MR images, which may be induced by a number of factors such as poor radio frequency coil uniformity, static field inhomogeneity and radio frequency penetration.

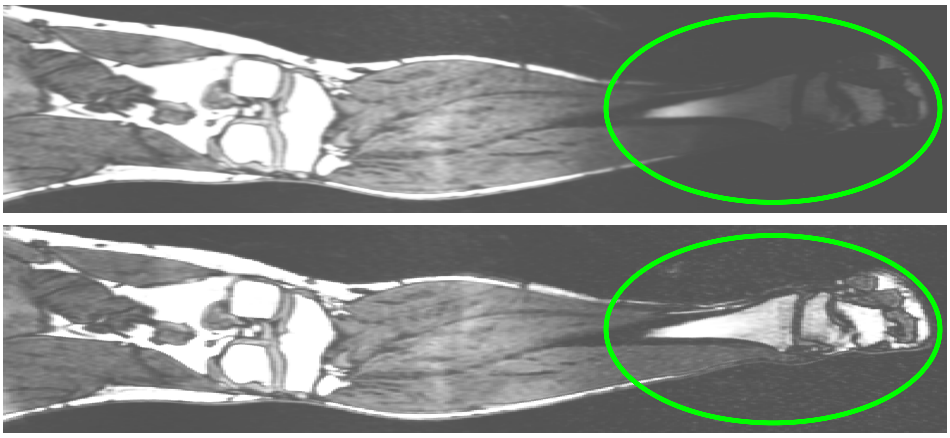


Figure 7.7: MRI saggital view of the lower limbs before (top) and after (bottom) intensity correction between slabs using histogram normalization. Note how intensity has been corrected at foot level.

7.3.2.2 Fusion of MRI slabs

An MRI acquisition of the whole lower limb cannot be performed in one scan due to a limited field of view (FOV) of the machine. Consecutive scans, known as *slabs*, are necessary to cover its entire length. The registration of these slabs is then required to generate a complete MR image of the lower limbs. The registration is computed thanks to a sufficient overlap, which is used to compute the transformation matrix between slabs. In our case, we use a rigid registration based on the manual placement of landmarks.

However, the registration is not enough to create a satisfactory MR image of the lower limbs. Indeed, intensity distribution may vary between registered slabs (see Figure 7.7, top), which means the intensity histogram for the same structure may slightly differ between slabs. This is because the intensity range of the MR image may change between slabs (*i.e.* the minimum and maximum intensity value is different). Also, the presence of strong artifacts at the image boundaries is likely to affect the intensity distribution (in addition to the common bias field). These artifacts are unpredictable and may differ in their number and importance between slabs.

Our solution consists in putting artifact intensity values into the background, so that when normalizing intensity between slabs, artifacts do not bias the correction. The normalization is performed using an intensity scaling factor. To this end, we calculate for each slab the histogram of the main structures (*i.e.* bones, muscles and fat) and compute the scaling factor so that histograms are similar. Finally, a bias correction is applied on the generic MR image of the lower limbs to correct intensity inhomogeneities [Styner *et al.* 2000], *i.e.* to correct the bias field and to smooth the intensity scaling between slabs (see Figure 7.7, bottom).

7.3.3 Creation of appearance and shape priors

In this paper, we propose to exploit priors built from both the appearance and shape of structures in model-based image segmentation. A set of P training shapes $\{\mathbf{S}_1, \ldots, \mathbf{S}_P\}$ with corresponding images $\{I_1, \ldots, I_P\}$ is necessary to model the priors.

Each shape $\mathbf{S}$ is modeled as a 2-simplex mesh [Delingette 1999] defined by N vertices $\mathbf{p}_i$ with normals $\mathbf{n}_i$:

$$\mathbf{S} = \{(\mathbf{x}_1, \mathbf{n}_1), \ldots, (\mathbf{x}_N, \mathbf{n}_N)\} \tag{7.1}$$

The shapes are produced by a supervised segmentation approach in which the point correspondence is established (*i.e.* landmarks on all P training shapes are located at corresponding positions). This point correspondence, which is the first necessary step when building shape models with PDM, is ensured through a registration between training shapes. Then, the P training shapes are aligned in a common coordinate frame. The most popular method to solve this problem is the General Procrustes Analysis (GPA) [Goodall 1991, Gower 1975], which aligns the set of P training shapes to their unknown mean by minimizing the mean squared distance between two shapes in an iterative procedure. After alignment, dimensionality of the training set is reduced to find a small set of modes that best describes the observed variation. This is accomplished using PCA [Jolliffe 2002]. More information on the issues of building training sets may be found in [Heimann 2009].

7.3.3.1 Shape priors

For the shape prior construction, we propose to use PCA-based priors (PCA) and shape memory (SMEM), which are both described in section 3.3.4.2.

7.3.3.2 Appearance priors

To take appearance into account, intensity profiles $\mathbf{x}_i$ are built by sampling the image intensity at each point $\mathbf{p}_i$ along the normal direction. From these profiles, various approaches to build a prior have been presented in the literature (see section 3.3.5.4). In this study, we consider multivariate Gaussian distributions of intensity profiles (MGD) and multimodal profiles (MPAM), whose construction is described in section 3.3.5.5 and chapter 4, respectively. In addition, we consider mean intensity profiles (PROF), whose construction is described in the following paragraph.

PROF Mean intensity profiles constitute the simplest appearance prior. At each corresponding point through all datasets, a mean intensity profile is computed as:

$$\boldsymbol{\mu}_i = \sum_{j=1}^{P} \mathbf{x}_i / P \tag{7.2}$$

Though faster to compute due to its simplicity, this prior does not make any assumption about the variance. As a result, mean intensity profiles are rather poor priors. However, they have been successfully exploited in previous works [Gilles & Magnenat-Thalmann 2010, Schmid & Magnenat-Thalmann 2008], in which they were combined with robust similarity measures such as the Normalized Cross Correlation (NCC) [Holden *et al.* 2000].

7.3.4 Segmentation based on priors

The proposed segmentation is based on dynamic deformable models. A deformable template evolves until reaching an equilibrium [Gilles & Magnenat-Thalmann 2010, Schmid & Magnenat-Thalmann 2008]. Internal forces regulate its evolution while external forces drive it towards anatomical boundaries. This section discusses the effect of above-mentioned priors into this deformable model framework.

7.3.4.1 Evolution

Dynamic deformable models behave like a particle system, in which each particle corresponds to a lumped-mass vertex subject to internal and external forces. The dynamics of the system follow the Newtonian laws of motion, which express particle state (*i.e.* position and velocity) with respect to forces. The resulting time-discretized differential equation system is solved by an implicit Euler scheme.

A multiresolution scheme is used to produce various levels of detail (LOD) of the shapes [Gilles & Magnenat-Thalmann 2010] (see Figure 7.9). The LOD are then exploited in a coarse-to-fine fashion, improving the robustness and accuracy of the segmentation evolution. In case of simultaneous segmentation of more than one structure, efficient collision response and detection are applied to prevent interpenetrations [Gilles & Magnenat-Thalmann 2010]. Alternatively, a post-processing correction method may be used [Schmid *et al.* 2009]. Forces are expressed based on the image information, the current model configuration and the pre-computed priors.

Forces at point $\mathbf{p}_i$ are expressed as the force of a Hookean spring:

$$f_i = \varepsilon * (\tilde{\mathbf{p}}_i - \mathbf{p}_i) \tag{7.3}$$

where $\tilde{\mathbf{p}}_i$ denotes the *target* point and ε is a weighting coefficient specific to each type of force. We will see in the following how the target point is computed given the different forces. This procedure is hereupon referred to as *source-to-target* approach.

7.3.4.2 Internal forces based on shape priors

PCA As depicted in section 7.3.3.1, the shape priors are expressed by a SSM built on a PCA. At each iteration, a closest shape $\hat{\mathbf{S}}$ is found by projecting the current deformable shape $\mathbf{S}$ into the PCA space. An iterative procedure computes the adequate transformation T and appropriate constrained shape parameters $\mathbf{b} = b_1, \ldots, b_R$ [Cootes & Taylor 1993]. Hard or soft constraints are applied to discard illegal configurations.

Then, $\hat{\mathbf{S}} = \{\hat{\mathbf{p}}_1, \ldots, \hat{\mathbf{p}_N}\}$ eventually becomes the target shape and the source-to-target approach is applied:

$$f_i^{\text{pca}} = \varepsilon^{\text{pca}}(\hat{\mathbf{p}}_i - \mathbf{p}_i) \tag{7.4}$$

SMEM In case of a single reference shape used as a prior, both precomputed and current simplex parameters are used to derive new target point positions. A force f_i^{smem} is then produced at each point $\mathbf{p}_i$.

7.3.4.3 External forces based on appearance priors

At each iteration of the evolution, a number W of intensity profiles $\{\mathbf{x}_i^1, \ldots, \mathbf{x}_i^W\}$ are sampled along the normal $\mathbf{n}_i$ at vertex $\mathbf{p}_i$. Among them, a target profile $\tilde{\mathbf{x}}_i$ is chosen, whose corresponding position is $\tilde{\mathbf{p}}_i$. Usually, this target profile is chosen so that it maximizes a similarity criterion or minimizes a distance.

PROF In case of mean intensity profiles, a target profile $\tilde{\mathbf{x}_i}$ is selected if it maximizes the Normalized Cross Correlation NCC with the mean intensity profile $\boldsymbol{\mu}_i$:

$$\tilde{\mathbf{x}_i} = \text{argmax } NCC(\mathbf{x}_i^j, \boldsymbol{\mu}_i) \tag{7.5}$$

where $j \in [1, W]$.

MGD When using a multivariate Gaussian distribution of intensity profiles, the information from the covariance matrix $\boldsymbol{\Sigma}_i$ is also taken into account. In this case, a target profile $\tilde{\mathbf{x}_i}$ is selected if it minimizes the Mahalanobis distance d_M derived from the computed distribution [Cootes & Taylor 1993]:

$$\tilde{\mathbf{x}_i} = \text{argmin } d_M(\mathbf{x}_i^j) = \sum_{l=1}^{m} \frac{b_{il}^2}{\lambda_{il}} \tag{7.6}$$

where $j \in [1, W]$ and b_i is computed from equation 3.41.

MPAM The target profile $\tilde{\mathbf{x}}_i$ of a given point $\mathbf{p}_i$ is chosen as one of the W profiles sampled along the normal that has the smallest Mahalanobis distance d_M with one of its K_i associated clusters:

$$\tilde{\mathbf{x}}_i = \operatorname{argmin} \, d_M(\mathbf{x}_i^{j,k}) = (\mathbf{x}_i - \boldsymbol{\mu}_k)^T \mathrm{diag}(1/\sigma_{k,1}, \ldots, 1/\sigma_{k,d})(\mathbf{x}_i - \boldsymbol{\mu}_k) \tag{7.7}$$

where $j \in [1, W]$, $k \in [1, K_i]$ and $\mathrm{diag}(1/\sigma_{k,1}, \ldots, 1/\sigma_{k,d})$ is the diagonal of covariance matrix $\boldsymbol{\Sigma}_k$ (as explained in section 5.2.2).

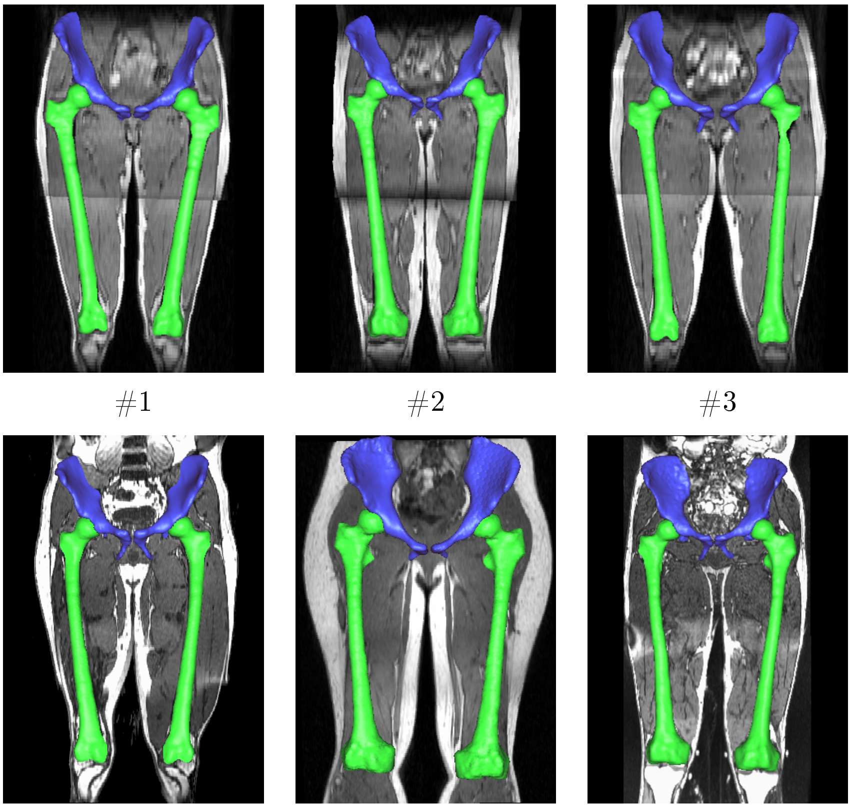
7.3.5 Experimental setup

Two training sets, D3 and D6 with, respectively, three and six lower limb MRI datasets (see Fig. 7.8 for more details), are used with Leave-One-Out (LOO) cross validation. A multiresolution scheme is used to produce four levels of detail for each structure. For femur bone, the four resolutions consist in $N = 514$, $N = 2056$, $N = 8224$ and $N = 32896$ vertices, respectively. For hip bone, they consist in $N = 814$, $N = 3256$, $N = 13024$ and $N = 52096$ vertices, respectively (see Fig. 7.9).

