Diffusion Tensor Imaging of the Human Skeletal Muscle: Contributions and Applications

Radhouène Neji

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

centre de recherche SACLAY - ÎLE-DE-FRANCE

PHD THESIS

to obtain the title of

Doctor of Ecole Centrale Paris

Specialty: Applied Mathematics

presented by

Radhouène Neji

Diffusion Tensor Imaging of the Human Skeletal Muscle: Contributions and Applications

prepared at Ecole Centrale Paris, MAS Laboratory and Supélec, SSE Department. Defended on March 9, 2010.

Committee :

Chairman:	James Duncan	-	Yale University
Reviewers:	Christos Davatzikos	-	University of Pennsylvania
	Jean-Baptiste Poline	-	CEA
	Carl-Fredrik Westin	-	Harvard Medical School
Advisors:	Nikos Paragios	-	Ecole Centrale Paris
	Gilles Fleury	-	Supélec
Examiners:	Jean-François Deux	-	CHU Henri Mondor
	Jean-Philippe Thiran	-	EPFL

2010ECAP0005

THÈSE DE DOCTORAT

pour obtenir le grade de

Docteur de l'Ecole Centrale Paris

Mention: Mathématiques Appliquées

présentée par

Radhouène NEJI

IRM du tenseur de diffusion du muscle squelettique humain: contributions et applications

préparée à l'Ecole Centrale Paris, laboratoire MAS et à Supélec, département SSE. Soutenue le 9 Mars 2010.

Jury :

Président :	James DUNCAN	-	Université Yale
Rapporteurs :	Christos DAVATZIKOS	-	Université de Pennsylvanie
	Jean-Baptiste POLINE	-	CEA
	Carl-Fredrik WESTIN	-	Harvard Medical School
Directeurs de thèse :	Nikos PARAGIOS	-	Ecole Centrale Paris
	Gilles FLEURY	-	Supélec
Examinateurs :	Jean-François DEUX	-	CHU Henri Mondor
	Jean-Philippe THIRAN	-	EPFL

2010ECAP0005

Abstract

In this thesis, we present several techniques for the processing of diffusion tensor images. They span a wide range of tasks such as estimation and regularization, clustering and segmentation, as well as registration.

The variational framework proposed for recovering a tensor field from noisy diffusion weighted images exploits the fact that diffusion data represent populations of fibers and therefore each tensor can be reconstructed using a weighted combination of tensors lying in its neighborhood.

The segmentation approach operates both at the voxel and the fiber tract levels. It is based on the use of Mercer kernels over Gaussian diffusion probabilities to model tensor similarity and spatial interactions, allowing the definition of fiber metrics that combine information from spatial localization and diffusion tensors. Several clustering techniques can be subsequently used to segment tensor fields and fiber tractographies. Moreover, we show how to develop supervised extensions of these algorithms.

The registration algorithm uses probability kernels in order to match moving and target images. The deformation consistency is assessed using the distortion induced in the distances between neighboring probabilities. Discrete optimization is used to seek an optimum of the defined objective function.

The experimental validation is done over a dataset of manually segmented diffusion images of the lower leg muscle for healthy and diseased subjects. The results of the techniques developed throughout this thesis are promising.

Key-words: DTI, Diffusion Tensor, Fiber, Kernels, Clustering, Registration, Denoising, Human Skeletal Muscle, Myopathies

Résumé

Cette thèse propose des techniques pour le traitement d'images IRM de diffusion. Les méthodes proposées concernent l'estimation et la régularisation, le groupement et la segmentation ainsi que le recalage.

Le cadre variationnel proposé dans cette thèse pour l'estimation d'un champ de tenseurs de diffusion à partir d'observations bruitées exploite le fait que les données de diffusion représentent des populations de fibres et que chaque tenseur peut être reconstruit à partir d'une combinaison pondérée de tenseurs dans son voisinage.

La méthode de segmentation traite aussi bien les voxels que les fibres. Elle est basée sur l'utilisation de noyaux défini-positifs sur des probabilités gaussiennes de diffusion afin de modéliser la similarité entre tenseurs et les interactions spatiales. Ceci permet de définir des métriques entre fibres qui combinent les informations de localisation spatiale et de tenseurs de diffusion. Plusieurs approches de groupement peuvent être appliquées par la suite pour segmenter des champs de tenseurs et des trajectoires de fibres. Un cadre de groupement supervisé est proposé pour étendre cette technique.

L'algorithme de recalage utilise les noyaux sur probabilités pour recaler une image source et une image cible. La régularité de la déformation est évaluée en utilisant la distortion induite sur les distances entre probabilités spatialement voisines. La minimisation de la fonctionnelle de recalage est faite dans un cadre discret.

La validation expérimentale est faite sur des images du muscle du mollet pour des sujets sains et pour des patients atteints de myopathies. Les résultats des techniques développées dans cette thèse sont encourageants.

Mots-clés : IRM de diffusion, Tenseur de diffusion, Fibre, Noyaux, Groupement, Recalage, Débruitage, Muscle squelettique, Myopathies

Acknowledgments

I would like to express my gratitude to my advisors Prof. Nikos Paragios and Prof. Gilles Fleury. They wished me the best during this experience and this work would not have been possible without their guidance.

My thanks go also to the members of my thesis committee. In particular, I am grateful to Prof. James Duncan and Prof. Christos Davatzikos for waking up early in the morning to listen to my presentation. Prof. Carl-Fredrik Westin, Dr. Jean-Baptiste Poline and Prof. Christos Davatzikos reviewed this manuscript and provided me with pertinent comments to improve its content despite their busy schedule. Thanks to Prof. Westin for traveling from the US to attend the presentation.

I am also grateful to Prof. Jean-Philippe Thiran for attending the thesis presentation and for having welcomed me at his laboratory at the Swiss Federal Institute of Technology Lausanne (EPFL) for three of the best and most productive months of my PhD experience. Thanks also extend to all the members of the Signal Processing laboratory of EPFL for their hospitality and for their kindness, in particular Meryem, Leila, Benoît, Daniele, Francesca, Gelareh, Ashkan, Ivan, Peter, Gorthi, Meritxell, etc. It took me very few days to find my bearings at EPFL and I will keep a very high opinion of the people that I met there.

This thesis would not have been possible without the collaboration with Dr. Jean-François Deux who kindly provided the data and the manual segmentations that were used for the experiments and accepted to be a member of my thesis committee. I enjoyed working with him and he provided insightful feedback and advice to guide the research project. Thanks also to Dr. Nikos Komodakis for kindly providing the optimization and clustering code that were used to obtain some of the results of this thesis. I would

like also to acknowledge the outstanding work done by Dr. Lilla Zöllei and Christian Wachinger that helped me a lot when I started my PhD.

I would like also to thank Noura for her kindness and for her excellent scientific input during the collaboration on the denoising part of this thesis. The registration part is a collaboration with Aris and I am deeply grateful to him for putting up with my nervousness and for his dedication and the funny moments we shared.

I would like to thank my very dear friend and colleague Ahmed for his unconditional support during this experience. In my most difficult and desperate moments, when support and understanding were scarce, his unselfishness and his patience, his ideas and his ideals, his very high stature as a person and as a scientist instilled in me hope and confidence. Thanks Ahmed for your team spirit, for your help and for listening to my complaints without ever complaining. I take so much pride in our friendship.

I am also grateful to Sylvie for her overwhelming kindness, her understanding and for her valuable help during these years. She helped me circumvent the heavy bureaucratic process that a PhD thesis may entail. Karine and Annie were also of valuable help as far as administrative issues are concerned.

Thanks also to my office mates Salma, Aris, Mickaël, Panos, Loïc, Olivier Teboul and Olivier Juan for the nice moments spent together, for the delightful atmosphere they were able to create and for listening to my complaints and to my rattling. I do measure with hindsight the privilege that I had to be surrounded by them.

I am grateful as well to the members (and former members) of the computer vision and medical imaging group of the MAS laboratory including Charlotte, Régis, Maxime, Fabrice, Katerina, Chaohui, Martin, Georg, Daniel, Siddharth, etc. It was a pleasure to work and interact with them.

I do not forget Prof. Florian de Vuyst and Prof. Erick Herbin with whom I shared lengthy moments of witty and passionate discussions and I appreciate their outspokenness and their openmindedness. Thanks also to Prof. Christophe Laux for encouraging me in my research.

Prof. Florence Mayo Quenette had a tremendous influence on me with her cheerfulness and the sheer forcefulness and originality of her ideas and her teaching. It was a honor to be her student.

I wonder how dull and monotonous my life outside the MAS laboratory would have been without the company of my dear friends Noureddine, Zied, Benoit, Mathieu, Neder, Marta, Fatima, Aymen, Mourad, Raïed, etc. I always sought their insightful advice and I treasure the delightful moments spent with them.

Last but not least, it is a tautology to say that a tribute should be paid to my brother Dhia and sister Wafa for their kind support and to my parents, for the strength of their will, the tremendous sacrifice they made for my well-being , for valuing so highly education and knowledge and for so many things that I feel that duly acknowledging them goes beyond the scope of this thesis. To them goes my everlasting gratitude.

Contents

List of figures

Chapter 1

Introduction

Medical imaging technologies have undergone major developments in the last decades, making their use ubiquitous in daily clinical routines, including patient condition assessment and monitoring, disease identification and cure and treatment design. They are generally non-invasive and relatively fast, allowing to image the inside of the human body and to provide relevant information about its integrity while saving time for the physician and minimizing discomfort for the patient. The medical imaging field is continuously progressing towards more challenging and more informative technologies, as strong emphasis is put on revealing a more accurate and relevant information to the physician in a more demanding time frame. In this context, diffusion tensor imaging (DTI) has emerged as a relatively recent imaging technology that goes beyond standard structural techniques like T1 and T2 weighted magnetic resonance imaging (MRI) by providing quantitative information about the water diffusion process within the imaged anatomical regions. By measuring diffusion in several space directions, it allows to model this physical phenomenon and to relate it to the underlying fiber architecture of the tissue, since anatomical fibers represent privileged diffusive pathways for water molecules. The modality has attracted much interest due to its potential in estimating and reconstructing in-vivo, i.e. without the need for any invasive inspection technique such as biopsy, the paths connecting different regions of the white matter brain. This paved the way to a more thorough exploration of brain architecture and connectivity and to clinical studies involving quantitative comparisons of measures computed

from diffusion tensor images between control (normal) subjects and diseased patients.

The high dimensionality and non-linearity of the data provided by the modality imply that it conveys an important amount of information that should be adequately processed in order to exploit the acquired images, extract and compute the parameters of interest and explore the possibilities of including this technique in clinical procedures of disease diagnosis and cure design. In this thesis, we focus on the study of the human skeletal muscle and more specifically the lower leg using diffusion tensor imaging. Our motivation is three-fold:

- The human skeletal muscle has attracted less attention in diffusion tensor imaging than the human brain white matter, despite the fact that it presents a well known anatomy of fiber bundles, and the development and use of computer-based processing techniques was rather limited as far as the muscle is concerned.

- It is important to see whether the modality is in accordance with the anatomical knowledge of clinical experts about the human skeletal muscle, in particular it is of interest to assess the reliability of the fiber direction estimation provided by the computed images as well as the accuracy of the fiber bundling algorithms in separating different muscle groups.

- It is of high clinical relevance to study the muscle and to better evaluate the potential of DTI in yielding reliable information about the effect of neuromuscular diseases (myopathies) on the muscle fiber architecture, especially for early illness detection as it leads to a timely follow-up of the subject. This is a domain with very high stakes and stimulating technical and clinical challenges, since myopathies are generally highly damaging to the muscle and result in severe dysfunctions. Moreover, the most common diagnosis procedure is biopsy, which consists in taking a sample of muscle tissue for microscopic analysis and has the shortcoming of being invasive. Therefore diffusion tensor imaging can overcome this limitation by providing quantitative ways of evaluating the state of a patient, while being suitable for the study of regions exhibiting a fiber architecture.

In this thesis, we propose diffusion tensor image processing techniques with application to the study of the human skeletal muscle. In particular, we develop denoising, segmentation (both at the tensor and at the fiber tract level) and registration techniques for diffusion tensor images and we show promising quantitative and qualitative experimental results for a dataset of healthy and diseased subjects. A comparison with the ground-truth based on manual segmentations provided by a clinical expert is provided. The main contribution of this thesis is that it introduces more elaborate and informative models for interactions between tensors, both in a local and global context. In the context of tensor field smoothing, we propose to use a higher-order model to represent local tensor neighborhoods, allowing to take into account more complex tensor arrangements to reconstruct a tensor field from noisy measurements. As far as tensor and fiber tractography segmentation is concerned, we use a model based on diffusion probabilities that introduces the spatial context and hence defines a natural way to perform spatially-consistent clustering of tensors and fiber trajectories. Similar ideas are applied to diffusion tensor registration, where we introduce a connectivity information in the deformation of tensor fields. This approach yields a way to quantify the effect of a field of deformations belonging to the Euclidean motion group on the source tensor field by evaluating the change in the distance between Gaussian diffusion probabilities. The thesis is organized as follows:

1. **Chapter 2**: in this chapter, we briefly review the principles of diffusion tensor imaging. We discuss the physics of diffusion (Fick's Laws) and introduce tensors as models for the diffusion phenomenon. We explain the significance of the different scalar parameters that can be computed from a tensor field. Moreover, the visual representation of the data using fiber tracts or ellipsoids is also briefly discussed. The last section of this chapter is dedicated to the physics of the acquisition scheme of diffusion tensor imaging, where links between the diffusion propagator (probability of motion of water molecules), the tensor model and the fundamental acquisition equations (Stejskal-Tanner equations) are discussed.

2. **Chapter 3**: this second preliminary chapter presents the current find-

ings and insights provided by diffusion tensor imaging for the study of the human muscle. It gives an overview of the muscle anatomy and neuromuscular diseases and reviews the underlying challenges of the acquisitions. The major part of this chapter reviews the results and methodologies used in clinical DTI studies of the skeletal muscle.

3. **Chapter 4**: In this chapter we introduce a novel variational method for joint estimation and regularization of diffusion tensor fields from noisy raw data. To this end, we use the classic quadratic data fidelity term derived from the Stejskal-Tanner equation with a new smoothness term leading to a convex objective function. The regularization term is based on the assumption that the signal can be reconstructed using a weighted average of observations on a local neighborhood. The weights measure the similarity between tensors and are computed directly from the diffusion images. We preserve the positive semi-definiteness constraint using a projected gradient descent. Experimental validation and comparisons with a similar method using synthetic data with known noise model, as well as classification of tensors towards understanding the anatomy of human skeletal muscle demonstrate the potential of our method.

4. **Chapter 5**: This is the core chapter of the thesis, where we present a kernel-based approach to the clustering of diffusion tensors and fiber tracts. We propose to use a Mercer kernel over the tensor space where both spatial and diffusion information are taken into account. This kernel highlights implicitly the connectivity along fiber tracts. Tensor segmentation is performed using kernel-PCA compounded with a landmark-Isomap embedding and k-means clustering. Based on a soft fiber representation, we extend the tensor kernel to deal with fiber tracts using the multi-instance kernel that reflects not only interactions between points along fiber tracts, but also the interactions between diffusion tensors. This unsupervised method is further extended by way of an atlas-based registration of diffusion-free images, followed by a classification of fibers based on nonlinear kernel Support Vector Machines (SVMs). By tackling a more general case, we present a manifold clustering algorithm for the classification of fibers using an-

gular Hilbertian metrics. As far as clustering is concerned, we propose to use two methods. The first approach is based on diffusion maps and k-means clustering in the spectral embedding space. The second approach uses a linear programming formulation of prototype-based clustering. This formulation allows for classification over manifolds without the necessity to embed the data in low dimensional spaces and determines automatically the number of clusters. Promising experimental results of tensor and fiber classification of the human skeletal muscle over a significant set of healthy and diseased subjects are reported.

5. **Chapter 6**: Following the approach proposed in Chapter 5, we present a method to register diffusion tensor images based on a discrete optimization approach in a Reproducing Kernel Hilbert Space (RKHS) setting. The presented approach encodes both the diffusion information and the spatial localization of tensors in a probabilistic framework. The diffusion probabilities are mapped to a RKHS, where we define a registration energy that accounts both for target matching and deformation regularity in both translation and rotation spaces. The six-dimensional deformation space is quantized and discrete energy minimization is performed using efficient linear programming. We show that the algorithm allows for tensor reorientation directly in the optimization framework. Experimental validation is performed over several pairwise registrations of images from a manually annotated dataset of diffusion tensor images of the calf muscle and comparison is done with a recent technique using eight quantitative criteria.

6. **Chapter 7**: This chapter is dedicated to a review of the contributions of this thesis as well as to the perspectives and the possible improvements that can be undertaken.

Chapter 2

Diffusion Tensor MRI Principles

2.1 Introduction

Diffusion Tensor MRI (DT-MRI) is a modality that provides quantitative information about water diffusion in tissues. The basic intuition behind it is that in organized tissues, water will have a privileged direction of diffusion. Indeed, when tissues present a fiber architecture, the fibers will act as "highways" for water molecules, thus resulting in a higher amount of diffusion along the fiber direction than in the perpendicular directions. Therefore one can see that if we could measure in-vivo the privileged direction of diffusion in a local fashion, then we will be able to know the orientation of the fiber. This information is provided (among others) by DT-MRI. DT-MRI allows to estimate a field (3D volume) of 3×3 symmetric positive definite matrices called tensors and the processing of these tensors yields information about the structure of the imaged region, for example the eigenvector corresponding to the largest eigenvalue of a tensor indicates the local fiber direction at the considered voxel. The use of the tensor as a model for local diffusion amounts to assuming that locally the motion of water protons follow a Gaussian distribution law, i.e. the diffusion propagator is Gaussian and the covariance of this law is given by the tensor.

In this chapter we review the physical principles of water diffusion and the DT-MRI acquisition scheme. Given the high dimensionality of the data

acquired using this modality, several scalar values were proposed in the literature in order to visualize standard 3D volumes that map some particular information of the diffusion tensors. Another enticing way to visualize the data is to reconstruct the diffusive pathways, i.e. the fiber trajectories which are 3D curves. These provide a way of estimating connectivity between different areas of the imaged anatomical region.

This chapter is organized as follows: in section 2.2, we give a brief overview of the physical principles of fluid diffusion and explain the links to the diffusion propagator and the the tensor model. In section 2.3, we discuss the ellipsoid representation of a tensor and provide the definitions of the tensor attributes commonly used in DTI studies. We also review the basics of fiber tractography. Section 2.4 is dedicated to the acquisition scheme of DTI and the link relating the diffusion signals to the diffusion propagator. We also discuss the Stejskal Tanner equations that are the fundamental mathematical relationship between tensors and diffusion weighted images.

2.2 Physics of Diffusion and Tensors

In this section, we give a brief overview of the physics of diffusion. A more detailed analysis can be found in [15, 11, 10]. The process of macroscopic diffusion is governed by Fick's law. It is similar to the heat equation which is a partial differential equation that describes the heat diffusion in a medium, i.e. how the heat spreads and the temperature evolves as a function of both space and time. The process of fluid diffusion is analogous to heat diffusion. Indeed, when the medium presents heterogeneities of concentration (the analogous of temperature in the heat equation), the fluid will tend to diffuse from the region with higher concentration to the region with lower concentration. This is expressed through the concentration gradient which is proportional to the diffusive flux and corresponds to the heat flow in the heat equation. More explicitly, we have the following mathematical relationship (Fick's first law):

$$J = -D\nabla C \tag{2.1}$$

where J is the diffusive flux, D is the diffusion coefficient or diffusivity and ∇C is the concentration gradient. Note that the negative sign enforces the fact that the fluid should diffuse to lower-concentration regions, therefore in

a direction opposite to the concentration gradient.

Fick's second law is analogous to heat equation. It is a partial differential equation describing the spatio-temporal evolution of the concentration

$$\frac{\partial C}{\partial t} + div(J) = 0 \tag{2.2}$$

or equivalently

$$\frac{\partial C}{\partial t} = div(D\nabla C) \tag{2.3}$$

These equations describes the case where there is a concentration gradient. When self-diffusion is considered, it is no longer possible to use the concentration as a means to describe the diffusion process. One should resort instead to the diffusion probability or the diffusion propagator p. The diffusion propagator is the probability of motion of a particle initially at a position $\mathbf{x}$ to a position $\mathbf{y}$ after a given time t. In a similar way to equations (2.1) and (2.2), one can define the diffusive flux and the space-time evolution of the propagator as follows:

$$J = -D\nabla p \tag{2.4}$$

$$\frac{\partial p}{\partial t} = div(D\nabla p) \tag{2.5}$$

Note that the underlying assumptions behind the use of a constant D is that the fluid diffuses in an isotropic region, which means that there are no privileged space directions and that the amount of diffusion in all directions can be modeled using a single scalar. In this particular case, Fick's law admit a closed-form solution. Indeed the diffusion propagator can be shown to be a spherical Gaussian distribution centered around the initial position $\mathbf{x}$ and of standard deviation $\sigma = \sqrt{2Dt}$

$$p(\mathbf{y}|\mathbf{x}, t, \mathbf{D}) = \frac{1}{\sqrt{D4\pi t}} \exp\left(-\frac{(\mathbf{y}-\mathbf{x})^2}{4Dt}\right) \tag{2.6}$$

This is in concordance with the relationship established by Einstein and relating the the average (expected) distance moved in a given time during Brownian motion of particles with the diffusion coefficient:

$$E(||\mathbf{y}-\mathbf{x}||^2) = \sigma^2 = 2Dt \tag{2.7}$$

In the case where diffusion takes place in an anisotropic medium, it is no longer adequate to represent diffusion using a single scalar coefficient. In order to model a diffusion process that is direction-dependent and to take into account the constraints imposed by tissue organization, it is possible to use a 3×3 matrix $\mathbf{D}$, that is symmetric and positive definite called tensor, introduced in [11]. This implies that it can be written using six parameters as follows

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} & \mathbf{D}_{13} \\ \mathbf{D}_{12} & \mathbf{D}_{22} & \mathbf{D}_{23} \\ \mathbf{D}_{13} & \mathbf{D}_{23} & \mathbf{D}_{33} \end{pmatrix} \tag{2.8}$$

Therefore the use of a second order tensor is a more general model than the single diffusion coefficient, since it is more appropriate for the case of anisotropic diffusion. Note that Fick's laws can be adapted this case by simply considering D as a matrix operator in equations (2.1) and (2.2). The diffusion propagator is a spherical Gaussian distribution for the case of isotropic diffusion, it is an anisotropic Gaussian in the case of anisotropic diffusion, where the covariance information is provided by the tensor:

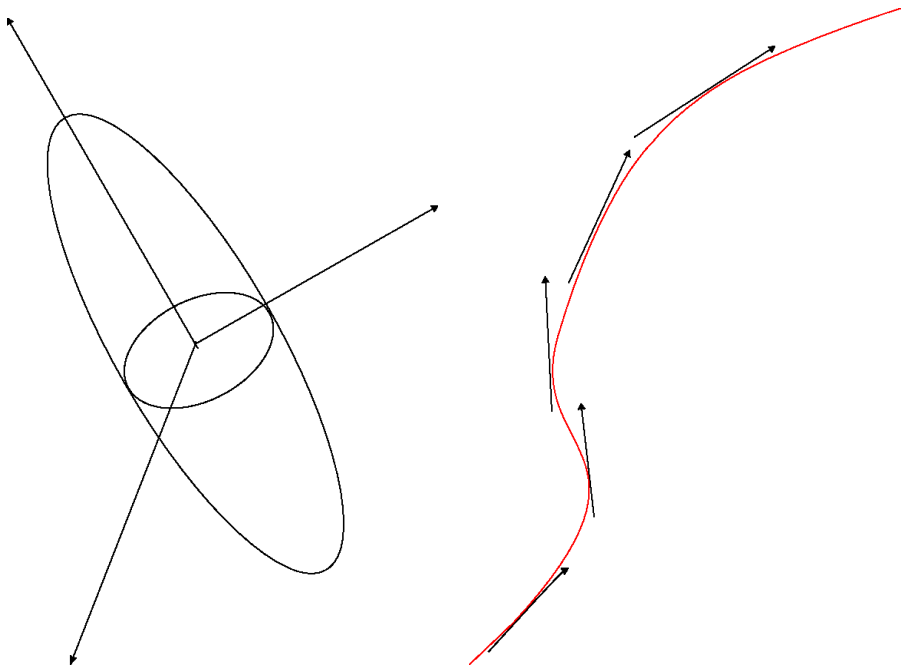
$$p(\mathbf{y}|\mathbf{x}, t, \mathbf{D}) = \frac{1}{\sqrt{\det(\mathbf{D})(4\pi t)^3}} \exp\left(-\frac{(\mathbf{y}-\mathbf{x})^t \mathbf{D}^{-1} (\mathbf{y}-\mathbf{x})}{4t}\right) \tag{2.9}$$

where $\mathbf{D}$ is the tensor at the initial position $\mathbf{x}$. This is the fundamental assumption behind the modeling of a diffusion process using second order tensors.

2.3 The Ellipsoid Representation of a Tensor and Tensor Indices

We now discuss the ellipsoid representation of tensors and some of the scalar parameters that can be estimated from a tensor field, following [11, 13, 123]. A tensor is a real symmetric matrix, so it has a diagonal representation in an orthonormal basis of eigenvectors $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ with corresponding eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$. One can write a tensor $\mathbf{D}$ as a sum of rank one matrices as follows:

$$\mathbf{D} = \lambda_1 \mathbf{e}_1 {\mathbf{e}_1}^t + \lambda_2 \mathbf{e}_2 {\mathbf{e}_2}^t + \lambda_3 \mathbf{e}_3 {\mathbf{e}_3}^t \tag{2.10}$$



(a) (b)

Figure 2.1: (a) Ellipsoid representation of a tensor, the long axis corresponds to the fiber direction locally. (b) Integration of principal directions of diffusion yields fiber trajectories.

The eigenvalues are called principal diffusivities and the eigenvectors principal directions of diffusion. An important property of a tensor is that one can derive from $\mathbf{D}$ a quadratic form $q(\mathbf{v}) = \mathbf{v}^t\mathbf{D}\mathbf{v}$ that corresponds physically to the amount of diffusion (diffusivity) in the direction $\mathbf{v}$, when we consider a unit-norm vector $\mathbf{v}$. Note that this is consistent with the definition of the principal diffusivities since it is straightforward to see that the eigenvalues are exactly the diffusivities in the direction of their respective eigenvectors. Another quadratic form that can be derived from the tensor is $\widehat{q} = \mathbf{v}^t\mathbf{D}^{-1}\mathbf{v}$. This form appears in the expression of the diffusion propagator in equation (2.9). It is positive definite because the tensor $\mathbf{D}$ (and therefore $\mathbf{D}^{-1}$) is positive definite, so its geometric representation in the 3D space is an ellipsoid. The axes of the ellipsoid are the eigenvectors and the radii are proportional to the square roots of the eigenvalues (see [Fig.2.1 (a)]). Indeed if we denote $\mathbf{v} = (x, y, z)^t$ the coordinates of $\mathbf{v}$ in the orthonormal basis of eigenvectors of $\mathbf{D}$ then the isosurface $\widehat{q}(\mathbf{v}) = 1$ corresponds mathematically

to the equation of an ellipsoid

$$\frac{x^2}{\lambda_1} + \frac{y^2}{\lambda_2} + \frac{z^2}{\lambda_3} = 1 \tag{2.11}$$

The long axis of the ellipsoid corresponds to the most likely direction of diffusion, or equivalently the local fiber orientation.

The high dimensionality of tensors is a tricky issue as far as visualization is concerned. In order to ease the interpretation of the data , some important scalar values that give an idea about the diffusion process are derived from tensors:

1. **Trace**: the trace is the sum of the eigenvalues:

$$Tr(\mathbf{D}) = \lambda_1 + \lambda_2 + \lambda_3 \tag{2.12}$$

   It is a rotationally independent value, i.e. independent of the choice of the basis in which the tensor is expressed. A related scalar value is the **mean diffusivity**, which is the average of the eigenvalues $\frac{Tr(\mathbf{D})}{3}$. It is the average amount of diffusion. An example of a trace map is provided in [Fig.2.2] where the axial slices of the 3D trace volume for the calf muscle are shown.

2. **Apparent Diffusion Coefficient** (**ADC**): this quantity is direction-dependent. Since the diffusion tensor model supposes that the diffusion process is free while it is actually restricted by barriers like cell membranes, what we can measure using DT-MRI are only apparent diffusion coefficients. Indeed the ADC is related to the intrinsic diffusion coefficient by

$$ADC = \frac{D}{\lambda^2} \tag{2.13}$$

   where the factor λ represents the restricting effects of barriers.

   Using the tensor model, the ADC in a direction $\mathbf{v}$ is $\mathbf{v}^t\mathbf{D}\mathbf{v}$. For example the apparent diffusion coefficients along the axes of the laboratory coordinate frame ADC_x, ADC_y and ADC_z are the diagonal tensor entries $\mathbf{D}_{11}$, $\mathbf{D}_{22}$ and $\mathbf{D}_{33}$.

3. **Fractional Anisotropy** (**FA**): an important property of the tensor is its anisotropy. If the ellipsoid representation of a tensor is close

in shape to a sphere, then the diffusion described by this tensor is spherical and there are no privileged directions of diffusion. On the other hand, if this representation is close to the shape of a cigar, this is an indicator of diffusion anisotropy and the long axis of the ellipsoid is likely to correspond to the local fiber orientation. Fractional anisotropy (FA) is a measure between 0 and 1 of the anisotropy of a tensor. It takes the value 0 for the case of isotropic diffusion and the value 1 for the limit case of maximum anisotropy (when the shape of the ellipsoid tends towards a cylindrical geometry with an infinite length). FA is given by the following equation:

$$FA = \frac{\sqrt{3}}{\sqrt{2}} \frac{\sqrt{(\lambda_1 - \frac{Tr(\mathbf{D})}{3})^2 + (\lambda_2 - \frac{Tr(\mathbf{D})}{3})^2 + (\lambda_3 - \frac{Tr(\mathbf{D})}{3})^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \tag{2.14}$$

Like trace, FA depends only on the eigenvalues of the tensor, so it is rotationally invariant. An example of an FA map is provided in [Fig.2.3] where the axial slices of the 3D FA volume for the calf muscle are shown.

4. **Orientations of principal directions of diffusion**: the principal directions of diffusion are three-dimensional vectors with coordinates in the laboratory frame. Therefore their orientation can be mapped using a spherical coordinate frame through the elevation angle θ and the azimuthal (in-plane) angle ϕ. The principal directions of diffusion can also be encoded using RGB colors for tensor orientation visualization.

5. **Tensor volume or determinant**: The tensor volume is given by its determinant, i.e. the product of the eigenvalues. It is also rotationally invariant.

$$\det(\mathbf{D}) = \lambda_1 \lambda_2 \lambda_3 \tag{2.15}$$

The above-mentioned scalar parameters that can be computed from tensors are helpful in extracting a specific information from highly-dimensional data. For instance, looking at variations of trace or FA between different populations of subjects may be useful to understand the effect of diseases on diffusion.

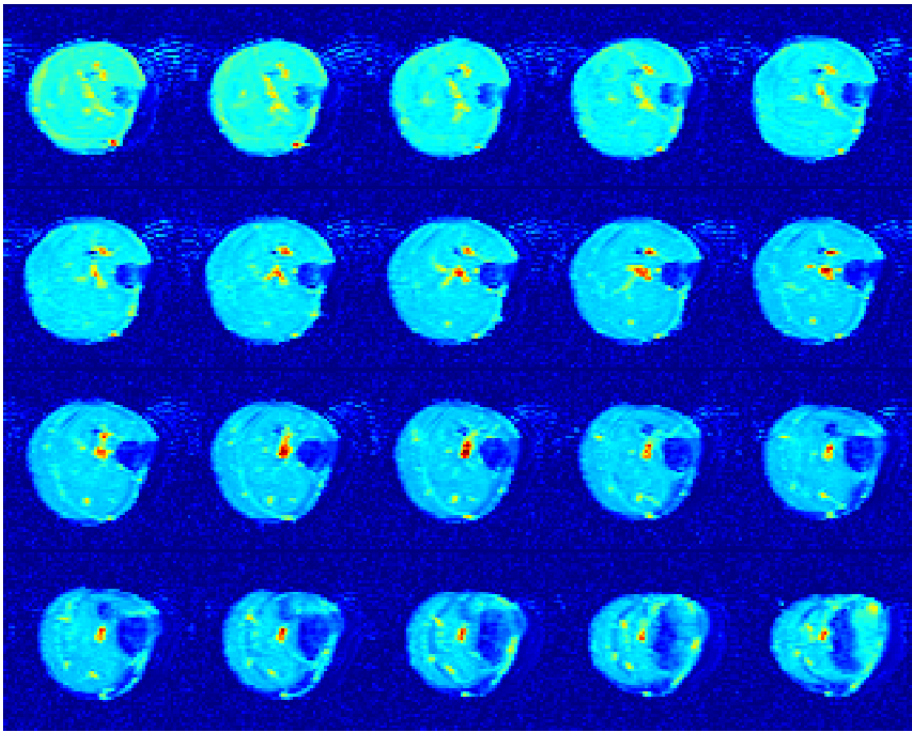


Figure 2.2: Axial slices of the trace map of the calf muscle. Diffusivities are low so the intensities are rescaled. Blue indicates lower mean diffusivity.

Another important way to recover visually appealing information from DTI data is fiber tractography. It consists in recovering 3D trajectories from tensors. The basic technique for reconstructing these curves is to integrate the field of principal directions of diffusion. This is analogous to a classical problem in mechanics where one considers a dynamical system where the dynamics are governed by a first order differential equation. By plotting the field of velocities one can recover the possible trajectories of the system by integrating this field, as can be seen schematically in [Fig.2.1 (b)]. In the case of fiber tractography, the velocities are analogous to the principal directions of diffusion. A tractography algorithm has several parameters including termination criteria, for example it makes sense to stop the fiber tracking if FA is too low, since the uncertainty on the principal direction of diffusion would be too high (in the limit case of isotropic diffusion, all the directions are plausible) or in the case where the curvature is too important as one expects the fiber to have a smooth trajectory. A review of such algorithms

can be found in [83]. Using the principal direction of diffusion may be a limitation in regions where anisotropy is not sufficiently high. In order to address this shortcoming, more evolved methods using the whole tensor information were proposed and are based on the deflection of the orientation of the fiber trajectory using the local tensor [71]. These are deterministic (streamline) approaches to fiber tractography. Different stochastic variants were proposed [45, 131] in order to take into account uncertainty in the fiber orientation by inferring fiber directions using both the observations and a probability model. This allows for more tracking robustness especially in areas with low anisotropy.

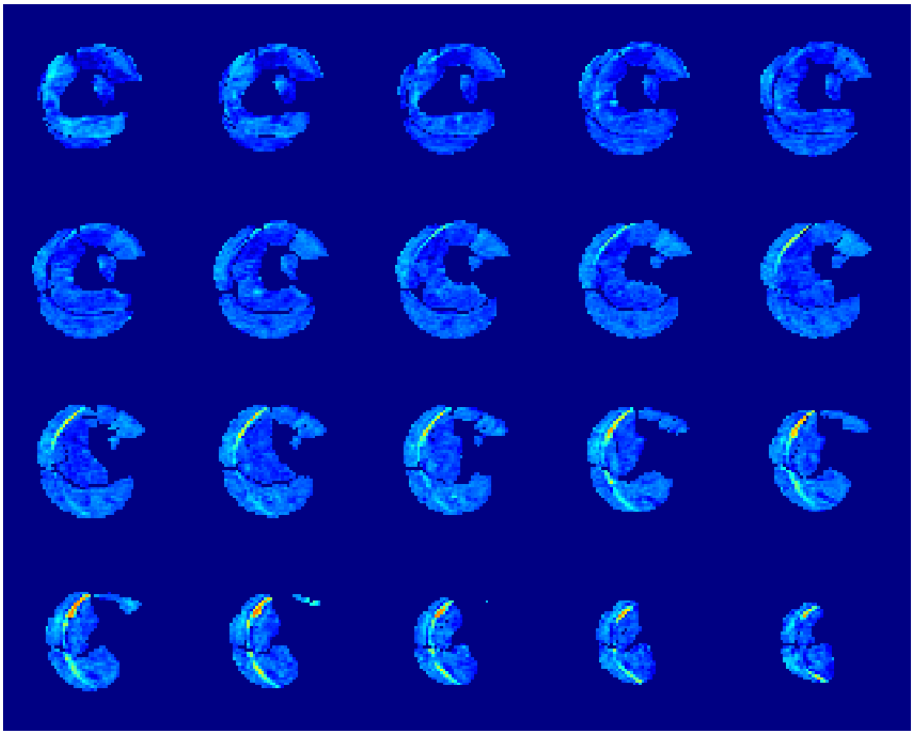


Figure 2.3: Axial slices of the FA map of the calf muscle. The intensities indicate true values of anisotropy (intensities are not rescaled). A mask is used to remove the background. Red indicates high anisotropy and blue indicates low anisotropy.