Three appearance models are compared: mean intensity profiles (PROF) with Normalized Cross Correlation measure, multivariate Gaussian distributions of intensity profiles (MGD) and multimodal profiles (MPAM), both with Mahalanobis distance. With MGD, PCA takes 95% of the total intensity into account to compute the approximated Mahalanobis distance. To regularize the covariance matrix during EM iterations, MPAM is created using *Constant Regularization* method coupled with parameter $h = 1.0$ (see section 4.3.1.3). Due to the limited number of training datasets, we prefer not to merge any mode (*i.e.* Jaccard index $\mathcal{J} = 1.0$).

Regarding intensity profiles, thirty-one intensities are sampled every 0.5 mm from 12.5 mm inside to 2.5 mm outside mesh surface at each vertex for PROF and MGD, since these values were successfully used for the segmentation of bones in MR images [Gilles & Magnenat-Thalmann 2010, Schmid & Magnenat-Thalmann 2008]. For MPAM, appropriate values need to be defined since the appearance prior construction method is different. Experiments showed that eleven intensities sampled every mm from 5 mm inside to 5 mm outside mesh surface give reasonable results in terms of clustering and segmentation. These values are thus used in the framework of this work.

To have a fair comparison between methods, same initialization and internal forces (*i.e.* PCA-based prior or shape memory) are used for every appearance model. The initialization is based on the manual placement of landmarks coupled with a shape interpolation approach [Gilles & Magnenat-Thalmann 2010]. Segmentation accuracy is assessed with DICE measure (DSC) [Dice 1945] based on reference manual segmentations. In the results, femur and hip bones are both considered as one structure, though they are both represented by two instances (*i.e.* left and right). To simplify the statistical analysis, the DICE measure on each structure is actually



#1 #2 #3

#4 #5 #6

Figure 7.8: The six datasets that are used for the experiments with their MR image and corresponding reference meshes (*i.e.* hip bones in blue and femur bones in green). The three datasets on the top constitute the D3 training set while all datasets together constitute the D6 training set. Note how rather linear is the intensity distribution between datasets of D3. Conversely, the three datasets on the bottom feature a quite heterogeneous intensity distribution. Among them, dataset #4 and dataset #6 are strongly subject to MRI artifacts.

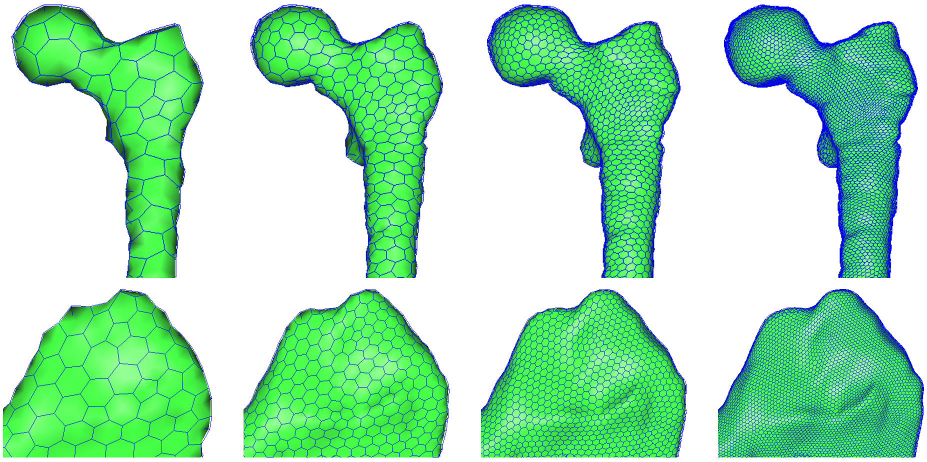


Figure 7.9: Detail of the four increasing resolutions (from left to right) for both femur (top) and hip (bottom) bone meshes. For femur bone, the four resolutions are $N = 514$, $N = 2056$, $N = 8224$ and $N = 32896$ vertices, respectively. For hip bone, the four resolutions are $N = 814$, $N = 3256$, $N = 13024$ and $N = 52096$ vertices, respectively.

D3	MGD	PROF	MPAM
SMEM	88.48	91.98	87.39
PCA	78.82	89.51	89.40

D6	MGD	PROF	MPAM
SMEM	87.64	90.91	87.19
PCA	70.37	89.14	88.08

Table 7.2: Mean DICE measure (*i.e.* on all structures and on all segmented datasets) when combining appearance and shape priors on both D3 and D6 training sets.

a mean on its both instances. Also, only datasets #1, #2 and #3 are segmented, since they feature a more homogeneous intensity distribution (see Figure 7.8).

7.3.6 Results

First results consist in averaging the DICE measure on all structures and on all segmented datasets (see Table 7.2). Using D3, PROF gives the best results, with both SMEM (DSC = 91.98) and PCA (DSC = 89.51), followed closely by MPAM with PCA (DSC = 89.40). Results are quite similar using D6, PROF performs the best with both SMEM (DSC = 90.91) and PCA (DSC = 89.14), followed by MPAM with PCA (DSC = 88.08). MGD coupled with PCA gives by far the worst results, when using both D3 and D6. Regardless of the appearance prior used, SMEM is more accurate than PCA except when the latter is coupled with MPAM. When comparing global results on D3 and D6, we notice that D3 always gives better results.

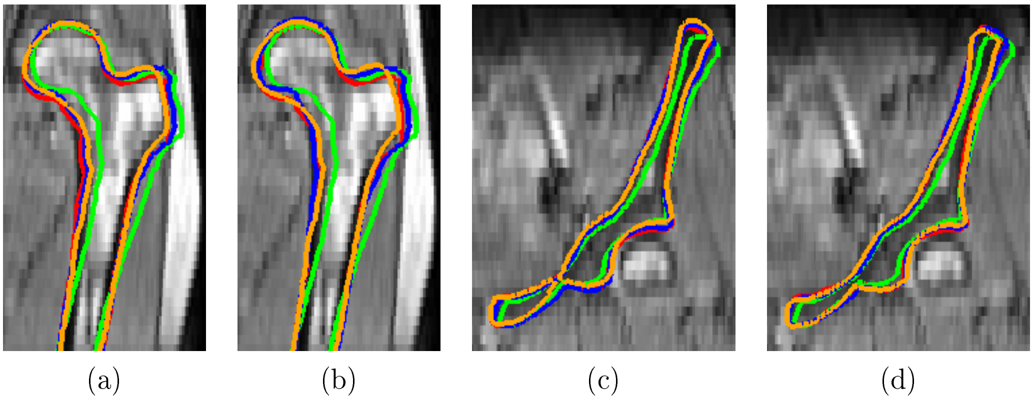

(a) (b) (c) (d)

Figure 7.10: Segmentation of both left femur {(a),(b)} and left hip {(c),(d)} bones from dataset #1 using D3 training set and: (a) PROF, (b) MPAM, (c) MGD, and (d) PROF appearance priors. Reference mesh is depicted in red, initialization in green, SMEM shape prior in blue and PCA shape prior in orange.

Femur	MGD	PROF	MPAM
SMEM	92.28 (0.63)	92.51 (1.36)	90.51 (0.51)
PCA	90.57 (1.36)	90.52 (0.91)	90.71 (0.90)
Hip	MGD	PROF	MPAM
SMEM	84.67 (9.84)	91.45 (2.46)	84.27 (6.58)
PCA	67.07 (30.6)	88.50 (2.07)	88.10 (3.22)

Table 7.3: Mean (SD) of DICE measure for each structure and on all segmented datasets when combining appearance and shape priors on D3 training set.

In a second step, we study in more details the results on the two separate structures (see some delineations in Figure 7.10). Using D3 (see Table 7.3), PROF gives the best results with SMEM for femur bone (DSC = 92.51). This combination performs also the best for hip bone (DSC = 91.45). Using D6 (see Table 7.4), MPAM coupled with PCA gives the most accurate segmentation for femur bone (DSC = 92.02). However, for hip bone, PROF with SMEM gives once again the best results (DSC = 90.09). A clear tendency shows that results for femur bone are clearly better than for hip bone, regardless of the training set, shape and appearance prior used.

When comparing the influence of shape priors on femur bone segmentation with D6, we notice very slight differences between appearance priors, except for MPAM. SMEM performs the best for hip bone segmentation though (with both D3 and D6), but once again except for MPAM. When comparing the influence of the training set, D3 gives better results on hip bone, regardless of the shape and appearance prior used. Conversely, D6 gives better results on femur bone when PCA is used with all appearance priors.

Femur	MGD	PROF	MPAM
SMEM	91.92 (0.89)	91.74 (1.72)	90.46 (0.37)
PCA	91.97 (1.42)	91.62 (1.71)	92.02 (0.96)
Hip	MGD	PROF	MPAM
SMEM	83.36 (5.24)	90.09 (2.22)	83.92 (5.65)
PCA	48.77 (33.1)	86.66 (2.29)	84.13 (7.20)

Table 7.4: Mean (SD) of DICE measure for each structure and on all segmented datasets when combining appearance and shape priors on D6 training set.

7.3.7 Discussion

Generating statistics from a segmentation framework using a small amount of training datasets is a quite challenging work and a difficult task. Indeed, statistics are usually relevant when numerous data are at disposal. However, results on both training sets D3 and D6 show some trends that we discuss in this section.

In general, using only three datasets with D3 gives better results than six datasets with D6. This may appear as counter-intuitive: more datasets should lead to better results. However, the three additional datasets from D6 are mostly corrupted by noise and artifacts. They also feature a quite heterogeneous intensity distribution. In this case, adding more information, which is corrupted, does not improve the appearance prior but rather weakens it. As a result, external forces have more difficulties to find the right boundaries.

In general, using shape memory with SMEM gives better results than PCA-based priors. This is because SMEM is a local approach. Given a shape that does not exhibit large inter-subject variation (*i.e.* like bones compared to soft organs) and that is sufficiently initialized, SMEM is a quite robust approach that can deal with few training datasets. However, PCA is known to give better results when using more datasets. In fact, a PCA-based evolution tends to be less sensitive to initialization when a sufficient number of datasets is available to provide a richer information about the shape [Schmid *et al.* 2010]. This is explained by the fact that PCA affects all the shape vertices in a global manner compared to the shape memory force, which confers a better robustness against local artifacts.

As depicted in both Table 7.3 and Table 7.4, MGD coupled with PCA gives very bad results when segmenting hip bone (DSC = 67.07 with D3 and 48.77 with D6). This would suggest that this combination is the worst: PCA as a shape prior (in case of few training datasets) and MGD as an appearance prior. External forces generated by MGD are thus less efficient than those generated by PROF and MPAM. This would explain the huge difference of DSC between D3 and D6, knowing that D6 contains datasets with images corrupted by noise and artifacts. As a result, those external forces are more sensitive to internal forces. But since internal forces based on PCA are also weak due to the few number of datasets, the segmentation is doomed to give bad results.

Regarding MPAM, the best results are for femur bone segmentation when combined with PCA and when using D6. Though in theory MPAM should need less datasets than MGD (*i.e.* a PCA-based appearance method), MPAM seems to feature a certain sensitivity to noise. Indeed, results for femur bone are clearly better than for hip bone (when using both D3 and D6). This is because intensity distribution is more heterogeneous at hip level, as hip bones are located close to image top border. This would suggest that the similarity measure (*i.e.* the Mahalanobis distance) should be optimized to cope with non linear intensity. Moreover, the intensity profile computation is different for MPAM than for PROF and MGD. Additional experiments should be performed to determine an optimal intensity profile length for the segmentation of bones in MR images.

In general, PROF gives the best results (*i.e.* compared to MGD and MPAM). We believe there are two explanations for that. First, reference meshes are produced by a supervised approach that uses a semi-automatic segmentation controlled by manually defined constraints [Schmid *et al.* 2010]. This semi-automatic segmentation exploits the same deformable model evolution coupled with the NCC similarity measure reported in this section. This creates a bias, which is likely to give an advantage to PROF-based appearance prior (*i.e.* PROF also uses NCC during the segmentation). Second, NCC similarity measure is known to be more robust to intensity change (*i.e.* NCC is affine invariant). Using NCC would thus help PROF to be more efficient in presence of intensity change (*i.e.* especially with the three additional datasets corrupted by noise and artifacts that are used with D6).