2.4 DT-MRI Acquisition Scheme

A DT-MRI acquisition consists of several 3D volumes according to different directions of diffusion-sensitizing magnetic field gradient. The fundamental relationship that links diffusion tensors to the acquired signals is called the Stejskal-Tanner equation. Given n magnetic field gradients $(\mathbf{g}_k)_{k=1...n}$ which are unit-norm vectors, the signal at a voxel $\mathbf{x}$ for the direction k as explained in [105] should respect the following condition

$$S_k(\mathbf{x}) = S_0(\mathbf{x}) \exp\big(- b\mathbf{g}_k^t \mathbf{D}(\mathbf{x})\mathbf{g}_k\big)$$

with the tensor $\mathbf{D}(\mathbf{x})$ being the tensor to estimate at the position $\mathbf{x}$, b a value that depends on the acquisition settings and S_0 the baseline or diffusion-free image (basically a T2-weighted image). The Stejskal-Tanner equation indicates that the signal shows an exponential decay with respect to the amount of diffusion in the direction $\mathbf{g}_k$ given by $\mathbf{g}_k^t \mathbf{D}(\mathbf{x})\mathbf{g}_k$. In order to estimate the six parameters of the tensor at each voxel of the 3D volume, at least six directions should be available, along with the baseline signal. We provide in [Fig.2.4] an axial slice of a diffusion-free image (S_0), of a diffusion-weighted image (S_k for a selected direction k) and the estimated tensors represented as a field of ellipsoids and colored according to their principal directions of diffusion. It can be seen that the diffusion-weighted image has a low-contrast, and a rather low quality, with a high amount of noise (clearly noticeable in the background) and a chemical shift artifact around the bone (the tibia is replicated).

Recall that MRI is based on the net magnetization of hydrogen nuclei in the human body that is aligned with the static magnetic field of the scanner magnet when the latter is applied. Using a radiofrequency pulse at the Larmor frequency (resonance frequency), the net magnetization can be flipped from the equilibrium position and the characteristic times (T1 and T2 respectively) of the longitudinal magnetization recovery and transerve magnetization decay provide contrast between tissues. For a detailed explanation of MRI physics, we refer the reader to the textbook [80]. The diffusion MRI pulse sequence exploits two magnetic field gradient pulses in order to be sensitive to the protons motion.

The physical intuition behind the modality is that for moving spins, ap-

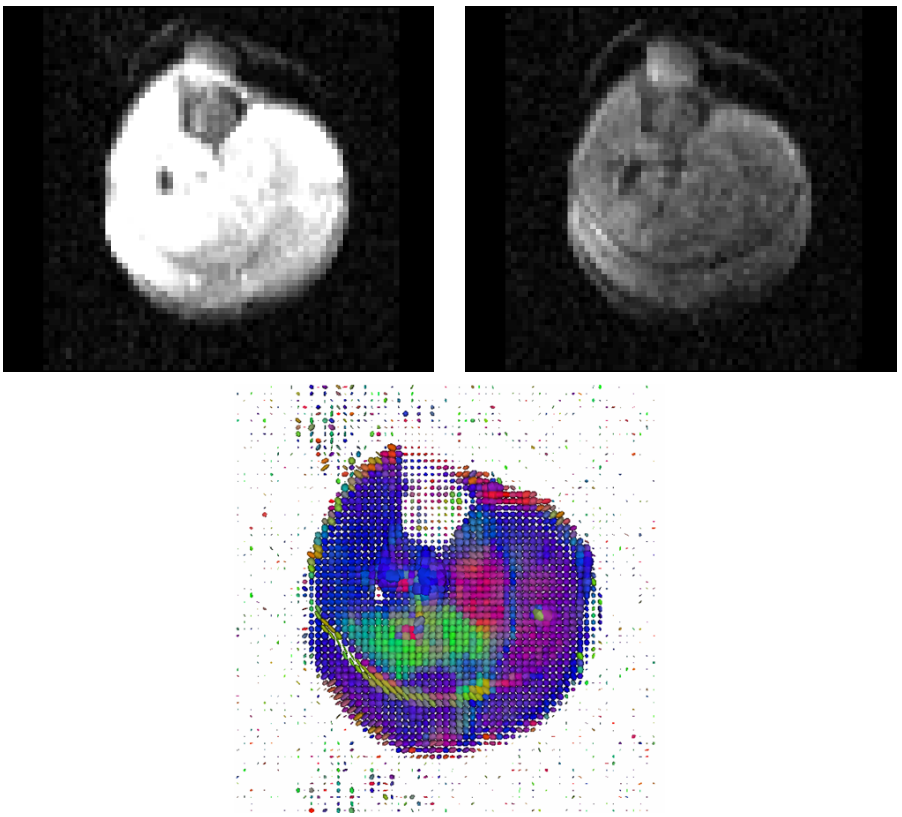


Figure 2.4: Axial slice of the calf muscle. Top left: a diffusion-free image. Top right: a diffusion-weighted image. Bottom: estimated tensors represented as ellipsoids with the RGB colors encoding the principal directions of diffusion

plying the diffusion gradient according to a precise scheme will introduce a phase shift in the signal, while for static spins no phase shift will be observed. The acquisition scheme is explained schematically in [Fig.2.5]. First a 90° RF (Radio-Frequency) pulse is applied so that the longitudinal spin magnetization is removed. The result of this pulse is that the magnetization lies in the transverse plane. Then a first gradient pulse is applied to introduce a phase shift for a duration δ. This is followed by a 180° pulse that inverses the first phase shift and a second gradient pulse that introduces another phase shift.

It can be shown that the total phase shift Φ resulting from applying this

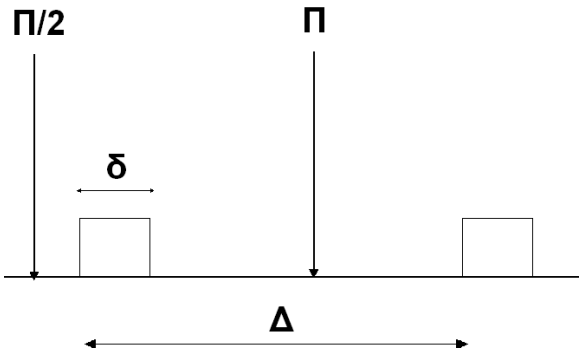


Figure 2.5: Pulse sequence of a DTI acquisition. Arrows indicate Radio-Frequency (RF) pulses, "landings" are diffusion gradient pulses. The $\frac{\pi}{2}$ RF pulse flips the net magnetization into the transverse plane, the π pulse inverts the magnetization. The second magnetic field gradient pulse rephases only static spins, resulting in a lower signal for moving spins. δ is the duration of a gradient pulse and Δ is the time separating the application of the two gradients.

sequence is

$$\Phi = \gamma \delta \mathbf{g}^t \mathbf{r} \tag{2.16}$$

where γ is the gyromagnetic ratio, $\mathbf{g}$ is the direction of the diffusion gradient and $\mathbf{r}$ is the spin displacement in the time interval between the two diffusion gradient pulses. This equation relates the signal phase shift to diffusion and one can see that in the particular case of a static spin ($\mathbf{r} = 0$) there is no phase shift ($\Phi = 0$). Intuitively, the first gradient pulse has a dephasing effect on spins. If a spin is static, then the second pulse compounded with the phase inversion of the 180° pulse compensates the effect of the first pulse. However for a moving spin, there is a phase shift and the produced signal has a smaller amplitude.

The measured diffusion signal $S_{\mathbf{g}}$ at a voxel is related to the average value of $\exp(i\Phi)$ for the spins within the voxel:

$$S_{\mathbf{g}} = S_0 E(\exp(i\Phi)) = S_0 E\left(\exp(i\gamma\delta\mathbf{g}^t\mathbf{r})\right) \tag{2.17}$$

The expectation of $\exp(i\Phi)$ may be rewritten using the diffusion propagator $p(\mathbf{r}|t)$ that gives the probability of a displacement $\mathbf{r}$ given a diffusion time

t. This leads to the following expression:

$$S_{\mathbf{g}} = S_0 \int_{\mathbb{R}^3} p(\mathbf{r}|t) \exp(i\gamma\delta\mathbf{g}^t\mathbf{r}) d\mathbf{r} \tag{2.18}$$

By introducing $\mathbf{q} = \gamma\delta\mathbf{g}$, one can rewrite this equation as

$$S_{\mathbf{q}} = S_0 \int_{\mathbb{R}^3} p(\mathbf{r}|t) \exp(i\mathbf{q}^t\mathbf{r}) d\mathbf{r} \tag{2.19}$$

This means that $\frac{S_{\mathbf{q}}}{S_0}$ is the Fourier transform of the diffusion propagator, or equivalently that the diffusion propagator is the inverse Fourier transform of the quantity $\frac{S_{\mathbf{q}}}{S_0}$.

DT-MRI makes the fundamental assumption that the diffusion propagator is Gaussian. This particular case allows to relate the diffusion signal to the tensor $\mathbf{D}$ that corresponds to the covariance information of the propagator through the Stejskal-Tanner equations, and the b value which is the diffusion-weighting factor introduced in [17] and given by:

$$b = \gamma^2\delta^2||\mathbf{g}||^2 \left(\Delta - \frac{\delta}{3}\right) \tag{2.20}$$

where Δ is the time between the two diffusion gradient pulses. The b value depends therefore on the acquisition parameters.

The Stejskal Tanner equations show that DTI allows to relate in a tractable way the diffusion signals with the tensor at each spatial location, or equivalently to recover in appealing way the diffusion propagator by estimating the tensor from the available measurements. Even if the underlying diffusion model is simple, it allows for fairly easy computations.

Chapter 3

A Survey of DTI-based Muscle Studies

3.1 Introduction

In this chapter, we present a survey of the existing work using DTI for a clinical study of the skeletal muscle. We also detail the different challenges and difficulties that are inherent to muscle diffusion imaging. The different findings that DTI led to in the study of the skeletal muscle span a wide range. These include injury and tissue damage assessment due to tears, changes due to contraction and elongation, influence of exercise load, estimation of pennation angles and fiber orientations, aging effects on water diffusion, gender differences as well as the evaluation of physiological criteria that give an indication about muscle strength. The variability of different DTI parameters across multiple muscle regions and for bundles of specific functional roles was also studied. The experiments were done on humans as well as on small animals.

Theses studies show that DTI is a very promising modality with various clinical applications for the assessment of the human skeletal muscle and a better understanding of the effect of disease-induced damage on fiber architecture. It is also noticed that diffusion tensor images carry information about the microstructure of the muscle tissue, such as the diameter of the myofibers which can be related to diffusion, as well as about the arrangement of these fibers and their sheet organization. Moreover, it may be

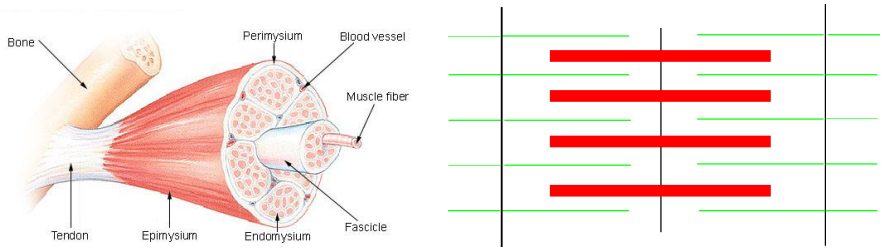
informative about water diffusion in both the intracellular and extracellular compartments. Indeed, DTI has several advantages compared to other modalities that may be used for the study of the muscle, such as ultrasound [27, 96]. Even though ultrasound can provide measurements of pennation (orientation) angles and fiber length, the majority of the experiments done in the field use 2D images which is a limitation that is overcome in DTI. 3D information is important since the pennation angles as well as fiber length may vary spatially. However the reliability of the obtained DTI findings is questionable due to the high amount of noise, the partial volume effect and the presence of chemical shift artifacts that may introduce some uncertainty regarding the estimated DTI parameters and the inter-subject comparison.

This chapter is organized as follows: in section 3.2, we provide a brief overview of skeletal muscle architecture. Neuromuscular diseases are presented in section 3.3. In section 3.4, we explain the difficulties of diffusion tensor imaging of the skeletal muscle. Current findings and experimental results of DTI studies of the muscle are discussed in 3.5. Section 3.6 is dedicated to the open issues and to the motivations of this work with respect to existing studies.

3.2 A Brief Overview of Muscle Architecture

The information of this section were collected from [102]. The human skeletal muscles are attached to the bones and are essential for movement through their contraction and elongation. One of the main differences between the skeletal muscles and the other types of muscles (cardiac muscle, smooth muscle) is that we can control the motion of the skeletal muscles voluntarily. Muscle tissue is composed of long and cylindrical cells which are separated by a connective tissue, called the endomysium. The striations that one observes in muscles are due to the arrangement of proteins within cells.

Muscle fiber bundles have a highly organized and hierarchical structure, as can be seen in [Fig.3.1 (a)]. This structure is modular since substructures are assembled into more sophisticated units. In fact, several muscle fibers compose a fascicle, which is surrounded by a tissue layer called the perimysium. Multiple fascicles constitute a fiber bundle. Each muscle fiber



(a) (b)

Figure 3.1: (a) Skeletal muscle structure [1]. (b) The basic unit composing a myofibril (sarcomere). The thick filaments (in red) are composed of Myosin protein and the the thin filaments (in green) are composed of Actin.

is composed of myofibers, which can be decomposed into an arrangement of successive basic units called sarcomeres (shown schematically in [Fig.3.1 (b)] which are joined along the long axis of the fiber. Since myofibers can have different numbers of sarcomeres across fiber bundles, the length of the fibers can vary. Indeed, for example the soleus fibers are shorter than the tibialis anterior fibers. It is important to note that this variability in the fiber length generates a variability in the maximum length change, called maximum fiber excursion, whether by shortening or lengthening. A longer fiber will contain more sarcomeres and hence has a wider excursion and motion range.

There are two different kinds of fibers depending on their activity response: the slow twitch fibers and the fast twitch fibers. The slow twitch fibers are aerobic and used for an effort that takes an important amount of time. The fast twitch fibers are anaerobic and are used to produce force over a short period of time.

Fibers may differ not only in the length, but also in their orientation or pennation with respect to the line of action (pull) of the muscle, as can be seen in [Fig.3.2]. For example a fiber bundle can be composed of fibers that run parallel to the line of action. In other cases, the fibers can have a pennation angle with respect to the line of action. Different geometries are possible: in a unipennate muscle, all the fibers have the same oblique direction. In a bipennate muscle, two different directions are present. A multipennate muscle has several fiber orientations. Fiber pennation is highly informative as far as muscle mechanics are concerned, and more particularly

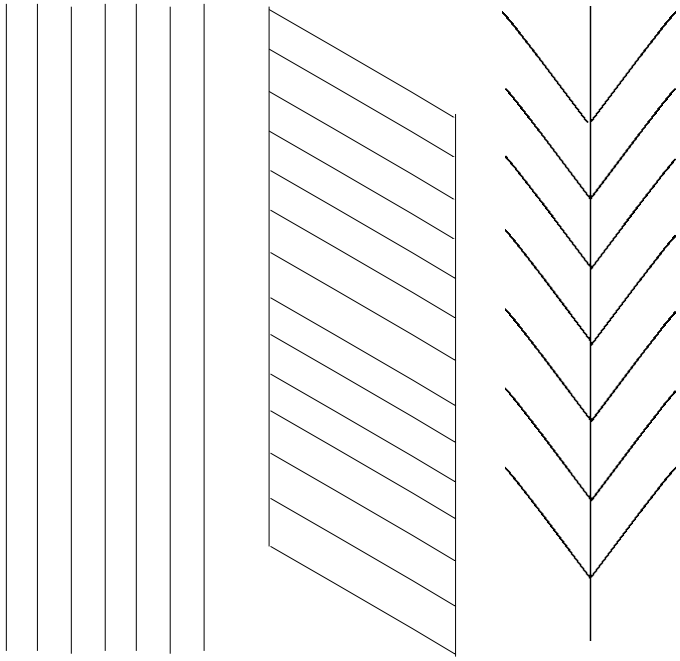


Figure 3.2: Different types of fiber organization. From left to right: parallel arrangement, unipennate muscle and bipennate muscle. The line of pull (action) of the muscle is supposed to be vertical.

force generation. The actual transmitted force along the line of action of the muscle is indeed less important than the force generated in the direction of the muscle fibers, and the two force values are linked through the cosine of the pennation angle as can be seen in the following relationship:

$$F_{line} = F_{fiber} * \cos(\theta) \tag{3.1}$$

where F_{line} is the transmitted force, F_{fiber} is the generated force and θ is the pennation angle of the fiber. Note how this mathematical relationship corroborates the fact that there is a loss in force generation due to pennation with respect to the case where the fibers run parallel to the line of action ($\theta = 0$). Despite this disadvantage, pennation is a space-saving arrangement because it allows to pack short fibers in a small cross-sectional area. This compensates for the loss of force due to pennation by an increase in the number of fibers that consequently develops the muscle strength. Pennation implies that one should consider a factor called Physiological Cross-Sectional

Area (PCSA) instead of the Anatomical Cross Sectional Area (ACSA) to measure muscle force. The ACSA is the sum of the cross-section areas of the fibers that compose the bundle when it is cut in a direction perpendicular to the line of pull of the muscle. This measure can be effective and informative in the case of parallel arrangement, however it is insufficient in the case of pennate muscles because it understates the packing effect due to pennation. One should therefore consider the PCSA which is defined as follows:

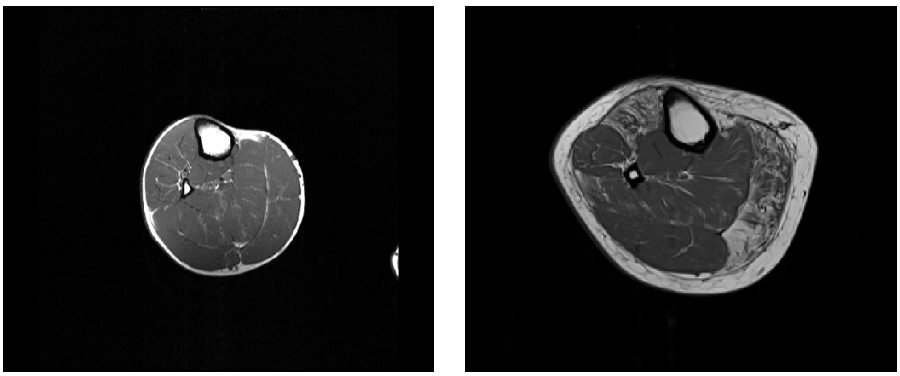
$$\text{PCSA} = \frac{V * \cos(\theta)}{L} \tag{3.2}$$

where V is the muscle volume and L is the fiber length. Note how this formula still indicates that pennation reduces the generated force, while taking into account the fact that numerous short fibers can be arranged in a pennate muscle, thus compensating the force transmission loss. It is known for example that the soleus, despite being composed of relatively short fibers, is one of the strongest muscles in the human body.

One of the most damaging diseases that affect the architecture of the skeletal muscle are myopathies. The next section provides the reader with basic information about these diseases.

3.3 Myopathies

Myopathies are neuromuscular diseases that can be subdivided in two groups: myopathies of genetic origin like muscular dystrophies (Duchenne Dystrophy, FacioScapuloHumeral Muscular Dystrophy , etc.) which are primarily caused by an alteration of fibers which degenerate progressively, and myopathies that affect a muscle previously healthy like inflammatory myopathies. These diseases can affect individuals at all ages, from infants to adults. They concern 30000 to 40000 persons in France and 400 new cases are ascertained each year according to [2]. The most frequent myopathy among children is the Duchenne myopathy, with nearly one diseased case over 3500 male neonates. The clinical symptoms vary from a subject to another, and the disease effect vary also across different muscle groups for the same subject. They result in functional anomalies of the muscle including fat infiltration, weakness, atrophy, paralysis, loss in muscle strength,



(a) (b)

Figure 3.3: Axial slice of the calf muscle. (a) healthy subject. (b) diseased subject

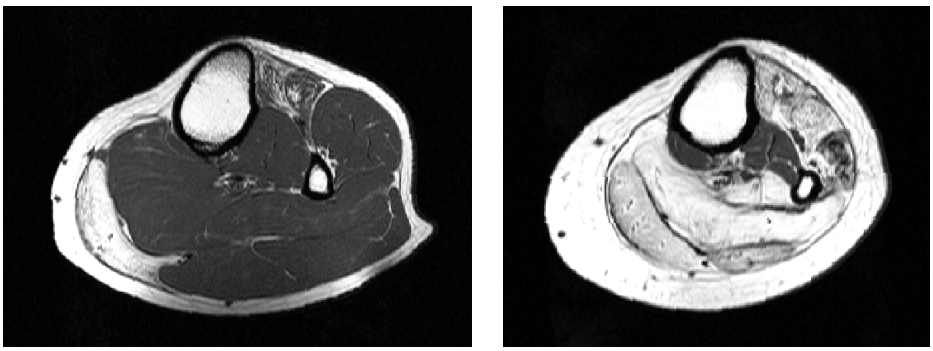


Figure 3.4: Axial slice of the calf muscle for two diseased subjects. Fat replaces the muscle tissue.

muscle pain due to exercise, etc. These various signs of dysfunction may help a physician in establishing his diagnosis. An example of axial slices of T1-weighted images for a healthy and a diseased subject are presented in [Fig.3.3] where one can see that fat replaces the muscle. Note that in this case the infiltration is localized (affects a specific region of the muscle) and that this phenomenon is not well understood. More examples of T1 images of diseased subjects are shown in [Fig.3.4] where one can see cases of severe fat infiltration.

There are a few procedures available for disease identification. For example muscle tissue biopsy can be performed and consists in extracting a small sample of tissue that can be analysed using the microscope in order to

detect anomalies. An electromyogram (EMG) measurement may be helpful since the assessment of the electrical activity of the muscle may ascertain a muscle weakness. T1 and T2 weighted MRI acquisitions are also helpful to detect signal heterogeneities due to the presence of affected regions and localize them spatially. Moreover, for inherited myopathies blood tests can be carried out for a molecular study of the DNA for myopathies with known genes and mutations. They can also be performed to look for an elevated protein level that may reflect muscle fiber necrosis.

Overall there is a lot of room for progress both in the areas of diagnosis and treatment of myopathies. Detection of the disease at an early stage is important so that appropriate measures can be undertaken. Some myopathies like Duchenne muscular dystrophy provoke serious damage and are life-threatening. Therefore, it is crucial to better understand the effect of these diseases on the muscle and to provide qualitative and quantitative tools that may help in designing a cure and in detecting the disease. In this context, diffusion tensor imaging can be a suitable modality for the evaluation of muscle condition. However such an acquisition can be very challenging as explained in the next section.

3.4 Challenges of Muscle DTI

Diffusion tensor acquisitions of the human skeletal muscle are technically challenging, as reported in [70]. One of the main underlying difficulties of such an acquisition is an intrinsic property of the muscle tissue: it has a short spin-spin (T2) relaxation time. Moreover, muscle has a lower water proton density than brain white matter. This results in images of overall poor quality, with a relatively low SNR.

Another major problem is the presence of chemical shift artifacts caused by the presence of fat. The importance of the artifacts varies with the amount of fat in the tissue, and may degrade dramatically the quality of the acquired images. Other artifacts are caused by the subject motion. Moreover, in imaging techniques like EPI (Echo-Planar imaging), important field gradient eddy currents result in image distortion. It is also known that diffusion images are affected by a partial volume artifact, i.e. in a specific voxel, the signal is actually the resulting signal from multiple components

in the voxel. Given the size of the voxel usually used (each dimension is of the order some millimeters) and the radius of a myofiber ($\approx 50\mu m$), we are unable to reconstruct the actual anatomical myofibers. Instead we compute an "average" of a bundle of several fibers in each voxel. Thus the fibers obtained from DTI tractography differ from the fibers that one can observe from histology. This is a major impediment in the way of replacing invasive techniques like biopsy by non-invasive DTI.

In [32], it is emphasized that the presence of noise, the effects of the setting of the acquisition parameters and the partial volume artifacts pose important difficulties for the use of DTI in clinical studies of the human skeletal muscle. A model that combines both muscle tissue and adipose tissue for tensor simulation is proposed and tensor estimation is done in a weighted least-square fashion. It is shown that in some conditions the findings of DTI parameters, including not only diffusivities by also directions of diffusions, may become unreliable. The study also confirms the stringent requirements of DTI studies as far as SNR is concerned compared to brain white matter DTI.

Overall finding an optimal diffusion acquisition scheme of the muscle is still an open issue, although several possible ways to improve the quality and the reliability of the obtained images were proposed as will be discussed later.

We now review the literature of clinical studies of the skeletal muscle using DTI.

3.5 Current Findings and Insights

There has been a growing interest in the application of DTI for the analysis of the skeletal muscles. In particular several studies focused on extracting tensor-related quantitative parameters and visual qualitative results in order to better characterize and understand the muscle tissue.

For instance, the evaluation of the Physiological Cross Sectional Area (PCSA) of muscles in a non-invasive way using DTI of six volunteers is discussed in [48]. PCSA is an important parameter that allows to measure the strength of the muscle, i.e. the maximum force that it can generate. The study shows both analytically and experimentally that there is a linear

relationship between PCSA and the third principal diffusivity. This corroborates the interpretation that the latter accounts for diffusion across the cross-sections of the fibers and that it is related to the radius of the fiber. In particular, the significant differences among some muscle groups in the third eigenvalue can be explained by differences in the average radii of the fibers composing the group.

In [57], a method for quantitative evaluation of fiber tracking in the human skeletal muscle is proposed. The work focuses on the accuracy of tractography results for the tibialis anterior muscle over nine healthy subjects. Data were acquired using a 3T MRI scanner with 10 gradient directions and surface coils to obtain a high Signal-to-Noise Ratio (SNR). It is based on the knowledge of the arrangement of fibers in muscles, in particular ending and starting points. For example the tibialis presents a bipennate structure and the fibers are expected to go from the aponeurosis to the muscle border. The conformity of fiber tracking results to this structure is a criterion of assessment of the quality of the tractography. Other criteria are proposed and used to provide an evaluation of the stopping criteria (FA and curvature in particular) of the tractography algorithm, however the issue of inter-subject variability is still open and the evaluation is sensitive to variations in image quality and to the importance of the fat artifact.

In [36], a study of the effect of muscle contraction on principal diffusivities, FA and ADC is presented. Diffusion tensor images of the calf muscle of twenty healthy volunteers were acquired using a 1.5T MRI scanner. The muscle was imaged at three states: rest, dorsal flexion and plantar flexion. Two muscle groups in the calf were investigated: the medial gastrocnemius and the tibialis anterior. The findings of the study is that the diffusion tensor parameters undergo a change due to the change of the state of the muscle (for example muscle contraction results in an increase of principal diffusivities), and that this change is related the functional roles of the different muscle groups. It is also reported that a visual inspection of fiber tractography can lead to the assessment of the state of the muscle.

In [89], the variations of muscle FA induced by exercise are studied. The images are acquired for ten healthy male volunteers with a 1.5 T MRI scanner and for both the right and the left calves, three muscles are considered: the anterior tibialis , the gastrocnemius and the soleus. The right calves of

two subjects were exercise-loaded and the evolution of FA after exercise over time are reported. The conclusions of the study is that the FA values are higher for the tibialis anterior and that FA decreases for the gastrocnemius and the soleus after exercise but reaches the initial levels a week later.

In [70], an approach for the study of the pennation angles (the orientation of muscle fibers with respect to the aponeurosis) of the tibialis anterior muscle is discussed. Data were acquired from eight healthy subjects using a 3T MRI scanner. The tibialis anterior was segmented manually and points were manually selected on the aponeurosis. The points were then meshed to obtain a surface representation of the aponeurosis. Polynomial fitting allows to compute a tangent plane at specific locations and to obtain the pennation angles. The main finding of the study is that the pennation angles vary with respect to the location due to a spatial variation in the direction of the aponeurosis.

An earlier work on the effect of exercise on DWI [84] suggests an increase in ADC values due to muscle stimulation. The experiments were done on the forearm muscle for nine subjects and considers a bi-exponential diffusion signal to model two compartments in muscle microstructure. For both compartments, the diffusion coefficients increase with respect to the rest state.

In [130], the possibility of using DTI in order to evaluate muscle injury and damage is explored. The study is based on data acquired from twelve subjects of whom eight are healthy and four suffer from muscle injury. Muscles of interest are the gastrocnemius and the soleus. The hypothesis that injuries introduce more disorder in the muscle are confirmed by significant changes in DTI indices. Due to muscle tears, FA is decreased and principal diffusivities are increased. Fiber tractography in damaged areas reveals shorter and less ordered fibers.

The work in [47] studied the influence of aging on DTI parameters. DTI using a 1.5 T MRI scanner of the calf muscle for 38 healthy subjects from different age groups showed an alteration for some of the indices, and this alteration varied among muscle groups. The focus was on the plantar flexors (gastrocnemius and soleus) and the tibialis anterior. While it is shown that the eigenvalues decreased significantly with increasing age for the plantar flexors, reflecting a smaller fibers, they did not vary significantly with age

for the tibialis anterior. However an opposite behavior was noticed for FA. This implies that even if all the muscle groups undergo some changes with increasing age, these changes may differ from a muscle group to another.

The study of the arrangement of fiber fascicles using DTI was explored in [19] on a healthy subject with a 1.5T DT-MRI scan. SPGR (based on an incoherent gradient echo pulse sequence) images were acquired at the same location and the soleus and tibialis anterior were manually delineated. It is noticed that the arrangement of the fibers was found to be in agreement with cadaveric measurements. Indeed, the soleus fibers were found to be short and oblique with respect to the aponeurosis, while the tibialis anterior fibers were found to be long and with a low pennation angle.

A qualitative and quantitative comparison between Magnetic Resonance Spectrum Imaging (MRSI) and DTI in the evaluation of fiber orientation in skeletal muscle is provided in [115, 114]. MRSI spectra of the muscle reveal orientation dependent peaks and therefore are useful to estimate fiber orientation in each muscle group. It is shown that for several muscle groups there is an agreement between the findings of MRSI and DTI and that significant differences in fiber orientation across some muscle groups can be established.

Another study of fiber orientation in the rat tibialis anterior was provided in [38] in order to evaluate the possibility of using DTI for the analysis and the simulation of muscle mechanics where fiber directions provide relevant information for the construction of a finite-element mesh model of the muscle. Comparison with fascicle directions visible directly on high resolution MRI and with an actual longitudinal section through the muscle provides a good agreement with DTI findings. A muscle contraction was simulated using the generated mesh.

Similarly in [109], DTI experiments on the cat semimembranous muscle is performed to estimate the orientation of the fibers and explore their possible use in biomechanical studies. The results obtained from DTI showed that the principal direction of diffusion coincided with the directions of muscle fibers determined through invasive visual inspection. This confirms the anisotropic nature of the muscle tissue, which is important for the analysis of its mechanical properties such as strain and stress.

The joint use of DTI and other MR imaging techniques is discussed

in [18]. The goal is to build models of the musculoskeletal system with two main advantages: overcoming oversimplified assumptions about muscle assumptions and obtaining case-specific models based on these imaging techniques. The work emphasizes the important and promising role of DTI as a non-invasive tool for the study of muscle architecture, in particular the fascicle arrangement. Combined with the other imaging techniques that provide information about muscle and joint dynamics and anatomy, DTI may be useful to design treatments for movement disorders.

In [103] a tetrahedral diffusion gradient scheme is proposed and compared to a standard orthogonal scheme over five volunteers, showing an improvement in terms of SNR. Acquisitions of the calf muscle are done along 4 different unit vectors (in both directions). It is shown that while this scheme does not allow to recover the entire diffusion tensor, it is possible to obtain the off-diagonal elements. Diagonal terms are not recovered, instead their sum (trace) can be computed. By making the hypothesis of axial symmetry, it is possible to approximate the orientations of the principal directions of diffusion in spherical coordinates and to compute two indices of anisotropy. An important finding of this study is that trace seems to have low variance across muscle groups and across subjects, so a significant change in trace may be an indication of tissue damage. It is also reported that there is low variability of the elevation angle θ between muscle groups while an opposite behavior is noticed for the azimuthal angle ϕ.

Another optimized imaging sequence was proposed in [104]. The sequence has the advantage of both limiting the distortion effects induced by Eddy currents and of providing good quality images in terms of SNR using a stimulated echo preparation instead of a standard spin-echo preparation. The acquisition scheme was validated using both phantoms and real data obtained from four healthy subjects. It yielded satisfactory maps of DTI parameters (orientation of principal directions of diffusion, FA, ADC and eigenvalues) for the two heads of the gastrocnemius, the soleus and the tibialis anterior.

The work in [98] is about the optimization of the b parameter in a DTI acquisition sequence at 1.5 T. Using a range of reasonable values of b, the acquisition is done on five volunteers, with the muscles of interest being the lateral gastrocnemius and the tibialis anterior. The study used both

qualitative criteria (the quality of fiber tractography is assessed visually) and quantitative criteria (SNR and fiber density). It is emphasized that a balance should be found between SNR and sensitivity to diffusion. Indeed, for relatively low b values, the SNR is acceptable but the quality of the tractography is degraded and at high b values the SNR is low. The work suggests that the use of a value $b = 625\text{s/mm}^2$ is optimal for a DTI muscle acquisition at 1.5T.

Another parameter that can be optimized is the magnetic field strength, and in [99], it is investigated whether a 1.5 T or a 3 T MRI scanner should be used for muscle imaging. Experiments are done on five subjects and with different b values. Muscles of interest are the two heads of the gastrocnemius and the soleus. The quantitative assessment criterion is the SNR and the study looked for differences of DTI parameters (anisotropy indices, eigenvalues and trace). Different field strengths did not affect DTI parameters apart from the third eigenvalue. This means that to a certain extent diffusion in muscle does not vary between the two values of field strengths used in the experiments. However as expected, imaging at 3T gave better SNR than a 1.5T acquisition.

A validation of fiber tractography in the lateral head of the gastrocnemius of five rats is provided in [33], with images acquired using a 4.7 T scanner. A comparison is done with direct anatomical inspection (dissection). A method for fiber bundling is proposed using a thresholding on a similarity measure and a nearest-neighbor grouping. The similarity measure is based on finding corresponding segments in two fiber tracts and estimating their average amount of overlap. The medial axis is computed along with the pennation angles over the entire length of each bundle. A strong correlation between pennation angles obtained from DTI and from dissection is found.

Another similar evaluation of fiber tractography in the skeletal muscle of six mice is provided in [58]. In order to segment the dorsal flexor muscles from the other muscle groups, the former were exercise-loaded, resulting in an enhancement in their T2-weighted signal. This segmentation was validated using dissection. This allowed to recover the muscle boundaries and to assess the amount of fiber tracts with erroneous trajectories, i.e. not respecting the detected boundaries. The study reports that fiber tracking

yielded rather accurate results.

A theoretical model of water diffusion in muscle is proposed in [49] based on cylindrical geometry. It accounts both for the muscle component and for the extracellular component (the endomysium) with different volume fractions. The assumption of cylindrical geometry implies that diffusion in the fiber component can be modeled using two coefficients, one for the fiber direction and another for the perpendicular direction. This model was fitted to diffusion data from twenty-four subjects to study gender differences. It is ascertained that there are statistically significant differences in DTI parameters between males and females. The diffusion model shows that the proportion of the muscle fibers with respect to the extracellular component is more important for males than for females. This may be an explanation for the noticed gender differences.

Another theoretical model for diffusion in the muscle was proposed in [97]. It is based on approximating the cross-section of the fiber by the simple model of a square and using the explicit expression relating the diffusion signal decay to the length of the square edges. This is used to estimate the radii of the fibers. Experiments were done on a group of control rats and denervated rats, the latter exhibiting subsequent muscle atrophy. It is shown that the estimated fiber diameter for the denervated group is significantly less important than the one of the normal group. An opposite behavior is noticed for FA. This shows that atrophy changes the structure of muscle fibers in a way that can be detected using DTI. A similar study was done on mice in [134] and showed that denervation atrophy resulted in a similar change in DTI parameters (increase of FA and decrease of second and third eigenvalues). It is therefore suggested that the changes in these parameters are good indicators of myofiber atrophy.

While most of DTI studies of the human skeletal muscle focused on the lower leg, other muscle regions, such as the forearm, the thigh and the tongue were studied. For instance, fiber tracking in the human forearm is shown to reflect the anatomical architecture of the forearm and to reveal its different components in [46]. Experiments are done on a healthy volunteer with a 3T scanner. Segmentation of different muscle groups was achieved by using ROI seeding and running fiber tractography.

The tongue has also a complex and heterogeneous structure of skeletal

muscle fibers. It was studied in [85] using DTI for a macroscopic analysis as well as two-photon microscopy for microscopic analysis. For two photon microscopy, an autocorrelation image was built to determine repetitive structures and finding the fiber orientation was done by computing the Hessian and taking the direction of the eigenvector corresponding to the smallest eigenvalue. The experiments were done on three mammalian (porcine) tongues. Both modalities reveal two populations of fibers in the lingual core with nearly orthogonal directions. However DTI is unable to resolve crossings of fibers with different orientations. The feasibility of in-vivo imaging of the human tongue was demonstrated in [50] by extending the method in [85] and validating it on three healthy subjects.