7.4 Conclusion

In this chapter, we presented two collaborative works that have been done within the framework of the 3D Anatomical Human project (3DAH).

First, we presented a complete pipeline describing how to create subject-specific models for kinematic simulations, from image and motion data acquisitions to anatomical model reconstruction and musculoskeletal model generation. This work is astride two fields, namely medical image analysis and biomechanics, and is the result of the collaboration with both University of Geneva and EPFL. This pipeline is meant to be a realistic proposition on how to create subject-specific models for kinematic simulations. Results are quite preliminary though, and a lot of work remains to fully validate the framework. The objective is to use this framework in a medical context where some specific motions on pathological subjects would be examined, with the aim of giving some highlights into the mechanisms of pathologies and their treatments.

Then, we compared the performance of different statistical models in the context of lower limb bone segmentation using MR images when only a small number of datasets is available for training. This work is the result of a collaboration with University of Geneva. Results tend to show that local and simple methods perform the best. As a shape prior, shape memory (SMEM) gives very good results. As an appearance prior, mean intensity profiles (PROF) gives the best results. We believe these good results are partly due to the Normalized Cross Correlation (NCC) similarity measure, which is more robust to intensity change in MR images. When increasing the number of training datasets, results tend to show that statistical methods feature the best results: PCA-based shape prior (PCA) and multimodal profiles (MPAM) as an appearance prior. Both methods capture more and more information while the number of training datasets increases and we believe that better segmentation results would be obtained when increasing this number of datasets. However, this hypothesis only holds when the training data is of enough quality to produce meaningful and efficient priors.

Chapter 8
Conclusion

Contents

8.1 Contributions

8.1.1 Main contributions

Multimodal Prior Appearance Model In this thesis, we propose a novel appearance prior, denoted as *Multimodal Prior Appearance Model (MPAM)*, for the segmentation of anatomical structures using explicit models (*i.e.* 2-simplex meshes in this thesis). MPAM is first built upon an EM clustering of intensity profiles with model order selection to automatically select the number of profile classes, or *modes*, in an attempt to study the appearance around anatomical structures.

Unlike classical PCA-based Appearance Priors (PCAP), the clustering is considered as *regional* because intensity profiles are classified for each mesh, and not for each vertex (*i.e.* over a population of meshes). With the MPAM, each point may be associated with several profile modes and each profile mode is estimated on each subject and not across subjects. All mesh instances are then projected on a common reference in which similar modes are possibly merged in an attempt to be more efficient and to reduce the complexity of the MPAM. Segmentation results from comparative tests in chapter 6 show that our MPAM outperforms PCAP despite the fact that less profile modes are used.

***OSI* criterion** Selecting the number of modes from intensity profiles in an automatic fashion is a quite difficult task. Classical model order selection criteria, such as *AIC* and *BIC* give unsatisfactory results because they are affected by the high dimensional space featured by intensity profiles.

We propose the *OSI* criterion, which is solely based on the EM posterior probabilities and penalizes the overlap between modes while encouraging their separation.

Spectral clustering using Gray code To overcome several issues featured by the classical EM algorithm (*i.e.* lack of robustness when changing the resolution, local maxima and initialization), we propose to use spectral clustering. Instead of directly classifying original data points, they are transformed into a graph that takes similarity between all pairs of data points into account, which leads to a similarity matrix. Top eigenvectors derived from this similarity matrix are used as input data points by the EM algorithm.

We propose to initialize centers with the Gray code to ensure them to be orthogonal from each other. This is a required step for the clustering of intensity profiles. Indeed, due to their nature (*i.e.* consequent overlap between profiles and not enough profiles with respect to the high dimensional space), spectral clustering may end up with similar eigenvalues (and thus similar eigenvectors), leading to similar centers.

Boosted clustering We propose the boosted clustering to optimize the classification of profiles so that generated external forces are optimal for segmentation purposes. A localization criterion, which may be used with three similarity measures, decides whether the clustering associated to a vertex is able to localize well the boundary along the profile. If yes, the corresponding profile is kept. Otherwise, the corresponding profile is rejected. We also propose a cascading boosted clustering, which classifies rejected profiles in an iterative way in an attempt to increase the number of kept profiles. At the end of this iterative procedure, if some profiles are still not able to localize well the boundary, associated vertices are considered as outliers. The objective is to avoid having misguided deformations during segmentation. We finally present the cascading boosted clustering with hierarchical approach, which gives the best results in terms of percentage of kept profiles. This last version of the boosted clustering has the advantage of enforcing a minimum mode size.

Subject-specific anatomical models for kinematic simulations Within the framework of the 3D Anatomical Human project, we present a complete pipeline for kinematic simulations in collaboration with our partners from both University of Geneva and EPFL. All steps are described, from image and motion data acquisitions to anatomical model reconstruction and musculoskeletal model generation.

Our contributions stand in a pipeline of software tools for the generation of subject-specific anatomical models from MR images of the lower limbs.

This pipeline includes:

- solutions to correct images from MRI acquisitions (*i.e.* intensity scaling and normalization using histogram equalization).
- a smoothing procedure for generated meshes using simplex mesh topological properties.

- a semi-automatic method to segment tendons using both an automatic method based on maximum intensity ridge tracing and a manual placement of missing centerline points.
- a post-processing method for interpenetration removal.

8.1.2 Other contributions

Covariance matrix regularization The classification of intensity profiles with the EM algorithm may lead to singularities due to the inversion of covariance matrices in the E-step. This is because of the coarse sampling of meshes combined with profiles featuring a high dimensional space (*i.e.* curse of dimensionality).

To overcome this issue, we propose three regularization approaches for the covariance matrix:

- Spectral Regularization, which uses a PCA-based approach.
- Diagonal Regularization, which diagonalizes the covariance matrix.
- Constant Regularization, which uses the trace of the covariance matrix.

The choice of the covariance matrix regularization method depends on the nature of the data. The proposed methods may be used on any data based on feature vectors presenting the same features (*i.e.* consequent overlap between data points and not enough data points to represent a high dimensional space).

Spatial regularization during classification Since EM algorithm does not take the neighborhood information of profiles into account (*i.e.* every data point is given an independent posterior probability to belong to modes, which is not correlated with its neighbors), the resulting clustering leads to non smooth posterior probability maps, which may impair the fusion of appearance regions.

To have a spatial regularization of EM classification, we present a spatial regularization approach based on the Neighborhood EM algorithm. This approach needs a function to compute the spatial distance between data points and we propose to use the topological properties of simplex vertices coupled with the correlation coefficient. Each vertex has exactly three neighbors, and thus three non-zero neighboring values to take into account, which substantially speeds-up the computation.

Similarly, when performing spectral clustering on profiles, we propose to add a neighborhood weight on the computation of the affinity matrix. Should two vertices be neighbors on the simplex mesh, we decrease the pairwise distance of their corresponding profiles, which leads to an increased local neighborhood.

Projection of modes into a reference mesh To build the MPAM, appearance information from several datasets needs to be projected into a common reference framework. For that, a geometric embedding must be provided so that the classification of profiles from several datasets are registered into a reference mesh, on which each vertex has a probability to belong to several modes.

First, we have proposed to register the mesh of each dataset towards a binary image with globally-constrained deformations. Due to the required manual interventions and inaccuracy of the method, we rather propose to use a mesh-to-mesh registration based on currents that gives very good results in terms of accuracy. After the mesh registration, we propose to register the classification from several datasets into the reference mesh using an interpolation based on a closest point approach. To ensure that the appearance information from every vertex is taken into account, we suggest reference mesh to be finer that the finest mesh from training datasets. Since the classification from datasets is projected into a reference mesh, meshes from the datasets may keep their original resolution. This means that no point correspondence needs to be ensured. These are two important advantages of our proposed projection, and thus of our MPAM, compared to PCA-based methods.

Fusion of modes Among the modes from different datasets that are projected into the common reference mesh, some of them are expected to be similar (*i.e.* tissues around the same structure in different images are expected to be mainly the same) and could be thus merged in an attempt to reduce the final number of modes. This fusion step is not a requirement though, as we could just keep all the modes from the datasets and add them directly to the MPAM. However, a mode fusion has the advantage to reduce the complexity of the MPAM, by making sure that every mode is as unique as possible.

We propose to use the Jaccard index to control the final number of modes associated with the MPAM. Each mode is represented by a specific region spanned by its mean and standard deviation. The Jaccard index measures the similarity between two modes by measuring the overlap between their respective region.

8.1.3 Publications

1. Journal and book chapter

    - **François Chung**, Jérôme Schmid, Nadia Magnenat-Thalmann and Hervé Delingette. Comparison of statistical models performance in case of segmentation using a small amount of training datasets. **The Visual Computer**, volume 27, issue 2, pages 141-151, February 2011.
    - Jérôme Schmid, Anders Sandholm, **François Chung**, Daniel Thalmann, Hervé Delingette and Nadia Magnenat-Thalmann. Musculoskeletal simulation model generation from MRI datasets and motion capture data. In **Recent Advances in the 3D Physiological Human**, pages 3-19. Springer, February 2009.

2. International conference

   - **François Chung** and Hervé Delingette. Multimodal Prior Appearance Models based on Regional Clustering of Intensity Profiles. In Guang-Zhong Yang, David Hawkes, Daniel Rueckert, Alison Noble, and Chris Taylor, editors, **MICCAI 2009 - Medical Image Computing and Computer-Assisted Intervention**, Part II, volume 5762 of Lecture Notes in Computer Science, London, UK, pages 1051-1058, September 2009. Springer.
   - Tobias Heimann, **François Chung**, Hans Lamecker and Hervé Delingette. Subject-Specific Ligament Models: Towards Real-Time Simulation of the Knee Joint. In **Computational Biomechanics for Medicine IV Workshop 2009**, September 2009.

3. National conference

   - **François Chung**, Jérôme Schmid, Olivier Clatz, Nadia Magnenat-Thalmann and Hervé Delingette. Reconstruction 3D des structures anatomiques des membres inférieurs. In **ORASIS'09**, June 2009.

8.2 Perspectives

Segmentation using MPAM combined with robust shape priors Our proposed MPAM is aimed at improving the segmentation of medical images. However, we compared its performance with PCA-based Appearance Models (PCAP) using only external forces. Indeed, for our tests in chapter 6, internal forces were generated using simple shape priors (*i.e.* shape memory associated with simplex meshes). The next step of this work would be to test our proposed MPAM when combined with robust shape priors, *e.g.* using a statistical analysis of shape based on PCA [Heimann 2009, Schmid & Magnenat-Thalmann 2008].

Intensity profiles Though we have studied a series of profile length and inward/outward ratio for the liver segmentation, these values may be extended to find better combinations. Also, for computational time, we have decided to keep a sampling step of 1 mm. Changing this step may give better results. Since appearance is clearly different between the inner part of the liver and its neighboring structures, a different step between the inside and outside could be foreseen. Also, other types of profiles are still to be tested, such as gradient profiles, normalized intensity profiles, and normalized gradient profiles. Finally, the feature vector *per se* could be changed, *e.g.* using other intensity features or texture features.

Spatial regularization Though the spatial regularization approach based on both the NEM algorithm (for EM clustering) and the neighborhood weight on the computation of the affinity matrix (for spectral clustering) gives interesting results,

results may be improved. In both cases, we use the direct neighborhood local at each vertex to spatially regularize the classification of profiles (*i.e.* using the three neighbors). This neighborhood may be increased to improve the spatial regularization in a multi-resolution fashion. To do so, the vertex neighborhood, which is a specific topological property of simplex mesh, could be increased/decreased in an iterative procedure.

Another improvement could be to change the similarity measure (for NEM algorithm) and the similarity function (for spectral clustering). For NEM algorithm, other similarity measures could be tested, such as (normalized) cross-correlation, sum of squared intensity differences and ratio image uniformity. For spectral clustering, other similarity functions from the literature could be tested [Balcan *et al.* 2007]. Generally speaking, it would be interesting to compare our proposed methods with Markov Random Fields (MRF), which are a more common approach for spatial regularization.

Fusion of modes The fusion of modes using the Jaccard index aims at reducing the complexity of the MPAM. Though modes may be easily merged by changing the threshold on Jaccard index, our experiments show that segmentation results are affected by non-restrictive thresholds. This is because final modes created by the fusion of multiple modes have an increased variance, leading to less discriminant modes during local search. To overcome this issue, the increased variance featured by merged modes should be somehow compensated by other discriminant features. Otherwise, other fusion indices could be foreseen to replace the Jaccard index. In any case, the objective should be the same: reducing the final number of modes while keeping/improving the segmentation results when no modes are merged into the MPAM.

Anatomical structures We have decided to test our regional modeling approach on livers from CT images thanks to the free availability of two liver datasets on Internet. However, our approach could be tested on any other anatomical structures from CT images, such as prostate, spleen, kidney and lungs. From MR images, we have not thoroughly tested our approach for the segmentation of lower limb structures (*i.e.* without significant statistical results), as we did not have enough datasets at disposal. It would be interesting to see how our approach performs for the segmentation of lower limb structures in MR images though. Generally speaking, testing our approach on other anatomical structures from other modalities would give more insight about the robustness of our approach.

List of Abbreviations

AAM	Active Appearance Models
AIC	Akaike Information Criterion
AICc	second-order AIC
ASM	Active Shape Models
ASSD	Average Symmetric Surface Distance
BIC	Bayesian Information Criterion
CT	Computed Tomography
EM	Expectation-Maximization
FCM	Fuzzy C-Means
FFD	Free Form Deformations
FLAIR	Fluid Attenuated Inversion Recovery
FOV	field of view
FVQ	Fuzzy Vector Quantization index
GCD	Globally Constrained Deformations
GMM	Gaussian Mixture Model
GPA	General Procrustes Analysis
LOD	level of detail
MPAM	Multimodal Prior Appearance Model
MRF	Markov Random Fields
MRI	Magnetic Resonance Imaging
MSD	Maximum Surface Distance
NCC	Normalized Cross Correlation
NEM	Neighborhood EM
OSI	Overlap Separation Index
PCA	Principal Component Analysis
PCAP	PCA-based Appearance Prior
PDM	Point Distribution Models
RF	radio frequency
SAM	Statistical Appearance Models
SD	standard deviation
SRVD	Signed Relative Volume Difference
SSM	Statistical Shape Models
SVD	Singular Value Decomposition
TE	echo time
TR	repetition time
VOE	Volumetric Overlap Error

List of Notations

$\mathbf{p}_i$	i-th vertex
$\mathbf{x}_i$	intensity profile at vertex $\mathbf{p}_i$
$\mathbf{y}_i$	current intensity profile sampled at vertex $\mathbf{p}_i$ during local search
$\mathbf{n}_i$	normal at vertex $\mathbf{p}_i$
$\mathbf{b}_i$	binormal vector at vertex $\mathbf{p}_i$
$\mathbf{t}_i$	tangent at vertex $\mathbf{p}_i$
φ_i	simplex angle at vertex $\mathbf{p}_i$
α	weight to control the strength of the internal forces
α_i	optimal offset during local search at vertex $\mathbf{p}_i$
β	weight to control the strength of the external forces
β_n	term weighting the spatial homogeneity in NEM algorithm
v_{ij}	similarity value between neighboring data points in NEM algorithm
V	neighborhood matrix in NEM algorithm
h	parameter for the regularization of EM covariance matrices
η	η-th order neighborhood associated with 2-simplex meshes
$\mathcal{J}$	Jaccard index
$\mathcal{R}$	ratio on eigenvalues
N	number of intensity profiles (*i.e.* data points)
d	number of dimensions featured by intensity profiles
K	number of modes (*i.e.* intensity profile classes)
P	number of datasets
$\mathcal{M}_p$	mesh from the p-th dataset
$\mathcal{M}^\star$	reference mesh
$\mathcal{M}^\star_j$	j-th resolution of reference mesh $\mathcal{M}^\star$
M^p_k	k-th mode associated with the p-th dataset
$\boldsymbol{\mu}^k_p$	mean of mode M_k from mesh $\mathcal{M}_p$
$\boldsymbol{\Sigma}^k_p$	covariance matrix of mode M_k from mesh $\mathcal{M}_p$
π^k_p	mixing coefficient of mode M_k from mesh $\mathcal{M}_p$
${}^p\gamma^k_i$	posterior probability of $\mathbf{x}_i$ from mesh $\mathcal{M}_p$ to belong to mode M_k

List of Algorithms

List of Figures

List of Tables

Bibliography

[Acharya & Ray 2005] Tinku Acharya and Ajoy K. Ray. Image processing - principles and applications. Wiley-Interscience, September 2005. (Cited on page 73.)

[Akaike 1974] Hirotsugu Akaike. *A new look at the statistical model identification.* IEEE Transactions on Automatic Control, vol. 19, no. 6, pages 716–723, December 1974. (Cited on page 67.)

[Alexa & Adamson 2004] Marc Alexa and Anders Adamson. *On normals and projection operators for surfaces defined by point sets.* In Proceedings of the Eurographics Symposium on Point-based Graphics 2004, pages 149–156, June 2004. (Cited on page 144.)

[Alpert & Yao 1994] Charles J. Alpert and So-Zen Yao. *Spectral partitioning: The more eigenvectors, the better.* In Proceedings of the 32nd annual ACM/IEEE Design Automation Conference, pages 195–200, 1994. (Cited on page 88.)

[Ambroise *et al.* 1997] Christophe Ambroise, Mo Dang and Gérard Govaert. *Clustering of spatial data by the EM algorithm.* Quantitative Geology and Geostatistics, vol. 9, pages 493–504, 1997. (Cited on pages 60 and 71.)

[Ambrosio & Soner 1996] Luigi Ambrosio and Halil Mete Soner. *Level set approach to mean curvature flow in arbitrary codimension.* Journal of Differential Geometry, vol. 43, no. 4, pages 693–737, 1996. (Cited on page 35.)

[Aylward & Bullitt 2002] Stephen R. Aylward and Elizabeth Bullitt. *Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction.* IEEE Transactions on Medical Imaging, vol. 21, no. 2, pages 61–75, February 2002. (Cited on page 143.)

[Balcan *et al.* 2007] Maria-Florina Balcan, Avrim Blum and Santosh Vempala. *A theory of similarity functions for clustering.* Technical report, Carnegie Mellon University, July 2007. (Cited on pages 95 and 166.)

[Bandyopadhyay & Maulik 2002a] Sanghamitra Bandyopadhyay and Ujjwal Maulik. *An evolutionary technique based on K-Means algorithm for optimal clustering in RN.* Information Sciences—Applications: An International Journal, vol. 146, no. 1, pages 221–237, 2002. (Cited on page 69.)

[Bandyopadhyay & Maulik 2002b] Sanghamitra Bandyopadhyay and Ujjwal Maulik. *Genetic clustering for automatic evolution of clusters and application to image classification.* Pattern Recognition, vol. 35, no. 6, pages 1197–1208, June 2002. (Cited on page 69.)

[Bandyopadhyay & Pal 2007] Sanghamitra Bandyopadhyay and Sankar Kumar Pal. Classification and learning using genetic algorithms: Applications in bioinformatics and web intelligence (Natural Computing Series). Springer, June 2007. (Cited on page 69.)

[Bandyopadhyay *et al.* 1998] Sanghamitra Bandyopadhyay, C.A. Murthy and Sankar K. Pal. *Pattern classification using genetic algorithms: Determination of H.* Pattern Recognition Letters, vol. 19, no. 13, pages 1171–1181, November 1998. (Cited on page 69.)

[Barr 1981] Alan H. Barr. *Superquadrics and angle-preserving transformations.* IEEE Computer Graphics and Applications, vol. 1, no. 1, pages 11–23, January 1981. (Cited on page 23.)

[Behiels *et al.* 2002] Gert Behiels, Frederik Maes, Dirk Vandermeulen and Paul Suetens. *Evaluation of image features and search strategies for segmentation of bone structures in radiographs using Active Shape Models.* Medical Image Analysis, vol. 6, no. 1, pages 47–62, March 2002. (Cited on page 44.)

[Bezdek 1973] James C. Bezdek. *Fuzzy mathematics in pattern classification.* Phd thesis, Center for Applied Mathematics, Cornell University, 1973. (Cited on page 20.)

[Bezdek 1974a] James C. Bezdek. *Cluster validity with fuzzy sets.* Journal of Cybernetics and Systems, vol. 3, no. 3, pages 58–73, 1974. (Cited on page 65.)

[Bezdek 1974b] James C. Bezdek. *Numerical taxonomy with fuzzy sets.* Journal of Mathematical Biology, vol. 1, no. 1, pages 57–71, May 1974. (Cited on page 65.)

[Bezdek 1981] James C. Bezdek. Pattern recognition with fuzzy objective function algorithms. Plenum Press, January 1981. (Cited on page 54.)

[Bigün *et al.* 1991] Josef Bigün, Gösta H. Granlund and Johan Wiklund. *Multidimensional orientation estimation with applications to texture analysis and optical flow.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 13, no. 8, pages 775–790, August 1991. (Cited on page 42.)

[Bishop 1995] Christopher M. Bishop. Neural networks for pattern recognition. Oxford University Press, November 1995. (Cited on page 67.)

[Bishop 2007] Christopher M. Bishop. Pattern recognition and machine learning. Springer-Verlag, October 2007. (Cited on pages 63 and 104.)

[Blum 2007] Avrim Blum. *A theory of similarity functions for learning and clustering.* In DS'07 - Proceedings of the 10th International Conference on Discovery Science, pages 39–39, October 2007. (Cited on page 95.)

[Bondiau *et al.* 2005] Pierre-Yves Bondiau, Grégoire Malandain, Stéphane Chanalet, Pierre-Yves Marcy, Jean-Louis Habrand, Francois Fauchon, Philippe Paquis, Adel Courdi, Olivier Commowick, Isabelle Rutten and Nicholas Ayache. *Atlas-based automatic segmentation of MR images: Validation study on the brainstem in radiotherapy context.* International Journal of Radiation Oncology*Biology*Physics, vol. 61, no. 1, pages 289–298, January 2005. (Cited on page 22.)

[Borsboom *et al.* 2003] Denny Borsboom, Gideon J. Mellenbergh and Jaap Van Heerden. *The theoretical status of latent variables.* Psychological Review, vol. 110, no. 2, pages 203–219, April 2003. (Cited on page 57.)

[Boudraa 1999] Abdel-Ouahab Boudraa. *Dynamic estimation of number of clusters in data sets.* Electronics Letters, vol. 35, no. 19, pages 1606–1608, September 1999. (Cited on page 65.)

[Brejl & Sonka 2000] Marek Brejl and Milan Sonka. *Object localization and border detection criteria design in edge-based image segmentation: Automated learning from examples.* IEEE Transactions on Medical Imaging, vol. 19, no. 10, pages 973–985, October 2000. (Cited on pages 43 and 44.)

[Brox & Weickert 2004] Thomas Brox and Joachim Weickert. *A TV Flow based local scale measure for texture discrimination.* In ECCV 2004 - Proceedings of the 8th European Conference on Computer Vision, volume 3022 of *Lecture Notes in Computer Science*, pages 578–590. Springer Berlin / Heidelberg, May 2004. (Cited on page 42.)

[Buades *et al.* 2005] Antoni Buades, Bartomeu Coll and Jean-Michel Morel. *A non-local algorithm for image denoising.* In CVPR '05 - Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, volume 2, pages 60–65, June 2005. (Cited on page 141.)

[Bugeau & Pérez 2007] Aurélie Bugeau and Patrick Pérez. *Bandwidth selection for kernel estimation in mixed multi-dimensional spaces.* Rapport de Recherche 6286, INRIA - Institut National de Recherche en Informatique et Automatique, September 2007. (Cited on page 68.)

[Burnham & Anderson 1998] Kenneth P. Burnham and David R. Anderson. Model selection and multi-model inference: A practical information-theoretic approach. Springer, October 1998. (Cited on page 68.)

[Buzug & Weese 1998] T.M. Buzug and J. Weese. *Voxel-based similarity measures for medical image registration in radiological diagnosis and image guided surgery.* Journal of Computing and Information Technology, vol. 6, no. 2, pages 165–179, 1998. (Cited on page 45.)

[Celeux & Govaert 1995] Gilles Celeux and Gérard Govaert. *Gaussian parsimonious clustering models.* Pattern Recognition, vol. 28, no. 5, pages 781–793, May 1995. (Cited on page 57.)

[Chuang *et al.* 2006] Keh-Shih Chuang, Hong-Long Tzeng, Sharon Chen, Jay Wu and Tzong-Jer Chen. *Fuzzy C-Means clustering with spatial information for image segmentation.* Computerized Medical Imaging and Graphics, vol. 30, no. 1, pages 9–15, January 2006. (Cited on page 60.)

[Chung & Delingette 2009] François Chung and Hervé Delingette. *Multimodal prior appearance models based on regional clustering of intensity profiles.* In MICCAI 2009 - Proceedings of the 12th International Conference on Medical Image Computing and Computer Assisted Intervention, volume 5762 of *Lecture Notes in Computer Science*, pages 1051–1058, September 2009. (Cited on pages 52, 56 and 64.)

[Chung *et al.* 2009] François Chung, Jérôme Schmid, Olivier Clatz, Nadia Magnenat-Thalmann and Hervé Delingette. *Reconstruction 3D des structures anatomiques des membres inférieurs.* In ORASIS'09, June 2009. (Cited on page 142.)