We note also a recent study [64] of the thigh (quadriceps muscle, in particular the vastus lateralis oblique and the vastus medialis oblique) using DTI in order to compare healthy subjects and patients affected by chronic lateral patellar dislocation. Both anatomical and physiological cross-sectional areas as well as pennation angles are computed. Using data from eight subjects, half of whom with patellar dislocation, it is ascertained that the muscle force vectors are more laterally oriented in the subjects with dislocation. This study corroborates the usefulness of DTI in biomechanical studies of the muscles. Moreover, the work in [22] ascertained the feasibility of tractography in human thigh muscles on five volunteers. Using ROI seeding, fiber tracts for several muscle groups of the thigh were reconstructed, showing a concordance with anatomical knowledge.

Inflammatory myopathies were studied using DWI of the thigh muscle in [90]. Results are given for 19 subjects and for two types of inflammatory myopathies, one of which is accompanied by a fat replacement. Typically it is established that muscle inflammation result in increasing values of ADC. In contrast, areas with fat replacement have lower ADC than normal muscle, even if they still show an anisotropic structure. This implies that a fiber structure was partially preserved. A recent study [92] used a more evolved diffusion model that uses a bi-exponential signal in order to quantify capillary perfusion. It is shown that in inflammated regions, the volume fraction of perfusion is lower than that of normal muscle.

3.6 Conclusion

DTI has recently found a wide range of applications for a better understanding of the human skeletal muscle. The rich information carried by tensor images are of high clinical relevance. Indeed several studies mentioned above have shown a modification of DTI-derived scalar indices or fiber tracts following injuries or structural disorder in the muscle. DTI offers also possibilities for a better modeling of the mechanical behavior of fiber bundles, which allows computer-based simulations to help physicians create suitable solutions and treatments for muscle dysfunction.

We note however that the use of medical image analysis algorithms for processing of muscle DTI has been tentative and limited. Apart from the algorithm in [33], no fiber clustering algorithm was proposed in the field and the studies are heavily based on manual segmentations. Moreover, diffusion tensor segmentation of muscle images did not attract much attention, when compared to brain studies. The same applies for other areas such as denoising and registration.

In the following chapters of this thesis, we propose denoising, segmentation and registration algorithms for DTI of the human skeletal muscle and show their usefulness for a better understanding of the muscle architecture.

Chapter 4

A Convex Framework for DTI Estimation and Regularization

4.1 Introduction

Diffusion tensors are computed from diffusion signal measurements. DTI is therefore based on inferring the estimation of a field of 3×3 symmetric positive definite matrices from these measurements. However, this inference is a tricky issue as the DTI experimental protocol yields noisy observations due to the diffusion-sensitizing magnetic gradient. Furthermore, the clinical protocols refer to relatively low magnet strength, or a rather low signal-to-noise ratio. Therefore, signal reconstruction is crucial to obtain an appropriate estimate of the tensor field and for subsequent use of this estimate in applications like fiber tractography.

Several methods have been proposed to address diffusion tensor estimation and regularization. The method presented in [123] presents an explicit estimate of diffusion tensors using a linear decomposition over a dual-tensor basis. In [30], a two-step regularization was proposed consisting of the restoration of the principal diffusion directions using a total variation-model followed by the smoothing of the eigenvalues using an anisotropic tensor-driven formulation. In [14], the maximization of a log-posterior probability based on the Rician noise model is considered to smooth directly the

diffusion-weighted images. A heat diffusion on a graph is proposed in [132]. This graph diffusion leads to the regularization of the tensor field and the positive definiteness of the regularized field is enforced by construction. A Bayesian model based on a Gaussian Markov Random Field was used in [79] to smooth the diffusion tensors. In [25], the authors consider the tensors as lying on a Riemannian manifold and use the corresponding distance to derive a local weighted averaging for DTI denoising. Tensors are assumed to be positive-definite matrices which was taken into account in [35] where an anisotropic filtering of the L^2 norm of the gradient of the diffusion tensor was considered and their proposed PDE scheme constrains the estimation to lie on this space. A non-spectral variant where matrix-based manifold flows are defined is also discussed. Such a concept was further developed in [118] where a variational method was proposed that aimed to minimize the L^p norm of the spatial gradient of the diffusion tensor under a constraint involving the non-linear form of Stejskal-Tanner equation. A non linear diffusion scheme is described in [121] where smoothing is made direction-dependent using a diffusion matrix in the PDE system. More recently, in [43] a joint reconstruction and regularization was proposed in the context of an energy minimization in a Log-Euclidean framework. The existing variational methods focused disproportionately on enforcing the positive-definiteness constraint, with the regularization term usually chosen as a function of the norm of the gradient. The main limitation of most of the above-mentioned methods is the nature of the cost function (non-convex) that entails a preliminary initialization step, while little attention was paid to defining appropriate smoothness components that account for the expected nature of tensors.

In this chapter we propose a new variational approach to jointly estimate and regularize diffusion tensor images. We use a convex energy functional which combines the linearized form of Stejskal-Tanner equation as a data fidelity term and a new regularization term involving precalculated weights which measure the similarity between neighboring tensors. We show the results of our method both on synthetic datasets and real data of diffusion tensor muscle images.

4.2 DTI Estimation and Regularization

Let us assume that n DTI acquisitions $(S_k)_{k=1...n}$ with respect to different magnetic gradient directions $(\mathbf{g}_k)_{k=1...n}$ are available. Ideally, the expected signal at a voxel $\mathbf{x}$ for the direction k as explained in [105] should respect the following condition

$$S_k(\mathbf{x}) = S_0(\mathbf{x}) \exp\big(-b\mathbf{g}_k^t \mathbf{D}(\mathbf{x})\mathbf{g}_k\big) \tag{4.1}$$

with the tensor $\mathbf{D}$ being the unknown variable and b a value that depends on the acquisition settings. The estimation of the tensors in the volume domain Ω can be done through direct inference (6 acquisitions are at least available), which is equivalent to minimizing:

$$E_{data}(\mathbf{D}) = \int_\Omega \sum_{k=1}^{n} \Big(\log\big(S_k(\mathbf{x})/S_0(\mathbf{x})\big) + b\mathbf{g}_k^t \mathbf{D}(\mathbf{x})\mathbf{g}_k\Big)^2 d\mathbf{x}$$

This energy is based on the linearized diffusion tensor model which is reasonable for moderate values of SNR [95]. Such a direct estimation is quite sensitive to noise, on the other hand, it refers to a convex term, which is rather convenient when seeking its lowest potential. The most common approach to account for noise is through the use of an additional regularization term which constrains the estimation of $\mathbf{D}$ to be locally smooth. One of the most prominent uses of DTI is fiber extraction. Therefore it is natural to assume that locally these fibers do have similar orientations. In such a context, the tensor can be expressed as a linear combination of the tensors lying in its neighborhood since they are likely to represent the same population of fibers. Such a regularization constraint was introduced in the case of image restoration in [9]. This assumption still holds at the boundaries between different groups of fibers as long as the linear combination is thoroughly chosen to ensure that the contribution of tensors belonging to a different fiber population is negligible. Indeed, in order to reconstruct a tensor at a specific location, one should assign a more important weight to neighbors that are similar to the considered tensor. This leads us to define

the following regularization component:

$$E_{smooth}(\mathbf{D}) = \int_{\Omega} \left|\left| \mathbf{D}(\mathbf{x}) - \frac{1}{Z(\mathbf{x})} \int_{\mathbf{y}\in\mathcal{N}_{\mathbf{x}}} w(\mathbf{x},\mathbf{y})\mathbf{D}(\mathbf{y})d\mathbf{y} \right|\right|_F^2 d\mathbf{x}$$

where $w(\mathbf{x},\mathbf{y})$ reflects the similarity between tensors $\mathbf{D}(\mathbf{x})$ and $\mathbf{D}(\mathbf{y})$, $||A||_F$ being the Frobenius norm $||A||_F = \sqrt{tr(A^t A)}$ and $Z(\mathbf{x})$ is a normalization factor, i.e $Z(\mathbf{x}) = \int_{\mathbf{y}\in\mathcal{N}_{\mathbf{x}}} w(\mathbf{x},\mathbf{y})d\mathbf{y}$.

We can note that we consider here a higher-order model than the gradient of the tensor field, i.e. a weighted Laplacian. Indeed one can define a discrete Laplacian operator by considering that the tensors can be represented as the nodes of a graph, with edges linking neighboring nodes. The weighted Laplacian is simply the sum over the neighbors of the differences between tensors weighted by the values affected to the graph edges, as can be seen in the following equation:

$$\triangle_{\mathbf{x}}(\mathbf{D}) = \frac{1}{Z(\mathbf{x})} \sum_{\mathbf{y}\in\mathcal{N}_{\mathbf{x}}} w(\mathbf{x},\mathbf{y})(\mathbf{D}(\mathbf{y}) - \mathbf{D}(\mathbf{x})) \tag{4.2}$$

where $\triangle_{\mathbf{x}}(\mathbf{D})$ is the discrete Laplacian of the tensor field in the position $\mathbf{x}$. Since the regularization energy is based on the norm of the Laplacian, the model is also more accurate than the underlying assumption of total-variation based approaches where the tensor field is considered piecewise constant

The most critical aspect of such an approximation model is the definition of weights, measuring the similarity between tensors within the local neighborhood. The use of Gaussian weights is a common weight's selection, i.e $\left[w(\mathbf{x},\mathbf{y}) = e^{\frac{-d^2(\mathbf{D}(\mathbf{x}),\mathbf{D}(\mathbf{y}))}{2\sigma^2}}\right]$, where $d(.;.)$ is a distance between tensors and σ a scale factor. Hence, with a suitable choice of σ, the regularization energy is edge-preserving. Indeed, high values of σ entail an important regularization while low values imply a more stringent choice of the neighbors contributing to the tensor $\mathbf{D}(x)$.

In the context of simultaneous estimation and regularization it is more appropriate to define similarities directly on the observation space rather than the estimation space. Such a choice will lead to a tractable estimation, while preserving the convexity of the cost function. Our distance defini-

tion as well as our minimization step are based on the representation of symmetric positive semi-definite matrices S^3_+ as a convex closed cone in the Hilbert space of symmetric matrices S^3, where the standard scalar product is defined by $\langle A, B\rangle_F = tr(A^t B)$ which induces the corresponding Frobenius norm.

4.2.1 Measuring Similarities from Diffusion Weighted Images

We aim at simultaneously estimating and smoothing the tensor field, therefore the weights $w(\mathbf{x}, \mathbf{y})$ in E_{smooth} should be precalculated using the raw data. The most straightforward estimation of the distances can be done through the algebraic distance between the $\log(S_k/S_0)$ for two neighborhood voxels in any direction

$$d\big(\mathbf{D}(\mathbf{x}), \mathbf{D}(\mathbf{y})\big) = \frac{1}{b}\sqrt{\sum_{k=1}^{N}\Big(\log\big(S_k(\mathbf{x})/S_0(\mathbf{x})\big) - \log\big(S_k(\mathbf{y})/S_0(\mathbf{y})\big)\Big)^2}$$

One can easily show that such an expression does not reflect similarity between tensors according to the norm $||.||_F$. In fact, based on [123], this leads to

$$d\big(\mathbf{D}(\mathbf{x}), \mathbf{D}(\mathbf{y})\big) = \sqrt{\sum_{k=1}^{N}\Big(\mathbf{g}_k^t\big(\mathbf{D}(\mathbf{x}) - \mathbf{D}(\mathbf{y})\big)\mathbf{g}_k\Big)^2} = \sqrt{\sum_{k=1}^{N} < \mathbf{D}(\mathbf{x}) - \mathbf{D}(\mathbf{y}), \mathbf{G}_k >_F^2}$$

where $\mathbf{G}_k = \mathbf{g}_k\mathbf{g}_k^t$ do not form necessarily an orthonormal basis. We use a Gram-Schmidt orthogonalization scheme to calculate an orthonormal basis $\widetilde{\mathbf{G}_k}$ such that $\widetilde{\mathbf{G}_k} = \sum_l \alpha_{kl}\mathbf{G}_l$ (each new vector of the new basis is a linear combination of the vectors of the initial basis). This procedure allows us to have an approximation of $||\mathbf{D}(\mathbf{x}) - \mathbf{D}(\mathbf{y})||_F$ directly from the raw data S_k and S_0 as follows

$$\begin{aligned} ||\mathbf{D}(\mathbf{x}) - \mathbf{D}(\mathbf{y})||_F &= \sqrt{\sum_{k=1}^{N} < \mathbf{D}(\mathbf{x}) - \mathbf{D}(\mathbf{y}), \widetilde{\mathbf{G}_k} >_F^2} \\ &= \frac{1}{b}\sqrt{\sum_{k=1}^{N}\Big(\sum_l \alpha_{kl}\big(\log\big(S_k(\mathbf{x})/S_0(\mathbf{x})\big) - \log\big(S_k(\mathbf{y})/S_0(\mathbf{y})\big)\big)\Big)^2} \end{aligned}$$

4.2.2 Semi-Definite Positive Gradient Descent

One now can seek the lowest potential of the cost function towards recovering the optimal solution on the tensor space. The present framework consists of a convex energy (a quadratic function of the tensor field) with a global minimum which can be reached using a projected gradient descent on the space of semi-definitive positive matrices. The projection from S^3 onto S^3_+ denoted by $\Pi_{S^3_+}$ is well defined and has an explicit expression. Indeed, projecting M amounts to replacing the negative eigenvalues in its spectral decomposition by 0 [35, 59]. Note that we minimize over the set of semi-definite positive matrices because it is topologically closed, as opposed to the set of definite positive matrices. For a suitable choice of a time step dt that shall be discussed later, the projected gradient of a convex functional over a closed, non-empty convex set of a Hilbert space is convergent. More important is the fact that in areas where the flow is decreased in some directions (for example due to pathologies), the amount of diffusion can be very low. Therefore, for an adequate estimation in regions with modified or anomalous diffusion properties we do not penalize low eigenvalues with priors like the one used in the Log-Euclidean framework [43]. In the current setting, the problem is well posed and the projected gradient descent algorithm is convergent for a suitable choice of the time step dt. Using a weighting factor λ between the data attachment term and the regularization energy, the gradient descent can be expressed using the derivatives of the data and smoothness energies $\frac{\partial E_{data}}{\partial \mathbf{D}(\mathbf{x})}$ and $\frac{\partial E_{smooth}}{\partial \mathbf{D}(\mathbf{x})}$ respectively as shown in the following equation

$$\begin{aligned}\mathbf{D}^{t+1}(\mathbf{x}) &= \Pi_{S^3_+}\left(\mathbf{D}^t(\mathbf{x}) - dt\frac{\partial E}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}^t)\right) \\ &= \Pi_{S^3_+}\left(\mathbf{D}^t(\mathbf{x}) - dt\lambda\frac{\partial E_{smooth}}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}^t) - dt\frac{\partial E_{data}}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}^t)\right)\end{aligned}$$

The expressions of the gradients of the data term and the smoothness energy with respect to the tensor field can be computed explicitly as follows

$$\begin{aligned}
\frac{\partial E_{smooth}}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}) &= 2\mathbf{D}(\mathbf{x}) - 2\int_{\mathbf{y}\in\mathcal{N}_\mathbf{x}} \frac{w(\mathbf{x},\mathbf{y})}{Z(\mathbf{x})}\mathbf{D}(\mathbf{y})d\mathbf{y} \\
&\quad - 2\int_{\mathbf{y}\in\mathcal{N}_\mathbf{x}} \frac{w(\mathbf{x},\mathbf{y})}{Z(\mathbf{y})}\Big(\mathbf{D}(\mathbf{y}) - \int_{\mathbf{z}\in\mathcal{N}_\mathbf{y}} \frac{w(\mathbf{z},\mathbf{y})}{Z(\mathbf{y})}\mathbf{D}(\mathbf{z})d\mathbf{z}\Big)d\mathbf{y} \\
&= 2\left(\triangle_\mathbf{x}(\mathbf{D}) - \int_{\mathbf{y}\in\mathcal{N}_\mathbf{x}} \frac{w(\mathbf{x},\mathbf{y})}{Z(\mathbf{y})}\triangle_\mathbf{y}(\mathbf{D})d\mathbf{y}\right) \\
&= 2\int_{\mathbf{y}\in\mathcal{N}_\mathbf{x}} w(\mathbf{x},\mathbf{y})\left(\frac{\triangle_\mathbf{x}(\mathbf{D})}{Z(\mathbf{x})} - \frac{\triangle_\mathbf{y}(\mathbf{D})}{Z(\mathbf{y})}\right)d\mathbf{y} \\
\frac{\partial E_{data}}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}) &= 2b\sum_{k=1}^{N}\Big(\log\big(S_k(\mathbf{x})/S_0(\mathbf{x})\big) + b\mathbf{g}_k^t\mathbf{D}(\mathbf{x})\mathbf{g}_k\Big)\mathbf{G}_k
\end{aligned}$$

Note that the gradient of the smoothness energy results in a force that takes into account higher-order neighborhood correlation between tensors by considering both the close neighborhood of the tensor as well as the neighborhood of the closest neighbors.

Let us define the norm $||.||_{TF}$ over the whole tensor field $\mathbf{D}$ as $||\mathbf{D}||_{TF} = \int_\Omega ||\mathbf{D}(\mathbf{x})||_F d\mathbf{x}$. Considering two tensor fields $\mathbf{D}_1$ and $\mathbf{D}_2$, we show in the following that the gradient of our energy functional is L-Lipschitz. The constant L will allow us to choose automatically a time step that insures the convergence of the algorithm.

$$\begin{aligned}
\left\|\frac{\partial E_{data}}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}_1) - \frac{\partial E_{data}}{\partial \mathbf{D}(\mathbf{x})}(\mathbf{D}_2)\right\|_F &= 2b^2\sum_{k=1}^{N} < \mathbf{G}_k, \mathbf{D}_1(\mathbf{x}) - \mathbf{D}_2(\mathbf{x}) >_F \\
&\leq 2b^2\sum_{k=1}^{N} ||\mathbf{G}_k||_F ||\mathbf{D}_1(\mathbf{x}) - \mathbf{D}_2(\mathbf{x})||_F
\end{aligned}$$

Therefore $\|\nabla E_{data}(\mathbf{D}_1) - \nabla E_{data}(\mathbf{D}_2)\|_{TF} \leq 2b^2\sum_{k=1}^{N}||\mathbf{G}_k||_F||\mathbf{D}_1 - \mathbf{D}_2||_{TF}$. Besides, we can easily show the following inequality

$$\|\nabla E_{smooth}(\mathbf{D}_1) - \nabla E_{smooth}(\mathbf{D}_2)\|_{TF} \leq 2(1 + 2|\mathcal{N}_\mathbf{x}| + |\mathcal{N}_\mathbf{x}|^2)||\mathbf{D}_1 - \mathbf{D}_2||_{TF}$$

where $|\mathcal{N}_\mathbf{x}|$ is the number of the considered neighbors. Thus the gradient of the objective function is L-Lipschitz with $L = 2b^2\sum_{k=1}^{N}||\mathbf{G}_k||_F + 2\lambda(|\mathcal{N}_\mathbf{x}| + 1)^2$. Choosing $0 < dt < \frac{1}{b^2\sum_{k=1}^{N}||\mathbf{G}_k||_F + \lambda(|\mathcal{N}_\mathbf{x}|+1)^2}$ makes the projected gradi-

ent descent convergent [16].

We can give an interpretation of our regularization energy in terms of diffusion-weighted images smoothing. It can be easily verified that for each direction k

$$\int_\Omega < \mathbf{D}(\mathbf{x}) - \int_{\mathbf{y}\in\mathcal{N}_\mathbf{x}} \frac{w(\mathbf{x},\mathbf{y})}{Z(\mathbf{x})}\mathbf{D}(\mathbf{y})d\mathbf{y}, \mathbf{G}_k >_F^2 d\mathbf{x} = \frac{1}{b^2}\int_\Omega \left[\log\left(\frac{S_k(\mathbf{x})}{S_0(\mathbf{x})}\right) - \log\left(\prod_{\mathbf{y}\in\mathcal{N}_\mathbf{x}}\left(\frac{S_k(\mathbf{y})}{S_0(\mathbf{y})}\right)^{\frac{w(\mathbf{x},\mathbf{y})}{Z(\mathbf{x})}}\right)\right]^2 d\mathbf{x} \tag{4.3}$$

Using Cauchy-Schwartz inequality we obtain:

$$\frac{1}{b^2}\int_\Omega \left[\log\left(\frac{S_k(\mathbf{x})}{S_0(\mathbf{x})}\right) - \log\left(\prod_{\mathbf{y}\in\mathcal{N}_\mathbf{x}}\left(\frac{S_k(\mathbf{y})}{S_0(\mathbf{y})}\right)^{\frac{w(\mathbf{x},\mathbf{y})}{Z(\mathbf{x})}}\right)\right]^2 d\mathbf{x} \leq E_{smooth}||\mathbf{G}_k||_F^2$$

This inequality provides a lower bound for the regularization energy E_{smooth} that involves the acquired diffusion signals. We can see that minimizing E_{smooth} has a direct implication on the normalized diffusion weighted images $\frac{S_k}{S_0}$. Reconstructing the tensors using a linear combination of the tensors in its neighborhood leads to the reconstruction of the normalized signals using a weighted geometric mean of the neighboring signals where the weights are not calculated only with a single volume S_k but also with the volumes obtained from the other magnetic gradient directions.

4.3 Experimental Validation

In order to validate the performance of the method we (i) have generated artificial tensors volumes corrupted with synthetic noise, (ii) used manual segmentation on T1 muscle images and tried to improve the separability of classes in the DTI space after regularization.

4.3.1 Artificially Corrupted Tensors

Let us consider two volumes, one that consists of two classes with orthogonal axes on a $20 \times 20 \times 20$ lattice and a helix in which the internal voxels are anisotropic and the external ones are spheric [Fig.4.1-b]. For the first volume, the tensor fields for each region are $\mathbf{D_1} = 0.001 \times [1\ 0.5\ 0.5\ 0\ 0\ 0]$

and $\mathbf{D_2} = 0.001 \times [0.2\ 0.4\ 0.2\ 0\ 0\ 0]$ where $\mathbf{D}$ is presented in the form of $\mathbf{D} = [D_{xx}\ D_{yy}\ D_{zz}\ D_{xy}\ D_{xz}\ D_{yz}]$. The helix dataset can be found at [3]. We considered for both datasets a field strength $b = 700\, s.mm^{-2}$, a constant value for $S_0 = 60$ for all volume voxels and twelve directions for diffusion gradient, which are used to generate the DTI corresponding to such tensor estimations. The chosen directions are the following:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 0.41 & 0.41 & 0.41 & 0.41 & 0.41 & 0.41 & 0.41 & 0.41 \\ 0.41 & -0.41 & -0.41 & 0.41 & 0.41 & 1 & 1 & 0.41 & -0.41 & -1 & -1 & -0.41 \\ -0.41 & -0.41 & 0.41 & 0.41 & 1 & 0.41 & -0.41 & -1 & -1 & -0.41 & 0.41 & 1 \end{pmatrix}$$

The images were corrupted with a white zero-mean Gaussian noise forming a data set where ground-truth on the tensor are available. An estimation of the tensor field relative to the noisy images provides the noisy tensors data.

Then, to perform comparisons we considered the regularization algorithm on noisy tensors presented in [35]. The following parameters were used for our method: $\lambda = 50$, $\mathcal{N}_{\mathbf{x}} = 3 \times 3 \times 3$, $dt = 10^{-7}$ with 50 iterations. To evaluate the performance of these methods, we considered the average sum of squared differences (**SSD**) between the regularized tensors and ground truth ones. In [Table 4.1], we can see that our estimation and regularization approach achieves better results and produces a tensor close to the ground truth. Our method performs better than the algorithm in [35] when the level of noise is relatively important. In fact, our method considers a more robust resemblance degree between voxels. Such a criterion insures a better selection of neighboring tensors involved in the estimation of a given tensor. On the other hand, the anisotropic diffusion based regularization relies on gradient information which is not robust in case of high noise. In order to assess qualitatively our algorithm, we reported in [Fig. 4.1] the resulting tensors using our regularization method and the constrained anisotropic one. We can observe that our method achieves a better direction preservation, even in the presence of a strong noise.

4.3.2 Real Data Experiments

In order to perform validation using real data, the following experiment was considered. DTI acquisitions of human skeletal muscle (calf) using

	Helix dataset			Homogeneous regions		
σ_n	0.5	1.2	3	1.5	4	9
Noisy Tensor	1.08	6.24	39.54	9.82	71.25	393.38
Method in [35]	0.33	1.60	10.57	3.32	22.47	120.70
Our Method	0.41	1.38	3.78	0.44	4.23	18.30

Table 4.1: Average Sum of Square Differences (SSD)$\times 10^4$. Comparisons between our method and the one in [35]

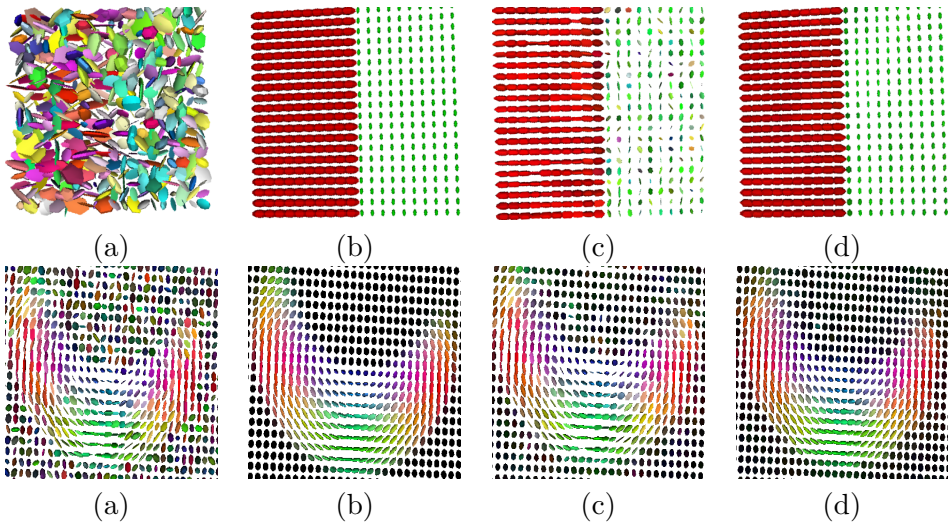


Figure 4.1: Tensors on a volume slice: (a) Noisy tensors (b) Ground-truth (c) Result obtained with [35] (d) Result obtained with our method

12 directions were carried out on a 1.5 T MRI scanner with the following parameters: repetition time (TR)= 3600 ms, echo time(TE) = 70 ms, slice thickness = 7mm and b value of 700$s.mm^{-2}$. In order to improve the signal-to-noise ratio, the acquisition was repeated thirteen times (one can use the average of the measurements) while a high resolution T1-weighted volume was also obtained and manually segmented [Fig. 4.2]. The muscles that were considered in our study were the soleus (SOL), lateral gastrocnemius (LG), medial gastrocnemius (MG), posterior tibialis (PT), anterior tibialis (AT), extensor digitorum longus (EDL), and the peroneus longus (PL).

In order to proceed with an evaluation of the proposed method, the following scenario was considered: Using the manual segmentation, and

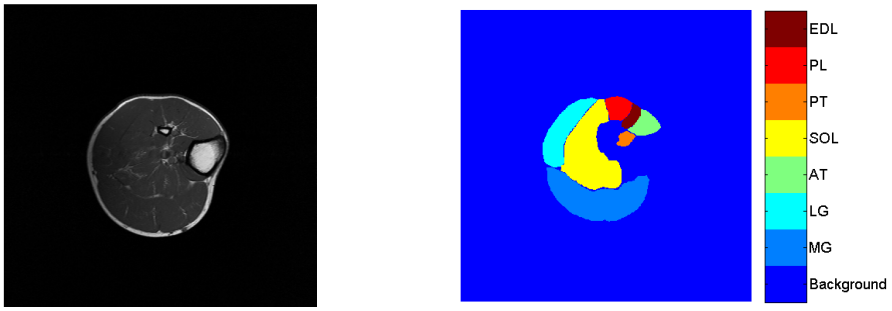


Figure 4.2: A slice of the T1-weighted volume, different muscle groups segmented manually

the observed measurements of a given acquisition (12 directions), we have constructed seven weak linear classifiers (in our case a multi-class linear SVM[62]) separating each class of muscle versus all others. Then, the success rate (percentage of voxels being attributed to the right class) from the classifier with respect to the ground truth was determined. We remark that linear separation is hardly achieved for PT, PL, EDL and AT while it yields quite satisfactory results for the MG, LG and to a lesser extent SOL which form the major part of the muscle. We have performed this test thirteen times for: (i) direct estimation (**DE**), (ii) direct estimation and regularization (**DER**), as well as using direct estimation of the average measurements of the thirteen acquisitions (**ADE**). One would expect that since muscles consist of myo-fibers of the same nature, the classification should be improved if the estimation of the tensors is properly done, i.e. with appropriate regularization. However, it is important to note that the aim of this chapter is not automatic classification of voxels in different muscle regions using DTI (in such a case more advanced classifiers can be used).

In [Table 4.2], we present quantitative validation of the present framework for the linearly separable muscles. One can see that our method leads to an improvement in the correct classification rates with respect to a plain direct estimation. We also obtain better results when compared to the averaging+estimation method.

For qualitative evaluation, we show the result of our regularization on a slice of the volume in [Fig. 4.3], where one can see that the regularization

	Overall	MG	LG	SOL
DE	78.1%	86.16%	51.1 %	84.43%
ADE	84.46%	90.47%	65.72%	88.43%
DER	86.45%	91.82%	69.76%	89.97%

Table 4.2: Correct classification rates for the different methods and for each muscle group. The first and third row show the average correct classification rates for the set of 13 volumes

algorithm smoothes some irregularities in the tensor field. We also run a fiber tracking algorithm on the tensor field without regularization and after the application of our method. In [Fig. 4.4], we can see that the obtained fiber tractography is of better visual quality for the regularized tensor field.

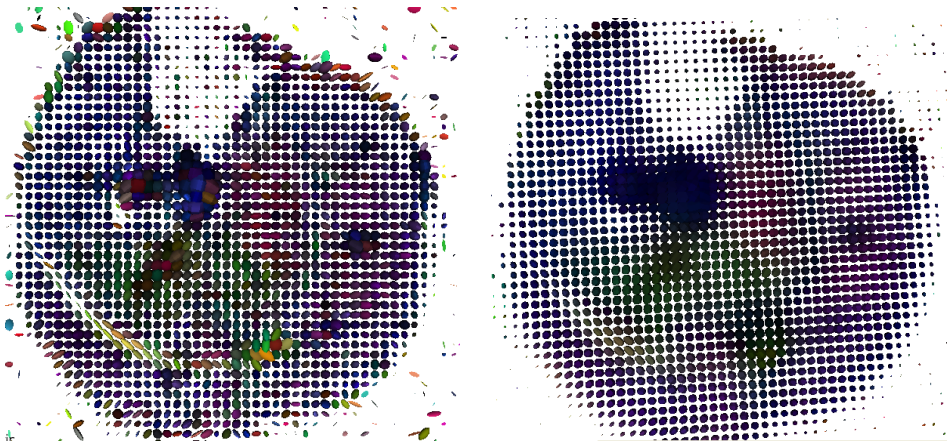


Figure 4.3: Estimated tensors without regularization, tensors obtained with our method

4.4 Discussion

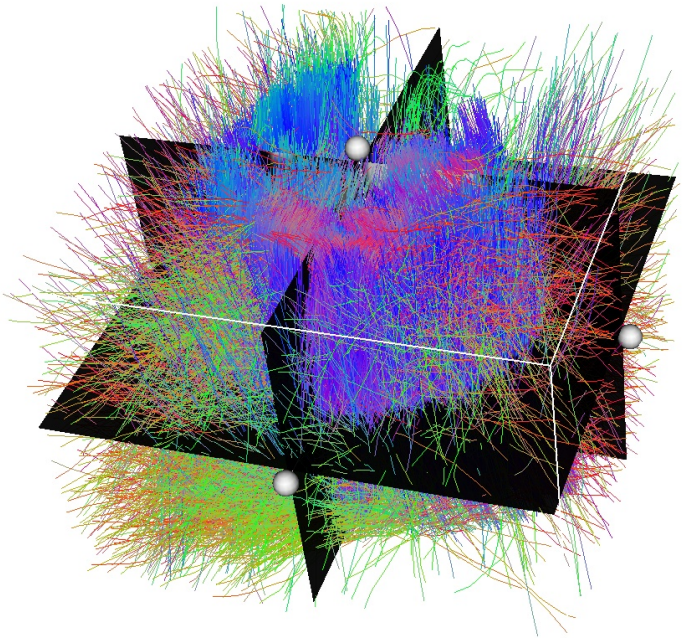
In this chapter a novel approach to direct estimation and regularization of diffusion tensor images was proposed. The main strength of our approach is the novel regularization term that assumes linear approximation of neighborhood tensors as well as the convex nature of the proposed cost function which can be easily optimized. Our method was compared to

the anisotropic constrained regularization using generated data with known noise model, and significantly improved human skeletal muscle segmentation/classification through DTI using real data.

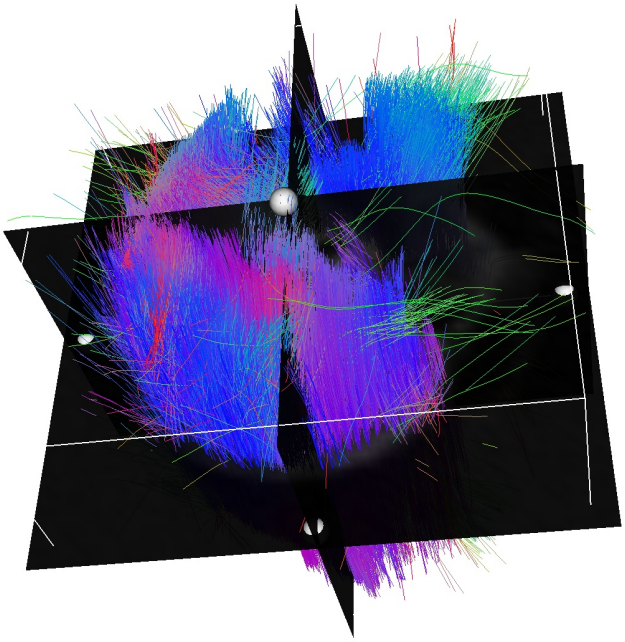
The selection of the bandwidth σ is a critical parameter of the process. Data-driven variable bandwidth models is a natural extension of the method. One would expect that the optimal bandwidth depends on the form of the observed anatomical structure which varies spatially. This suggests a a possible improvement to the method can be obtained by using a spatially dependent σ. Another possible extension of this work is to replace the Frobenius norm in the energy functional by the Riemannian metric [25] or the Log-Euclidean metric [43]. However this will be done at the expense of the convexity of the function and the computational time.

The noise model for MRI images is known to be Rician. This model can be used with the proposed regularization prior especially for images with very low SNR but will result in a non-convex optimization framework.

The next chapter will focus on the segmentation of diffusion tensors and fiber tractographies of the calf muscle.



(a)



(b)

Figure 4.4: Fiber tractography on the tensor field (a) without regularization (b) after the application of the proposed algorithm. Fibers which are generated due to background noise are removed.

Chapter 5

Kernel-based Clustering of Diffusion Tensors and Fiber Tracts

5.1 Introduction

5.1.1 Context and Motivation

Segmentation of diffusion data is important for a better understanding of the potential of the modality in providing relevant information and criteria that allow to discriminate different components (fiber bundles) of the muscle tissue. The segmentation task presents difficult challenges due to the high dimensionality of the data and the presence of rather important amounts of noise and artifacts. One has also to mention that segmentation either at the tensor or the fiber level requires adequate choices of data representation, a metric that quantifies the similarity between data appropriately and a clustering algorithm that aims at minimizing a relevant criterion of class assignment. Moreover, separating the different muscle groups of the skeletal muscles of the lower leg (calf) in regions consistent with anatomical knowledge is a crucial step for a localized quantitative study and comparison of healthy and diseased fiber bundles.

As seen in chapter 3, DTI previous studies of the human skeletal muscle provided a comparative study between subjects and different muscle regions of scalar values derived from tensors like trace, fractional anisotropy, etc.

They also evaluated experimentally important quantitative parameters like PCSA. However, little emphasis was put on muscle segmentation (apart from the method in [33]) in comparison with brain white matter, where several approaches were proposed. In the following, we review the existing methods for tensor and fiber clustering.

5.1.2 Previous Work

Diffusion Tensor Segmentation and Clustering

The existing diffusion tensor clustering algorithms can be subdivided in three groups. The first class of methods uses a variational approach along with an adequate distance over the manifold of 3×3 symmetric positive definite matrices. For instance in [74], a 3D surface is evolved using an implicit level set representation to segment a region of interest where the spatial gradient is computed using the geodesic distance and the distributions of the tensors in each region are modeled as Gaussian. In [117], the Mumford-Shah functional is minimized using a distance between tensors derived from the Burg divergence. A level set technique is also used in [81] to extract the cingulum based on the Finsler metric. Similarly in [63], several similarity measures are investigated and guide the evolution of coupled level sets.