[Chung *et al.* 2011] François Chung, Jérôme Schmid, Nadia Magnenat-Thalmann and Hervé Delingette. *Comparison of statistical models performance in case of segmentation using a small amount of training datasets.* The Visual Computer, vol. 27, no. 2, pages 141–151, February 2011. (Cited on pages 139 and 147.)

[Chung 1997] Fan R. K. Chung. Spectral graph theory (CBMS Regional Conference Series in Mathematics, No. 92). American Mathematical Society, February 1997. (Cited on pages 97 and 98.)

[Claussen 1985] Martin Claussen. *Estimation of the Monin-Obukhov similarity functions from a spectral model.* Boundary-Layer Meteorology, vol. 33, no. 3, pages 233–243, November 1985. (Cited on page 95.)

[Cocquerez & Philipp 1995] Jean-Pierre Cocquerez and Sylvie Philipp. Analyse d'images: Filtrage et segmentation. Masson, 1995. (Cited on page 42.)

[Cohen & Cohen 1993] Laurent D. Cohen and Isaac Cohen. *Finite Element Methods for Active Contour Models and balloons for 2D and 3D images.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, no. 11, pages 1131–1147, November 1993. (Cited on page 42.)

[Cohen 1991] Laurent D. Cohen. *On active contour models and balloons.* Computer Vision, Graphics and Image Processing : Image Understanding, vol. 53, no. 2, pages 211–218, March 1991. (Cited on pages 24, 35 and 36.)

[Cohen 1992] Isaac Cohen. *Modèles déformables 2D et 3D, application à la segmentation d'images médicales.* Phd thesis, Université Paris-IX Dauphine, February 1992. (Cited on page 41.)

[Collier *et al.* 2003] Dawn C. Collier, Stuart S. C. Burnett, Mayankkumar Amin, Stephen Bilton, Christopher Brooks, Amanda Ryan, Dominique Roniger, Danny Tran and George Starkschall. *Assessment of consistency in contouring of normal-tissue anatomic structures.* Journal of Applied Clinical Medical Physics, vol. 4, no. 1, pages 781–793, Winter 2003. (Cited on page 20.)

[Comaniciu & Meer 2002] Dorin Comaniciu and Peter Meer. *Mean Shift: A robust approach toward feature space analysis.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 5, pages 603–619, May 2002. (Cited on page 68.)

[Commowick 2007] Olivier Commowick. *Création et utilisation d'atlas anatomiques numériques pour la radiothérapie (Design and use of anatomical atlases for radiotherapy).* Phd thesis, Université de Nice Sophia-Antipolis, February 2007. (Cited on page 22.)

[Cootes & Taylor 1993] Timothy F. Cootes and Christopher J. Taylor. *Active Shape Model search using local grey-level models: A quantitative evaluation.* In BMVC 1993 - Proceedings of the 4th British Machine Vision Conference, pages 639–648. BMVA Press, 1993. (Cited on pages 24, 39, 40, 44, 47, 146 and 152.)

[Cootes & Taylor 1994] Timothy F. Cootes and Christopher J. Taylor. *Using grey-level models to improve Active Shape Model search.* In ICPR 1994 - Proceedings of the 12th IAPR International Conference on Pattern Recognition, volume 1, pages 63–67, October 1994. (Cited on pages 1, 43, 44 and 51.)

[Cootes & Taylor 1996] Timothy F. Cootes and Christopher J. Taylor. *Data driven refinement of Active Shape Model search.* In BMVC 1996 - Proceedings of the 7th British Machine Vision Conference, 1996. (Cited on page 147.)

[Cootes & Taylor 2001] Timothy F. Cootes and Christopher J. Taylor. *Constrained Active Appearance Models.* In ICCV-01 - Proceedings of the 8th International Conference on Computer Vision, volume 1, pages 484–498, July 2001. (Cited on pages 24, 39 and 146.)

[Cootes & Taylor 2004] Timothy F. Cootes and Christopher J. Taylor. *Statistical models of appearance for computer vision.* Technical report, University of Manchester, March 2004. (Cited on pages 39 and 146.)

[Cootes *et al.* 1993] Timothy F. Cootes, Andrew Hill, Christopher J. Taylor and Jane Haslam. *The use of Active Shape Models for locating structures in medical images.* In IPMI '93 - Proceedings of the 13th International Conference on Information Processing in Medical Imaging, pages 33–47. Springer-Verlag, June 1993. (Cited on pages 20, 43 and 44.)

[Costa *et al.* 2007] Jimena Costa, Hervé Delingette and Nicholas Ayache. *Automatic segmentation of the bladder using deformable models.* In ISBI'07 - Proceedings of the 4th IEEE International Symposium on Biomedical Imaging, pages 904–907, April 2007. (Cited on page 24.)

[Costa 2008] Jimena Costa. *Segmentation of anatomical structures of the lower abdomen using 3D deformable models.* Phd thesis, École Nationale Supérieure des Mines de Paris, March 2008. (Cited on page 22.)

[Coupé *et al.* 2006] Pierrick Coupé, Pierre Yger and Christian Barillot. *Fast non local means denoising for 3D MR images.* In MICCAI'06 - 9th International Conference on Medical Image Computing and Computer-Assisted Intervention, volume 4191 of *LNCS*, pages 33–40, October 2006. (Cited on page 141.)

[Cremers *et al.* 2007] Daniel Cremers, Mikael Rousson and Rachid Deriche. *A review of statistical approaches to Level Set segmentation: Integrating color, texture, motion and shape.* International Journal of Computer Vision, vol. 72, no. 2, pages 195–215, April 2007. (Cited on pages 24 and 42.)

[Cristinacce & Cootes 2008] David Cristinacce and Tim Cootes. *Automatic feature localisation with constrained local models.* Pattern Recognition, vol. 41, no. 10, pages 3054–3067, October 2008. (Cited on page 1.)

[Csiszár & Tusnády 1984] Imre Csiszár and Ge Tusnády. *Information geometry and alternating minimization procedures.* Statistics and Decisions, vol. 1, no. 1, pages 205–237, 1984. (Cited on page 59.)

[Dam *et al.* 2008] Erik B. Dam, P. Thomas Fletcher and Stephen M. Pizer. *Automatic shape model building based on principal geodesic analysis bootstrapping.* Medical Image Analysis, vol. 12, no. 2, pages 136–151, 2008. (Cited on page 52.)

[Dang & Govaert 1998] Mo Dang and Gérard Govaert. *Spatial fuzzy clustering using EM and Markov Random Fields.* International Journal of System Research and Information Science, pages 183–202, 1998. (Cited on page 71.)

[Daugman 1988] John G. Daugman. *Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression.* IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 36, no. 7, pages 1169–1179, July 1988. (Cited on page 1.)

[Dawant *et al.* 1999] Benoit M. Dawant, Steven L. Hartmann, Jean-Philippe Thirion, Frederik Maes, Dirk Vandermeulen and Philippe Demaerel. *Automatic 3-D segmentation of internal structures of the head in MR images using a combination of similarity and Free-Form Transformations: Part I,* *methodology and validation on normal subjects*. IEEE Transactions on Medical Imaging, vol. 10, no. 18, pages 909–916, October 1999. (Cited on page 22.)

[Delingette *et al.* 2006] Hervé Delingette, Xavier Pennec, Luc Soler, Jacques Marescaux and Nicholas Ayache. *Computational models for image-guided robot-assisted and simulated medical interventions*. Proceedings of the IEEE, vol. 94, no. 9, pages 1678–1688, September 2006. (Cited on page 19.)

[Delingette 1994] Hervé Delingette. *Modélisation, déformation et reconnaissance d'objets tridimensionnels à l'aide de maillages simplexes*. Thèse de sciences, Ecole Centrale de Paris, July 1994. (Cited on page 38.)

[Delingette 1998] Hervé Delingette. *Initialization of deformable models from 3D data*. In ICCV'98 - Proceedings of the 6th International Conference on Computer Vision, pages 311–316, January 1998. (Cited on page 34.)

[Delingette 1999] Hervé Delingette. *General object reconstruction based on simplex meshes*. International Journal of Computer Vision, vol. 32, pages 111–146, September 1999. (Cited on pages 20, 24, 25, 31, 41, 71, 142 and 150.)

[Dempster *et al.* 1977] Arthur P. Dempster, Nan M. Laird and Donald B. Rubin. *Maximum likelihood from incomplete data via the EM algorithm*. Journal of the Royal Statistical Society, vol. 39, pages 1–38, 1977. (Cited on pages 54 and 56.)

[Dice 1945] Lee R. Dice. *Measures of the amount of ecologic association between species*. Ecology, vol. 26, no. 3, pages 297–302, July 1945. (Cited on page 153.)

[Dindoyal *et al.* 2007] Irving Dindoyal, Tryphon Lambrou, Jing Deng and Andrew Todd-Pokropek. *Level Set Snake algorithms on the fetal heart*. In ISBI 2007 - 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2007, pages 864–867, April 2007. (Cited on page 24.)

[Donath & Hoffman 1973] Wilm E. Donath and Alan J. Hoffman. *Lower bounds for the partitioning of graphs*. IBM Journal of Research and Development, vol. 17, no. 5, pages 420–425, September 1973. (Cited on page 88.)

[Dunn 1973] J. C. Dunn. *A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters*. International Journal of Cybernetics and Systems, vol. 3, pages 32–57, 1973. (Cited on page 54.)

[Durrleman *et al.* 2009] Stanley Durrleman, Xavier Pennec, Alain Trouvé and Nicholas Ayache. *Statistical models on sets of curves and surfaces based on currents*. Medical Image Analysis, vol. 13, no. 5, pages 793–808, October 2009. (Cited on page 76.)

[Ecabert *et al.* 2008] Olivier Ecabert, Jochen Peters, Hauke Schramm, Christian Lorenz, Jens von Berg, Matthew J. Walker, Mani Vembar, Mark E. Olszewski, Skrishna Ubramanyan, Guy Lavi and Jürgen Weese. *Automatic model-based segmentation of the heart in CT images.* IEEE Transactions on Medical Imaging, vol. 27, no. 9, pages 1189–1201, September 2008. (Cited on page 24.)

[Feremans *et al.* 2004] Corinne Feremans, Martine Labbe and Gilbert Laporte. *The Generalized Minimum Spanning Tree problem: Polyhedral analysis and branch-and-cut algorithm.* Networks, vol. 43, no. 2, pages 71–86, March 2004. (Cited on page 95.)

[Figueiredo & Jain 2002] Mário A. T. Figueiredo and Anil K. Jain. *Unsupervised learning of finite mixture models.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 3, pages 381–396, March 2002. (Cited on page 68.)

[Filler 2009] Aaron G. Filler. *The history, development, and impact of computed imaging in neurological diagnosis and neurosurgery: CT, MRI, DTI.* Nature Precedings, vol. 7, no. 1, July 2009. (Cited on pages 13 and 14.)

[Fiorino *et al.* 1998] C. Fiorino, M. Reni, A. Bolognesi, G. M. Cattaneo and R. Calandrino. *Intra- and inter-observer variability in contouring prostate and seminal vesicles: Implications for conformal treatment planning.* Radiotherapy and Oncology, vol. 47, no. 3, pages 285–292, 1998. (Cited on page 20.)

[Fitzpatrick & Sonka 2000] Michael Fitzpatrick and Milan Sonka. Handbook of medical imaging, volume 2. Medical image processing and analysis, volume PM80. SPIE - The International Society for Optical Engineering, June 2000. (Cited on page 19.)

[Flandin 2004] Guillaume Flandin. *Utilisation d'informations géométriques pour l'analyse statistique des données d'IRM fonctionnelle.* Phd thesis, Université de Nice-Sophia Antipolis, April 2004. (Cited on page 59.)

[Forsyth & Ponce 2002] David A. Forsyth and Jean Ponce. Computer vision: A modern approach. Prentice Hall, August 2002. (Cited on page 56.)

[Fripp *et al.* 2005] Jurgen Fripp, Stuart Crozier, Simon Warfield and Sébastien Ourselin. *Automatic initialization of 3D deformable models for cartilage segmentation.* In DICTA'05 - Proceedings of the Australian Conference on Digital Image Computing: Techniques and Applications, December 2005. (Cited on page 34.)

[Fukuyama & Sugeno 1989] Y. Fukuyama and M. Sugeno. *A new method of choosing the number of clusters for the Fuzzy C-Means method.* In Proceedings of the 5th Fuzzy Systems Symposium, pages 247–250, 1989. (Cited on page 65.)

[Gabriel & Sokal 1969] Ruben K. Gabriel and Robert R. Sokal. *A new statistical approach to geographic variation analysis.* Systematic Zoology, vol. 18, no. 3, pages 259–278, September 1969. (Cited on page 61.)

[Gérard *et al.* 2002] Olivier Gérard, Antoine Collet Billon, Jean-Michel Rouet, Marie Jacob, Maxim Fradkin and Cyril Allouche. *Efficient model-based quantification of left ventricular function in 3-D echocardiography.* IEEE Transactions on Medical Imaging, vol. 21, no. 9, pages 1059–1068, September 2002. (Cited on page 19.)

[Ghanei *et al.* 1998] Amir Ghanei, Hamid Soltanian-Zadeh and Joe P. Windham. *A 3D deformable surface model for segmentation of objects from volumetric data in medical images.* Computers in Biology and Medicine, vol. 28, no. 3, pages 239–253, May 1998. (Cited on pages 36 and 37.)

[Gheissari & Bab-Hadiashar 2008] Niloofar Gheissari and Alireza Bab-Hadiashar. *A comparative study of model selection criteria for computer vision applications.* Image Vision Computing, vol. 26, no. 12, pages 1636–1649, December 2008. (Cited on page 65.)

[Gilles & Magnenat-Thalmann 2010] Benjamin Gilles and Nadia Magnenat-Thalmann. *Musculoskeletal MRI segmentation using multi-resolution simplex meshes with medial representations.* Medical Image Analysis, vol. 14, no. 3, pages 291–302, June 2010. (Cited on pages 151 and 153.)

[Gilles *et al.* 2006] Benjamin Gilles, Laurent Moccozet and Nadia Magnenat-Thalmann. *Anatomical modelling of the musculoskeletal system from MRI.* In MICCAI 2006 - Proceedings of the 9th International Conference on Medical Image Computing and Computer Assisted Intervention, volume 4190 of *Lecture Notes in Computer Science*, pages 289–296, October 2006. (Cited on pages 20, 40, 48 and 145.)

[Gilles 2007] Benjamin Gilles. *Anatomical and kinematical modelling of the musculoskeletal system from MRI.* Phd thesis, University of Geneva, August 2007. (Cited on pages 25, 32, 34, 44, 51, 112 and 144.)

[Glaunes 2005] Joan Glaunes. *Modeles déformables en appariement de formes. Transport par difféomorphismes de points, de mesures et de courants pour la comparaison de formes et l'anatomie numérique.* Phd thesis, Université Paris 13, September 2005. (Cited on pages 76 and 110.)

[Golub & Loan 1996] Gene H. Golub and Charles F. Van Loan. Matrix computations. Johns Hopkins University Press, October 1996. (Cited on page 98.)

[Gonzalez & Woods 2008] Rafael C. Gonzalez and Richard E. Woods. Digital image processing. Pearson Prentice Hall, August 2008. (Cited on page 73.)

[Goodall 1991] Colin Goodall. *Procrustes methods in the statistical analysis of shape.* Journal of the Royal Statistical Society. Series B (Methodological), vol. 53, no. 2, pages 285–339, 1991. (Cited on pages 39 and 150.)

[Gower 1975] John C. Gower. *Generalized Procrustes Analysis.* Psychometrika, vol. 40, pages 33–51, 1975. (Cited on pages 39 and 150.)

[Green & Sibson 1978] P. J. Green and R. Sibson. *Computing Dirichlet tessellations in the plane.* The Computer Journal, vol. 21, no. 2, pages 168–173, 1978. (Cited on page 61.)

[Grünwald 2007] Peter D. Grünwald. The Minimum Description Length principle. MIT Press, June 2007. (Cited on page 68.)

[Guest *et al.* 2001] Elizabeth Guest, Elizabeth Berry, Richard A. Baldock, Márta Fidrich and Mike A. Smith. *Robust point correspondence applied to two- and three-dimensional image registration.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 2, pages 165–179, February 2001. (Cited on page 52.)

[Hamarneh *et al.* 2002] Ghassan Hamarneh, Tim Mcinerney and Demetri Terzopoulos. *Deformable organisms for automatic medical image analysis.* Medical Image Analysis, vol. 6, no. 3, pages 251–266, September 2002. (Cited on page 34.)

[Han *et al.* 1993] S. Han, D.B. Goldgof and K.W. Bowyer. *Using hyperquadrics for shape recovery from range data.* In ICCV'93 - Proceedings of the 4th International Conference on Computer Vision, pages 492–496, May 1993. (Cited on page 23.)

[Hartigan & Wong 1979] John A. Hartigan and M. A. Wong. *Algorithm AS 136: A K-Means clustering algorithm.* Journal of the Royal Statistical Society, Series C (Applied Statistics), vol. 28, no. 1, pages 100–108, 1979. (Cited on page 20.)

[Hastie *et al.* 2009] Trevor Hastie, Robert Tibshirani and Jerome Friedman. The elements of statistical learning: Data mining, inference, and prediction, second edition (Springer Series in Statistics). Springer, February 2009. (Cited on page 54.)

[Hathaway 1986] Richard J. Hathaway. *Another interpretation of the EM algorithm for mixture distributions.* Statistics and Probability Letters, vol. 4, no. 2, pages 53–56, March 1986. (Cited on page 59.)

[Hebert & Leahy 1989] Tom Hebert and Richard Leahy. *A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors.* IEEE Transactions on Medical Imaging, vol. 8, no. 2, pages 194–202, June 1989. (Cited on page 56.)

[Heimann & Delingette 2010] Tobias Heimann and Hervé Delingette. Recent advances in biomedical image processing and analysis, chapter Model-based segmentation. Springer, 2010. (Cited on pages 23 and 25.)

[Heimann *et al.* 2007] Tobias Heimann, Sascha Münzing, Hans-Peter Meinzer and Ivo Wolf. *A shape-guided deformable model with evolutionary algorithm initialization for 3D soft tissue segmentation.* In IPMI'07 - Proceedings of the 20th International Conference on Information Processing in Medical Imaging, pages 1–12. Springer-Verlag, July 2007. (Cited on pages 34 and 146.)

[Heimann *et al.* 2009] T. Heimann, B. van Ginneken, M.A. Styner, Y. Arzhaeva, V. Aurich, C. Bauer, A. Beck, C. Becker, R. Beichel, G. Bekes, F. Bello, G. Binnig, H. Bischof, A. Bornik, P. Cashman, Ying Chi, A. Cordova, B.M. Dawant, M. Fidrich, J.D. Furst, D. Furukawa, L. Grenacher, J. Hornegger, D. Kainmuller, R.I. Kitney, H. Kobatake, H. Lamecker, T. Lange, Jeongjin Lee, B. Lennon, Rui Li, Senhu Li, H.-P. Meinzer, G. Nemeth, D.S. Raicu, A.-M. Rau, E.M. van Rikxoort, M. Rousson, L. Rusko, K.A. Saddi, G. Schmidt, D. Seghers, A. Shimizu, P. Slagmolen, E. Sorantin, G. Soza, R. Susomboon, J.M. Waite, A. Wimmer and I. Wolf. *Comparison and evaluation of methods for liver segmentation from CT datasets.* IEEE Transactions on Medical Imaging, vol. 28, no. 8, pages 1251–1265, August 2009. (Cited on page 125.)

[Heimann 2009] Tobias Heimann. Statistical Shape Models for 3D medical image segmentation. VDM Verlag Dr. Muller Aktiengesellschaft & Co. KG, January 2009. (Cited on pages 1, 34, 40, 44, 51, 75, 88, 112, 117, 147, 150 and 165.)

[Heitz *et al.* 2004] Geremy Heitz, Torsten Rohlfing and Calvin R. Maurer. *Automatic generation of shape models using nonrigid registration with a single segmented template mesh.* In VMV 2004 - Proceedings of the 9th International Fall Workshop on Vision, Modeling and Visualization, November 2004. (Cited on page 52.)

[Holden *et al.* 2000] Mark Holden, Derek L.G. Hill, Erika R.E. Denton, Jo M. Jarosz, Tim C.S. Cox, Torsten Rohlfing, Joanne Goodey and David J. Hawkes. *Voxel similarity measures for 3-D serial MR brain image registration.* IEEE Transactions on Medical Imaging, vol. 19, pages 94–102, February 2000. (Cited on pages 46 and 151.)

[Huttenlocher *et al.* 1993] Daniel P. Huttenlocher, Gregory A. Klanderman and William A. Rucklidge. *Comparing images using the Hausdorff distance.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, no. 9, pages 850–863, September 1993. (Cited on page 127.)

[Jain *et al.* 1998] Anil K. Jain, Yu Zhong and Marie-Pierre Dubuisson-Jolly. *Deformable template models: A review.* Signal Processing, vol. 71, no. 2, pages 109–129, December 1998. (Cited on page 25.)

[Jain *et al.* 1999] Anil K. Jain, M. Narasimha Murty and Patrick J. Flynn. *Data clustering: A review.* ACM Computing Surveys (CSUR), vol. 31, no. 3, pages 264–323, September 1999. (Cited on page 54.)

[Jian & Vemuri 2005] Bing Jian and Baba C. Vemuri. *A robust algorithm for point set registration using mixture of Gaussians.* In ICCV'05 - Proceedings of the 10th IEEE International Conference on Computer Vision, volume 2, pages 1246–1251, October 2005. (Cited on page 76.)

[Jiang & Kececioglu 1992] Siyuan Jiang and Dimitri Kececioglu. *Maximum likelihood estimates, from censored data, for mixed-Weibull distributions.* IEEE Transactions on Reliability, vol. 41, no. 2, pages 248–255, June 1992. (Cited on page 56.)

[Jiang 1991] Siyuan Jiang. *Mixed Weibull distributions in reliability engineering: Statistical models for the lifetime of units with multiple modes of failure.* Phd thesis, Arizona University, August 1991. (Cited on page 56.)

[Jolliffe 2002] Ian T. Jolliffe. Principal Component Analysis. Springer, second edition, October 2002. (Cited on pages 39 and 150.)

[Josephson *et al.* 2005] Klas Josephson, Anders Ericsson and Johan Karlsson. *Segmentation of medical images using three-dimensional Active Shape Models.* In Heikki Kalviainen, Jussi Parkkinen and Arto Kaarna, editors, Image Analysis, volume 3540 of *Lecture Notes in Computer Science*, pages 719–728. Springer, 2005. (Cited on page 52.)

[Karayiannis & Pai 1995] Nicolaos B. Karayiannis and Pin-I Pai. *Fuzzy vector quantization algorithms and their application in image compression.* IEEE Transactions on Image Processing, vol. 4, no. 9, pages 1193–1201, September 1995. (Cited on page 56.)

[Kass *et al.* 1988] Michael Kass, Andrew Witkin and Demetri Terzopoulos. *Snakes: Active Contour Models.* International Journal of Computer Vision, vol. 1, no. 4, pages 321–331, 1988. (Cited on pages 24, 35 and 40.)

[Kaus *et al.* 2007] Michael R. Kaus, Kristy K. Brock, Vladimir Pekar, Laura A. Dawson, Alan M. Nichol and David A. Jaffray. *Assessment of a model-based deformable image registration approach for radiation therapy planning.* International Journal of Radiation Oncology*Biology*Physics, vol. 68, no. 2, pages 572–580, June 2007. (Cited on page 19.)

[Kim *et al.* 2001] Do-Jong Kim, Yong-Woon Park and Kong-Jo Park. *A novel validity index for determination of the optimal number of clusters.* IEICE Transactions on Information and Systems, vol. E84-D, no. 2, pages 281–285, February 2001. (Cited on page 65.)

[Kim *et al.* 2004] Dae-Won Kim, Kwang H. Lee and Doheon Lee. *On cluster validity index for estimation of the optimal number of fuzzy clusters.* Pattern Recognition, vol. 37, pages 2009–2025, October 2004. (Cited on pages 65, 66, 69 and 70.)

[Kimia *et al.* 1992] Benjamin B. Kimia, Allen Tannenbaum and Steven W. Zucker. *On the evolution of curves via a function of curvature. I. The classical case.* Journal of Mathematical Analysis and Applications, vol. 163, no. 2, pages 438–458, January 1992. (Cited on pages 35 and 36.)

[Kotsiantis & Pintelas 2004] Sotiris Kotsiantis and Panayiotis Pintelas. *Recent advances in clustering: A brief survey.* WSEAS Transactions on Information Science and Applications, vol. 1, no. 1, pages 73–81, 2004. (Cited on page 54.)

[Kriegel *et al.* 2009] Hans-Peter Kriegel, Peer Kröger and Arthur Zimek. *Clustering high-dimensional data: A survey on subspace clustering, pattern-based clustering, and correlation clustering.* ACM Transactions on Knowledge Discovery from Data (TKDD), vol. 3, no. 1, pages 1–58, March 2009. (Cited on page 54.)

[Kwon 1998] S.H. Kwon. *Cluster validity index for fuzzy clustering.* Electronics Letters, vol. 34, no. 22, pages 2176–2177, October 1998. (Cited on page 65.)

[Lachaud & Montanverta 1999] Jacques-Olivier Lachaud and Annick Montanverta. *Deformable meshes with automated topology changes for coarse-to-fine three-dimensional surface extraction.* Medical Image Analysis, vol. 3, no. 2, pages 187–207, June 1999. (Cited on page 32.)

[Lauterbur 1973] Paul C. Lauterbur. *Image formation by induced local interactions: Examples of employing nuclear magnetic resonance.* Nature, vol. 242, pages 190–191, March 1973. (Cited on page 14.)