The second class of methods uses common clustering algorithms to achieve segmentation of the tensor field. In [8], the authors propose to use the Log-Euclidean metric to obtain a kernel density estimate of the probability distribution of tensors and include it in a fuzzy k-means framework where the spatial interactions are handled using local Gaussian kernels with a fixed bandwidth. In [39], mean-shift clustering is applied for the segmentation of the Thalamus using Gaussian kernels both in the tensor and in the position space.

The third class of methods consists of graph theoretical approaches and manifold learning techniques that try to capture the structure of the tensor field and propagate the information between neighbors. In [122], a graph-cut approach is used with seed-point initialization. Spectral clustering is performed in [139] through the eigenanalysis of an affinity matrix between tensors based on a selected similarity measure. In [53], several manifold learning techniques are extended to deal with non-Euclidean spaces and

spatial connectivity is ensured by considering isotropic neighborhoods. The particular case of Locally Linear Embedding (LLE) is further discussed in [54] with different choices of tensor metrics.

Fiber Clustering

Graph theoretical approaches have been particularly popular for fiber clustering. For instance, a Gaussian model of the fibers and a normalized-cut approach based on the Euclidean distance between the moments is presented in [20]. In [86], spectral clustering along with the Hausdorff distance between fibers is considered. The method presented in [21] relies on Laplacian Eigenmaps and similarity between fibers is determined using their end points. In [107], the authors suggest another manifold learning technique by constructing a graph-based distance that captures local and global dissimilarities between fibers and use LLE for clustering of the tracts. Curve modeling has attracted attention and was handled in [76] by defining a spatial similarity measure between curves and using the Expectation-Maximization algorithm for clustering. The method proposed in [119] considers the simultaneous use of medoid-shift clustering and isomap-like manifold learning and proposed to include prior knowledge in the segmentation process using a white matter fiber atlas. Mean-shift was also used in [140] where each fiber is first embedded in a high dimensional space using its sequence of points, and kernels with variable bandwidths are considered in the mean-shift algorithm. Fibers were represented in [100] using their differential geometry and frame transportation and a consistency measure was used for clustering. Another fiber modeling approach was proposed recently in [116] where fiber bundles were represented using a probabilistic framework based on a hierarchical Dirichlet processes mixture (HDPM), therefore allowing to avoid computations of pairwise distances over an important number of fiber tracts.

Other clustering approaches use agglomerative techniques based on a defined similarity criterion between fibers. For example in [37], the corresponding segments between each pair of fiber tracts is found and a similarity measure is defined so that both the amount of overlap and the Euclidean distance between the corresponding segments are taken into account. The clustering is performed by thresholding of the similarity measure and grouping of the nearest neighbors. In the same spirit, the method in [29] uses a

threshold over a distance between fibers like the Chamfer or the Hausdorff distances and ensures that the distance between every pair of fibers within a class is below the predefined threshold. Agglomerative hierarchical clustering was also investigated in [135, 136] by starting from singleton clusters and merging the closest clusters based on a distance between paths. The clustering is made consistent across subjects using registration and matching of the centers of starting, middle and end points as feature vectors. This allows to find correspondences between clusters.

Of particular interest in the field are the supervised methods that try to achieve a segmentation consistent with a predefined atlas. Registration of B0 images and a hierarchical classification of fibers is performed in [78] using the B-spline representation of fibers. The method proposed in [86] is further extended in [87] by means of a Nystrom approximation of the out-of-sample extension of the spectral embedding. The atlas provided by this technique is used as an initialization for an Expectation-Maximization algorithm for groupwise registration and labeling in [137]. This is particularly useful in order to remove the noise in tractography labeling and the outlier tracts. Similarly, an atlas provides anatomical knowledge in [124] in order to initialize the classification and is followed by several steps that improve the clustering and reject the outlier tracts.

5.1.3 Contributions of this Work

We address several issues in this chapter. While the previous fiber clustering methods discard the tensor information and rely on the obtained tracts as sets of spatial positions, the method we propose handles tensors and is easily extended to segment fibers while taking into account the initial tensor field. We also bridge the gap between tensor and fiber clustering. To this end, we propose a kernel over the tensor space that is consistent with the physical intuition of diffusion tensors as representing the covariance of the probability distribution of water protons positions. Unlike the existing similarity measures and the use of isotropic Gaussian kernels or neighborhoods for spatial interaction, the proposed kernel quantifies not only the dissimilarity between tensors, but takes into account their localization in space in a tractable way and enhances implicitly the connectivity along fiber tracts, i.e. in the feature space provided by the kernel embedding, tensors which

are aligned will be closer than tensors which do not lie on the same fiber tract, which is not guaranteed by spatially isotropic Gaussian kernels. We use kernel-PCA for embedding in a Euclidean space and landmark-Isomap for information propagation along diffusive pathways to obtain the final embedding. This is important as it allows to reflect the diffusion flow along the muscle fibers which are characterized by an elongated structure. The clustering is done in the embedding space by a plain k-means algorithm.

The proposed kernel is extended to deal with fiber tracts as input of the clustering algorithm by way of the summation kernel, which is a handy way to define kernels over sets. Not only we take into account in this kernel the interactions between the points (as spatial positions) but also the information provided by the whole tensor field. Given that the fiber pathways are provided, only kernel-PCA and k-means are sufficient for clustering in the fiber domain. We give an interpretation of the fiber kernel as a comparison of soft representations of the fiber tracts and show that it provides a natural generalization for Gaussian kernel Correlation.

Moreover, we show how this unsupervised algorithm can be made supervised in an efficient way. Given that the kernel satisfies the Mercer conditions, we can look at the problem from a discriminative perspective and use kernel Support Vector Machine (SVM) pairwise classifiers in order to segment the calf muscle in regions consistent with a previously defined atlas.

We further develop this kernel-based viewpoint and build Hilbertian angular metrics between fibers. These are derived from their counterparts between tensors, providing a more general framework. We investigate two additional approaches for clustering purposes. The first approach resorts to k-means in the diffusion maps embedding space. Note that diffusion maps were used for Orientation Distribution Function (ODF) segmentation in Q-ball images in [120], where spatial coherence was imposed using the Markovian relaxation of the affinity matrix. However the fiber domain provides no straightforward spatial neighborhood relationships like those given by the nearest neighbors in the 3D image grid. We show that the proposed metrics impose spatial coherence in the fiber domain while taking into account the information provided by the tensor field.

The second clustering approach is motivated by the limits of manifold

embedding methods. Indeed, the use of embeddings and common clustering techniques like k-means requires to choose the dimension of the embedding and the number of clusters. It would be preferable to obtain the number of clusters as a result of the clustering algorithm, especially when the inter-subject variability (which is rather important for skeletal muscles) may require the use of different number of clusters across subjects. Moreover, selecting the embedding dimension is an issue since a too low dimension will result in information loss and a too high dimension will include an important dispersion in the data. Furthermore, clustering on the manifold directly is a tricky issue since one has to compute intrinsic means on submanifolds where an explicit expression of geodesic distances is not necessarily available. Another issue is the sensitivity of methods like k-means to initialization. Therefore, we propose a method that performs manifold clustering of fibers without resorting to manifold embeddings or computations of intrinsic means. It is based on linear programming (LP) and uses the geodesic distances in a way similar to [119] from the fibers to a reduced set of landmark fibers to perform the clustering. Unlike k-means, the algorithm provides automatically the number of clusters, is not sensitive to initialization and the class centers are chosen as exemplars from the dataset.

The remainder of this chapter is organized as follows: In section 5.2, we discuss the kernel over the tensor space and provide an analysis of its properties and advantages. A more general case is also explored. Moreover, we discuss its use in a kernel-PCA and landmark-Isomap setting for clustering of diffusion tensors. In section 5.3 we propose several fiber clustering strategies based on extensions of the similarity measures defined in the tensor space, both in a supervised and unsupervised fashion. Section 5.4 is dedicated to the experimental results and we discuss the perspectives of this work in section 5.5.

5.2 Diffusion Tensor Clustering

5.2.1 A Probability Kernel on Tensors

Diffusion tensors measure the motion distribution of water molecules. More explicitly, they refer to the covariance of a Gaussian probability over the

displacements $\mathbf{r}$ of the water protons given a diffusion (mixing) time t:

$$p(\mathbf{r}|t,\mathbf{D}) = \frac{1}{\sqrt{\det(\mathbf{D})(4\pi t)^3}} \exp\left(-\frac{\mathbf{r}^t\mathbf{D}^{-1}\mathbf{r}}{4t}\right) \tag{5.1}$$

Given a diffusion tensor $\mathbf{D}$ localized at voxel $\mathbf{x}$, we can obtain the probability of the position y of the water molecule previously localized at $\mathbf{x}$ in a straightforward way:

$$p(\mathbf{y}|\mathbf{x},t,\mathbf{D}) = \frac{1}{\sqrt{\det(\mathbf{D})(4\pi t)^3}} \exp\left(-\frac{(\mathbf{y}-\mathbf{x})^t\mathbf{D}^{-1}(\mathbf{y}-\mathbf{x})}{4t}\right) \tag{5.2}$$

Therefore, a natural way to define a kernel over the tensor space where position is taken into account is to consider the expected likelihood kernel [61]. Let us consider two tensors $\mathbf{D}_1$ and $\mathbf{D}_2$ localized at $\mathbf{x}_1$ and $\mathbf{x}_2$ respectively, and a diffusion time t. The expected likelihood kernel $k_t((\mathbf{D_1},\mathbf{x}_1);(\mathbf{D_2},\mathbf{x}_2))$ between the pairs $(\mathbf{D_1},\mathbf{x}_1)$ and $(\mathbf{D_2},\mathbf{x}_2)$ is defined as the expectation of Gaussian probability $p_2(\mathbf{y}|\mathbf{x_1},t,\mathbf{D}_1)$ under the probability law of $p_1(\mathbf{y}|\mathbf{x_2},t,\mathbf{D}_2)$ and is given by the following expression:

$$\begin{aligned} k_t((\mathbf{D_1},\mathbf{x}_1);(\mathbf{D_2},\mathbf{x}_2)) &= E_{p_2(\mathbf{y}|\mathbf{x_2},t,\mathbf{D}_2)}(p_1(\mathbf{y}|\mathbf{x_1},t,\mathbf{D}_1)) &(5.3)\\ &= \int p_1(\mathbf{y}|\mathbf{x_1},t,\mathbf{D}_1)p_2(\mathbf{y}|\mathbf{x_2},t,\mathbf{D}_2)d\mathbf{y} &(5.4)\end{aligned}$$

Note that the diffusion time t is naturally a parameter for this kernel and that to be more precise mathematically, this kernel is not actually defined on the space of tensors S^3_+ but rather over the product space $S^3_+ \times \mathbb{R}^3$. Based on the derivation in [61] and using the expression provided in 5.2, one can obtain the following closed-form expression of this kernel:

$$k_t((\mathbf{D_1},\mathbf{x}_1);(\mathbf{D_2},\mathbf{x}_2)) = \frac{1}{\sqrt{(4\pi t)^3}} k_1(\mathbf{D_1},\mathbf{D_2})k_2((\mathbf{D_1},\mathbf{x}_1);(\mathbf{D_2},\mathbf{x}_2)) \tag{5.5}$$

where

$$k_1(\mathbf{D_1}, \mathbf{D_2}) = \frac{1}{\sqrt{\det(\mathbf{D}_1 + \mathbf{D}_2)}}$$
$$k_2((\mathbf{D_1}, \mathbf{x}_1); (\mathbf{D_2}, \mathbf{x}_2)) = \exp\left(-\frac{1}{4t}(\mathbf{x}_1^t \mathbf{D}_1^{-1} \mathbf{x}_1 + \mathbf{x}_2^t \mathbf{D}_2^{-1} \mathbf{x}_2)\right) \times$$
$$\exp\left(\frac{1}{4t}(\mathbf{D}_1^{-1}\mathbf{x}_1 + \mathbf{D}_2^{-1}\mathbf{x}_2)^t (\mathbf{D}_1^{-1} + \mathbf{D}_2^{-1})^{-1} (\mathbf{D}_1^{-1}\mathbf{x}_1 + \mathbf{D}_2^{-1}\mathbf{x}_2)\right) \quad (5.6)$$

We notice that k_2 has a much simpler expression. Indeed, using the following inversion properties

$$(\mathbf{D}_1 + \mathbf{D}_2)^{-1} = \mathbf{D}_1^{-1} - \mathbf{D}_1^{-1}(\mathbf{D}_1^{-1} + \mathbf{D}_2^{-1})^{-1}\mathbf{D}_1^{-1} \quad (5.7)$$
$$(\mathbf{D}_1 + \mathbf{D}_2)^{-1} = \mathbf{D}_2^{-1} - \mathbf{D}_2^{-1}(\mathbf{D}_1^{-1} + \mathbf{D}_2^{-1})^{-1}\mathbf{D}_2^{-1} \quad (5.8)$$

we obtain the following compact expression for k_2:

$$k_2((\mathbf{D_1}, \mathbf{x}_1); (\mathbf{D_2}, \mathbf{x}_2)) = \exp\left(-\frac{1}{4t}(\mathbf{x}_1 - \mathbf{x}_2)^t (\mathbf{D}_1 + \mathbf{D}_2)^{-1} (\mathbf{x}_1 - \mathbf{x}_2)\right) \quad (5.9)$$

5.2.2 Properties of the Tensor Kernel

The kernel stems from an L^2 inner product defined on the Hilbert space of square-integrable functions, to which Gaussian probability densities belong. Therefore the kernel verifies the Mercer conditions, i.e. it is positive definite. We hereafter provide an analysis of this kernel:

- The first term $k_1(\mathbf{D_1}, \mathbf{D_2})$ may be rewritten as follows:

$$k_1(\mathbf{D_1}, \mathbf{D_2}) = \frac{1}{\sqrt{\det(\mathbf{D}_2)}} \frac{1}{\sqrt{\det(\mathbf{Id} + \mathbf{D}_1\mathbf{D}_2^{-1})}} = \frac{1}{\sqrt{\det(\mathbf{D}_2)}} \frac{1}{\sqrt{\prod_{i=1}^{3}(1+\lambda_i)}} \quad (5.10)$$

  where $\mathbf{Id}$ is the 3×3 identity matrix and λ_i are the generalized eigenvalues of the pair of matrices $(\mathbf{D}_1, \mathbf{D}_2)$. This is reminiscent of the geodesic distance on the manifold of 3×3 symmetric positive definite matrices $d = \sqrt{\sum_{i=1}^{3}(\log(\lambda_i))^2}$ [91] which is also based on the generalized eigenvalues: the distance (respectively the kernel) is increasing (respectively decreasing) with increasing (respectively decreasing)

generalized eigenvalues, which is a reasonable behavior (recall that the kernel reflects similarity). The symmetry is ensured in the geodesic distance by the squared logarithm function because the latter is invariant with respect to the inverse transformation $\lambda_i \mapsto \frac{1}{\lambda_i}$ (recall that the generalized eigenvalues of the pair $(\mathbf{D}_2, \mathbf{D}_1)$ are the inverse of those of the pair $(\mathbf{D}_1, \mathbf{D}_2)$), while it can be seen that the factor $\frac{1}{\sqrt{\det(\mathbf{D}_2)}}$ has a similar role since it also preserves the symmetry. Note that the original expression is clearly symmetric.

- Three special cases of the second factor k_2 are interesting:

  1. When $\mathbf{D}_1 = \mathbf{D}_2 = \mathbf{D}$

  $$k_2((\mathbf{D}, \mathbf{x}_1); (\mathbf{D}, \mathbf{x}_2)) = \exp\left(-\frac{1}{8t}(\mathbf{x}_1 - \mathbf{x}_2)^t \mathbf{D}^{-1}(\mathbf{x}_1 - \mathbf{x}_2)\right) \tag{5.11}$$

  As expected, when the tensors are equal, what appears is the Mahalanobis distance between positions $\mathbf{x}_1$ and $\mathbf{x}_2$ with respect to $\mathbf{D}$. In particular when the tensor $\mathbf{D}$ is isotropic, i.e. $\mathbf{D} = \mu\mathbf{Id}$, $k_2((\mathbf{D}, \mathbf{x}_1); (\mathbf{D}, \mathbf{x}_2)) = \exp\left(-\frac{1}{8\mu t}||\mathbf{x}_1 - \mathbf{x}_2)||^2\right)$, which is plainly a Gaussian kernel between the points $\mathbf{x}_1$ and $\mathbf{x}_2$.

  2. When $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}$, $k_2((\mathbf{D}_1, \mathbf{x}); (\mathbf{D}_2, \mathbf{x})) = 1$, which means that the kernel k reduces to k_1. Again this was expected since the kernel will rely only on tensor similarity if there is no difference in spatial positions.

  3. When $t \to \infty$, this corresponds to the limit case where the spatial interaction is not taken into account ($k_2 = 1$). As in the previous case, the kernel will rely only on tensor similarity as an infinite time of diffusion will remove the information of spatial arrangement. Hence the diffusion time t is important to weight the contribution of each term in the kernel definition.

- The first special case is of particular interest in diffusion tensor analysis to enhance the connectivity between tensors which are aligned on the same fiber tract. Let us consider the tensor configuration in [Fig.5.1] where all tensors are equal to $D = \mu(\vec{\mathbf{e}_1}\vec{\mathbf{e}_1}^t + \vec{\mathbf{e}_2}\vec{\mathbf{e}_2}^t) + \nu\, \vec{\mathbf{e}_3}\vec{\mathbf{e}_3}^t$, where $(\vec{\mathbf{e}_i})_{i=1\ldots3}$ are the canonical basis of $\mathbb{R}^3$ and $\nu > \mu$ are the eigenvalues of

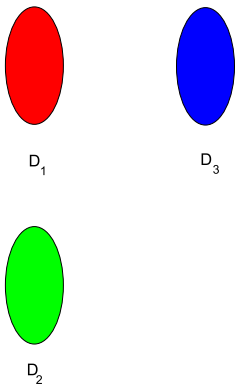


Figure 5.1: A configuration where the tensor kernel implicitly puts more weight on the connection between $\mathbf{D}_1$ and $\mathbf{D}_2$ than between $\mathbf{D}_1$ and $\mathbf{D}_3$, reflecting their alignment.

D. The tensors have therefore a principal direction of diffusion along $\vec{\mathbf{e}}_3$. The tensors are all equal yet the second term k_2 allows to affect more affinity between tensors 1 and 2 than between tensors 1 and 3. Indeed, we can compute the kernel values to obtain

$$k_2((\mathbf{D}, \mathbf{x}_1); (\mathbf{D}, \mathbf{x}_2)) = \exp\left(-\frac{d^2}{8\nu t}\right) \quad (5.12)$$

$$k_2((\mathbf{D}, \mathbf{x}_1); (\mathbf{D}, \mathbf{x}_3)) = \exp\left(-\frac{d^2}{8\mu t}\right) \quad (5.13)$$

where $d = ||\mathbf{x}_1 - \mathbf{x}_2|| = ||\mathbf{x}_1 - \mathbf{x}_3||$. Since $\nu > \mu$, we can see that the similarity between the tensors 1 and 2 is higher than the similarity the tensors 1 and 3, despite the fact that the tensors are all equal. The kernel captures locally the fiber structure and takes into account the spatial context by introducing a coupling between tensors and their spatial positions.

Now we discuss the more general case where powers of diffusion probabilities are considered with a normalized L^2 inner product.

5.2.3 A More General Case: Multivariate Normals as a Subset of the Exponential Distributions Family

The structure of the set of multivariate normal distributions $\mathcal{M}$ as a statistical manifold endowed with the Fisher information geometry was discussed in [35], where a closed-form solution of the geodesic distance [91] over this manifold is available for the particular case of Gaussian distributions with common mean. Here we view the multivariate normal distributions as a subset of the exponential distributions family. Let us consider a normal probability density p. In this context, given the exponential decay of the distribution, it is interesting to notice that not only p is an element of the Hilbert space L^2 of square integrable functions (the case discussed above) but any power p^α, with α a strictly positive real number is also square integrable. Also the L^2 inner product between probabilities is not normalized as can be seen from the factor $\frac{1}{\sqrt{\det(\mathbf{D}_2)}}$ that appears in the expression of the tensor similarity term k_1. This motivates the use of normalized probability product kernels [61] to define a family of angular similarities between multivariate normal distributions. We identify the multivariate normal probabilities with their couples of parameters.

$$k_\alpha(p_1, p_2) = \frac{\int p_1(\mathbf{x})^\alpha p_2(\mathbf{x})^\alpha d\mathbf{x}}{\sqrt{\int p_1(\mathbf{x})^{2\alpha} d\mathbf{x}} \sqrt{\int p_2(\mathbf{x})^{2\alpha} d\mathbf{x}}} \tag{5.14}$$

where p_1 and p_2 are two multivariate normal distributions. k_α is simply the normalized L^2 inner product between p_1^α and p_2^α. The geometric interpretation of this quantity is that it is the cosine of the angle between p_1^α and p_2^α. It defines a Mercer kernel over the space of multivariate normal distributions, i.e. for any subset $p_{1\ldots N}$ of $\mathcal{M}$, the Gram matrix G of C_α with entries $G_{ij} = k_\alpha(p_i, p_j)$ is semi-definite positive. The Mercer property allows the construction of a mapping ϕ_α associated with the kernel C_α that provides an embedding of $\mathcal{M}$ in the Reproducing Kernel Hilbert Space (RKHS) $\mathcal{H}_\alpha$ such that $k_\alpha(p_1, p_2) =< \phi_\alpha(p_1), \phi_\alpha(p_2) >_{\mathcal{H}_\alpha}$, where $< ., . >_{\mathcal{H}_\alpha}$ is the inner product of $\mathcal{H}_\alpha$. This allows to have the following Hilbertian metric $d_{\alpha|\mathcal{H}_\alpha}$:

$$d_{\alpha|\mathcal{H}_\alpha}(p_1, p_2) = \sqrt{k_\alpha(p_1, p_1) - 2k_\alpha(p_1, p_2) + k_\alpha(p_2, p_2)} \tag{5.15}$$

Given that k_α is a normalized scalar product, i.e. $k_\alpha(p,p) = 1$, we obtain the following expression:

$$d_{\alpha|\mathcal{H}_\alpha}(p_1,p_2) = \sqrt{2 - 2k_\alpha(p_1,p_2)} \tag{5.16}$$

Note that k_α is the generalization of the normalized Expected Likelihood kernel since it corresponds to the special case $\alpha = 1$. Similar computations to those in 5.2.1 lead to very similar equations:

$$k_\alpha(p_1,p_2) = k^{tensor}(\mathbf{D}_1,\mathbf{D}_2)k_\alpha^{spatial}(p_1,p_2) \tag{5.17}$$

where

$$\begin{aligned} k^{tensor}(\mathbf{D}_1,\mathbf{D}_2) &= 2\sqrt{2}\,\frac{\det(\mathbf{D}_1)^{\frac{1}{4}}\det(\mathbf{D}_2)^{\frac{1}{4}}}{\sqrt{\det(\mathbf{D}_1+\mathbf{D}_2)}} \\ k_\alpha^{spatial}(p_1,p_2) &= \exp\left(-\frac{\alpha}{4t}(\mathbf{x}_1-\mathbf{x}_2)^t(\mathbf{D}_1+\mathbf{D}_2)^{-1}(\mathbf{x}_1-\mathbf{x}_2)\right) \end{aligned} \tag{5.18}$$

We can see that k_α^{tensor} is a tensor similarity term and is independent of the parameter α. The difference with respect to k_1 is that k_α^{tensor} is normalized. The second term $k_\alpha^{spatial}$ is also very similar to k_2. Again it is a spatial connectivity term where appears the Mahalanobis distance between the locations $\mathbf{x}_1$ and $\mathbf{x}_2$ with respect to the sum of tensors $(\mathbf{D}_1+\mathbf{D}_2)$. The power α operates as a scale parameter in the exponential function, so changing the parameter α amounts to a rescaling of the diffusion time t.

In the next subsection, we discuss how to embed the tensors in a Euclidean space based on the kernel information as a preliminary step to clustering.

5.2.4 Embedding of the Tensors through Kernel PCA and Landmark Isomap

Let us consider the N pairs $(\mathbf{D}_i,\mathbf{x}_i)_{i=1\ldots N}$ representing a tensor field. We construct the $N \times N$ kernel matrix $\mathbf{K}$ of entries $\mathbf{K}_{ij} = k_t((\mathbf{D}_i,\mathbf{x}_i);(\mathbf{D}_j,\mathbf{x}_j))$ for a fixed diffusion time t and normalize it to obtain $\widetilde{\mathbf{K}}$ such that $\widetilde{\mathbf{K}}_{ij} = \frac{\mathbf{K}_{ij}}{\sqrt{\mathbf{K}_{ii}\mathbf{K}_{jj}}}$. These pairs are then embedded in a k-dimensional Euclidean space using an eigenvalue decomposition of $\widetilde{\mathbf{K}} = \mathbf{USU}^t$ where $\mathbf{U}$ is an orthogonal

$N \times k$ matrix and $\mathbf{S}$ is a $k \times k$ diagonal matrix, or equivalently by using a kernel PCA. The coordinates of the embedded tensors are given by the $N \times k$ matrix $\mathbf{X} = \mathbf{U}\sqrt{\mathbf{S}}$ where $\sqrt{\mathbf{S}}$ is obtained by setting the diagonal elements of $\mathbf{S}$ to their square roots [101]. Each row m of $\mathbf{X}$ holds the coordinates in the feature space of the m-th pair.

Given the k-dimensional representation $\mathbf{X}$ of the tensor field, one has to propagate the local interaction between the tensors and take into account the distance along diffusive pathways, i.e. simulate the water flow along these trajectories. This is done using the Isomap algorithm which is based on three steps [106]:

1. Consider the data points as vertices of a graph with edges linking neighbors. The notion of neighborhood can be defined using a threshold over distances (two points are neighbors if the distance between them is smaller than a threshold) or using k-neighborhoods (two points are neighbors if one is among the k closest points to the other).

2. Between every two points of the dataset, find the shortest path over the graph using the Dijkstra algorithm based on the Euclidean distance in the feature space and compute a new distance matrix that holds the lengths of the path. This step is important in diffusion tensor analysis in order to propagate diffusion information along fiber tracts.

3. Perform Multidimensional Scaling (MDS) to obtain the new embedding, i.e. a configuration of points that respects approximately the distance matrix computed in the previous step. Since the kernel enhances fiber connectivity, we expect the new configuration to reflect the diffusion flow in the tissues.

Kernel PCA and Isomap are techniques that were previously used for statistical analysis of diffusion tensors in [65, 113], however here they are applied in the context of clustering in the space of Gaussian diffusion probabilities. Note that in practice we use a faster version of this algorithm called landmark-Isomap [34] that reduces the computational time of the first step by computing the distance of the points to a reduced set of landmarks chosen randomly in the dataset. The clustering is done afterwards using a plain k-means algorithm. Note that the kernel PCA step amounts to denoising

in the feature space and that from a theoretical point of view (with perfect data), we could have used from the outset the Euclidean distance (L^2 norm) implied by the kernel k_t in the Isomap algorithm without the kernel PCA projection, since the latter requires only the distances between points given by the metric based on normalized kernels in [Eq. 5.16].

We now propose a supervised algorithm for tensor segmentation based on Support Vector Machine (SVM) classifiers and probability kernels.

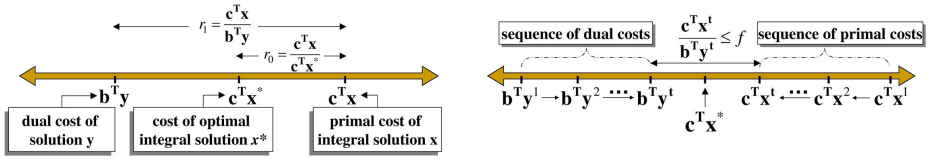
5.2.5 Supervised SVM Tensor Clustering

We briefly review the principles of two class SVMs [110]. Given N points $\mathbf{x}_i$ with known class information y_i (either $+1$ or -1), SVM training consists in finding the optimal separating hyperplane described by the equation $\mathbf{w}^t\mathbf{x} + b = 0$ with the maximum distance to the training examples. It amounts to solving a dual convex quadratic optimization problem and each data point $\mathbf{x}$ is classified using the SVM output function $f(x) = (\sum_i^N \alpha_i y_i \mathbf{x}\mathbf{x}_i) + b$. The algorithm is extended to achieve non linear separation using a kernel function $K(\mathbf{x}, \mathbf{y})$(symmetric, positive semi-definite) that is used instead of the standard inner product. For the case of multi-class learning, several pairwise SVM classifiers are built and the classification is done in a one against all fashion.

In our case the data points are tensors and the used kernel function is given by the normalized probability kernel $k^{tensor}(\mathbf{D}_1, \mathbf{D}_2)$ that verifies the Mercer property and is therefore suitable for SVM learning. Note that the learning and testing procedures make the assumption that the global rotation between the used learning and testing datasets is negligible, which is the case when the DT-MRI acquisitions are done in approximately the same conditions.

In order to impose spatial regularity in the segmentation, we use a Markov Random Field framework. It also allows to use all the scores given by the SVMs, instead of making the labeling decision by simply taking the maximum score. Therefore, we define the following energy E to minimize:

$$E = \sum_{i \in \Omega} u^s(l(i)) + \lambda \sum_{i \in \Omega, j \in \mathcal{N}(i)} u^p(l(i), l(j)) \tag{5.19}$$



(a) The primal-dual principle (b) The primal-dual schema

Figure 5.2: (a) The optimal cost $\mathbf{c}^T\mathbf{x}^*$ is between the costs $\mathbf{b}^T\mathbf{y}$ (cost of a dual feasible solution) and $\mathbf{c}^T\mathbf{x}$ (the cost of an integral-primal and dual feasible solutions). The intuitive idea that is exploited is that if the gap between the primal and dual costs is reduced, the gap between the optimal cost and the primal cost is also reduced. (b) The primal-dual technique consists in building a sequence of dual and integral-primal feasible solutions so that the approximate solution is within a certain range from the optimal solution. The primal-dual paradigm can be subsequently applied.

where Ω is the image domain, $l(i)$ is the label of the voxel i, $\mathcal{N}(i)$ is the considered neighborhood, u^s is the potential given by the SVM scores and u^p is a pairwise potential that imposes spatial regularization. The parameter λ is a trade-of factor between the two terms. We choose $u^s(l(i)) = \log(1 + \exp(-\alpha f_l(\mathbf{D}(i))))$ which is a decreasing potential in the score given by a one-against-all SVM classifier f_l. The following pairwise potential is chosen:

$$u^p(l(i), l(j)) = (1 - \delta_{l(i),l(j)}) \tag{5.20}$$

where δ is the Kronecker delta. This means that we penalize a configuration where different labels are assigned to neighboring voxels, resulting in a less important amount of noise in the segmentation result. In order to minimize the energy, we use the Fast-PD algorithm proposed in [67, 68].

We now provide a description of the Fast-PD algorithm that is used to optimize the Markov Random Field energy. This description will be helpful in the remainder of the thesis, in particular in subsection 5.3.4 and in chapter 6. The algorithm is based on introducing auxiliary variables and reformulate the energy minimization as a linear programming problem, i.e. the optimization consists now in minimizing a linear function in the auxiliary variables under linear equality and inequality constraints. The optimization is based on the primal-dual technique, which is a popular approach in the Linear Programming literature. In order to have a clearer picture of

the mechanisms used in the algorithm , let us consider the following linear programs, which are called primal and dual Linear Programs (LPs):

$$\begin{array}{lll} \text{PRIMAL:} & \min \mathbf{c}^T\mathbf{x} & \text{DUAL:} \quad \max \mathbf{b}^T\mathbf{y} \\ & \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} & \text{s.t.} \quad \mathbf{A}^T\mathbf{y} \leq \mathbf{c} \end{array} \tag{5.21}$$

where $\mathbf{A}$ is a matrix that provides with the vectors $\mathbf{b}$ and $\mathbf{c}$ the (linear) nequality constraints of the primal and dual problems. The optimization variables of the problems are $\mathbf{x}$ and $\mathbf{y}$. Note however that the variable $\mathbf{x}$ considered in the primal problem is continuous while the solution to the original MRF energy should be integral. This is the at the root of major difficulties in solving the MRF problem since it is in general NP-hard. Therefore the sought solution is but an approximation of the solution of the original problem. A primal-dual f-approximation technique refers to the following paradigm (illustrated also in Fig. 5.2(a)):

Primal-Dual Paradigm. *If* $\mathbf{x}$ *and* $\mathbf{y}$ *are integral-primal and dual feasible solutions having a primal-dual gap less than* f*, i.e.:*

$$\mathbf{c}^T\mathbf{x} \leq f \cdot \mathbf{b}^T\mathbf{y}, \tag{5.22}$$

then $\mathbf{x}$ *is an* f*-approximation to the optimal integral solution* $\mathbf{x}^*$*, i.e.* $\mathbf{c}^T\mathbf{x}^* \leq \mathbf{c}^T\mathbf{x} \leq f \cdot \mathbf{c}^T\mathbf{x}^*$.

This paradigm paves the way to the primal-dual strategy, since it suggests reducing the duality gap in order to have an approximate solution that is closer to the optimum. The strategy consists in computing iteratively an f-approximate solution (this strategy is explained in [Fig.5.2(b)]):

Primal-Dual Strategy. *Iteratively computing pairs of integral-primal and dual solutions* $\{(\mathbf{x}^k, \mathbf{y}^k)\}_{k=1}^t$*, until the elements* $\mathbf{x}^t$*,* $\mathbf{y}^t$ *of the last primal and dual solutions are both feasible and have a primal-dual gap which is less than* f*, i.e. the paradigm* (5.22) *can be applied.*

One the main strengths of the Fast-PD algorithms is its versatility since it makes very few hypotheses about the MRF energy function. In particular, the only stringent requirement is that the MRF pairwise potential function is

nonnegative. Moreover, note that bounds on the optimality of the computed solution are given. The global optimum is therefore computed up to certain precision that can be input by the user. Also, this technique has a favorable computation time, resulting in a significant efficiency improvement with respect to exisiting methods in the literature[67, 68].

In the following section, we extend in a tractable way the proposed kernel defined over the tensor field to the fiber domain.

5.3 Fiber Clustering

5.3.1 Probability Kernel in the Fiber Domain

The fiber trajectories are obtained through the integration of the vector field of principal directions of diffusion. Based on the continuous tensor field approximation (by means of interpolation), we represent each fiber tract as a sequence of tensors localized in spatial positions, i.e. is a set of pairs $\tau_i = (\mathbf{D}_i, \mathbf{x}_i)_{i=1...n}$ where n is the number of points lying on the fiber. Note that the tractography already requires tensor interpolation and that the interpolated tensors are therefore kept for kernel computation. So it is natural to extend the tensor kernel using a kernel over sets. We simply use the summation kernel [56] to obtain the following Mercer kernel K_t between two fibers $\mathbf{F}_1$ and $\mathbf{F}_2$:

$$K_t(\mathbf{F}_1, \mathbf{F}_2) = \frac{1}{n_1}\frac{1}{n_2} \sum_{(\mathbf{D}_i,\mathbf{x}_i)\in\mathbf{F}_1} \sum_{(\mathbf{D}_j,\mathbf{x}_j)\in\mathbf{F}_2} k_t((\mathbf{D}_i, \mathbf{x}_i); (\mathbf{D}_j, \mathbf{x}_j)) \qquad (5.23)$$

where n_1 (resp. n_2) is the number of points of the fiber $\mathbf{F}_1$ (resp. $\mathbf{F}_2$). This kernel sums the interactions between tensors belonging to the fiber tracts. It captures the diffusion and spatial links between diffusive pathways. It is important to notice that while all the interactions are summed, the diffusion time t acts as a scale parameter. Therefore for a suitable choice of t, a tensor interacts only with tensors lying in a local neighborhood and far-away tensors have a negligible impact on the summation. While this formulation does not take into account geometric shape information like pointwise curvature and torsion, it allows to avoid finding point correspondences that is needed in the case where Hausdorff or Chamfer distances are used. It has

also the advantage of yielding directly a symmetric similarity between fiber tracts. Indeed given the separability of the tensor kernel in a tensor similarity term and a spatial connectivity term, the fiber kernel can be seen as a similarity between fibers that tries to match tensors as local features within an anisotropic neighborhood. As in the case of tensors, the segmentation of the fiber tracts is achieved using kernel PCA and k-means clustering.

5.3.2 A Physical Interpretation of the Fiber Kernel

One can see that the summation kernel is simply the expected likelihood kernel between distributions providing a soft representation for fibers. More explicitly, we consider a dynamical system where a particle can be initially at a position $\mathbf{x}_i$ on the fiber tract and moves to a position y with the following probability

$$p(y|t,(\mathbf{D}_i)_{i=1\ldots n}) = \sum_{i=1}^{n} p(\mathbf{x}_i)p(\mathbf{y}|\mathbf{x}_i,t,\mathbf{D}_i) \tag{5.24}$$

With a uniform prior distribution on $\mathbf{x}_i$, $p(y|t,(\mathbf{D}_i)_{i=1\ldots n}) = \frac{1}{n}\sum_{i=1}^{n} p(\mathbf{y}|\mathbf{x}_i,t,\mathbf{D}_i)$. If the initial positions were independent (which is not the case because they are the result of the tractography), this would have amounted exactly to an *adaptive* kernel density estimation of the position of the water molecules along the fiber tracts where the point-dependent Gaussian kernels use the diffusion tensors as covariance matrices to model the uncertainty. However the distributions in equation 5.24 still provide a soft representation of the fibers and measure the compactness of the spatial configuration of the fiber tract. By bilinearity of the expected likelihood kernel, it is straightforward to see that the expected likelihood kernel of the distributions given in Equation 5.24 is exactly the summation kernel.