[Lazarsfeld & Henry 1968] Paul F. Lazarsfeld and Neil W. Henry. Latent structure analysis. Houghton Mifflin Company, 1968. (Cited on page 56.)

[Leemput *et al.* 2001] Koen Van Leemput, Frederik Maes, Dirk Vandermeulen, Alan Colchester and Paul Suetens. *Automated segmentation of multiple sclerosis lesions by model outlier detection.* IEEE Transactions on Medical Imaging, vol. 20, no. 8, pages 677–688, August 2001. (Cited on page 60.)

[Leung & Malik 2001] Thomas Leung and Jitendra Malik. *Representing and recognizing the visual appearance of materials using three-dimensional textons.* International Journal of Computer Vision, vol. 43, no. 1, pages 29–44, June 2001. (Cited on page 42.)

[Leventon *et al.* 2000] Michael E. Leventon, W. Eric L. Grimson and Olivier Faugeras. *Statistical shape influence in Geodesic Active Contours.* In CVPR 2000 - Proceedings of the IEEE Computer Society Conference on Computer

Vision and Pattern Recognition, volume 1, pages 316–323, June 2000. (Cited on page 24.)

[Li *et al.* 2004] Shuyu Li, Litao Zhu and Tianzi Jiang. *Active Shape Model segmentation using local edge structures and AdaBoost.* In MIAR 2004 - 2nd International Workshop on Medical Imaging and Augmented Reality, volume 3150, pages 121–128, August 2004. (Cited on page 1.)

[Likar *et al.* 2004] Boštjan Likar, Max A. Viergever and Franjo Pernuš. *Retrospective correction of MR intensity inhomogeneity by information minimization.* In MICCAI 2000 - Proceedings of the 3rd International Conference on Medical Image Computing and Computer-Assisted Intervention, volume 1935, pages 177–201, February 2004. (Cited on pages 73 and 142.)

[Liu *et al.* 1995] W.Y. Liu, I.E. Mangnin and G. Gimenez. *Un nouvel opérateur pour la détection de ruptures dans des signaux bruités.* Traitement du Signal, vol. 12, no. 3, pages 225–236, 1995. (Cited on page 42.)

[Lorensen & Cline 1987] William E. Lorensen and Harvey E. Cline. *Marching cubes: A high resolution 3D surface construction algorithm.* ACM SIGGRAPH Computer Graphics, vol. 21, no. 4, pages 163–169, July 1987. (Cited on pages 21, 110 and 142.)

[Lorenzo-Valdés *et al.* 2002] Maria Lorenzo-Valdés, Gerardo I. Sanchez-Ortiz, Raad H. Mohiaddin and Daniel Rueckert. *Atlas-based segmentation and tracking of 3D cardiac MR images using non-rigid registration.* In MICCAI'02 - Proceedings of the 5th International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 642–650. Springer-Verlag, September 2002. (Cited on page 22.)

[Lovász 1993] Laszlo Lovász. *Random walks on graphs: A survey.* In Combinatorics, Paul Erdos is Eighty (Bolyai Society Mathematical Studies), volume 2, pages 1–46, 1993. (Cited on page 98.)

[Luxburg 2007] Ulrike Luxburg. *A tutorial on spectral clustering.* Statistics and Computing, vol. 17, no. 4, pages 395–416, December 2007. (Cited on pages 88, 94, 95, 97 and 98.)

[MacQueen 1967] J. B. MacQueen. *Some methods for classification and analysis of multivariate observations.* In Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, volume 1, pages 281–297, 1967. (Cited on pages 55 and 59.)

[Maes *et al.* 1997] Frederik Maes, André Collignon, Dirk Vandermeulen, Guy Marchal and Paul Suetens. *Multimodality image registration by maximization of mutual information.* IEEE Transactions on Medical Imaging, vol. 16, no. 2, pages 187–198, April 1997. (Cited on page 44.)

[Maintz & Viergever 1998] J.B. Antoine Maintz and Max A. Viergever. *A survey of medical image registration.* Medical Image Analysis, vol. 2, no. 1, pages 1–36, March 1998. (Cited on page 22.)

[Malik *et al.* 2001] Jitendra Malik, Serge Belongie, Thomas Leung and Jianbo Shi. *Contour and texture analysis for image segmentation.* International Journal of Computer Vision, vol. 43, no. 1, pages 7–27, June 2001. (Cited on page 88.)

[Malladi *et al.* 1995] Ravikanth Malladi, James A. Sethian and Baba C. Vemuri. *Shape modeling with front propagation: A level set approach.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 17, no. 2, pages 158–175, February 1995. (Cited on page 24.)

[Mansi 2010] Tommaso Mansi. *Image-based physiological and statistical models of the heart: Application to tetralogy of fallot.* Phd thesis, Mines ParisTech, September 2010. (Cited on page 76.)

[Maulik & Bandyopadhyay 2000] Ujjwal Maulik and Sanghamitra Bandyopadhyay. *Genetic algorithm-based clustering technique.* Pattern Recognition, vol. 33, no. 9, pages 1455–1465, September 2000. (Cited on page 69.)

[McInerney & Terzopoulos 1993] Tim McInerney and Demetri Terzopoulos. *A finite element model for 3D shape reconstruction and nonrigid motion tracking.* In ICCV'93 - Proceedings of the 4th International Conference on Computer Vision, pages 518–523, May 1993. (Cited on page 41.)

[McInerney & Terzopoulos 1995a] Tim McInerney and Demetri Terzopoulos. *A dynamic finite element surface model for segmentation and tracking in multidimensional medical images with application to cardiac 4D image analysis.* Computerized Medical Imaging and Graphics, vol. 19, no. 1, pages 69–83, February 1995. (Cited on page 34.)

[McInerney & Terzopoulos 1995b] Tim McInerney and Demetri Terzopoulos. *Topologically adaptable snakes.* In ICCV'95 - Proceedings of the 5th International Conference on Computer Vision, pages 840–845, June 1995. (Cited on page 41.)

[McInerney & Terzopoulos 1996] Tim McInerney and Demetri Terzopoulos. *Deformable models in medical image analysis: A survey.* Medical Image Analysis, vol. 1, no. 2, pages 91–108, 1996. (Cited on page 25.)

[McInerney & Terzopoulos 1999] Tim McInerney and Demetri Terzopoulos. *Topology adaptive deformable surfaces for medical image volume segmentation.* IEEE Transactions on Medical Imaging, vol. 18, no. 10, pages 840–850, October 1999. (Cited on page 24.)

[McLachlan & Krishnan 1996] Geoffrey J. McLachlan and Thriyambakam Krishnan. The EM algorithm and its extensions. Wiley-Interscience, November 1996. (Cited on page 56.)

[McLachlan & Peel 2000] Geoffrey McLachlan and David Peel. Finite mixture models. Wiley-Interscience, September 2000. (Cited on page 56.)

[Meila & Shi 2000] Marina Meila and Jianbo Shi. *Learning segmentation by Random Walks*. In NIPS 2000 - Proceedings of the 14th Annual Conference on Neural Information Processing Systems, pages 873–879, December 2000. (Cited on page 95.)

[Mohar 1997] Bojan Mohar. *Some applications of Laplace eigenvalues of graphs*. In Graph Symmetry: Algebraic Methods and Applications, volume 497, pages 227–275, 1997. (Cited on pages 97 and 98.)

[Montagnat & Delingette 1998] Johan Montagnat and Hervé Delingette. *Globally constrained deformable models for 3D object reconstruction*. Signal Processing, vol. 71, no. 2, pages 173–186, December 1998. (Cited on pages 30 and 71.)

[Montagnat *et al.* 2001] Johan Montagnat, Hervé Delingette and Nicholas Ayache. *A review of deformable surfaces: Topology, geometry and deformation*. Image and Vision Computing, vol. 19, pages 1023–1040, December 2001. (Cited on pages 25, 35 and 36.)

[Montagnat 1999] Johan Montagnat. *Modèles déformables pour la segmentation et la modélisation d'images médicales 3D et 4D*. Phd thesis, Université de Nice-Sophia Antipolis, December 1999. (Cited on pages 24, 25, 31, 32, 34, 38, 39, 41, 42 and 45.)

[Neal & Hinton 1998] Radford M. Neal and Geoffrey E. Hinton. *A new view of the EM algorithm that justifies incremental, sparse and other variants*. In Learning in Graphical Models, pages 355–368. Kluwer Academic Publishers, 1998. (Cited on page 59.)

[Ng *et al.* 2001] Andrew Y. Ng, Michael I. Jordan and Yair Weiss. *On spectral clustering: Analysis and an algorithm*. In Advances in Neural Information Processing Systems 14, pages 849–856. MIT Press, 2001. (Cited on pages 88, 89, 90, 95, 98 and 99.)

[Ng-Thow-Hing 2000] V. Ng-Thow-Hing. *Anatomically-based models for physical and geometric reconstruction of humans and other animals*. Phd thesis, Department of Computer Science, University of Toronto, 2000. (Cited on page 24.)

[Nielson 2003] Gregory M. Nielson. *On Marching Cubes.* IEEE Transactions on Visualization and Computer Graphics, vol. 9, no. 3, pages 283–297, September 2003. (Cited on page 21.)

[Nuyts *et al.* 2001] Johan Nuyts, Christian Michel and Patrick Dupont. *Maximum-likelihood Expectation-Maximization reconstruction of sinograms with arbitrary noise distribution using NEC-transformations.* IEEE Transactions on Medical Imaging, vol. 20, no. 5, pages 365–375, May 2001. (Cited on page 56.)

[Osher & Sethian 1988] Stanley Osher and James A. Sethian. *Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations.* Journal of Computational Physics, vol. 79, no. 1, pages 12–49, November 1988. (Cited on page 24.)

[Pal & Bezdek 1995] Nikhil R. Pal and James C. Bezdek. *On cluster validity for the Fuzzy C-Means model.* IEEE Transactions on Fuzzy Systems, vol. 3, no. 3, pages 370–379, August 1995. (Cited on page 65.)

[Paragios & Deriche 2000] Nikos Paragios and Rachid Deriche. *Coupled geodesic active regions for image segmentation: A Level Set approach.* In ECCV 2000 - Proceedings of the 6th European Conference on Computer Vision, pages 224–240. Springer, July 2000. (Cited on page 42.)

[Park *et al.* 2001] Joo-Young Park, Tim McInerney, Demetri Terzopoulos and Myoung-Hee Kim. *A non-self-intersecting adaptive deformable surface for complex boundary extraction from volumetric images.* Computers & Graphics, vol. 25, no. 3, pages 421–440, June 2001. (Cited on page 36.)

[Parveen *et al.* 2006] Runa Parveen, Cliff Ruff and Andrew Todd-Pokropek. *Three dimensional tissue classifications in MR brain images.* In Computer Vision Approaches to Medical Image Analysis, volume 4241 of *Lecture Notes in Computer Science*, pages 236–247. 2006. (Cited on page 56.)

[Pedrycz & Hirota 2007] Witold Pedrycz and Kaoru Hirota. *Fuzzy vector quantization with the particle swarm optimization: A study in fuzzy granulation-degranulation information processing.* Signal Processing, vol. 87, no. 9, pages 2061–2074, September 2007. (Cited on page 67.)

[Penney *et al.* 1998] G. Penney, J. Weese, J. Little, P. Desmedt, D. Hill and D. Hawkes. *A comparison of similarity measures for use in 2D-3D medical image registration.* In MICCAI'98 - Proceedings of the 1st International Conference on Medical Image Computing and Computer Assisted Intervention, volume 1496, pages 1153–1161, October 1998. (Cited on page 45.)

[Petrou & Bosdogianni 1999] Maria Petrou and Panagiota Bosdogianni. Image processing: The fundamentals. Wiley, October 1999. (Cited on page 22.)

[Pitiot *et al.* 2004] Alain Pitiot, Hervé Delingette, Paul M. Thompson and Nicholas Ayache. *Expert knowledge guided segmentation system for brain MRI.* NeuroImage, vol. 23, pages S85–S96, 2004. Special Issue: Mathematics in Brain Imaging. (Cited on page 22.)

[Pock *et al.* 2005] Thomas Pock, Christian Janko, Reinhard Beichel and Horst Bischof. *Multiscale medialness for robust segmentation of 3D tubular structures.* In CVWW 2005 - Proceedings of the 10th Computer Vision Winter Workshop, pages 93–102, February 2005. (Cited on page 143.)

[Press *et al.* 1992] William H. Press, Brian P. Flannery, Saul A. Teukolsky and William T. Vetterling. Numerical recipes in C: The art of scientific computing, second edition. Cambridge University Press, October 1992. (Cited on page 144.)

[Rezaee *et al.* 1998] M.R. Rezaee, B.P.F Lelieveldt and J.H.C Reiber. *A new cluster validity index for the Fuzzy C-Mean.* Pattern Recognition Letters, vol. 19, no. 3-4, pages 237–246, March 1998. (Cited on page 65.)