5.3.3 Extension of the General Case to the Fiber Domain

We can extend the normalized kernels to the fiber domain though the use of a Hilbert space fiber tract model. A fiber is again represented as a set of probability measures $(p_i)_{i=1\ldots N}$ where N is the number of points of the fiber. Every probability measure (p_i) has a pair of parameters $(\mathbf{x}_i,\mathbf{D}_i)$ where $\mathbf{x}_i$. When considering the mapping ϕ_α of these measures in the RKHS $\mathcal{H}_\alpha$, we can represent $\mathbf{F}$ as a weighted average of $\phi_\alpha(p_i)_{i=1\ldots N}$, i.e. $\mathbf{F} =$

$\sum_{i=1}^{N} w_i \phi_\alpha(p_i)$. A straightforward choice of weights is $\forall i, w_i = \frac{1}{N}$.

Let us consider a fibers $\mathbf{F}_1$ (resp. $\mathbf{F}_2$) represented using a set of probabilities $(p_i)_{i=1\ldots N_1}^{(1)}$ (resp. $(p_i)_{i=1\ldots N_2}^{(2)}$) and weights $w_i^{(1)}$ (resp. $w_i^{(2)}$). The angular similarity $\widehat{C}_\alpha$ between $\mathbf{F}_1$ and $\mathbf{F}_2$ points is defined as follows:

$$K_\alpha(\mathbf{F}_1, \mathbf{F}_2) = \frac{< \sum_{i=1}^{N_1} w_i^{(1)} \phi_\alpha(p_i^{(1)}), \sum_{j=1}^{N_2} w_j^{(2)} \phi_\alpha(p_j^{(2)}) >_{\mathcal{H}_\alpha}}{\left\| \sum_{i=1}^{N_1} w_i^{(1)} \phi_\alpha(p_i^{(1)}) \right\|_{\mathcal{H}_\alpha} \left\| \sum_{j=1}^{N_2} w_j^{(2)} \phi_\alpha(p_j^{(2)}) \right\|_{\mathcal{H}_\alpha}} \tag{5.25}$$

Using the bilinearity of the inner product $< .,. >_{\mathcal{H}_\alpha}$, we can express K_α using k_α :

$$\widehat{C}_\alpha(\mathbf{F}_1, \mathbf{F}_2) = \frac{\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} w_i^{(1)} w_j^{(2)} k_\alpha(p_i^{(1)}, p_j^{(2)})}{\left\| \sum_{i=1}^{N_1} w_i^{(1)} \phi_\alpha(p_i^{(1)}) \right\|_{\mathcal{H}_\alpha} \left\| \sum_{j=1}^{N_2} w_j^{(2)} \phi_\alpha(p_j^{(2)}) \right\|_{\mathcal{H}_\alpha}} \tag{5.26}$$

where $\left\| \sum_{i=1}^{N_k} w_i^{(k)} \phi_\alpha(p_i^{(k)}) \right\|_{\mathcal{H}_\alpha} = \sqrt{\sum_{i=1}^{N_k} \sum_{j=1}^{N_k} w_i^{(k)} w_j^{(k)} k_\alpha(p_i^{(k)}, p_j^{(k)})}$ for $k = \{1, 2\}$. Again this allows to derive a Hilbertian metric between fibers in a similar way to equation 5.16.

5.3.4 Alternative Unsupervised Clustering Techniques

In 5.3.1, we suggested the use of a kernel PCA technique for fiber clustering. In the following, we discuss two alternative strategies: diffusion maps and linear programming clustering. Both are used for manifold clustering, the first gives an estimation of a diffusion distance and the second overcomes some limitations of manifold embeddings.

Diffusion Maps

Diffusion maps [28] are a spectral embedding of a set X of n nodes, for which *local geometries* are defined by a kernel $\kappa : X \times X \to \mathbb{R}$. The kernel κ satisfies $\kappa(x, y) \geq 0$, and $\kappa(x, y) = \kappa(y, x)$. This kernel can be interpreted as an *affinity* between nodes. The resulting graph (an edge between x and y carries the weights $\kappa(x, y)$) can be transformed into a reversible Markov chain by the so called normalized graph Laplacian construction. In [69] a related construction was used to define a geometry on a set of observations,

or trajectories. We define

$$s(x) = \sum_y \kappa(x, y) \quad \text{and} \quad p(x, y) = \frac{\kappa(x, y)}{s(x)}. \tag{5.27}$$

This new kernel is no longer symmetric, but it satisfies

$$\forall x, \ \sum_y p(x, y) = 1. \tag{5.28}$$

Therefore it can be interpreted as the probability of the transition from node x to node y in one time step, or a *transition kernel* of a Markov chain. It gives a diffusion operator

$$Pf(x) = \sum a(x, y) f(y) d\mu(y), \tag{5.29}$$

P is the Markov matrix with the entries $p(x, y)$ and its powers P^τ allow to propagate information through the Markov chain in τ timesteps according to the transition kernels. According to P^τ we can define a family of *diffusion distances* parameterized by τ on the set of nodes

$$D_\tau(x, y) = \sum_{l=1,\ldots,m} \frac{(p_\tau(x, l) - p_\tau(y, l))^2}{\pi(l)} \tag{5.30}$$

where $\pi(l) = s(l)/\sum_j s(j)$ is the probability of the node x in the unique stationary distribution (the uniqueness is fulfilled if the graph is connected). D_τ is an L^2 distance between the posterior distributions of reaching x or y from all points l in the graph. It captures the connectivity in the Markov chain, summing over all possible paths from x to y. It is low if there is a large number of paths of length τ with high transition probabilities between the nodes x and y.

The operator P defines a geometry which can be mapped to an Euclidean geometry by an eigenvalue decomposition of P. The latter results in a sequence of eigenvalues $\lambda_1, \lambda_2 \ldots$ and corresponding eigenfunctions $\Psi_1, \Psi_2, \ldots$ that fulfill $P\Psi_i = \lambda_i \Psi_i$. The diffusion map after τ timesteps $\mathbf{\Psi}_\tau : X \to \mathbf{R}^w$ embeds each node $i = 1, \ldots, n$ in the Markov chain into a w dimensional Euclidean space where the clustering of the data points can be done using

k-means

$$i \mapsto \mathbf{\Psi}_\tau(i) \triangleq \begin{pmatrix} \lambda_1^\tau \Psi_1(i) \\ \lambda_2^\tau \Psi_2(i) \\ \vdots \\ \lambda_w^\tau \Psi_w(i) \end{pmatrix} \tag{5.31}$$

In this space, the Euclidean distance reflects the distances (parameterized by τ) defined by the Diffusion distance D_τ.

$$\|\mathbf{\Psi}_\tau(i) - \mathbf{\Psi}_\tau(j)\| = D_\tau(i,j). \tag{5.32}$$

Note that a common choice for the kernel $\kappa(.,.)$ is the Gaussian kernel, i.e. $\kappa(x,y) = \exp\left(-\frac{d^2(x,y)}{2\sigma^2}\right)$, where d is a distance over the set X and σ a scale factor. In our case, d corresponds to the fiber metric defined in section 5.3.

Manifold Clustering via Linear Programming

Clustering refers to the process of organizing a set of objects into groups such that the members of each group are as similar to each other as possible. A common way of tackling this problem is to formulate it as the following optimization task: given a set of objects $\mathcal{V} = \{p_1, \ldots, p_n\}$, endowed with a distance function $d(\cdot,\cdot)$ that measures dissimilarity between objects, the goal of clustering is to choose K objects from $\mathcal{V}$, say, $\{q_1, \ldots, q_K\}$ (these will be referred to as cluster centers hereafter) such that the obtained sum of distances between each object and its nearest center is minimized, or:

$$\min_{q_1,\ldots,q_K \in \mathcal{V}} \sum_{p \in \mathcal{V}} \min_i d(p, q_i) \ . \tag{5.33}$$

An important drawback of the above formulation is that it requires the number of clusters K to be provided beforehand, which is problematic as this number is very often not known in advance. Note that a wrong value for K may have a very negative effect on the final outcome. One would thus prefer K to be automatically estimated by the algorithm as a byproduct of the optimization process. To address this issue, we will let K be a variable here, and, instead of (5.33), we will use the following modified objective function, which additionally assigns a penalty $g(q_i)$ to each one of the chosen

cluster centers q_i:

$$\min_{K} \min_{q_1,\ldots,q_K \in \mathcal{V}} \left(\sum_{p \in \mathcal{V}} \min_{i} d(p, q_i) + \sum_{i=1}^{K} g(q_i) \right) \quad . \tag{5.34}$$

But, even if K is known, another serious drawback of many of the existing optimization-based techniques for clustering is that they are particularly sensitive to initialization and thus may get easily trapped in bad local minima. For instance, K-means (one of the most commonly used clustering methods) is doomed to fail if its initial cluster centers happen not to be near the actual cluster centers. To deal with that, here we will rely on a recently proposed clustering algorithm [66], which has been shown to yield approximately optimal solutions to the NP-hard problem (5.34). This algorithm relies on reformulating (5.34) as an equivalent integer program, whose LP-relaxation (denoted as PRIMAL hereafter) has the following form:

$$\text{PRIMAL} \equiv \min_{\mathbf{x}} \sum_{p,q \in \mathcal{V}, p \neq q} d(p, q) x_{pq} + \sum_{q \in \mathcal{V}} g(q) x_{qq} \tag{5.35}$$

$$\text{s.t.} \sum\nolimits_{q \in \mathcal{V}} x_{pq} = 1,\ x_{pq} \leq x_{qq},\ x_{pq} \geq 0 \tag{5.36}$$

If constraints $x_{pq} \geq 0$ are replaced with $x_{pq} \in \{0, 1\}$, then the resulting integer program is equivalent to clustering problem (5.34). In this case, each binary variable x_{pq} with $p \neq q$ indicates whether object p has been assigned to cluster center q or not, while binary variable x_{qq} indicates whether object q has been chosen as a cluster center or not. Constraints $\sum_{q \in \mathcal{V}} x_{pq} = 1$ simply express the fact that each object must be assigned to exactly one center, while constraints $x_{pq} \leq x_{qq}$ require that if p has been assigned to q then object q must obviously be chosen as a center. The most crucial issue for tackling this integer LP is setting the variables x_{qq} correctly, i.e, deciding which objects will be chosen as centers. To this end, the so-called *stability* of an object has been introduced in [66]. This is a measure which tries to quantitatively answer the following question: *how much does one need to further penalize an object to ensure that it will never be selected as an optimal cluster center?* Intuitively, the greater the stability of an object, the more appropriate that object is to become a cluster center. For having a practical algorithm based on object stabilities, an efficient way of

estimating them is required. It turns out that this can indeed be done very fast by moving to the dual domain and appropriately updating a solution of a dual relaxation to PRIMAL. Since each dual cost provides a lower bound to the cost of the optimal clustering, an additional advantage of working in the dual domain is the ability to provide online optimality guarantees and to avoid bad local minima. This concept is formalized with the help of the LP relaxation PRIMAL, thus leading to the following definition for the stability $S(q)$ of an object q:

$$S(q) = \inf\{\text{perturbation } s \text{ applied to penalty } g(q) \text{ (i.e., } g(q) \leftarrow g(q) + s) \\ \text{such that PRIMAL has no optimal solution } \mathbf{x} \text{ with } x_{qq} > 0\}$$

An object q can be stable or unstable depending on whether it holds $S(q) \geq 0$ or $S(q) < 0$. Furthermore, a stable object with a large value of $S(q)$ has high probability to actually be an optimal center in (5.34). The reason is that the assumption of a large $S(q) \geq 0$ is essentially a very strong requirement (much stronger than simply requiring q to be an active center in the relaxed problem PRIMAL): it further requires that q is an active center for a whole family of problems. Let us denote $\text{PRIMAL}_q(s)$ a modified instance of problem PRIMAL where the value s has been added to the penalty for object q, i.e, $g(q) \leftarrow g(q) + s$. This entails that, for all problems $\text{PRIMAL}_q(s)$ with $s \leq S(q)$. Based on this observation, the algorithm in [66] uses the following strategy for choosing its set of cluster centers $\mathcal{Q}$: it adds stable objects to $\mathcal{Q}$ in a sequential manner, each time selecting an object of approximately maximum stability. Furthermore, whenever it inserts a stable object q in $\mathcal{Q}$ (i.e, whenever it sets $x_{qq} = 1$), it also reestimates the stabilities of the remaining objects (this is because, for example, an object may become unstable if another object q gets selected as a new cluster center). To this end, the additional constraint $x_{qq} = 1$ is included in problem PRIMAL, and this updated LP is then used for the reestimation of stabilities. Of course, the whole process repeats until no more stable objects can be found. We refer the reader to [66] for more details.

We now discuss the case where the objects lie on a manifold. This implies the use of the geodesic distance as a similarity measure. Ideally this distance should correspond to the pairwise cost $d(p, q)$ for $p \neq q$ in the linear program-

ming formulation proposed in (5.35). A first possible choice is to compute the geodesic distances between all the pairs of points using the Dijkstra algorithm in an Isomap-like fashion, as suggested in [119]. The shortest path is found using a local approximation of the geodesic distance, for example a Euclidean distance. The pairwise cost $d(p,q)$ is set to $d(p,q) = d_g(p,q)$ where d_g is the corresponding geodesic distance. However, inspired by the landmark Isomap algorithm [34], we can compute the geodesic distances from all the data points to a reduced set of randomly selected landmarks. This will reduce the computational load that a full computation of the geodesic distances between every pair of data points would entail. Let $(l_m)_{m=1\ldots n_l}$ be a set of such chosen n_l landmarks. We would like to replace $d_g(p,q)$ by a reasonable approximation. Given that the geodesic distance between two points is the length of the shortest path linking these points, we note the following $\forall m \in [1 \ldots n_l]$, $|d_g(p,l_m) - d_g(q,l_m)| \leq d_g(p,q) \leq d_g(p,l_m) + d_g(q,l_m)$, which implies

$$\sup_m |d_g(p,l_m) - d_g(q,l_m)| \leq d_g(p,q) \leq \inf_m (d_g(p,l_m) + d_g(q,l_m)) \tag{5.37}$$

This provides a lower bound and an upper bound to the cost $d_g(p,q)$ in the case where only the geodesic distances to some landmarks are computed. Note that in the particular case where p and q are landmarks $d_g(p,q) = \sup_m |d_g(p,l_m) - d_g(q,l_m)| = \inf_m (d_g(p,l_m) + d_g(q,l_m))$. On the other hand we can also note that

$$\inf_m (d_g(p,l_m) + d_g(q,l_m)) - 2\eta \leq \quad d_g(p,q) \tag{5.38}$$

$$d_g(p,q) \leq \quad \sup_m |d_g(p,l_m) - d_g(q,l_m)| + 2\eta \tag{5.39}$$

where $\eta = \inf_m \min(d_g(p,l_m), d_g(q,l_m))$. Therefore it makes sense to replace the cost $d_g(p,q)$ whether by its upper bound or its lower bound, since both approximate the cost up to 2η. A byproduct of inequalities (5.38) and (5.39) is that both approximations are exact if p *or* q are landmarks, since in that case we have $\eta = 0$.

It is interesting to note in this setting that the lower bound is the L^∞ norm between the distance-to-landmarks representation of p and q. Indeed, let $\mathbf{u_p}$ (resp. $\mathbf{u}_q$) be the n_l-dimensional vector of geodesic distances of p

(resp. q) to the landmarks

$$\mathbf{u}_p = [d_g(p, l_1), \ldots, d_g(p, l_{n_l})]^t \quad , \quad \mathbf{u}_q = [d_g(q, l_1), \ldots, d_g(q, l_{n_l})]^t \tag{5.40}$$

By definition, $\sup_m |d_g(p, l_m) - d_g(q, l_m)| = ||\mathbf{u}_p - \mathbf{u}_q||_\infty$. Thus the lower bound approximation has the advantage of defining a metric cost. Intuitively, for a number of landmarks sufficiently larger than the intrinsic dimension of the manifold, the distance vector representation will provide a good characterization of the points on the manifold.

5.3.5 A Supervised Fiber Clustering Framework

When considering the special case where the tensor field is constant (equal to a tensor $\mathbf{D}$), we obtain the following equation for the fiber kernel:

$$K_t(\mathbf{F}_1, \mathbf{F}_2) \propto \frac{1}{n_1}\frac{1}{n_2} \sum_{\mathbf{x}_i \in \mathbf{F}_1} \sum_{\mathbf{x}_j \in \mathbf{F}_2} \exp\left(-\frac{1}{8t}(\mathbf{x}_i - \mathbf{x}_j)^t \mathbf{D}^{-1}(\mathbf{x}_i - \mathbf{x}_j)\right) \tag{5.41}$$

which is the anisotropic form of the Gaussian kernel correlation [108]:

$$KG_t(\mathbf{F}_1, \mathbf{F}_2) \propto \frac{1}{n_1}\frac{1}{n_2} \sum_{\mathbf{x}_i \in \mathbf{F}_1} \sum_{\mathbf{x}_j \in \mathbf{F}_2} \exp\left(-\frac{||\mathbf{x}_i - \mathbf{x}_j)||^2}{8\mu t}\right) \tag{5.42}$$

where $\mathbf{D} = \mu\mathbf{Id}$ which means that we suppose the tensors are isotropic. We can see that the proposed kernel deals with a generic tensor field and provides a generalization of the Gaussian kernel correlation. A by-product of this reasoning is that Gaussian kernel correlation is a kernel on fibers that considers only point positions. One could have seen from the outset that is a Mercer kernel since it is a summation of Mercer (Gaussian) kernels. This will be particularly useful to learn spatial interactions between fibers as will be detailed in the following.

Given an atlas of fibers segmented by an expert in R regions, we can learn the spatial interactions between these fibers using the Gaussian kernel correlation in Equation 5.42. Indeed, it can be used as an input in a kernel Support Vector Machines (SVMs) [110] to learn boundaries between the different segmented regions, including background fibers that are not localized in the region of interest. The kernel SVMs provide support vectors

(fibers), which are the fibers that define the decision boundaries. Note that the SVMs are used in a one-against-one fashion in order to deal with multiple regions. A point of interest here is that from the initial set of training fibers, we have only to keep a sparse subset of support fibers that will guide the classification process. The fact that the Gaussian kernel correlation defines a Mercer kernel allows to look at the problem from a discriminative classification perspective.

Given a new set of fibers that we wish to segment in a manner consistent with the already-defined atlas, we start by finding the affine transformation that maps the diffusion-free (B0) image of the training atlas to the corresponding one in the testing dataset, as in [78]. This tranformation is subsequently used to register the testing set of fibers to the space of the support fibers obtained from the atlas. We are therefore able to compute the scores of the $\frac{R(R-1)}{2}$ pairwise SVM classifiers on the testing dataset. In the following subsection, we show how to use the SVM classification results to obtain the segmentation while respecting the interactions provided by the fiber kernel.

5.4 Results and Experiments

Thirty subjects (twenty healthy patients and ten patients affected by myopathies) underwent a diffusion tensor imaging of the calf muscle using a 1.5 T MRI scanner. The following parameters were used : repetition time (TR)= 3600 ms, echo time(TE) = 70 ms, slice thickness = 7 mm and b value of 700 $s.mm^{-2}$ with 12 gradient directions and 13 repetitions. The size of the obtained volumes is $64 \times 64 \times 20$ voxels with a voxel resolution of $3.125\,mm \times 3.125\,mm \times 7\,mm$. We acquired simultaneously high resolution T1-weighted images that were segmented manually by an expert into 7 muscle groups to provide the ground truth and fiber trajectories were reconstructed using [44]. We recall the muscle architecture of the calf in [Fig.5.3 (a)], where we show a manual segmentation overlaid on an axial slice of a high-resolution T1-weighted image. As in chapter 4, the following muscle groups are considered: the soleus (SOL), lateral gastrocnemius (LG), medial gastrocnemius (MG), posterior tibialis (PT), anterior tibialis (AT), extensor digitorum longus (EDL), and the peroneus longus (PL).

(a) (b)

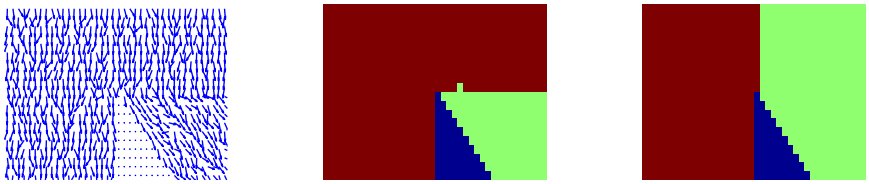
Figure 5.3: (a) An example of a manual segmentation of an axial slice of a high-resolution T1-weighted image showing different muscle groups in the calf. (b) An axial slice of a T1-weighted image of the calf of a diseased patient where the zone in hypertension is fat that replaced the muscle.

In the following, we present the obtained experimental results on a synthetic dataset and for tensor classification and fiber bundling of the lower leg muscles, both in supervised and unsupervised ways.

5.4.1 Tensor Clustering

Preliminary Experiments

We first generated a 20×40 lattice of synthetic tensors composed of two close fiber bundles. The first bundle has a vertical principal direction, the second starts with a vertical direction then deviates with a 45° angle [Fig.5.4 (a)]. We added a Gaussian noise of standard deviation 10° to these directions. The eigenvalues of the tensors were set to $\{2\ 10^{-3}, 1.5\ 10^{-3}, 10^{-3}\}$. We tested the following values of diffusion time t: $\{10^4, 10^5\}$. We compare the behavior of the kernel PCA + Isomap embedding with spectral clustering using the metric $d(\mathbf{D_1}, \mathbf{D_2}) = \arccos(<\vec{\mathbf{e}_1}, \vec{\mathbf{e}_2}>)$ where $\vec{\mathbf{e}_1}$ (resp. $\vec{\mathbf{e}_2}$) is the principal direction of diffusion of $\mathbf{D}_1$ (resp. $\mathbf{D}_2$), as in [139]. Following [139], the scale parameter in the affinity matrix is set as the sample variance of d between neighboring tensors. The clustering is obtained using k-means with 50 restarts in the spectral embedding space. The result is obtained by considering the clustering with the least distortion. Distortion is computed as the ratio of intra-class and inter-class variances. However, we do not



(a) (b) (c)

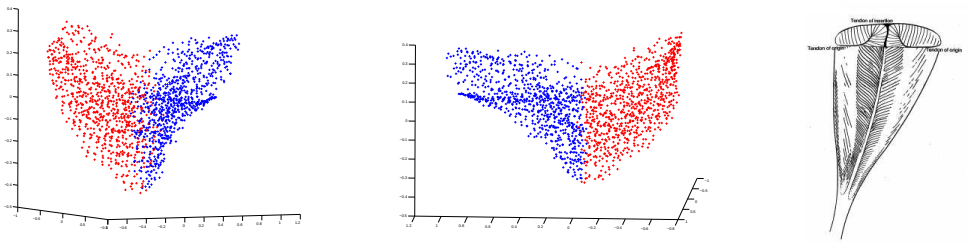
Figure 5.4: (a) Synthetic noisy field of principal directions of diffusion. (b) Result of spectral clustering based on angles between principal directions of diffusion without Markovian relaxation. (c) Result of our method

use a Markovian relaxation of the affinity matrix, hence removing spatial constraints. [Fig.5.4 (c)] shows the segmentation result obtained by our approach (stable across the tested set of diffusion times) and [Fig.5.4 (b)] shows its counterparts for spectral clustering without Markovian relaxation. We can notice that unlike spectral clustering, the proposed algorithm finds a clustering solution which is more compatible with the tensor arrangement. This is due to the fact that it captures both tensor similarity and spatial connectivity and that the spectral clustering considered in this experiment does not enforce spatial regularity.

To further assess qualitatively the method, we used the proposed kernel PCA + Isomap embedding to see if it is faithful to the known structure of the muscles. Of particular interest is the soleus which is a major part of the calf. It has a bipennate structure where oblique fibers converge towards a central aponeurosis [Fig.5.5 (c)]. In [Fig.5.5 (a), (b)], we show the (here three-dimensional) proposed embedding for the soleus muscle of one subject for $t = 2\,10^5$. The obtained points reveal the structure of the muscle, which means that the embedding is faithful to the diffusion flow in the tissues as the points are aligned along the diffusive pathways.

Unsupervised Classification of the Muscle Groups

For each subject, a region of interest (ROI) was manually delineated and we tested the performance of the tensor clustering algorithm both for healthy and diseased subjects. Recall that in regions affected by myopathies, the tensors have a relatively small volume since fat replaces the fibers (as can be



(a) (b) (c)

Figure 5.5: (a), (b) Two views of a three-dimensional embedding of the tensors of the soleus muscle, k-means clustering shows its bipennate structure. (c) Anatomy of the soleus [82].

seen in [Fig.5.3 (b)]). These were eliminated through simple thresholding over the determinant. In all the experiments, we set the diffusion time t to $t = 2\,10^5$ and we used a ten-dimensional embedding. For a quantitative evaluation of the method, we follow the validation protocol proposed in [139]: for each subject, the manually delineated ROI was segmented at two different levels (in 7 and 10 classes). The resulting clusters are then classified according to the labels given by the expert. As in [139], several clusters are allowed to have the same label. We also test the algorithm on a section near the knee which is characterized by a higher amount of noise and artifact obtained automatically by a threshold (set to 20) on the diffusion-free images. We report in [Fig.5.11] the boxplots of the dice volume overlap with the expert labels for all the subjects and its counterpart for the spectral clustering as described previously but with spatial regularity being enforced by the means of Markovian relaxation. We can see that our algorithm performs slightly better for the case of the manually-delineated ROI, however for the noisy automatic ROI, spectral clustering is misled by isolated points (see [Fig.5.7 (a), (b)]), whereas the performance of the presented algorithm is not significantly worsened (the problem with the Markovian relaxation is that it relies on the 3D grid structure and continuity of the image, this can be addressed using hierarchical clustering instead of k-means, as in [139]). Note also that the thresholding over the determinant removes some of the tensors originating from the chemical shift artifact that do not belong to the foreground. From a qualitative point of view, one can see in [Fig.5.6

(a), (b)] that the algorithm was able to segment correctly fine structures like AT and PL and the segmentation result is rather smooth. Additional examples of unsupervised tensor classification are provided in [Fig.5.8] and in [Fig.5.9]. We can observe that in [Fig.5.9 (b)] the algorithm did not achieve an optimal segmentation result since it chose to assign the tensors belonging to the artifact at the boundary between the soleus and the lateral gastrocnemius to a separate class.

Supervised Tensor Classification

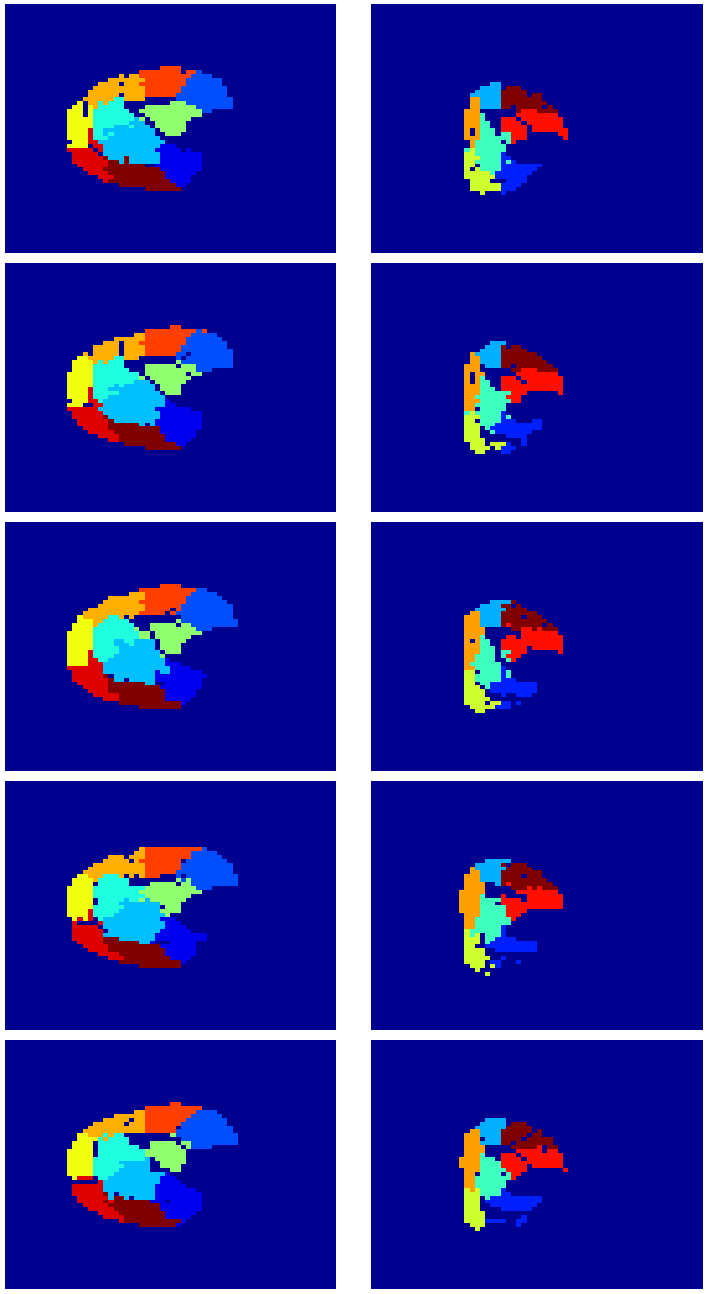
The major components of the calf muscle are the two heads of the gastrocnemius and the soleus. These three muscle groups have different orientations. This motivates the learning of the tensors of each class in order to have a prior information about the class membership of a tensor in a volume that we aim to segment. We consider four manually labeled volumes of randomly chosen healthy subjects as a training set (approximately 27500 tensors are used for training, note that the size of the training dataset is sufficiently larger than the dimension of the space of symmetric matrices which contains the tensors). The sixteen remaining images of healthy patients are used as the testing dataset. We build three one-against-all SVM classifiers based on the training set. The parameter α in the single potential u^s of the MRF energy is set to $\alpha = 10$ in all the experiments. The trade off factor λ in [Eq.5.19] is set to be $\lambda = 5\frac{n_{nodes}}{n_{pairs}}$ where n_{nodes} is the number of nodes (tensors) and n_{pairs} is the number of interacting pairs of tensors.

The boxplots of the computed dice coefficients are shown in [Fig.5.23] when only SVM scores are used and when these are combined with MRF regularization. We can observe that the MRF regularization improves slightly the classification results. Observe in [Fig.5.10] how the use of the MRF model results in smoother results and removes part of the artifact near the boundary between SOL and LG.

5.4.2 Fiber Clustering

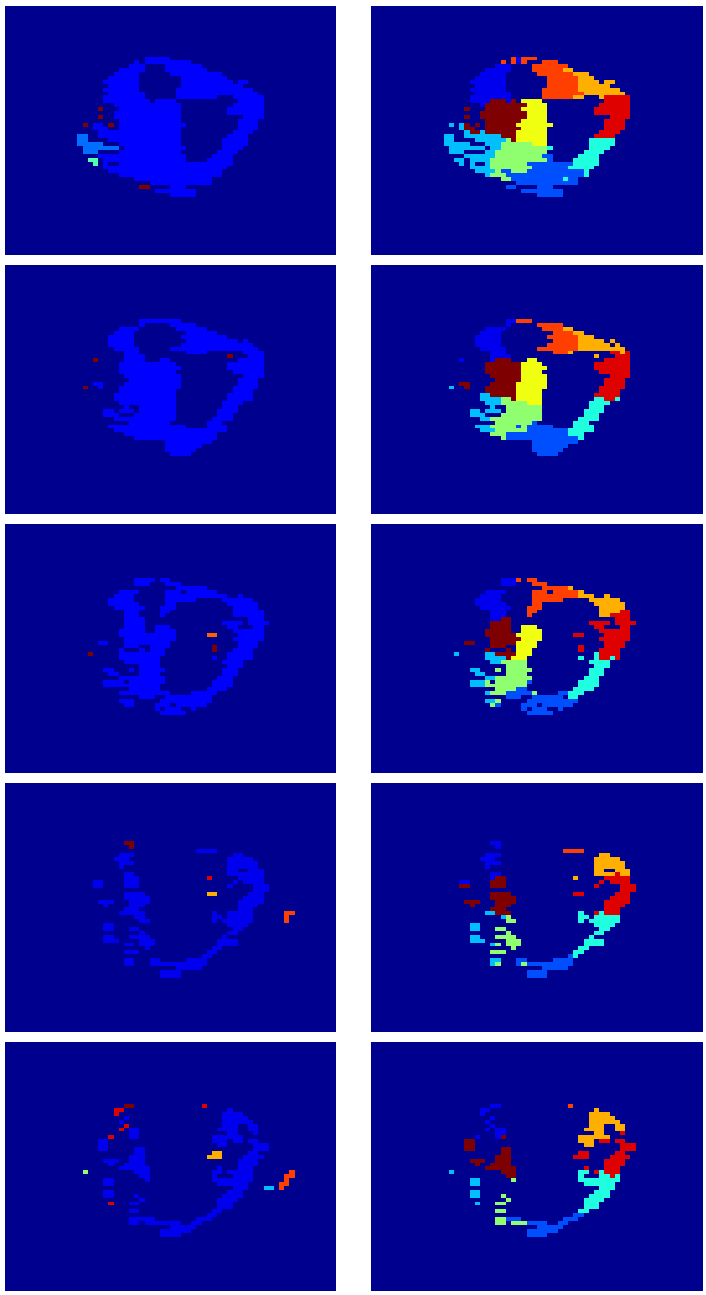
Unsupervised Fiber Clustering using Kernel PCA

To test the unsupervised kernel-PCA clustering algorithm, we only kept the fibers which have a majority of points lying in the manually delineated ROI.



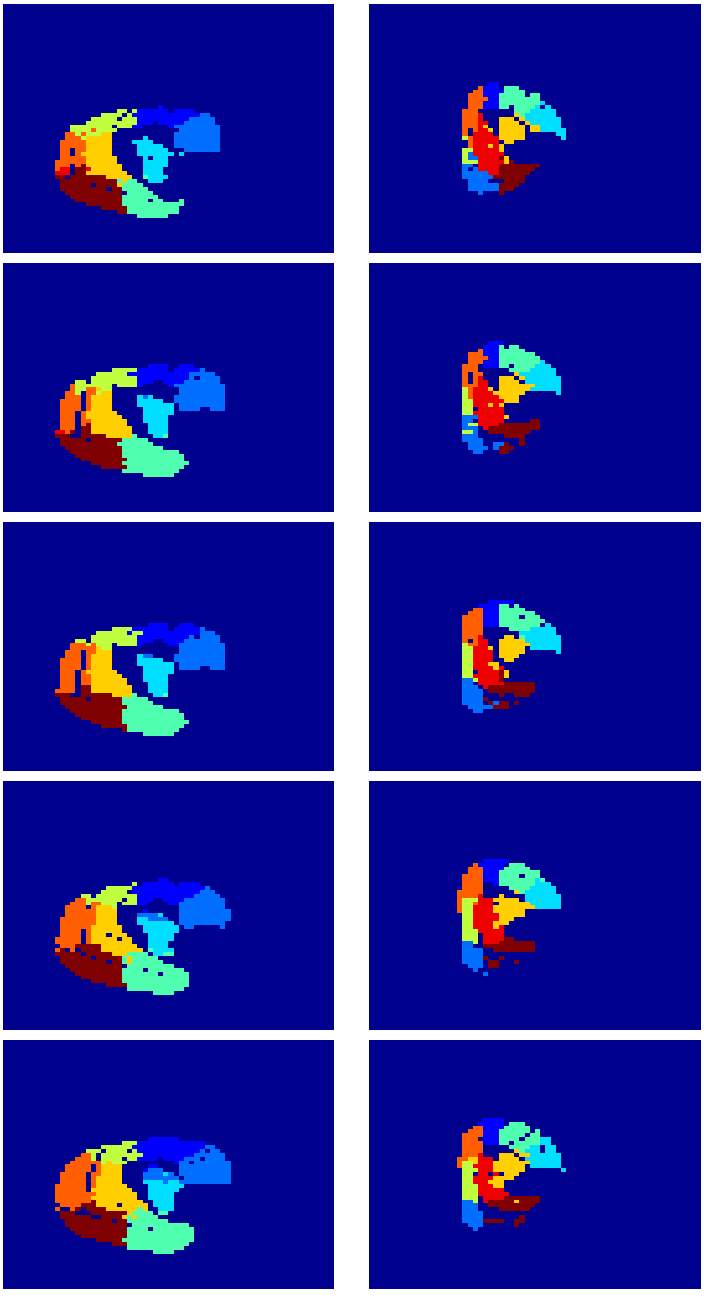
(a) (b)

Figure 5.6: Axial slices of the segmentation of the tensors for (a) a healthy subject in 10 classes, manual ROI (b) a diseased subject in 7 classes where the MG is partially affected, manual ROI.



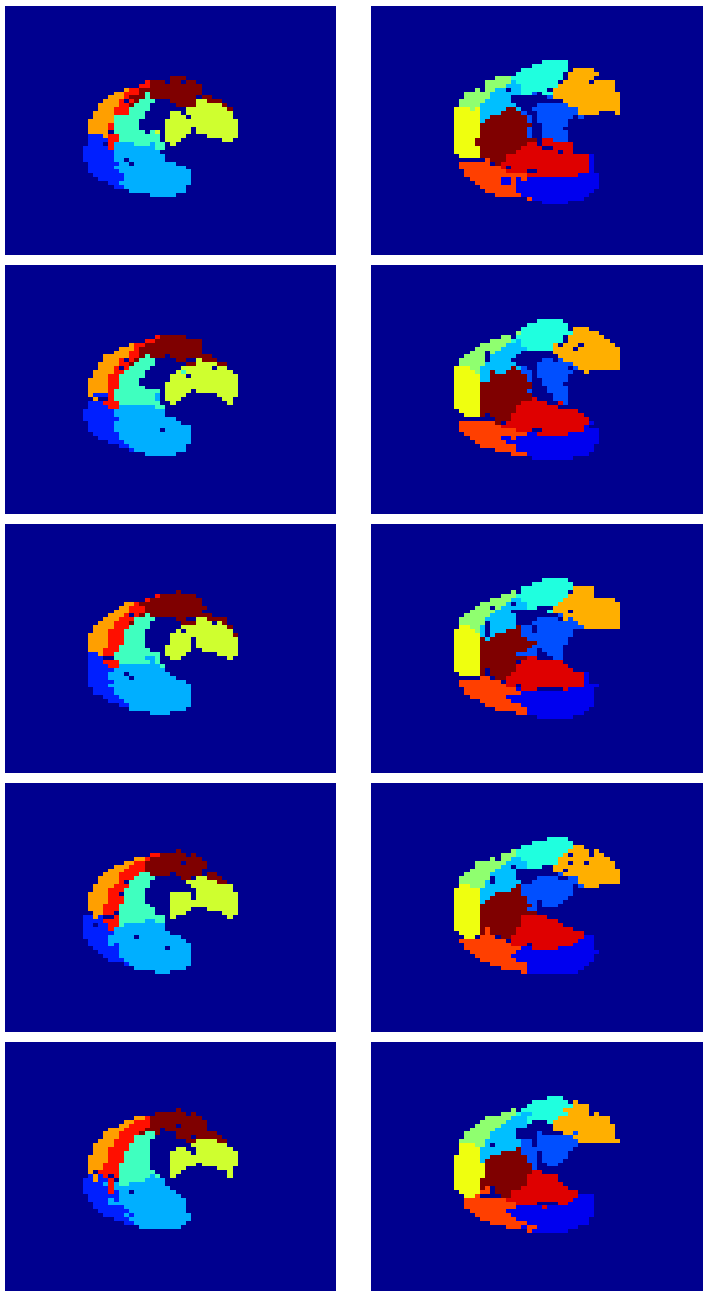
(a) (b)

Figure 5.7: Axial slices of the segmentation of the tensors for (a) noisy automatic ROI of a section near the knee using spectral clustering and k-means (b) noisy automatic ROI using our method.



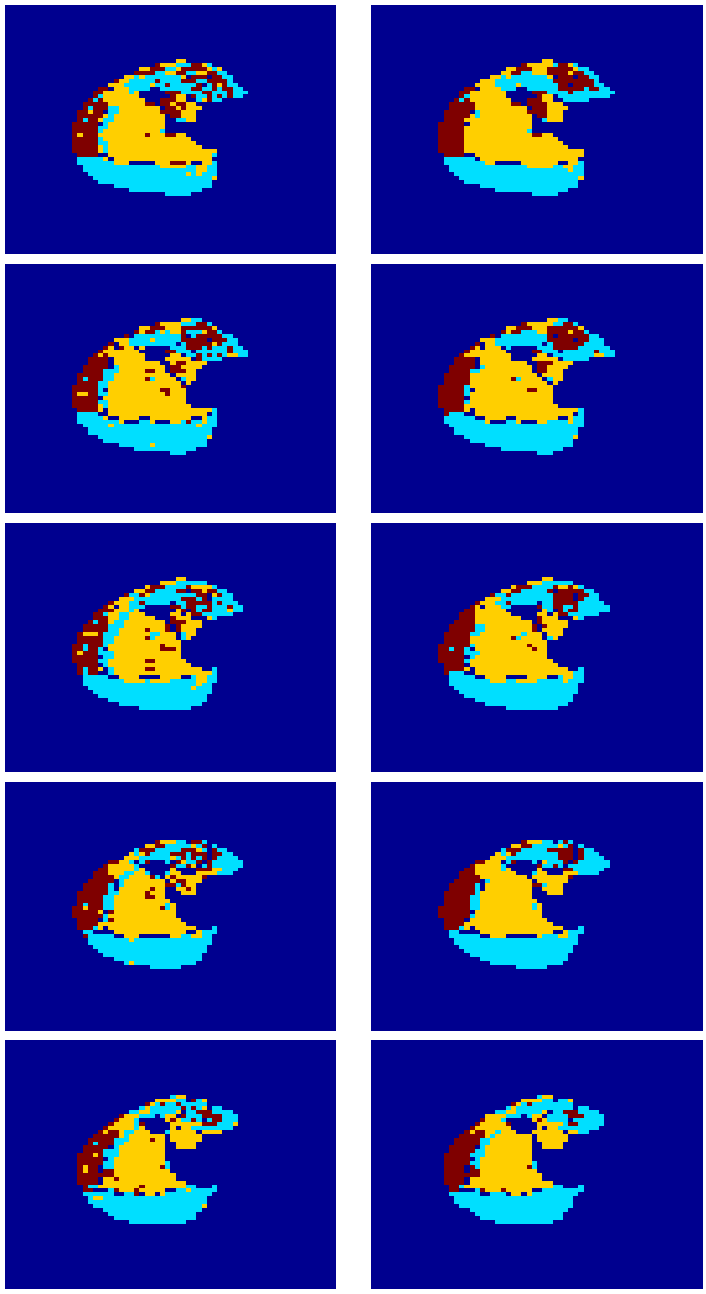
(a) (b)

Figure 5.8: Additional tensor clustering results for (a) a healthy subject, (b) a diseased subject.



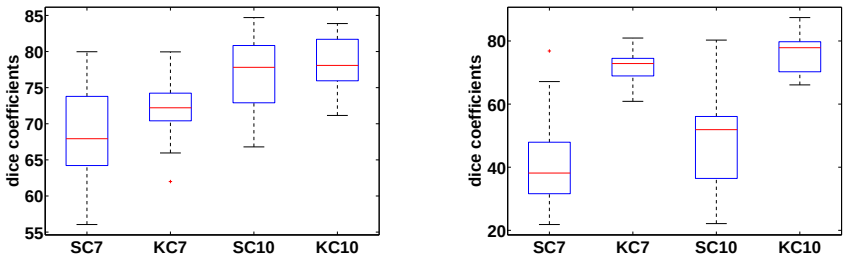
(a) (b)

Figure 5.9: Supplementary examples of tensor clustering. Note that in (a) the algorithm mistakes the artifact (in red) for a class.



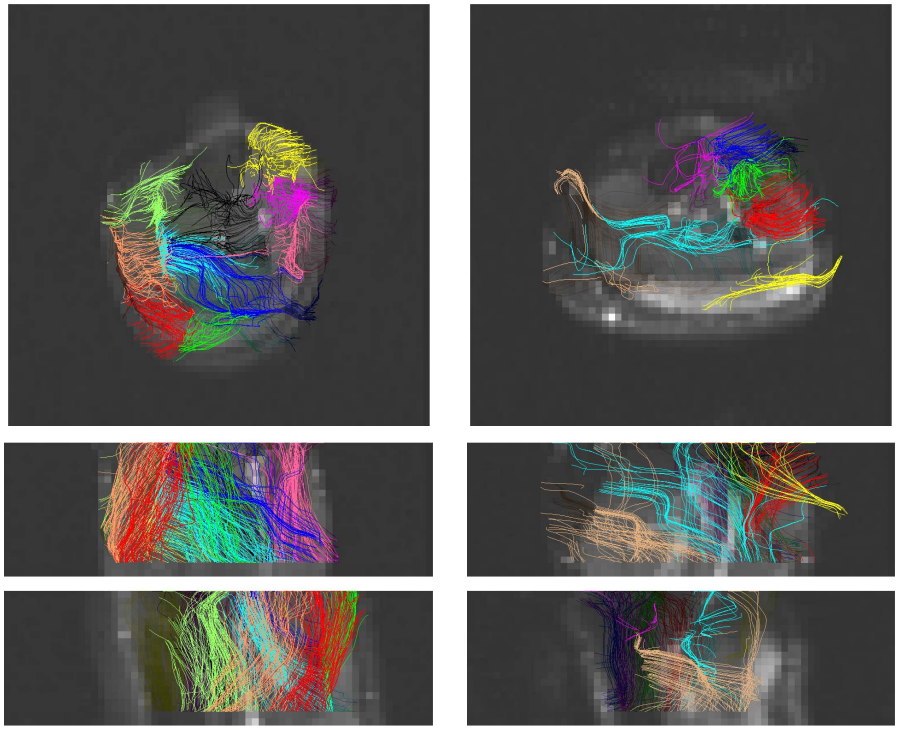
(a) (b)

Figure 5.10: Axial slices of the supervised SVM classification of tensors in three classes (MG, LG and SOL) (a) SVM scores only are used. (b) SVM scores are combined with MRF regularization.



(a) (b)

Figure 5.11: Boxplot of the dice overlap coefficients for tensor clustering in 7 and 10 classes for (a) manual ROI and (b) automatic noisy ROI. SC7 (resp. SC10) refers to spectral clustering in 7 (resp. 10) classes and KC7 (resp. KC10) refers to kernel clustering in 7 (resp. 10) classes. The box has lines at the lower quartile, median, and upper quartile values. The whiskers are lines extending from each end of the box to show the extent of the rest of the data. Outliers are data with values beyond the ends of the whiskers.



(a) (b)

Figure 5.12: Axial, coronal and sagittal views of kernel PCA fiber segmentation for (a) a healthy subject in 10 classes, (b) a diseased subject in 7 classes.

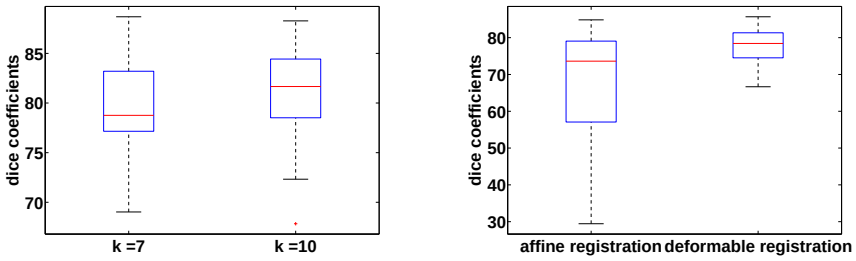


Figure 5.13: Boxplot of the dice overlap coefficients for (a) unsupervised kernel PCA fiber clustering in 7 and 10 classes, (b) supervised fiber clustering using affine and deformable registration.

The number of fibers in the different datasets ranged approximately from 1000 to 2500, which makes the computation and eigenanalysis of the kernel matrix achievable in a rather reasonable time. The diffusion time t was set to $t = 2\,10^4$ and we used a ten-dimensional embedding for kernel-PCA. Figure 5.12 shows the clustering results in 10 (resp. 7) classes for the fiber tracts of a healthy (resp. diseased) subject. The fibers in [Fig.5.12 (a)] correspond to the tensors segmented in [Fig.5.6 (a)]. It is interesting to notice that despite the fact that the tractography algorithm was unable to recover fiber tracts in the diseased regions due to the presence of degenerate tensors, the clustering algorithm could still segment the fiber tracts of the healthy region in anatomically relevant subgroups. Additional visual examples that illustrate the performance of the algorithm are shown in [Fig.5.14]. For quantitative assessment, we report in [Fig.5.13 (a)] the boxplot of the dice overlap measures of the fiber segmentation with the expert labeling for 7 and 10 classes. Overall, the algorithm performs well in separating the regions of the calf muscle with a mean dice coefficient of 79.5% (respectively 80.93%) and a standard deviation of 5.04% (respectively 5.14%) for 7 (respectively 10) classes.

Unsupervised Fiber Clustering Using Diffusion Maps

We tested the diffusions maps clustering method at two levels: for 7 and 10 clusters. The dimension of the diffusion maps embedding w was set to the number of clusters. The number of timesteps τ in (5.31) was set

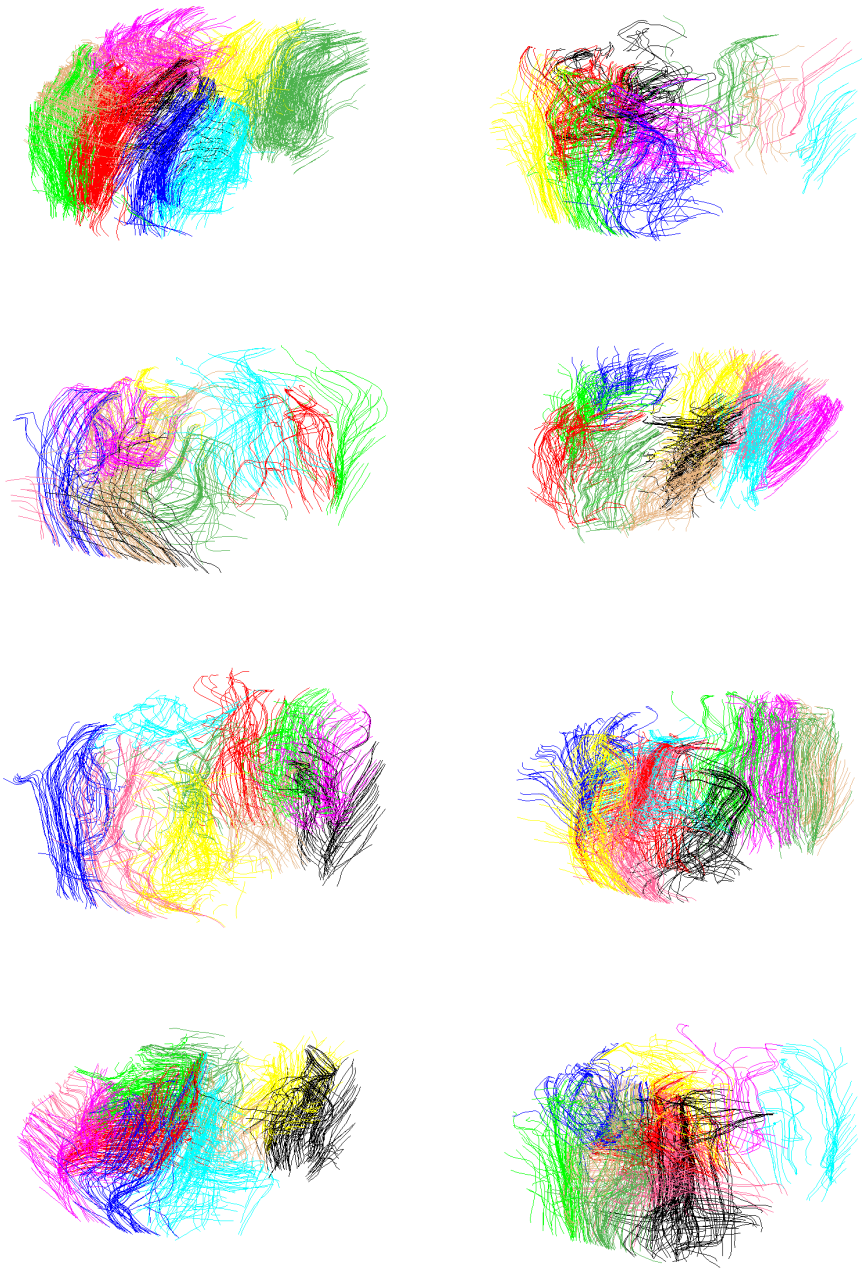


Figure 5.14: Some visual examples of Kernel-PCA fiber segmentation.

to $\tau = 1$. We tested two values for the scale parameter of the Gaussian kernel: $\sigma = \{0.5, 1\}$ and in all the experiments, the diffusion time t was set to $t = 2\ 10^4$. Again, the clustering in the embedding space is done using k-means with 50 restarts.

In [Fig.5.17], we present the boxplots of the dice overlap coefficients for the thirty subjects, using the different values of σ for 7 and 10 clusters. We can note that the quantitative results are rather satisfactory. For example, with the parameter σ set to $\sigma = 1$, we obtain a mean dice overlap coefficient of 0.78 (resp. 0.82) and a standard deviation of 0.05 (resp. 0.045) for 7 (resp. 10) classes . For a qualitative assessment, we show in [Fig.5.15, 5.16] the obtained fiber classification in 7 clusters for two healthy and two diseased subjects. The obtained fiber tracts are rather noisy, especially for diseased patients where the fat artifact is stronger. Despite the low quality of the tractography, the algorithm could still segment it in coherent fiber bundles.

Unsupervised Fiber Clustering Using Linear Programming

In all the linear programming clustering experiments we set the diffusion time to $t = 2\ 10^4$ and the parameter α in the fiber metric to $\alpha = 1$. To compute the Hilbertian metrics between fiber tracts, the weights w_i of each fiber $\mathbf{F}$ in (5.26) were chosen as the inverse of the number of points in $\mathbf{F}$. We selected 30% of the fibers as landmarks and for the computation of the geodesic distances using the Dijkstra algorithm, we considered a k-NN graph where k was set to $k = 12$. The cost $g(\mathbf{F})$ of choosing a fiber $\mathbf{F}$ as a class center in (5.35) was set to a constant $g = \beta\ \mu_{\frac{1}{2}}\left(d_g(\mathbf{F}_i, \mathbf{F}_j)_{i \neq j}\right)$ where $\mu_{\frac{1}{2}}$ is the statistical median. We tested the following values of β: $\{7, 10, 13\}$. For the sake of comparison, we evaluate also the performance of k-means clustering using the same metric and a manifold embedding. The dimensionality of the embedding is chosen to be the number of clusters obtained by our method, which is a common choice in embedding-based approaches. The k-means algorithm is run 50 times and each time we compute the dice overlap of the clustering result with the ground-truth segmentation. We consider both the average dice coefficients over the restarts of the k-means algorithm and the dice coefficient of the clustering with the least distortion. We run the following experiments:

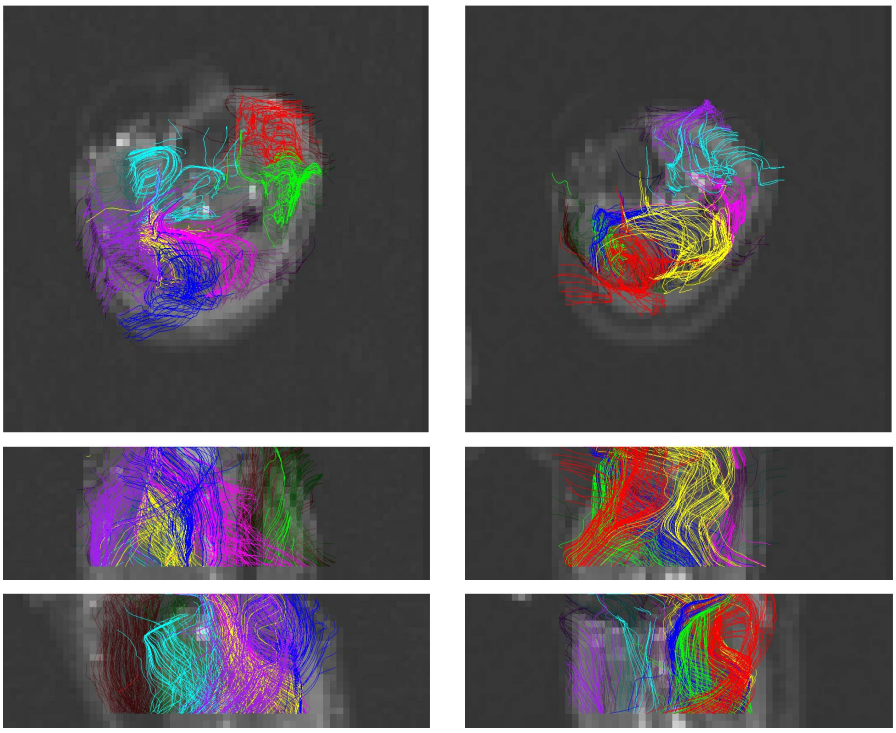


Figure 5.15: Axial, coronal and sagittal views of diffusion-maps based tractography segmentation in 7 classes overlaid on diffusion-free (B0) images for two healthy subjects

1. We compute all the geodesic distances between every pair of points and use them for linear programming clustering. We compare the obtained result with an Isomap embedding followed by k-means.

2. We compute the geodesic distances to a set of landmarks and use the lower (resp. upper) bound approximation for linear programming clustering. We compare the obtained result with a landmark-Isomap embedding followed by k-means.

We provide in [Fig.5.18, 5.19, 5.20] the boxplots showing the distributions of the dice coefficients for the thirty subjects using different values of β for our algorithm, compared with k-means after manifold embedding. We can note that linear programming clustering performs significantly better than the average score achieved by k-means both for a full and landmark-based computation of the geodesic distances. Furthermore, it achieves results equivalent

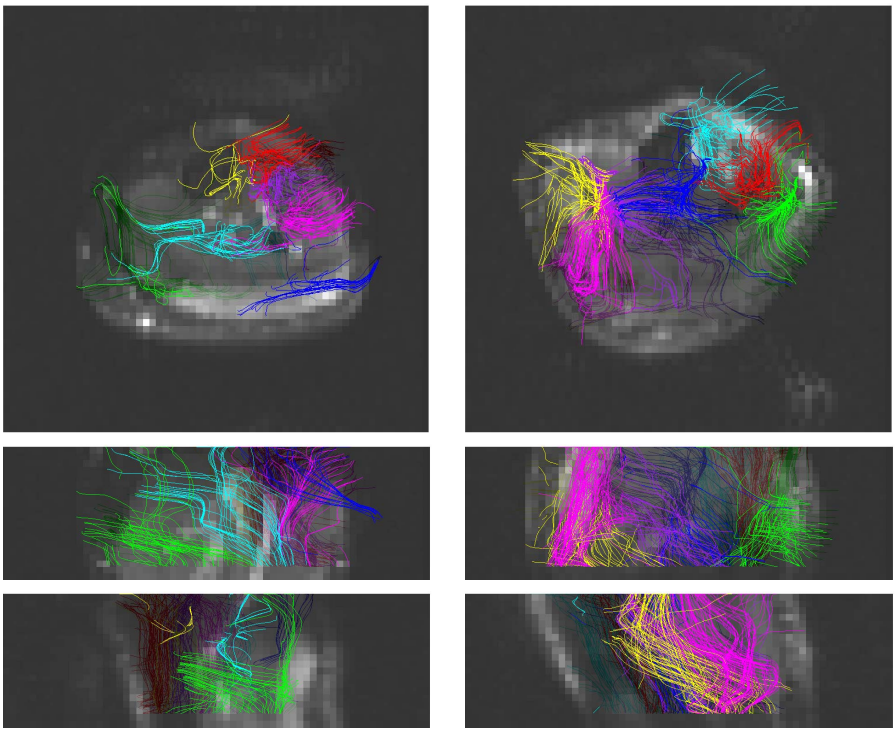


Figure 5.16: Axial, coronal and sagittal views of diffusion-maps based tractography segmentation in 7 classes overlaid on diffusion-free (B0) images for two diseased subjects

to the best k-means with an average dice coefficient of approximately 0.8 and in some cases it improves marginally the dice overlap. The advantage is that our result is reproducible, i.e. unlike k-means it is not sensitive to initialization. When comparing the three versions of linear programming clustering, we can see in [Fig.5.21] that the lower bound and upper bound approximations perform similarly apart from the case $\beta = 10$ where the lower bound approximation performed better, which may be explained by the metricity of the corresponding cost. The full computation yields slightly better results than the approximations. This corroborates the analysis provided is section 5.3.4. For qualitative evaluation, we show in [Fig.5.22 (a)] (resp. [Fig.5.22 (b)]) a clustering result obtained for a healthy (resp. diseased) subject for $\beta = 10$. Ground truth segmentation for the healthy subject is provided in [Fig.5.22 (c)]. There are too few fibers in [Fig.5.22 (b)] because the tractography fails to recover fibers through the manual region of interest. This is

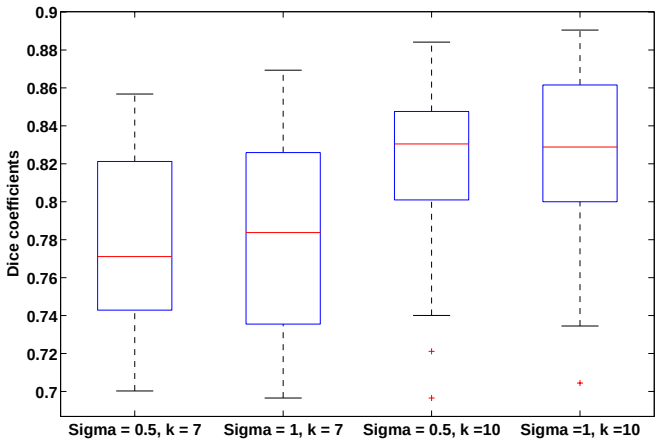


Figure 5.17: Boxplots of dice overlap coefficients for the thirty subjects with different values of σ and k (the number of clusters) for the diffusion-maps based clustering.

due to the presence of tensors with very low determinant (low diffusion). It is interesting to note that with the same parameter $\beta = 10$, the algorithm found ten clusters for the healthy subject while it found only three for the diseased patient, which seems to reflect the advantage of letting the number of clusters a variable of the optimization problem. Note also how the soleus (in cyan in [Fig.5.22 (c)]) is subdivided in an anterior and a posterior part in [Fig.5.22 (a)], which is consistent with its anatomy.

Supervised Fiber Clustering Using Kernel SVMs

A manually segmented volume was used as an atlas. Atlas fibers were assigned to a class based on a simple voting procedure: each fiber is classified according to the majority vote class of the voxels it crosses. We experimented both with affine [44] and deformable [51] registration to map the B0 images of the testing case to the B0 images of the atlas. We use kernel SVM classification to learn the fibers of the atlas as explained in subsection 5.3.5, using 21 one-against-one pairwise classifiers. The scale parameter in the Gaussian correlation kernel was set to $\sigma = 2\sqrt{\mu t} = 10$. We report in [Fig.5.13 (b)] the boxplot of the dice overlap coefficients both for deformable and affine registration. We can note that we obtain significantly

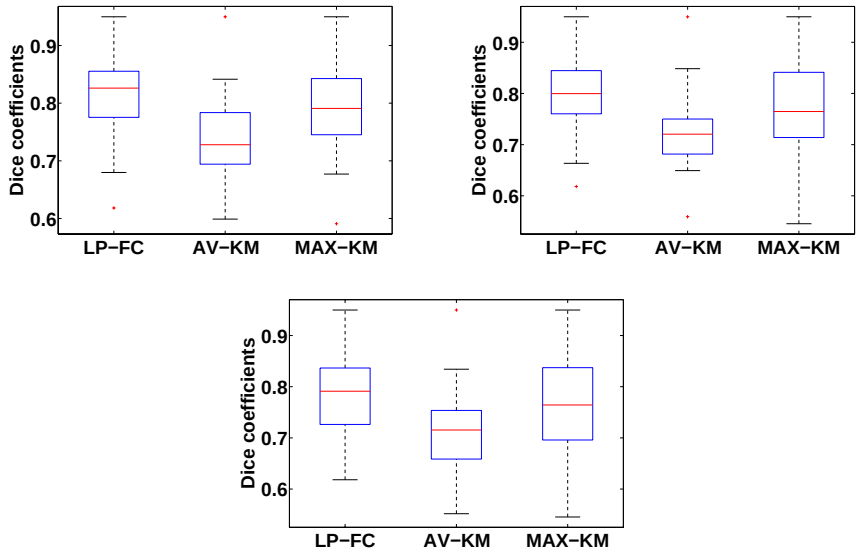


Figure 5.18: Boxplots of dice overlap coefficients for the thirty subjects and linear programming clustering using full computation of distances (LP-FC). Clockwise β takes the following values 7 , 10 and 13. Comparison is done with respect to the average score of k-means (AV-KM) and the score of the k-means clustering with least distortion (MAX-KM) after manifold embedding. (b) LP clustering using lower bound approximation (LP-LLWBD). (c) LP clustering using upper bound approximation (LP-LUPBD). (d) Comparison between LP-FC, LP-LLWBD and LP-LUPBD.

better results with deformable registration, which was expected given the relatively high inter-patient variability and that muscles are soft tissues, so the anatomy and shape are likely to vary significantly across patients. In [Fig.5.24 (a)], we show an example of supervised segmentation compared with the ground truth in [Fig.5.24 (b)]. We can observe that as opposed to the unsupervised setting ([Fig.5.12 (a)]), the MG is not oversegmented.

5.5 Conclusion

In this chapter, we proposed a kernel-based method for clustering of both tensors and fibers in diffusion tensor images. It exploits the physical interpretation behind the modality and offers a unified approach towards tensor and fiber grouping. The kernel defined over the tensor space en-

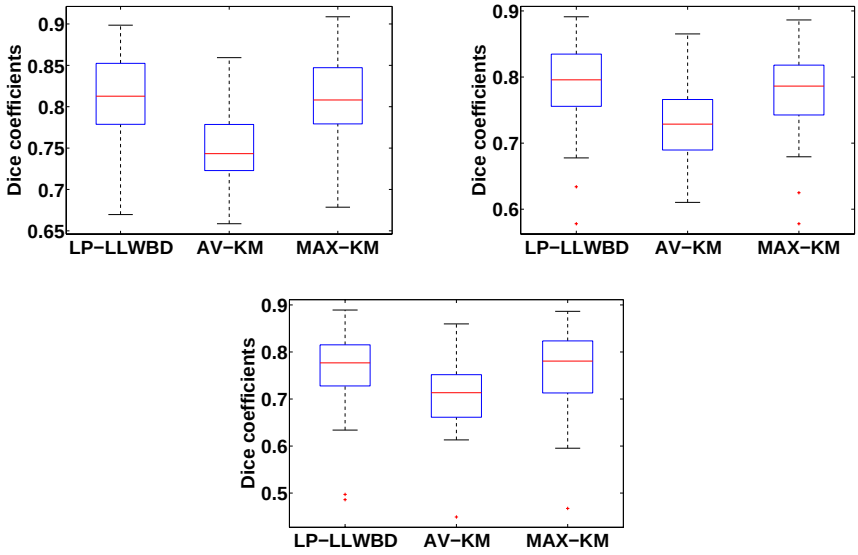


Figure 5.19: Boxplots of dice overlap coefficients for the thirty subjects and linear programming clustering using lower bound approximation (LP-LLWBD).

compasses both localization and diffusion information and naturally reflects tensor alignment along fiber tracts. We showed its flexibility by extending it to deal with fibers and gave the physical intuition behind its mathematical definition as a kernel over sets of tensors. Moreover, we investigated the use of diffusion maps and linear programming clustering using the derived metrices for unsupervised fiber clustering. We also showed how to include expert knowledge by means of kernel SVMs.

The next chapter of this thesis will focus on the use of the defined kernels for deformable registration of diffusion tensors using a discrete optimization framework in order to minimize an objective function that is based on the notions used in this chapter for segmentation and clustering.

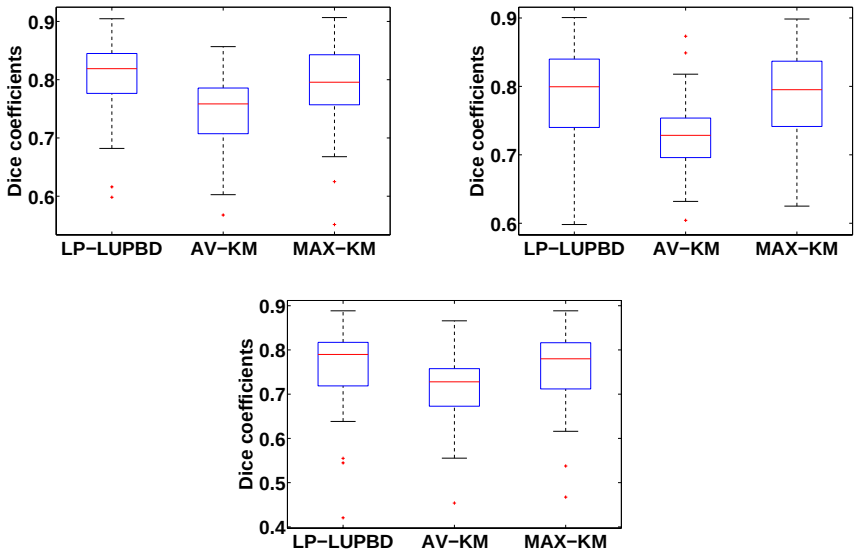


Figure 5.20: Boxplots of dice overlap coefficients for the thirty subjects and linear programming clustering using upper bound approximation (LP-LUPBD).

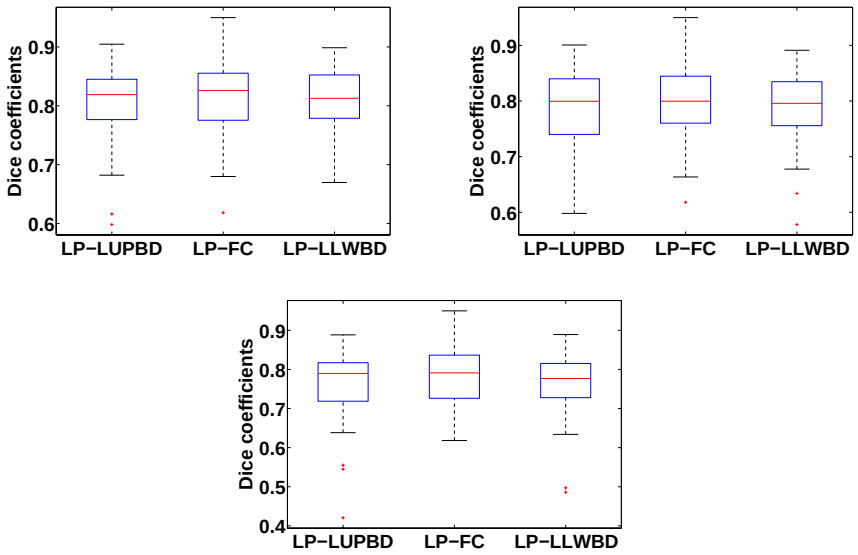


Figure 5.21: Boxplots of dice overlap coefficients for the thirty subjects and linear programming clustering. Comparison between the LP-FC, LP-LLWBD and LP-LUPBD versions.

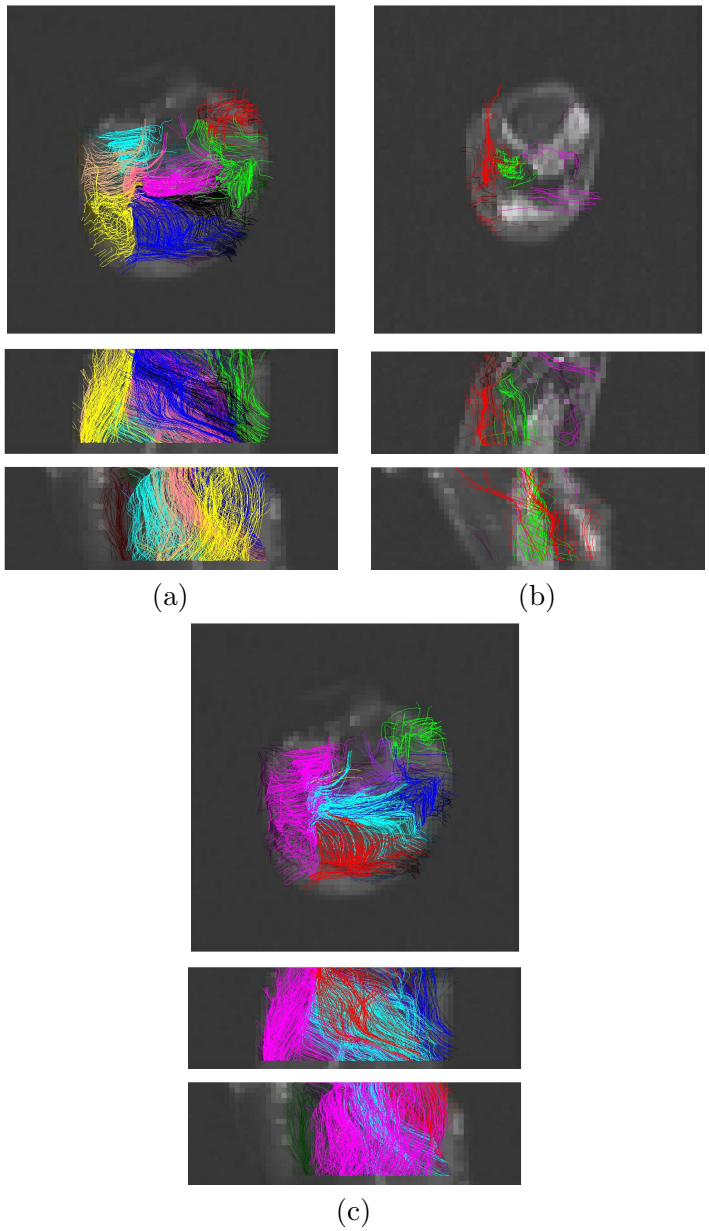


Figure 5.22: Fiber segmentation obtained with linear programming ($\beta = 10$) for (a) a healthy subject in 10 classes (b) a diseased subject in 3 classes.