[Rissanen 1978] Jorma Rissanen. *Modeling by shortest data description.* Automatica, vol. 14, pages 465–471, 1978. (Cited on page 68.)

[Rissanen 1989] Jorma Rissanen. Stochastic complexity in statistical inquiry. World Scientific, November 1989. (Cited on page 67.)

[Rissanen 2007] Jorma Rissanen. Information and complexity in statistical modeling. Springer, 2007. (Cited on page 68.)

[Roche *et al.* 1998a] Alexis Roche, Grégoire Malandain, Xavier Pennec and Nicholas Ayache. *The correlation ratio as a new similarity measure for multimodal image registration.* In MICCAI'98 - Proceedings of the 1st International Conference on Medical Image Computing and Computer Assisted Intervention, volume 1496, pages 1115–1124, October 1998. (Cited on page 45.)

[Roche *et al.* 1998b] Alexis Roche, Grégoire Malandain, Xavier Pennec and Nicholas Ayache. *Multimodal image registration by maximization of the correlation ratio.* Rapport de Recherche 3378, INRIA - Institut National de Recherche en Informatique et Automatique, August 1998. (Cited on page 45.)

[Ronfard 1994] Rémi Ronfard. *Region-based strategies for Active Contour Models.* International Journal of Computer Vision, vol. 13, no. 2, pages 229–251, October 1994. (Cited on page 42.)

[Rosenfeld & Kak 1982] Azriel Rosenfeld and Avinash C. Kak. Digital picture processing. Academic Press, Inc., 1982. (Cited on page 41.)

[Saarnak *et al.* 2000] A.E. Saarnak, M. Boersma, B.N. van Bunningen, R. Wolterink and M.J. Steggerda. *Inter-observer variation in delineation of bladder and* *rectum contours for brachytherapy of cervical cancer.* Radiotherapy and Oncology, vol. 56, no. 1, pages 37–42, July 2000. (Cited on page 20.)

[Saha & Bandyopadhyay 2007] Sriparna Saha and Sanghamitra Bandyopadhyay. *A new cluster validity index based on fuzzy granulation-degranulation criterion.* In ADCOM 2007 - Proceedings of the 15th International Conference on Advanced Computing, pages 353–358, December 2007. (Cited on page 67.)

[Savage 1997] Carla Savage. *A survey of combinatorial Gray codes.* SIAM Review, vol. 39, no. 4, pages 605–629, December 1997. (Cited on page 100.)

[Schäfer & Strimmer 2005] Juliane Schäfer and Korbinian 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, November 2005. (Cited on page 47.)

[Schmid & Magnenat-Thalmann 2008] Jérôme Schmid and Nadia Magnenat-Thalmann. *MRI bone segmentation using deformable models and shape priors.* In MICCAI 2008 - Proceedings of the 11th International Conference on Medical Image Computing and Computer Assisted Intervention, volume 5241 of *Lecture Notes in Computer Science*, pages 119–126, September 2008. (Cited on pages 20, 32, 48, 146, 151, 153 and 165.)

[Schmid *et al.* 2009] Jérôme Schmid, Anders Sandholm, François Chung, Daniel Thalmann, Hervé Delingette and Nadia Magnenat-Thalmann. *Musculoskeletal simulation model generation from MRI datasets and motion capture data.* In Recent Advances in the 3D Physiological Human, pages 3–19. Springer-Verlag, February 2009. (Cited on pages 20, 139, 140, 141, 142, 144 and 151.)

[Schmid *et al.* 2010] Jérôme Schmid, Jinman Kim and Nadia 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, pages 1–11, May 2010. (Cited on pages 157 and 158.)

[Schneider 2001] Tapio Schneider. *Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values.* Journal of Climate, vol. 14, no. 5, pages 853–871, March 2001. (Cited on page 64.)

[Schwarz 1978] Gideon Schwarz. *Estimating the dimension of a model.* Annals of Statistics, vol. 6, no. 2, pages 461–464, March 1978. (Cited on page 68.)

[Schwefel 1993] Hans-Paul Schwefel. Evolution and optimum seeking: The sixth generation. John Wiley & Sons, 1993. (Cited on page 34.)

[Sederberg & Parry 1986] Thomas W. Sederberg and Scott R. Parry. *Free-Form Deformation of solid geometric models.* ACM SIGGRAPH Computer Graphics, vol. 20, no. 4, pages 151–160, August 1986. (Cited on page 37.)

[Seim *et al.* 2008] Heiko Seim, Dagmar Kainmueller, Markus Heller, Hans Lamecker, Stefan Zachow and Hans-Christian Hege. *Automatic segmentation of the pelvic bones from CT data based on a Statistical Shape Model.* In Charl Botha, Gordon Kindlmann, Wiro Niessen and Bernhard Preim, editors, Eurographics Workshop on Visual Computing for Biomedicine, pages 93–100. Eurographics Association, April 2008. (Cited on page 146.)

[Sermesant *et al.* 2006] Maxime Sermesant, Hervé Delingette and Nicholas Ayache. *An electromechanical model of the heart for image analysis and simulation.* IEEE Transactions on Medical Imaging, vol. 25, no. 5, pages 612–625, May 2006. (Cited on page 24.)

[Sezgin & Sankur 2004] Mehmet Sezgin and Bülent Sankur. *Survey over image thresholding techniques and quantitative performance evaluation.* Journal of Electronic Imaging, vol. 13, no. 1, pages 146–168, January 2004. (Cited on page 21.)

[Shen *et al.* 2003] Dinggang Shen, Yiqiang Zhan and Christos Davatzikos. *Segmentation of prostate boundaries from ultrasound images using Statistical Shape Model.* IEEE Transactions on Medical Imaging, vol. 22, no. 4, pages 539–551, April 2003. (Cited on page 34.)

[Shi & Malik 2000] Jianbo Shi and Jitendra Malik. *Normalized cuts and image segmentation.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pages 888–905, August 2000. (Cited on page 98.)

[Sijbers *et al.* 1998] Jan Sijbers, Arjan den Dekker, Johan Van Audekerke, Marleen Verhoye and Dirk Van Dyck. *Estimation of the noise in magnitude MR images.* Magnetic Resonance Imaging, vol. 16, no. 1, pages 87–90, 1998. (Cited on page 141.)

[Soler *et al.* 2001] Luc Soler, Hervé Delingette, Grégoire Malandain, Johan Montagnat, Nicholas Ayache, Christophe Koehl, Olivier Dourthe, Benoit Malassagne, Michelle Smith, Didier Mutter and Jacques Marescaux. *Fully automatic anatomical, pathological, and functional segmentation from CT scans for hepatic surgery.* Computer Aided Surgery, vol. 6, no. 3, pages 131–142, 2001. (Cited on page 25.)

[Souplet 2009] Jean-Christophe Souplet. *Evaluation de l'atrophie et de la charge lésionnelle sur des séquences IRM de patients atteints de sclérose en plaques.* Thèse de sciences (PhD thesis), Université de Nice Sophia-Antipolis, January 2009. (Cited on page 22.)

[Styner *et al.* 2000] Martin Styner, Christian Brechbühler, Gábor Székely and Guido Gerig. *Parametric estimate of intensity inhomogeneities applied to MRI.* IEEE Transactions on Medical Imaging, vol. 19, no. 3, pages 153–165, March 2000. (Cited on pages 21, 142 and 149.)

[Taubin 1994] Gabriel Taubin. *Distance approximations for rasterizing implicit curves.* ACM Transactions on Graphics, vol. 13, no. 1, pages 3–42, January 1994. (Cited on page 23.)

[Terzopoulos *et al.* 1988] Demetri Terzopoulos, Andrew Witkin and Michael Kass. *Constraints on deformable models: Recovering 3D shape and nongrid motion.* Artificial Intelligence, vol. 36, no. 1, pages 91–123, August 1988. (Cited on page 24.)

[Tsai *et al.* 2003] Andy Tsai, Anthony Yezzi, William Wells, Clare Tempany, Dewey Tucker, Ayres Fan, W. Eric Grimson and Alan Willsky. *A shape-based approach to the segmentation of medical imagery using Level Sets.* IEEE Transactions on Medical Imaging, vol. 22, no. 2, pages 137–154, February 2003. (Cited on page 24.)

[Udupa & Saha 2003] Jayaram K. Udupa and Punam K. Saha. *Fuzzy connectedness and image segmentation.* Proceedings of the IEEE, vol. 91, no. 10, pages 1649–1669, October 2003. (Cited on page 60.)

[Vaillant & Glaunes 2005] Marc Vaillant and Joan Glaunes. *Surface matching via currents.* In IPMI 2005 - Proceedings of the International Conference on Information Processing in Medical Imaging, volume 3565, page 381, July 2005. (Cited on pages 76 and 110.)

[Vemuri *et al.* 2003] Baba C. Vemuri, J. Ye, Y. Chen and C.M. Leonard. *Image registration via level-set motion: Applications to atlas-based segmentation.* Medical Image Analysis, vol. 7, no. 1, pages 1–20, March 2003. (Cited on page 22.)

[Vovk *et al.* 2007] Uros Vovk, Franjo Pernuš and Boštjan Likar. *A review of methods for correction of intensity inhomogeneity in MRI.* IEEE Transactions on Medical Imaging, vol. 26, no. 3, pages 405–421, March 2007. (Cited on pages 73 and 142.)

[Wallace & Freeman 1987] Christopher S. Wallace and Patricia R. Freeman. *Estimation and inference by compact coding.* Journal of the Royal Statistical Society. Series B (Methodological), vol. 49, no. 3, pages 240–265, 1987. (Cited on page 67.)

[Wan *et al.* 2009] Li Wan, Wee Keong Ng, Xuan Hong Dang, Philip S. Yu and Kuan Zhang. *Density-based clustering of data streams at multiple resolutions.* ACM Transactions on Knowledge Discovery from Data (TKDD), vol. 3, no. 3, pages 1–28, July 2009. (Cited on page 54.)

[Wang *et al.* 2004] Hai Xian Wang, Bin Luo, Quan Bing Zhang and Sui Wei. *Estimation for the number of components in a mixture model using stepwise split-and-merge EM algorithm.* Pattern Recognition Letters, vol. 25, no. 16, pages 1799 – 1809, December 2004. (Cited on page 69.)

[Weisenfeld & Warfteld 2004] Neil L. Weisenfeld and Simon K. Warfteld. *Normalization of joint image-intensity statistics in MRI using the Kullback-Leibler divergence.* In ISBI 2004 - Proceedings of the 1st IEEE International Symposium on Biomedical Imaging: From Nano to Macro, volume 1, pages 101–104, April 2004. (Cited on pages 73 and 142.)

[Xie & Beni 1991] Xuanli L. Xie and Gerardo Beni. *A validity measure for fuzzy clustering.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 13, no. 8, pages 841–847, August 1991. (Cited on page 66.)

[Xu & Jordan 1996] Lei Xu and Michael I. Jordan. *On convergence properties of the EM algorithm for Gaussian mixtures.* Neural Computation, vol. 8, no. 1, pages 129–151, January 1996. (Cited on page 56.)

[Xu & Prince 1998] Chenyang Xu and Jerry L. Prince. *Snakes, shapes, and gradient vector flow.* IEEE Transactions on Image Processing, vol. 7, no. 3, pages 359–369, March 1998. (Cited on page 41.)

[Yang & Wu 2010] Fan Yang and Jin Wu. *An improved image contrast enhancement in multiple-peak images based on histogram equalization.* In ICCDA 2010 - Proceedings of the 2010 International Conference on Computer Design and Applications, volume 1, pages V1–346 –V1–349, June 2010. (Cited on page 21.)

[Zadeh 1965] Lotfi A. Zadeh. *Fuzzy sets.* Information and Control, vol. 8, no. 3, pages 338–353, 1965. (Cited on page 54.)

[Zadeh 1968] Lotfi A. Zadeh. *Fuzzy algorithms.* Information and Control, vol. 12, no. 2, pages 94–102, 1968. (Cited on page 54.)

[Zahid *et al.* 1999] N. Zahid, M. Limouri and A. Essaid. *A new cluster-validity for fuzzy clustering.* Pattern recognition, vol. 32, no. 7, pages 1089–1097, July 1999. (Cited on page 65.)

[Zhang *et al.* 2001] Yongyue Zhang, Stephen Smith and Michael Brady. *Hidden Markov Random Field model and segmentation of brain MR images.* IEEE Transactions on Medical Imaging, vol. 20, pages 45–57, 2001. (Cited on page 60.)

[Zhang 1994] Zhengyou Zhang. *Iterative point matching for registration of free-form curves and surfaces.* International Journal of Computer Vision, vol. 13, no. 2, pages 119–152, October 1994. (Cited on page 31.)

[Zhu & Yuille 1996] Song Chun Zhu and Alan Yuille. *Region competition: Unifying snakes, region growing, and Bayes/MDL for multi-band image segmentation.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, pages 884–900, 1996. (Cited on page 42.)