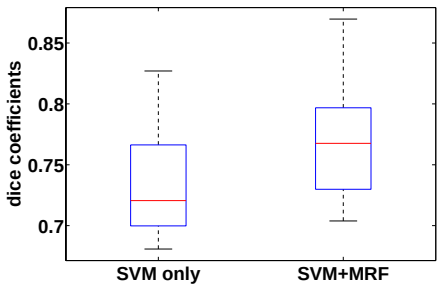
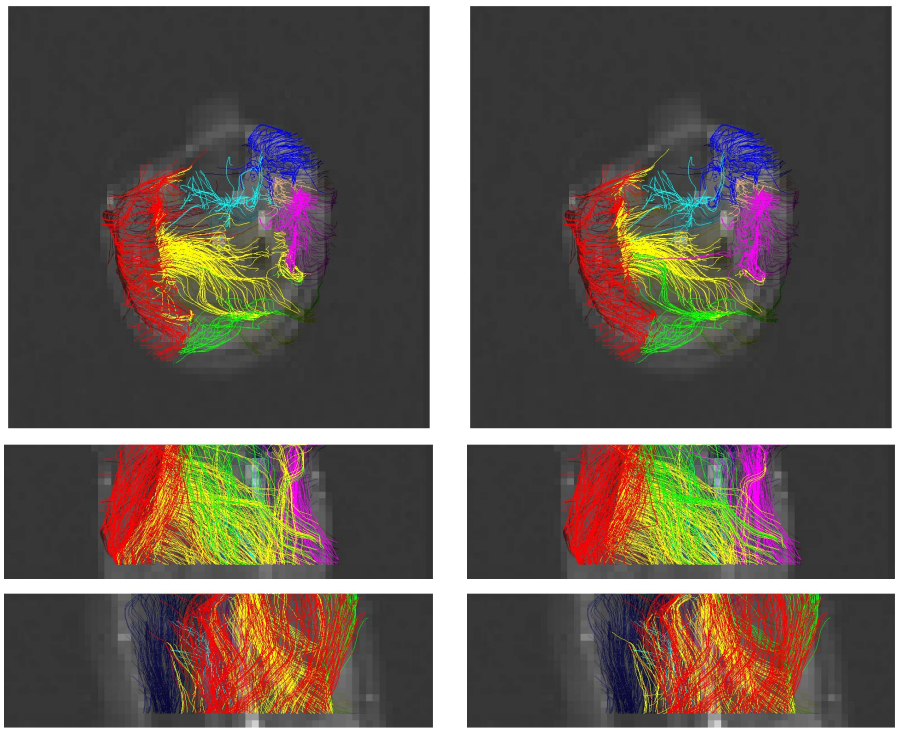


Figure 5.23: Boxplots of dice overlap coefficients for supervised SVM tensor classification when only SVM scores are used and when these are combined to an MRF model.



(a) (b)

Figure 5.24: Axial, coronal and sagittal views for (a) supervised kernel SVM fiber classification in 7 classes (b) the ground truth segmentation.

Chapter 6

DTI Registration Using Kernels and Discrete Optimization

6.1 Introduction

Medical image registration is an important step in many clinical studies. It consists in recovering the transformation that should be applied such that a correspondence is established between two images. Image registration has found many applications in medical imaging as it can permit the analysis of the temporal evolution of an illness, the fusion of information coming from different modalities and the inter-subject comparison to name some of them. This step is of great importance, as a lack of accuracy in spatial normalization can influence dramatically the quality of the drawn conclusions. In the same context, diffusion tensor registration can be helpful in clinical studies where the DTI characteristics can provide valuable information to differentiate between normal and diseased populations.

Diffusion tensor (DT) registration is an inherently intricate problem because of the directional and high-dimensional nature of the data. Indeed, this process not only requires spatial transformation, but also tensor reorientation to account for its rotational component [6]. The high-dimensional nature of the data poses challenges on the definition of the similarity measures that will drive the registration process. On the other hand, the di-

rectional nature of the data highly complicates the warping of DT-images, as no simple interpolation is sufficient to maintain the consistency between tensor orientations before and after the warping. Consequently, appropriate tensor reorientation that depends on the local rotations introduced by the spatial transformation is needed. The existing diffusion tensor registration algorithms can be subdivided in two classes.

The first class of methods transforms the diffusion tensor data in multi-channel feature images and uses vector-data registration algorithms. For instance, in [126], both geometric features describing the distribution characteristics of tensor geometry over an isotropic neighborhood and orientation features based on principal directions distributions are combined for vector registration. Similarly, tensor shape and orientation are both taken into account in [60] using the Geodesic Loxodrome-based distance and a modified version of Multidimensional Scaling. The approach proposed in [125] uses attributes from T1-weighted images to register the corresponding diffusion images along with the Hierarchical Attribute Matching Mechanism for Elastic Registration (HAMMER). The tensor reorientation scheme is based on the alignment on principal directions based on estimation of the fiber direction probability density at each voxel of the source image. The work in [55] investigated the use of multi-channel demons registration using the T2-weighted signal and tensor eigenvalues in a feature vector. Another multi-channel approach based on rotationally invariant Gabor features was presented in [112]. In [55], several scalar values derived from diffusion tensors are studied for feature selection and tensor components provide better experimental performance. The method in [94] proposes a multi-resolution scheme based on tensor-template matching through the extension of similarity measures from the scalar to the tensor case. In [93], it is reported that the use of rotationally-dependent components such as the elements of the diffusion tensor may improve in some cases the registration accuracy with respect to the case where only rotationally-invariant indices are used. However as mentioned in [31], the direct use of tensors can encounter important problems of bad local minima in the optimization procedure. Other rotationally dependent components may be used, for example the diffusion weighted images were matched in [72] using mutual information. More recently, the approach in [75] proposes to characterize a tensor in the image

by running a local fast marching on a neighborhood and use the arrival time for matching across images. In [127], the proposed algorithm includes regional tensor statistics and edge maps to capture relevant information in a multiscale fashion for hierarchical registration. In [138], the attributes used are derived from the fiber tracts instead of the tensors. 3D probability maps of fiber bundles are built and their matching is performed to find an optimal affine transformation between corresponding fiber bundles (correspondences are determined through a clustering-based atlas). The fusion of the different local affine transformations is done through the polyaffine framework, which provides an explicit expression of the Jacobian of the deformation field and paves the way to finite strain reorientation. Another algorithm based on fiber tract registration was proposed in [73] where a rigid transformation is found between a set of curves through the representation of fibers by the means of pointwise curvature and torsion followed by subcurve matching. More recently, in [40], a model of fiber tracts based on currents that defines a fiber metric is proposed to find a diffeomorphism between fiber bundles.

The second class of approaches opts for the use of a metric for tensor matching in a framework previously used for scalar image registration. For instance, the algorithm in [26] uses the symmetrized Kullback-Leibler divergence between probability distributions in the fluid registration framework, followed by a preservation of principal directions (PPD) tensor reorientation. In [5, 4], a multiscale elastic registration algorithm is considered and several possibilities for the similarity term of the registration energy are considered, in particular measures of similarities between tensors and correlation between some scalar indices derived from the tensor images, such as the lattice anisotropy. The methods proposed in [128] and [133] use an explicit reorientation scheme in the registration framework. The former suggests an exact Finite-Strain (FS) differential and includes it in the demons algorithm, the latter proposes a piecewise affine deformation model and optimizes over the available rotational component of each affine transformation. It also provides an affine registration framework based on an L^2 distance between diffusion profiles. An explicit PPD reorientation in the optimization framework is also proposed in [23], where the Large Deformation Diffeomorphic Metric Mapping (LDDMM) algorithm is extended to deal with tensor fields. To recover the displacement field, geodesics on the space of diffeomorphisms

linking two tensor images are recovered and the Frobenius norm is used for matching. LDDMM was also used for registration of fields of principal directions of diffusion considered as unit vectors in [24]. The work in [52] tackles the problem of linear tensor registration through an analogy with optical flow by making the assumption of diffusion tensor constancy and shows that recovering the transformation can be done simply by solving a linear system, and this is adapted to the case where point (and tensor) correspondences are available as constraints that guide the registration process.

We note that in a DT registration context, it is important to combine spatial and diffusion information and not to treat them separately, i.e. to include tensor connectivity where the spatial context is linked to the diffusion information. Unlike the existing literature, we propose a method that combines simultaneously both tensor and spatial information in a probabilistic framework. Using a probabilistic kernel, diffusion tensor images are mapped implicitly to a RKHS, where we define local smoothness properties of the registration deformation to account both for the transformation regularity and for the reorientation of tensors. This mapping defines also a closed-form metric for tensor matching. Furthermore, we extend the framework proposed in [51] to diffusion tensor images by minimizing the defined energy in a discrete setting using the fast Primal-Dual (fast-PD) algorithm [68]. This is done by considering a quantized six-dimensional deformation space where the quaternion representation of rotations allows for proper interpolation and discrete sampling.

The remainder of this chapter is organized as follows: in section 2, we recall the kernel definition and precise useful properties and notations. The concept of the proposed deformable registration framework is presented in section 3. Section 4 is dedicated to experimental results and the perspectives of this work are discussed in section 5.

6.2 Mapping Diffusion Probabilities to a Hilbert Space

Given the orientational nature of DT data, considering tensors independently from their particular spatial arrangement is a major drawback in a registration process. To overcome this issue, we rely on the tools used in

the segmentation framework which were detailed in chapter 5. We consider the normalized L^2 inner-product $k_t(p_1, p_2)$ between two multivariate normal probability densities p_1 and p_2 with parameters $(x_1, \mathbf{D}_1)$ and $(x_2, \mathbf{D}_2)$ respectively:

$$k_t(p_1, p_2) = \frac{\int p_1\left(\mathbf{y}|\mathbf{x_1}, t, \mathbf{D}_1\right) p_2\left(\mathbf{y}|\mathbf{x_2}, t, \mathbf{D}_2\right) d\mathbf{y}}{\sqrt{\int \left(p_1\left(\mathbf{y}|\mathbf{x_1}, t, \mathbf{D}_1\right)\right)^2 d\mathbf{y}} \sqrt{\int \left(p_2\left(\mathbf{y}|\mathbf{x_2}, t, \mathbf{D}_2\right)\right)^2 d\mathbf{y}}} \tag{6.1}$$

Recall that the kernel k_t has a closed form expression and that it verifies the Mercer property over the space of multivariate normal distributions. The Mercer property accounts for the existence of a mapping ϕ associated with the kernel k_t that provides an embedding from the space of Gaussian probabilities in the Reproducing Kernel Hilbert Space (RKHS) $\mathcal{H}$ such that $k_t(p_1, p_2) =< \phi(p_1), \phi(p_2) >_{\mathcal{H}}$, where $< ., . >_{\mathcal{H}}$ is the inner product of $\mathcal{H}$. This allows to define the Hilbertian metric δ_t:

$$\delta_t(p_1, p_2) = \sqrt{2 - 2k_t(p_1, p_2)} \tag{6.2}$$

It is important to note that k_t and δ_t are invariant with respect to translational and rotational transformations. Indeed, when considering a translation vector $\mathbf{t}$ and a rotation matrix $\mathbf{R}$, this property can be verified easily by replacing $\mathbf{x}_1$ (resp. $\mathbf{x}_2$) by $\mathbf{R}\mathbf{x}_1 + \mathbf{t}$ (resp. $\mathbf{R}\mathbf{x}_2 + \mathbf{t}$) and $\mathbf{D}_1$ (resp. $\mathbf{D}_2$) by $\mathbf{R}\mathbf{D}_1\mathbf{R}^t$ (resp. $\mathbf{R}\mathbf{D}_2\mathbf{R}^t$) in the closed form expression of the kernel). In order to ease the notation, in the remainder of the chapter we will identify a Gaussian probability distribution with its parameters and denote:

$$\delta_t(p_1(\mathbf{y}|\mathbf{x}_1, t, \mathbf{D}_1), p_2(\mathbf{y}|\mathbf{x}_2, t, \mathbf{D}_2)) = \delta_t((\mathbf{x}_1, \mathbf{D}_1), (\mathbf{x}_2, \mathbf{D}_2)) \tag{6.3}$$

6.3 Deformable Registration

Let us consider a source DT image $U : \Omega \mapsto S_+(3)$ and a target image V, where Ω is the source image domain and $S_+(3)$ is the space of symmetric positive definite matrices. We aim at computing a deformation field $\mathcal{T} : \Omega \mapsto \mathbb{R}^3 \times \mathrm{SO}(3)$ where $\mathrm{SO}(3)$ is the special orthogonal group. At each point $\mathbf{x} \in \Omega$, $\mathcal{T}(\mathbf{x}) = (\mathbf{t}(\mathbf{x}), \mathbf{R}(\mathbf{x}))$ is a pair composed of a translation vector $\mathbf{t}(\mathbf{x})$ and a rotation matrix $\mathbf{R}(\mathbf{x})$ that deforms U in an image W such that

$W(\mathbf{x}+\mathbf{t}(\mathbf{x})) = \mathbf{R}(\mathbf{x})U(\mathbf{x})\mathbf{R}(\mathbf{x})^t$. Hence W is the deformed (moved) DT image, i.e. the source U is transformed in an image W. The goal is to find a balance between the minimization of the mismatch of W with the target V and the regularity of the applied deformation.

6.3.1 Deformation Model

We consider a grid-based deformation model that can provide for one-to-one and invertible transformations. The basic idea of the deformation model is that by superimposing a grid $G : [1, K] \times [1, L] \times [1, M]$ (where K, L and M are smaller than the dimensions of the domain) onto the moving image, it is possible to deform the embedded image by manipulating the control points belonging to the grid. Consequently, the goal is to recover the deformation vector $\mathcal{T}_\mathbf{p} = (\mathbf{t}_\mathbf{p}, \mathbf{R}_\mathbf{p})$ that should be applied to the node $\mathbf{p}$ of the grid, in order for the images to be aligned. In such a framework, the deformation $\mathcal{T}(\mathbf{x}) = (\mathbf{t}(\mathbf{x}), \mathbf{R}(\mathbf{x}))$ that should be applied to an image position $\mathbf{x}$ can be obtained through interpolation of the deformations obtained at the control points:

$$\mathbf{t}(\mathbf{x}) = \sum_{\mathbf{p}\in G} \eta_s(|\mathbf{x}-\mathbf{p}|)\mathbf{t}_\mathbf{p}, \quad \mathbf{R}(\mathbf{x}) = \sum_{\mathbf{p}\in G} \eta_r(|\mathbf{x}-\mathbf{p}|)\mathbf{R}_\mathbf{p}. \tag{6.4}$$

η_s and η_r are functions that weight the influence of each control point of the grid to each point of the domain in relation to their spatial distance from it. The group of rotations matrices is not a linear space, i.e. a weighted average of rotations taken elementwise does not preserve orthogonality. Therefore the interpolation of rotations is not performed in SO(3) but in the quaternion space, i.e.

$$\mathbf{q}(\mathbf{x}) = \frac{\sum_{\mathbf{p}\in G} \eta_r(|\mathbf{x}-\mathbf{p}|)\mathbf{q}_\mathbf{p}}{\left\|\sum_{\mathbf{p}\in G} \eta_r(|\mathbf{x}-\mathbf{p}|)\mathbf{q}_\mathbf{p}\right\|}, \tag{6.5}$$

where $\mathbf{q}(\mathbf{x})$ (resp. $\mathbf{q}_\mathbf{p}$) is the quaternion representation of $\mathbf{R}(\mathbf{x})$ (resp. $\mathbf{R}_\mathbf{p}$). Note that this is an approximation for the spherical interpolation since the weighted average is normalized, i.e. backprojected to the hypersphere. The matrix representation $\mathbf{R}(\mathbf{x})$ is then obtained easily from $\mathbf{q}(\mathbf{x})$.

6.3.2 DT Image Registration: Continuous Domain

Given the above-defined deformation model, the DT images will be deformed in such a way that an appropriately defined dissimilarity criterion with respect to the distance δ_t implied by the kernel k_t is minimized:

$$E_{data} = \int_{\Omega} \delta_t((\mathbf{x}, W(\mathbf{x})), (\mathbf{x}, V(\mathbf{x})))d\mathbf{x}. \tag{6.6}$$

E_{data} is simply a data term that will drive the deformation towards a minimal mismatch between the deformed image W and the target image V. Note that in E_{data}, only the tensor similarity term in the closed form expression of the kernel is relevant, since we will compare tensors that share the same location. We can rewrite E_{data} using the the control points of the superimposed grid G. Indeed, each voxel $\mathbf{x}$ is back-projected to the points of the grid, in the following form:

$$E_{data} = \frac{1}{|G|} \sum_{\mathbf{p} \in G} \int_{\Omega} \eta_p^{-1}(|\mathbf{x} - \mathbf{p}|)\delta_t((\mathbf{x}, W(\mathbf{x})), (\mathbf{x}, V(\mathbf{x})))d\mathbf{x}. \tag{6.7}$$

The back-projection function η^{-1} computes the influence of the position $\mathbf{x}$ to the control point $\mathbf{p}$. If the nearest neighbor weighting scheme is considered, then each position $\mathbf{x}$ contributes to only one control point $\mathbf{p}$ with a weight equal to one. In the general case, it takes the following form

$$\eta_p^{-1} = \frac{\eta_p(|\mathbf{x} - \mathbf{p}|)}{\int_{\Omega} \eta_p(|\mathbf{x} - \mathbf{p}|)}. \tag{6.8}$$

It should be noted that, as the different indices imply, different weighting schemes can be used for the interpolation of the displacement field (η_s), the interpolation of the rotations (η_r) as well as for the back-projection to the nodes of the grid (η_p).

The minimization of E_{data} is ill-posed as there are fewer constraints than the number of variables to be determined. A common way to tackle such a limitation is to consider a regularization term that will smooth the deformation field, and more importantly take into account local structural information of the source image U. We suppose that the deformation field will be approximately, up to a suitable change in the diffusion time t to account for local scale, *locally isometric* in the RKHS $\mathcal{H}$, i.e. that it preserves the distance δ_t between spatially neighboring Gaussian probabilities when

deforming U and accounts for a possible change of scale. This leads us to define the following smoothness term:

$$E_{smooth} = \int_G \int_{\mathbf{z}\in\mathcal{N}_\mathbf{x}} |\delta_t((\mathbf{x}, U(\mathbf{x})), (\mathbf{z}, U(\mathbf{z}))) - \delta_{t_{\mathbf{xz}}}((\overline{\mathbf{x}}, W(\overline{\mathbf{x}})), (\overline{\mathbf{z}}, W(\overline{\mathbf{z}})))| \, d\mathbf{z}d\mathbf{x} \tag{6.9}$$

where $\overline{\mathbf{x}} = \mathbf{x} + \mathbf{t}(\mathbf{x})$, $\overline{\mathbf{z}} = \mathbf{z} + \mathbf{t}(\mathbf{z})$, $t_{\mathbf{xz}} = t\frac{||\overline{\mathbf{x}}-\overline{\mathbf{z}}||^2}{||\mathbf{x}-\mathbf{z}||^2}$ and $\mathcal{N}_\mathbf{x}$ is a local neighborhood of $\mathbf{x}$ on the grid G. We expect the minimization of E_{smooth} to favor tensor reorientation so that the local source image structure can be preserved in the deformed image. In other words, it is through the regularization term that the rotations and the translations are coupled in such a way that local structure in the deformed image, expressed by the distance between spatially neighboring Gaussian probability distributions, remains consistent with the local structure in the source image. An important underlying property is that the coupling provided by the smoothness term constrains the rotations with respect to displacements, so that tensors do not rotate independently from the translation they undergo with respect to their neighbors. More schematically, it is as if any pair of neighboring grid tensors $(\mathbf{x}, \mathbf{z})$ were linked by a spring of a known rest length $l_0(\mathbf{x}, \mathbf{z})$ provided by the Hilbert space distance in the source image:

$$l_0(\mathbf{x}, \mathbf{z}) = \delta_t((\mathbf{x}, U(\mathbf{x})), (\mathbf{z}, U(\mathbf{z}))) \tag{6.10}$$

Deforming the tensors (by means of both rotation and translation) will affect the length of the spring which is given by

$$l(\mathbf{x}, \mathbf{z}) = \delta_{t_{\mathbf{xz}}}((\overline{\mathbf{x}}, W(\overline{\mathbf{x}})), (\overline{\mathbf{z}}, W(\overline{\mathbf{z}}))) \tag{6.11}$$

It is straightforward to see that the minimization of the smoothness energy will amount to keeping the length of the spring l close enough to its rest length l_0. In order to recover the optimal deformation parameters, we have to minimize the registration energy

$$E = E_{data} + \alpha E_{smooth}, \tag{6.12}$$

where α is a trade-off factor.

A natural way to continue would be to apply a gradient-descent opti-

mization scheme. Such an approach is intuitively simple and easy to implement, however it suffers from some important disadvantages. It is prone to stuck in local minima and it is not modular as it depends both on the used deformation model and the objective criterion to be minimized. Moreover, it is computationally demanding. Ideally, one would prefer a method that would be able to provide a solution in an efficient way and at the same time guarantee that it will be "close" to the optimal one. Methods that comply with the previous characteristics can be found in the discrete optimization field.

Following recent ideas in scalar image registration [51] and recent advances in discrete optimization [68], we opt for the use of a discrete optimization technique called Fast-PD [68]. As seen in chapter 5, Fast-PD can provide an optimal solution (up to a user-defined bound) in an efficient way. Moreover, it allows for a gradient-free optimization thus permitting the use of different deformation methods. In the following, we detail the discretization of the deformation space as well as the Markov Random Field (MRF) formulation of the problem.

6.3.3 DT Image Registration: Discrete Domain

To be able to apply the Fast-PD optimization, it is obligatory to provide a quantized version of the deformation space. Let $\Theta = (\mathbf{d}^1, \ldots, \mathbf{d}^n)$ be a quantized version of the deformation space $\mathbb{R}^3 \times \mathrm{SO}(3)$, then to each quantized deformation $\mathbf{d}^i$, a label l^i can be assigned to it, thus defining a discrete set of labels $L = \{l^1, \ldots, l^n\}$. Then, assigning a label $l_\mathbf{p}$ to the node $\mathbf{p}$, where $l_\mathbf{p} \in L$, corresponds to applying the deformation $\mathbf{d}^{l_\mathbf{p}}$ to the node, that is translating it by $\mathbf{t}^{l_\mathbf{p}}$ and rotating the corresponding tensor by $\mathbf{R}^{l_\mathbf{p}}$. The quantization of spatial displacements is intuitive, the case of rotations is however less straightforward. In order to quantize the group of rotation matrices, we use their quaternion representation. The problem is equivalent to sampling points over the unit sphere $\mathbb{S}^3$ of $\mathbb{R}^4$. We use layered Sukharev Grid sequences [129] that offer a multi-resolution, deterministic and uniform sampling of $\mathbb{S}^3$ by back-projecting points sampled over a hypercube inscribed in $\mathbb{S}^3$ outward onto the spherical surface. The set Θ is therefore formed by the pairs of sampled translations and rotations.

Following [51], we cast the registration problem as a discrete multi-

labeling problem. In such a context, the goal is to recover the optimal individual label $l_{\mathbf{p}}$ that should be assigned to each node $\mathbf{p}$ of the grid. This can be done using the theory of MRFs, the general form of which is the following:

$$E_{\text{MRF}} = \sum_{\mathbf{p}\in G} V_{\mathbf{p}}(l_{\mathbf{p}}) + \alpha \sum_{\mathbf{p}\in G} \sum_{\mathbf{q}\in\mathcal{N}(\mathbf{p})} V_{\mathbf{pq}}(l_{\mathbf{p}}, l_{\mathbf{q}}) \tag{6.13}$$

where $V_{\mathbf{p}}(\cdot)$ are the unary potentials that encode the data term and $V_{\mathbf{pq}}(\cdot,\cdot)$ are the pairwise potentials that encode smoothness constraints. $\mathcal{N}(\mathbf{p})$ represents the neighborhood system of the node $\mathbf{p}$. The unary potentials will be defined according to the data term in (Eq.6.6):

$$V_{\mathbf{p}}(l_{\mathbf{p}}) \approx \int_{\Omega} \eta_p^{-1}(|\mathbf{x}-\mathbf{p}|)\delta_t((\mathbf{x}, W(\mathbf{x})), (\mathbf{x}, V(\mathbf{x})))d\mathbf{x}, \tag{6.14}$$

where W is deformed by applying the label $l_{\mathbf{p}}$ to the node $\mathbf{p}$, that is translating it and rotating the corresponding tensor accordingly. This is actually a local approximation of the similarity measure as the complex interaction of the neighboring control points is not taken into account. However this approximation allows for an efficient computation of the unary potentials while providing satisfactory results. The main idea of the efficient scheme used to compute the unary potentials is that as the labels should be applied to all nodes, the computations are done by applying global translations and rotations and then projecting back the cost to the respective nodes. Similarly the pairwise potentials are derived following (Eq.6.9):

$$V_{\mathbf{pq}}(l_{\mathbf{p}}, l_{\mathbf{q}}) = \left|\delta_t((\mathbf{p}, U(\mathbf{p})), (\mathbf{q}, U(\mathbf{q}))) - \delta_{t_{\mathbf{pq}}}((\overline{\mathbf{p}}, W(\overline{\mathbf{p}})), (\overline{\mathbf{q}}, W(\overline{\mathbf{q}})))\right| \tag{6.15}$$

where $(\overline{\mathbf{p}}, W(\overline{\mathbf{p}}))$ (resp. $(\overline{\mathbf{q}}, W(\overline{\mathbf{q}}))$) are obtained by applying the deformation parameters of the label $l_{\mathbf{p}}$ (resp. $l_{\mathbf{q}}$). To further clarify the idea behind the regularization term, it should be pointed out that this term is similar in spirit to what is usually done when trying to solve the problem of 3D articulated body matching [7], where in order to have consistent deformations the links between the articulations are constrained to deform coherently. Similarly, in this context the problem amounts to deforming articulations at the control points of the grid , where the local coordinate frame of the articulation is given by the tensor.

The main challenge of discrete optimization methods is the quantization of the search space. When quantizing the deformation space a compromise

between the computational complexity and the ability to capture a good minimum is sought. A great number of labels will allow us to be confident that the optimal solution will be approached but will result in high computational times and great memory demands. On the other hand, a small number of labels will keep the computational time small but will, in general, fail to approach the optimal solution with precision. A compromise can be achieved through a compositional approach, where the final solution is obtained through successive optimization problems [111, 51]. In each successive optimization problem finer grids (and consequently shorter diffusion times t) and label sets will be applied. Thus, by keeping the set of the labels to a reasonable size it becomes possible to approximate the optimal solution in an efficient way.

6.4 Experimental Validation

For validation purposes, we considered DT images of the calf muscle of 10 healthy subjects. Recall that the size of the obtained volumes is $64 \times 64 \times 20$ voxels with a voxel resolution of $3.125mm \times 3.125mm \times 7mm$ and that high-resolution T1-weighted images were simultaneously acquired and segmented in 7 muscle groups by an expert. In order to assess quantitatively our method, we consider several evaluation criteria that estimate both spatial normalization accuracy and angular alignment.

In order to measure the spatial mismatch between the deformed and the target images, we deform the ground-truth segmentations and compute the dice overlap, the sensitivity and the specificity of the deformed source segmentation with respect to the target segmentation. Four angular similarity criteria are also evaluated on the target mask: the mean difference in the azimuthal angle θ and the polar angle ϕ in spherical coordinates of the principal directions of diffusion, their average angular separation (AAS) as well as the average overlap of eigenvalue-eigenvector pairs (AOE). We also compute the mean difference in fractional anisotropy (FA). The AOE [12] is defined as follows

$$\frac{1}{N}\sum_{i=1}^{N}\frac{\sum_{j=1}^{3}\lambda_j^i\bar{\lambda}_j^i(e_j^i\cdot\bar{e}_j^i)^2}{\sum_{j=1}^{3}\lambda_j^i\bar{\lambda}_j^i}. \tag{6.16}$$

where λ_j^i, e_j^i and $\bar{\lambda_j^i}$, $\bar{e}_j^i$ denote the $j-$th eigenvalue-eigenvector pair that

is in the $i-$th voxel in each one of the images, and N is the number of the voxels of the region of interest. This measure evaluates the mean alignment of two tensors that are placed at corresponding positions in the pair of images. As far as the AAS [6] is concerned, it is defined as follows

$$\frac{\sum_{i=1}^{N} \sqrt{FA_i \bar{FA}_i} \arccos |e_i^1 \cdot \bar{e}_i^1|}{\sum_{i=1}^{N} \sqrt{FA_i \bar{FA}_i}}. \tag{6.17}$$

By $e_i^1, \bar{e}_i^1$ (resp. $FA_i, \bar{FA}_i$) we denote the principal eigenvector (resp. the FA) of the tensor which is situated at the $i-$th voxel in the pair of the images. This similarity criterion provides information on how well the principal eigenvectors are aligned.

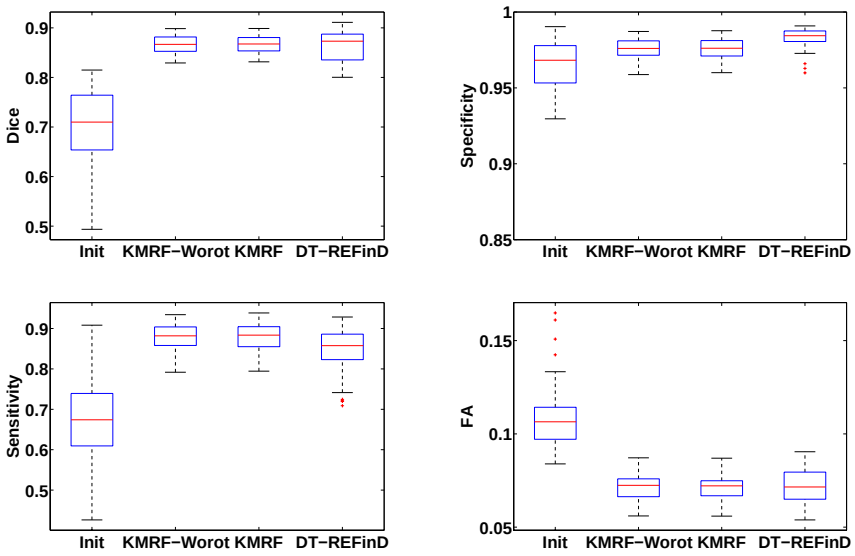


Figure 6.1: Boxplots of the evaluation criteria over the 50 registrations before registration (Init), with our method with a single identity rotation label (KMRF-Worot) and several rotation labels (KMRF), as well as the method in [128] (DT-REFinD).

Among the possible 90 registrations, we chose randomly a subset of 50 pairs of DT images. In all our experiments, we used a three-level multiresolution scheme. The grids used at the three levels were of size $6 \times 6 \times 5$, $12 \times 12 \times 10$ and $18 \times 18 \times 15$. The following diffusion times were used: $t = \{2\ 10^5, 5\ 10^4, 2\ 10^4\}$. The parameter α in (Eq.6.13) was set to $\alpha = 1$.

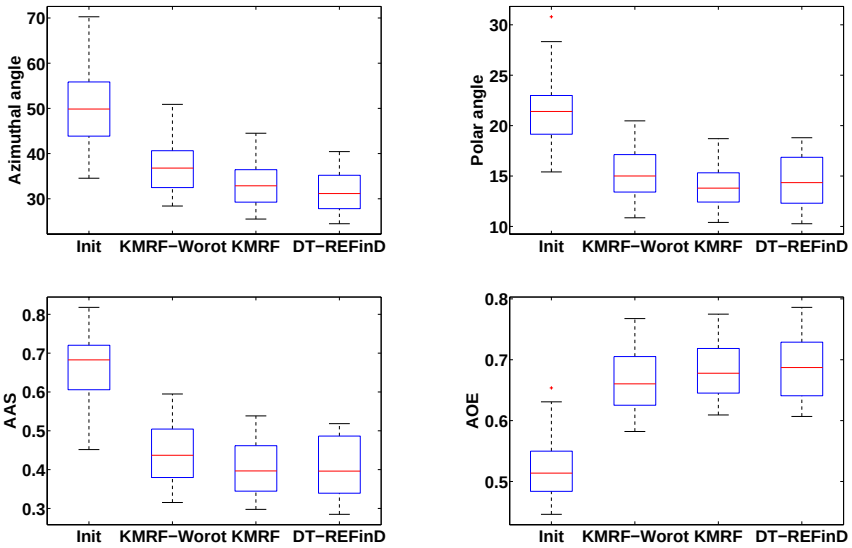


Figure 6.2: Boxplots of the evaluation criteria over the 50 registrations before registration (Init), with our method with a single identity rotation label (KMRF-Worot) and several rotation labels (KMRF), as well as the method in [128] (DT-REFinD).

A number of $73 = 18 \times 4 + 1$ translation labels were used per resolution level, sampled along the horizontal and vertical directions as well as the diagonals. For rotation sampling, we generated 10^3, 10^4 and 10^5 quaternions using Sukharev layered grids. Of these we selected 100, 50 and 25 for the three levels respectively. These samples were chosen as the closest with respect to the geodesic distance $\arccos(.,.)$ on $\mathbb{S}^3$ to the identity matrix (or equivalently with the smallest angle). Towards imposing the diffeomorphic property on the deformation field, we use a cubic B-spline interpolation of the displacement field, with the maximum displacement being restricted to 0.4 times the grid spacing. We used a simple trilinear scheme for tensor interpolation and a nearest-neighbor backprojection (η_p).

For the sake of comparison, we provide the values of the computed evaluation criteria before and after registration. We also compare our method to a reference algorithm recently proposed in [128] (the software is publicly available at http://www-sop.inria.fr/asclepios/software/MedINRIA/) and to the result of our method without a rotational component, i.e. with a sin-

gle rotation label equal to the identity matrix. For the reference algorithm, we considered a three-level multiresolution pyramid with a smoothing kernel of size 1 and a maximum displacement of 4. The main strength of [128] is its use of a true gradient derivation, allowing a more precise reorientation of tensors by differentiating the finite strain reorientation scheme.

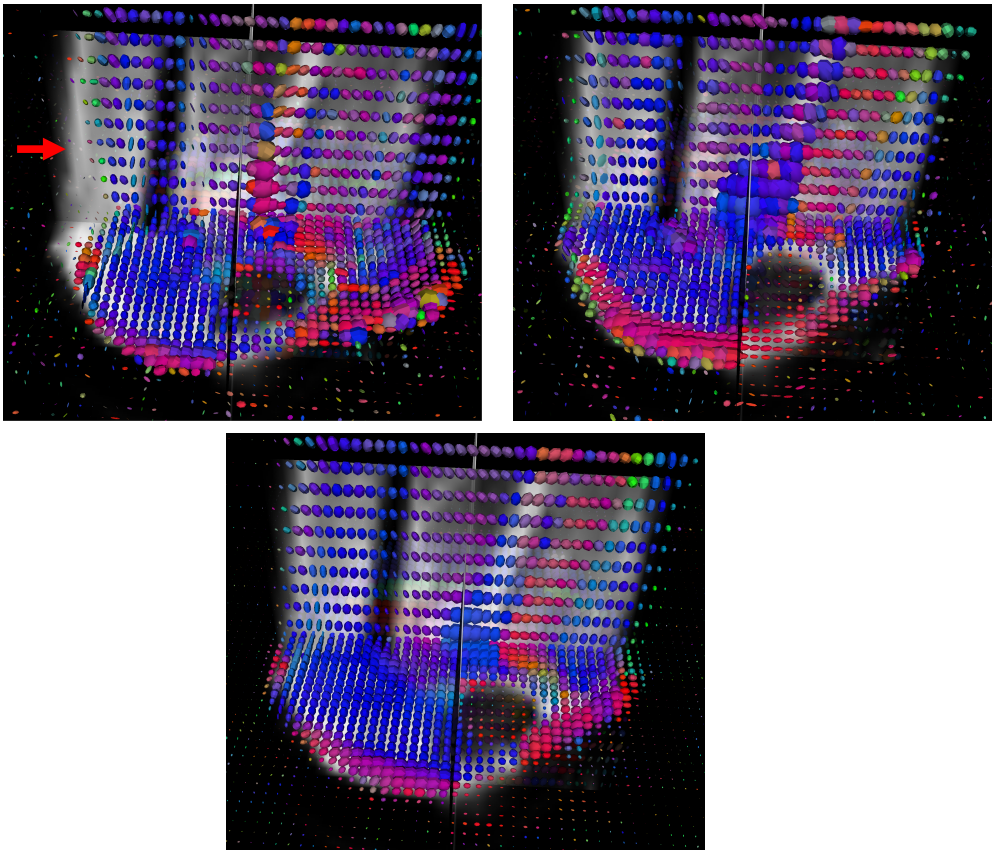


Figure 6.3: From left to right and top to bottom: moving, fixed and deformed tensors. All are overlaid on the B0-image of the target subject. RGB colors encode principal directions of diffusion. Red arrow indicates a region where spatial normalization is clearly suboptimal before registration.

We report in [Fig.6.1, 6.2] the boxplots of the evaluation criteria over the 50 registrations for our method and for the approaches described above. We can see from the boxplots that our approach improves significantly all the evaluation criteria with respect to the initial state (no registration) and that it achieves close results to [128]. We run a paired statistical Student t-test

with a significance level of 0.05 for comparison and we found that the two approaches performed equivalently for the dice and FA, that our method achieved better results for θ, AAS and sensitivity while [128] performed better in ϕ, AOE and specificity. The inclusion of rotation labels improved (in a significant way according to the t-test) all the four angular evaluation criteria with respect to the no-rotation experiments.

For qualitative evaluation, we report in [Fig.6.3] a view of moving tensors, target tensors and deformed tensors, all overlaid on the B0-image of the target subject. We can see that the spatial mismatch is minimized while the tensor field obtained is smooth and the directions of the deformed tensors are similar to the fixed ones. In order to further assess qualitatively the obtained results, we chose randomly a subject as the template and then registered the remaining images to it. The group-average image is computed before and after the registration by a voxel-wise averaging of the corresponding tensors in the individual subjects. The group-average image is then compared to the template by computing the absolute difference of the FA values of the tensors in the respective images. The color code is the following: red corresponds to high values, while blue corresponds to values close to zero. As it can be seen in [Fig.6.4] the great differences that exist before the spatial normalization have been decreased significantly. The algorithm runs in approximately 15 minutes on a standard PC.

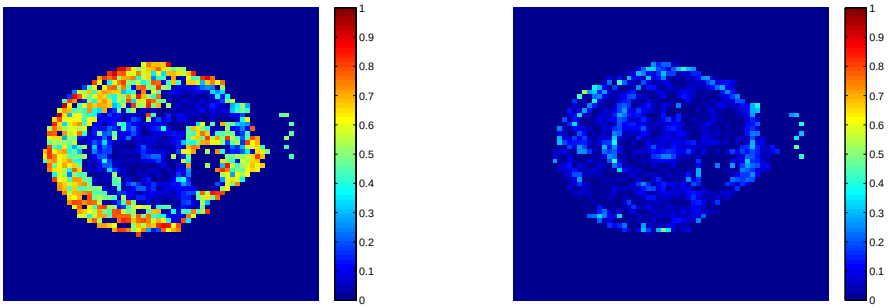


Figure 6.4: Difference between the FA values of the template and the group-average image before (left hand-side) and after (right hand-side) the spatial normalization.

6.5 Discussion

We introduced a novel approach to diffusion tensor registration. The main contribution is two-fold. On one hand we proposed to use a diffusion probability kernel that models both spatial and data dependencies in the tensor field in order to drive the registration process. The proposed formulation allows for the matching of the deformed and the target images while reorienting the tensors and taking into account the local structural information of the source image. On the other hand, we showed that the discrete MRF-based formulation for scalar images proposed in [51] can also be extended to the case of tensor images. A possible improvement of the proposed framework is to consider automatic and location-dependent adaptive quantization of the search space in the context of discrete optimization. This could improve significantly the performance of the method especially when seeking an adequate discretization of such high-dimensional spaces. We can also investigate the use of more evolved interpolation techniques on the Special Euclidean group $\mathrm{SE}(3) = \mathbb{R}^3 \times \mathrm{SO}(3)$ of 3D rigid poses like dual quaternions. These methods could provide better reconstruction of the displacement and rotation fields through the obtained solution at the control points.

Chapter 7

Conclusion

7.1 Contributions

Throughout this thesis, we investigated the use of novel models for processing of diffusion tensor images of the human skeletal muscle. We presented methods for joint estimation and smoothing of diffusion tensor fields, and techniques for segmentation of tensors and fiber tracts, both in supervised and unsupervised fashions. We also proposed a diffusion tensor registration framework based on discrete optimization, Markov Random Field modeling, and probability kernels.

As far as diffusion tensor estimation and denoising is concerned, the presented approach allows to jointly compute and regularize a tensor field based on noisy observations. The main strength of this technique is its ability to incorporate a prior model about local neighborhoods of tensors that overcomes the limitation of the piecewise constant tensor field assumption and the convexity of the defined objective function that allows for a tractable variational optimization. The backbone of the segmentation framework is the use of Gaussian diffusion probabilities that incorporate the spatial information and the definition of kernels on the product space of symmetric positive definite matrices and spatial positions. Hence the approach takes into account the spatial regularity requirement for the segmentation result. More importantly, this paves the way to the definition of similarity measures between fiber tracts that use the tensor information and do not require the computation of point correspondences, while using a single scale parame-

ter (the diffusion time) that stems from the Gaussian diffusion propagators. These similarity measures can be interpreted in an appealing way as affinities between soft probabilistic representations of the fiber trajectories. The Mercer kernels used for unsupervised segmentation were also exploited in a supervised setting, using Support Vector Machine classifiers in order to learn tensors or fiber tracts from a manually labeled template. This is done thanks to the fact that they verify the positive semi-definite (Mercer) conditions. This shows the flexibility and the versatility of such an approach. Various clustering techniques were tested on a significant dataset of the lower leg muscle for a population of healthy and diseased subjects, providing encouraging experimental results. Besides embedding techniques like kernel PCA (that is naturally used due to the Mercer property of the used kernels) and diffusion maps, the use of linear programming optimization combined with pairwise cost approximations that reproduce the behavior of landmark Isomap allows to obtain automatically the number of clusters in a way that is insensitive to the initial conditions. This provides robust and reproducible clustering results.

The registration framework is based on the kernels used in the segmentation process. They allow to define a notion of deformation regularity that couples the rotation applied to tensors with the translations of their spatial locations. The proposed objective function aims at optimizing the match between the moved image and the target image while respecting a constraint on the deformation regularity. The variables of the objective function are in the group of rigid motions which depend on six parameters. This group is quantized to generate likely discrete solutions and the optimization is done efficiently in a Markov Random Field framework.

The algorithmic developments proposed in this thesis offer new representations of the diffusion data, such as the use of local weights for tensor reconstruction based on the piecewise linear assumption and the use of a model of fiber tracts that includes the tensor information. These representations can be used handily for a wide range of applications like denoising, segmentation and registration. The experimental results on the human skeletal muscle show a practical case where these techniques can be useful.

7.2 Future Directions

A possible future direction is the development of an analysis tool that is able to extract relevant statistical information about the diffusion properties like diffusivity and anisotropy indices based on fiber clustering techniques, following the lines of tract-based statistics presented in [88, 77]. This can be useful to detect subtle differences between populations. One can also use the defined kernels in a variational framework for diffusion tensor segmentation. Another important perspective of this work is to fuse the diffusion tensor data with the T1 weighted MRI images that have the advantage of having a higher resolution but where one still observes a lack of contrast between the different muscle groups. Since these images were acquired along with the diffusion images in our experimental protocol, it is interesting to assess the possibility of combining the information provided by both modalities. This is eased by the natural alignment of these modalities when acquiring simultaneously the data, so that inter-modality registration is superfluous. For example, the fiber directions can be used as features to guide muscle segmentation when the gradient of T1 intensity is not reliable enough. Indeed while all muscle groups have the same graylevel intensity in the T1-weighted images and the contours are not clearly defined, the fiber orientations or the other principal directions of diffusion can be more discriminative.

Another interesting direction consists in using diffusion tensor information with an adequate shape prior such as the one proposed in [42, 41] in order to co-segment both T1 and diffusion data. The same applies for registration and atlas building, since co-registration of T1 weighted images and diffusion images is a challenging issue since one has to find a compromise between these two modalities and a way to improve registration with respect to the case where only a single modality is used.

As far as tensor estimation and smoothing is concerned, an important issue that has to be tackled is the regularization of diffusion images with an anomalous signal, for example originating from a tumor or a disease. In the case of diffusion tensor images of the human skeletal muscle, this is relevant in the areas where myopathies result in fat infiltration and alter fiber properties. Regularization of tensor fields should not smooth out the anomalous tensors since they are of crucial importance for diagnosis. Also the nature

itself of the tensor estimates (should they be positive definite or semi-definite positive ?) is not negligible since supposing that very low eigenvalues do not correspond to physical intuition can mislead tensor estimation in the case of the cessation of water flow in some directions due to a disease.

Diffusion tensor registration remains a very challenging issue. One possible improvement to the technique we proposed is to consider geodesics between control points and to consider an approximation of the geodesic distance over the field of Gaussian diffusion probabilities. Then the regularity of a deformation can be estimated using the distortion induced on the computed geodesic distances. However the optimization of the deformation in this context seems to be problematic.

Regarding the experiments, the potential of the fiber clustering techniques discussed in this thesis in segmenting fiber tracts of the brain white matter should be evaluated and compared to the performance of other approaches proposed in the literature, using various fiber metrics like Chamfer and Hausdorff distances. Also the experiments can be run on tractographies obtained from more evolved and robust techniques like those based on stochastic modeling. They can significantly improve the quality of fiber tracts and recover better trajectories, especially in the case of diseased subjects.

Another open issue is the correlation of diffusion data with histological images obtained using biopsy. In particular relationships between the fiber diameter estimated at a microscopic level and tensor parameters should be investigated. Moreover, for diseased subjects, possible effects of fiber diameter shrinking ascertained using histology on diffusion tensor data should be explored. An enticing perspective is to learn the histological patterns as a function of diffusion tensor information and to be able subsequently to reproduce in a non-invasive way the result of a biopsy based on diffusion data.

The ultimate goal is to provide the acquisition improvements and the algorithmic tools that can make diffusion tensor imaging a robust, reliable and clinically useful means for diagnosis and monitoring of myopathies. This requires the implementation and optimization of better acquisition sequences and an in-depth understanding of the effect of myopathies on diffusion. The use of more complex models like High Angular Resolution Diffusion Imaging

(HARDI) is to be envisaged, however this should be a consecutive step to an optimal use of the tensor model for the extraction of attributes that can be helpful for physicians dealing with neuromuscular diseases.

Publications of the Author

- **Book chapter**
  - A. Sotiras, R. Neji, J.-F. Deux, N. Komodakis, M. Maatouk, A. Rahmouni, G. Bassez, G. Fleury and N. Paragios: *Diffusion Tensor Registration Using Probability Kernels and Discrete Optimization*, Biomedical Image Analysis: Methodologies and Applications. Editors: N. Paragios, J. Duncan, and N. Ayache. Springer, 2010. To Appear.
- **International conferences**
  - A. Sotiras, R. Neji, J.-F. Deux, N. Komodakis, G. Fleury and N. Paragios: *A Kernel-based Graphical Model for Diffusion Tensor Registration*, IEEE International Symposium on Biomedical Imaging : from Nano to Macro (ISBI'10). **Oral presentation**.
  - R. Neji, A. Besbes, N. Komodakis, J.-F. Deux, M. Maatouk, A. Rahmouni, G. Bassez, G. Fleury and N. Paragios: *Clustering of the Human Skeletal Muscle Fibers Using Linear Programming and Angular Hilbertian Metrics*, International Conference on Information Processing in Medical Imaging (IPMI'09), Williamsburg (Virginia), USA, 2009. **Oral presentation**.
  - R. Neji, N. Paragios, G. Fleury, J.-P. Thiran and G. Langs: *Classification of Tensors and Fiber Tracts Using Mercer-Kernels Encoding Soft Probabilistic Spatial and Diffusion Information*, IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'09), Miami Beach (Florida), USA, 2009.
  - R. Neji, G. Langs, J.-F. Deux, M. Maatouk, A. Rahmouni, G. Bassez, G. Fleury and N. Paragios: *Unsupervised Classification*

of Skeletal Fibers Using Diffusion Maps, IEEE International Symposium on Biomedical Imaging : from Nano to Macro (ISBI'09), Boston (Massachusetts), USA, 2009. **Oral presentation**.

- E. Dittrich, R. Neji, T. Schmoll, S. Schriefl, C. Ahlers, R. A. Leitgeb and G. Langs: *Detection of Capillary Vessels in Optical Coherence Tomography Based on a Probabilistic Kernel*, Medical Image Understanding and Analysis (MIUA'09), London, UK, 2009. **Oral presentation**.
- R. Neji, G. Fleury, J.-F. Deux, A. Rahmouni, G. Bassez, A. Vignaud and N. Paragios: *Support Vector Driven Markov Random Fields Towards DTI Segmentation of the Human Skeletal Muscle*, IEEE International Symposium on Biomedical Imaging (ISBI'08), Paris, France, 2008.
- R. Neji, N. Azzabou, N. Paragios and G. Fleury: *A Convex Semi-Definite Positive Framework for DTI Estimation and Regularization*, International Symposium on Visual Computing (ISVC'07), Lake Tahoe (Nevada), USA, 2007. **Oral presentation**.

- **Abstract**

  - R. Neji, S. Essafi, G. Langs, G. Fleury, A. Sotiras, J.-F. Deux, A. Vignaud, A. Rahmouni, G. Bassez and N. Paragios: *Perspectives in Myopathy Assessment: Virtual Biopsy and the Potential of Medical Image Analysis using Emerging Modalities*, 6th International Workshop for Muscoskeletal and Neuronal Interactions (ISMNI), Cologne, Germany, 2008. **Oral presentation**.

- **Technical reports**

  - A. Sotiras, R. Neji, J.-F. Deux , N. Komodakis , M. Maatouk , A. Rahmouni , G. Bassez , G. Fleury and N. Paragios: *Diffusion Tensor Registration Using Probability Kernels and Discrete Optimization*, INRIA research report 6943, 2009.
  - R. Neji, J.-F. Deux, A. Besbes, N. Komodakis, G. Langs, M. Maatouk, A. Rahmouni, G. Bassez, G. Fleury and N. Paragios: *Manifold-driven Grouping of Skeletal Muscle Fibers*, INRIA research report 6825, 2009.

– R. Neji, J.-F. Deux, M. Maatouk, G. Fleury, J.-P. Thiran, G. Bassez, A. Rahmouni and N. Paragios: *A Kernel-based Approach to Diffusion Tensor and Fiber Clustering in the Human Skeletal Muscle*, INRIA research report 6686, 2008.

Bibliography

[1] http://training.seer.cancer.gov/anatomy/muscular/structure.html. License: public domain.

[2] www.afm-france.org.

[3] www.sci.utah.edu/~gk/DTI-data/.

[4] D. C. Alexander and J. C. Gee. Elastic matching of diffusion tensor images. *Computer Vision and Image Understanding*, 77(2):233–250, 2000.

[5] D. C. Alexander, J. C. Gee, and R. Bajcsy. Similarity measures for matching diffusion tensor images. In *BMVC*, 1999.

[6] D. C. Alexander, C. Pierpaoli, P. J. Basser, and J. C. Gee. Spatial transformations of diffusion tensor magnetic resonance images. *IEEE TMI*, 20(11):1131–1139, 2001.

[7] D. Anguelov, P. Srinivasan, H.-C. Pang, D. Koller, S. Thrun, and J. Davis. The correlated correspondence algorithm for unsupervised registration of nonrigid surfaces. In *NIPS*, 2005.

[8] S. P. Awate, H. Zhang, and J. C. Gee. A fuzzy, nonparametric segmentation framework for DTI and MRI analysis: With applications to DTI-tract extraction. *IEEE TMI*, 26(11):1525–1536, 2007.

[9] N. Azzabou, N. Paragios, F. Guichard, and F. Cao. Variable bandwidth image denoising using image-based noise models. In *CVPR*, 2007.

[10] P. J. Basser. Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. *NMR in Biomedicine*, 8:333–344, 1995.

[11] P. J. Basser, J. Mattiello, and D. Le Bihan. Estimation of the effective self-diffusion tensor from the nmr spin echo. *Journal of Magnetic Resonance, Series B*, 103(3):247–254, 1994.

[12] P. J. Basser and S. Pajevic. Statistical artifacts in diffusion tensor MRI (DT-MRI) caused by background noise. *Magnetic Resonance in Medicine*, 44:41–50, 2000.

[13] P. J. Basser and C. Pierpaoli. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor mri. *Journal of Magnetic Resonance, Series B*, 111(3):209 – 219, 1996.

[14] S. Basu, P. T. Fletcher, and R. T. Whitaker. Rician noise removal in diffusion tensor mri. In *MICCAI*, 2006.

[15] H. C. Berg. *Random walks in biology.* Princeton University Press, 1983.

[16] D. P. Bertsekas. *Nonlinear Programming.* Athena Scientific, Belmont, MA, 1999.

[17] D. Le Bihan, E. Breton, D. Lallemand, P. Grenier, E. Cabanis, and M. Laval-Jeantet. R imaging of intravoxel incoherent motions: Application to diffusion and perfusion in neurologic disorders. *Radiology*, pages 401–407, 1986.

[18] S. S. Blemker, D. Asakawa, G. Gold, and S. L. Delp. Image-based musculoskeletal modeling: Applications, advances, and future opportunities. *Journal of Magnetic Resonance Imaging*, 25:441451, 2007.

[19] S. S. Blemker, A. J. Sherbondy, D. L. Akers, R. Bammer, S. L. Delp, and G. E. Gold. Characterization of skeletal muscle fascicle arrangements using diffusion tensor tractography. ISMRM, 2005.

[20] A. Brun, H. Knutsson, H. J. Park, M. E. Shenton, and C.-F. Westin. Clustering fiber tracts using normalized cuts. In *MICCAI*, 2004.

[21] A. Brun, H.-J. Park, H. Knutsson, and C.-F. Westin. Coloring of DT-MRI fiber traces using Laplacian eigenmaps. In *EUROCAST*, 2003.

[22] J. F. Budzik, V. Le Thuc, X. Demondion, M. Morel, D. Chechin, and A. Cotten. In vivo MR tractography of thigh muscles using diffusion imaging: initial results. *European Radiology*, 17(12):3079–3085, 2007.

[23] Y. Cao, M. I. Miller, S. Mori, R. L. Winslow, and L. Younes. Diffeomorphic matching of diffusion tensor images. In *CVPRW '06: Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop*, 2006.

[24] Y. Cao, M. I. Miller, R. L. Winslow, and L. Younes. Large deformation diffeomorphic metric mapping of vector fields. *IEEE TMI*, 24(9), 2005.

[25] C. A. Castano-Moraga, C. Lenglet, R. Deriche, and J. Ruiz-Alzola. A Riemannian approach to anisotropic filtering of tensor fields. *Signal Processing [Special Issue on Tensor Signal Processing]*,, 87(2):263–276, 2007.

[26] M.-C. Chiang, A. D. Leow, A. D. Klunder, R. A. Dutton, M. Barysheva, S. E. Rose, K. McMahon, G. I. de Zubicaray, A. W. Toga, and P. M. Thompson. Fluid registration of diffusion tensor images using information theory. *IEEE TMI*, 27(4):442–456, 2008.

[27] R. S. Chow, M. K. Medri, D. C. Martin, R. N. Leekam, A. M. Agur, and N. H. McKee. Sonographic studies of human soleus and gastrocnemius muscle architecture: gender variability. *European Journal of Applied Physiology*, 82(3):236–244, 2000.

[28] R. R. Coifman and S. Lafon. Diffusion maps. *Appl. Comput. Harmon. Anal.*, 21:5–30, 2006.

[29] I. Corouge, G. Gerig, and S. Gouttard. Towards a shape model of white matter fiber bundles using diffusion tensor mri. *ISBI*, 2004.

[30] O. Coulon, D. C. Alexander, and S. Arridge. Diffusion tensor magnetic resonance image regularization. *Medical Image Analysis*, 8(1):47–67, March 2004.

[31] K. M. Curran and D. C. Alexander. Orientation coherence optimization in tensor image registration. In *MIUA*, 2004.

[32] B. M. Damon. Effects of image noise in muscle diffusion tensor (dt)-mri assessed using numerical simulations. *Magnetic Resonance in Medicine*, 60(4):934–44, 2008.

[33] B. M. Damon, Z. Ding, A. W. Anderson, A. S. Freyer, and J. C. Gore. Validation of diffusion tensor mri-based muscle fiber tracking. *Magnetic Resonance in Medicine*, 48:97–104, 2002.

[34] V. de Silva and J. B. Tenenbaum. Global versus local methods in nonlinear dimensionality reduction. In *NIPS*, 2002.

[35] R. Deriche, D. Tschumperle, C. Lenglet, and M. Rousson. Variational approaches to the estimation, regularization and segmentation of diffusion tensor images. In Faugeras Paragios, Chen, editor, *Mathematical Models in Computer Vision: The Handbook*. Springer, 2005 edition, 2005.

[36] J.-F. Deux, P. Malzy, N. Paragios, G. Bassez, A. Luciani, P. Zerbib, F. Roudot-Thoraval, A. Vignaud, H. Kobeiter, and A. Rahmouni. Assessment of calf muscle contraction by diffusion tensor imaging. *European Radiology*, 18(10):2303–2310, 2008.

[37] Z. Ding, J. C. Gore, and A. W. Anderson. Classification and quantification of neuronal fiber pathways using diffusion tensor MRI. *Magnetic Resonance in Medicine*, 49:716–721, 2003.

[38] C. C. Van Donkelaar, L. J. G. Kretzers, P. H. M. Bovendeerd, L. M. A. Lataster, K. Nicolay, J. D. Janssen, and M. R. Drost. Diffusion tensor imaging in biomechanical studies of skeletal muscle function. *Journal of Anatomy*, 194(01):79–88, 1999.

[39] Y. Duan, X. Li, and Y. Xi. Thalamus segmentation from diffusion tensor magnetic resonance imaging. *Journal of Biomedical Imaging*, 2007(2):1–1, 2007.

[40] S. Durrleman, P. Fillard, X. Pennec, A. Trouvé, and N. Ayache. A statistical model of white matter fiber bundles based on currents. In *IPMI*, 2009.

[41] S. Essafi, G. Langs, J.-F. Deux, A. Rahmouni, G. Bassez, and N. Paragios. Wavelet-driven knowledge-based mri calf muscle segmentation. In *ISBI*, 2009.

[42] S. Essafi, G. Langs, and N. Paragios. Hierarchical 3d diffusion wavelet shape priors. In *ICCV*, 2009.

[43] P. Fillard, V. Arsigny, X. Pennec, and N. Ayache. Clinical DT-MRI estimation, smoothing and fiber tracking with log-Euclidean metrics. In *ISBI*, 2006.

[44] P. Fillard, N. Toussaint, and X. Pennec. Medinria: DT-MRI processing and visualization software. Similar Tensor Workshop, 2006.

[45] O. Friman, G. Farneback, and C.-F. Westin. A Bayesian approach for stochastic white matter tractography. *IEEE TMI*, 25(8):965–978, 2006.

[46] M. Froeling, J. Oudeman, S. van den Berg, A. Vilanova, K. Nicolay, M. Maas, G. Strijkers, A. Nederveen, and M. Drost. Muscle fiber tracking and segmentation in the human forearm. ISMRM, 2009.

[47] C. J. Galban, S. Maderwald, F. Stock, and M. E. Ladd. Age-Related Changes in Skeletal Muscle as Detected by Diffusion Tensor Magnetic Resonance Imaging. *Journal of Gerontology*, 62(4):453–458, 2007.

[48] C. J. Galban, S. Maderwald, K. Uffmann, A. de Greiff, and M. E. Ladd. Diffusive sensitivity to muscle architecture: a magnetic resonance diffusion tensor imaging study of the human calf. *European Journal of Applied Physiology*, 93(3):253 – 262, 2004.

[49] C. J. Galban, S. Maderwald, K. Uffmann, and M. E. Ladd. A diffusion tensor imaging analysis of gender differences in water diffusivity within human skeletal muscle. *NMR in Biomedicine*, 2005.

[50] R. J. Gilbert and V. J. Napadow. Three-dimensional muscular architecture of the human tongue determined in vivo with diffusion tensor magnetic resonance imaging. *Dysphagia*, 20:1–7, 2005.

[51] B. Glocker, N. Komodakis, G. Tziritas, N. Navab, and N. Paragios. Dense image registration through MRFs and efficient linear programming. *Medical Image Analysis*, 12(6):731–741, 2008.

[52] A. Goh and R. Vidal. Algebraic methods for direct and feature based registration of diffusion tensor images. In *ECCV*, 2006.

[53] A. Goh and R. Vidal. Clustering and dimensionality reduction on Riemannian manifolds. In *CVPR*, 2008.

[54] A. Goh and R. Vidal. Segmenting fiber bundles in diffusion tensor images. In *ECCV*, 2008.

[55] A. Guimond, C. R. G. Guttmann, S. K. Warfield, and C.-F. Westin. Deformable registration of DT-MRI data based on transformation invariant tensor characteristics. In *ISBI*, 2002.

[56] D. Haussler. Convolution kernels on discrete structures. Technical report, 1999.

[57] A. M. Heemskerk, T. K. Sinha, K. J. Wilson, Z. Ding, and B. M. Damon. Quantitative assessment of DTI-based muscle fiber tracking and optimal tracking parameters. *Magnetic Resonance in Medicine*, 61(2):467–472, 2009.

[58] A. M. Heemskerk, G. J. Strijkers, A. Vilanova, M. R. Drost, and K. Nicolay. Determination of mouse skeletal muscle architecture using three-dimensional diffusion tensor imaging. *Journal of Magnetic Resonance Imaging*, 53:1333–40, 2005.

[59] J.-B. Hiriart-Urruty and C. Lemaréchal. *Fundamentals of Convex Analysis*. Springer Verlag, Heidelberg, 2001.

[60] M. O. Irfanoglu, R. Machiraju, S. Sammet, C. Pierpaoli, and M. V. Knopp. Automatic deformable diffusion tensor registration for fiber population analysis. In *MICCAI*, 2008.

[61] T. Jebara, R. Kondor, and A. Howard. Probability product kernels. *Journal of Machine Learning Research*, 5:819–844, 2004.

[62] T. Joachims. Making large-scale support vector machine learning practical. In A. Smola B. Schölkopf, C. Burges, editor, *Advances in Kernel Methods: Support Vector Machines*. MIT Press, Cambridge, MA, 1998.

[63] L. Jonasson, P. Hagmann, C. Pollo, X. Bresson, C. Richero Wilson, R. Meuli, and J.-P. Thiran. A level set method for segmentation of the thalamus and its nuclei in DT-MRI. *Signal Processing*, 87(2):309–321, 2007.

[64] J. H. Kan, A. M. Heemskerk, Z. Ding, A. Gregory, G. Mencio, K. Spindler, and B. M. Damon. DTI-based muscle fiber tracking of the quadriceps mechanism in lateral patellar dislocation. *Journal of Magnetic Resonance Imaging*, 29(3):663–670, 2009.

[65] P. Khurd, R. Verma, and C. Davatzikos. Kernel-based manifold learning for statistical analysis of diffusion tensor images. In *IPMI*, 2007.

[66] N. Komodakis, N. Paragios, and G. Tziritas. Clustering via LP-based stabilities. In *NIPS*, 2008.

[67] N. Komodakis, G. Tziritas, and N. Paragios. Fast, approximately optimal solutions for single and dynamic MRFs. In *CVPR*, 2007.

[68] N. Komodakis, G. Tziritas, and N. Paragios. Performance vs computational efficiency for optimizing single and dynamic MRFs: Setting the state of the art with primal-dual strategies. *Computer Vision and Image Understanding*, 112(1):14–29, 2008.

[69] G. Langs and N. Paragios. Modeling the structure of multivariate manifolds: Shape maps. In *CVPR*, 2008.

[70] A. D. Lansdown, Z. Ding, M. Wadington, J. L. Hornberger, and B. M. Damon. Quantitative diffusion tensor MRI-based fiber tracking of human skeletal muscle. *Journal of Applied Physiology*, 103(2):673–681, 2007.

[71] M. Lazar, D. M. Weinstein, J. S. Tsuruda, K. M. Hasan, K. Arfanakis, M. E. Meyerand, B. Badie, H. A. Rowley, V. Haughton, A. Field, , and A. L. Alexander. White matter tractography using diffusion tensor deflection. *Human Brain Mapping*, 18:306–321, 2003.

[72] A. Leemans, J. Sijbers, S. De Backer, E. Vandervliet, and P. M. Parizel. Affine coregistration of diffusion tensor magnetic resonance images using mutual information. In *ACIVS*, 2005.

[73] A. Leemans, J. Sijbers, S. De Backer, E. Vandervliet, and P. M. Parizel. Multiscale white matter fiber tract coregistration: a new feature-based approach to align diffusion tensor data. *Magnetic Resonance in Medicine*, 55(6):1414–1423, 2006.

[74] C. Lenglet, M. Rousson, and R. Deriche. DTI segmentation by statistical surface evolution. *IEEE TMI*, 25(06):685–700, 2006.

[75] H. Li, Z. Xue, L. Guo, and S. T. C. Wong. Simultaneous consideration of spatial deformation and tensor orientation in diffusion tensor image registration using local fast marching patterns. In *IPMI*, 2009.

[76] M. Maddah, W. E. L. Grimson, S. Warfield, and W. Wells. A unified framework for clustering and quantitative analysis of white matter fiber tracts. *Medical Image Analysis*, 12(2):191–202, 2008.

[77] M. Maddah, M. Kubicki, W. Wells III, C.-F. Westin, M. Shenton, and W. E. L. Grimson. Findings in schizophrenia by tract-oriented dt-mri analysis. In *MICCAI*, 2008.

[78] M. Maddah, A. U. J. Mewes, S. Haker, W. E. L. Grimson, and S. K. Warfield. Automated atlas-based clustering of white matter fiber tracts from DTMRI. In *MICCAI*, 2005.

[79] M. Martín-Fernández, C.-F. Westin, and C. Alberola-López. 3D Bayesian regularization of diffusion tensor MRI using multivariate Gaussian Markov random fields. In *MICCAI*, 2004.

[80] D. W. McRobbie, E. A. Moore, M. J. Graves, and M. R. Prince. *MRI from Picture to Proton*. Cambridge University Press, 2007.

[81] J. Melonakos, V. Mohan, M. Niethammer, K. Smith, M. Kubicki, and A. Tannenbaum. Finsler tractography for white matter connectivity analysis of the cingulum bundle. In *MICCAI*, 2007.

[82] N. Milne. Human functional anatomy 213. http://www.lab.anhb.uwa.edu.au/hfa213/week3/lec3unimuscle.pdf.

[83] S. Mori and P. C. M. van Zijl. Fiber tracking: principles and strategies - a technical review. *NMR in Biomedicine*, 15:468–480, 2002.

[84] D. Morvan. In vivo measurement of diffusion and pseudo-diffusion in skeletal muscle at rest and after exercise. *Magnetic Resonance Imaging*, 13(2):193 – 199, 1995.

[85] V. J. Napadow, Q. Chen, V. Mai, P. T.C. So, and R. J. Gilbert. Quantitative analysis of three-dimensional-resolved fiber architecture in heterogeneous skeletal muscle tissue using nmr and optical imaging methods. *Biophysical Journal*, 80:2968–2975, 2001.

[86] L. J. ODonnell and C.-F. Westin. White matter tract clustering and correspondence in populations. In *MICCAI*, 2005.

[87] L. J. ODonnell and C.-F. Westin. Automatic tractography segmentation using a high-dimensional white matter atlas. *IEEE TMI*, 26(11):1562–1575, 2007.

[88] L. J. ODonnell, C.-F. Westin, and A. J. Golby. Tract-based morphometry. In *MICCAI*, 2007.

[89] Y. Okamoto, A. Kunimatsu, S. Miki, M. Shindo, M. Niitsu, and M. Minami. Fractional anisotropy values of calf muscles in normative state after exercise: Preliminary results. *Magnetic Resonance in Medical Sciences*, 7(3):157–162, 2008.

[90] J. H. Park, J. Qi, R. Price, and N. Olsen. Diffusion-weighted imaging (dwi) for evaluation of muscle diseases: Dermatomyositis and polymyositis. ISMRM, 2003.

[91] X. Pennec, P. Fillard, and N. Ayache. A Riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41–66, 2006.

[92] J. Qi, N. J. Olsen, R. R. Price, J. A. Winston, and J. H. Park. Diffusion-weighted imaging of inflammatory myopathies: Polymyositis and dermatomyositis. *Journal of Magnetic Resonance Imaging*, 27(1):212–217, 2008.

[93] G. K. Rohde, S. Pajevic, and C. Pierpaoli. Multi-channel registration of diffusion tensor images using directional information. In *ISBI*, 2004.

[94] J. Ruiz-Alzola, C.-F. Westin, S. K. Warfield, C. Alberola, S. E. Maier, and R. Kikinis. Nonrigid registration of 3d tensor medical data. *Medical Image Analysis*, 6(2):143–161, 2002.

[95] R. Salvador, A. Pea, D. K. Menon, T. A. Carpenter, J. D. Pickard, and E. T. Bullmore. Formal characterization and extension of the linearized diffusion tensor model. *Human Brain Mapping*, 24(2):144–155, 2005.

[96] T. Saotome, M. Sekino, F. Eto, and S. Ueno. Comparing human skeletal muscle architectural parameters of cadavers with in vivo ultrasonographic measurements. *Journal of Anatomy*, 199(4):429 – 434, 2001.

[97] T. Saotome, M. Sekino, F. Eto, and S. Ueno. Evaluation of diffusional anisotropy and microscopic structure in skeletal muscles using magnetic resonance. *Magnetic Resonance Imaging*, 24(1):19 – 25, 2006.

[98] N. Saupe, L. M. White, J. Stainsby, G. Tomlinson, and M. S. Sussman. Diffusion Tensor Imaging and Fiber Tractography of Skeletal Muscle: Optimization of b Value for Imaging at 1.5 T. *American Journal of Roentgenology*, 192(6):W282–290, 2009.

[99] N. Saupe, L. M. White, M. S. Sussman, A. Kassner, G. Tomlinson, and M. D. Noseworthy. Diffusion tensor magnetic resonance imaging of the human calf: comparison between 1.5 T and 3.0 T-preliminary results. *Investigative radiology*, 43(9):612–8, 2008.

[100] P. Savadjiev, J. S. W. Campbell, G. B. Pike, and K. Siddiqi. Streamline flows for white matter fibre pathway segmentation in diffusion MRI. In *MICCAI*, 2008.

[101] B. Scholkopf, A. Smola, and K.-R. Muller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comp.*, 10(5):1299–1319, 1998.

[102] B. Sellers. Muscle mechanics. http://homepage.mac.com/wis/Personal/lectures/musculoskeletal/MuscleMechanics.pdf.

[103] U. Sinha and L. Yao. In vivo diffusion tensor imaging of human calf muscle. *Journal of Magnetic Resonance Imaging*, 15(1):87–95, 2002.

[104] G. Steidle and F. Schick. Echoplanar diffusion tensor imaging of the lower leg musculature using eddy current nulled stimulated echo preparation. *Magnetic Resonance in Medicine*, 55(3):541–548, 2006.

[105] E. O. Stejskal and J. E. Tanner. Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. *The Journal of Chemical Physics*, 42(1):288–292, 1965.

[106] J. B. Tenenbaum, V. de Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.

[107] A. Tsai, C.-F. Westin, A. O. Hero, and A. S. Willsky. Fiber tract clustering on manifolds with dual rooted-graphs. In *CVPR*, 2007.

[108] Y. Tsin. *Kernel Correlation as an Affinity Measure in Point-Sampled Vision Problems*. PhD thesis, Robotics Institute, Carnegie Mellon University, 2003.

[109] A. van Doorn, P. H. Bovendeerd, K. Nicolay, M. R. Drost, and J. D. Janssen. Determination of muscle fibre orientation using diffusion-weighted mri. *European Journal of Morphology*, 34(1):5–10, 1996.

[110] V. Vapnik. *Statistical Learning Theory*. Wiley, 1998.

[111] O. Veksler. *Efficient graph-based energy minimization methods in computer vision*. PhD thesis, Cornell University, 1999.

[112] R. Verma and C. Davatzikos. Matching of diffusion tensor images using gabor features. In *ISBI*, 2004.

[113] R. Verma, P. Khurd, and C. Davatzikos. On analyzing diffusion tensor images by identifying manifold structure using isomaps. *IEEE TMI*, 26(6):772–778, 2007.

[114] P. Vermathen, C. Boesch, S. E. Maier, and R. Kreis. Comparison of fiber orientation in human muscle by short te mrsi and diffusion weighted imaging. ISMRM, 2002.

[115] P. Vermathen, R. Kreis, and C. Boesch. Mapping fiber orientation in muscle by short te mrsi. regional variations of 1h spectra of human muscle due to orientation dependent dipolar splittings. ISMRM, 2001.

[116] X. Wang, W. E. L. Grimson, and C.-F. Westin. Tractography segmentation using a hierarchical dirichlet processes mixture model. In *IPMI*, 2009.

[117] Z. Wang and B. C. Vemuri. DTI segmentation using an information theoretic tensor dissimilarity measure. *IEEE TMI*, 24(10):1267–1277, 205.

[118] Z. Wang, B. C. Vemuri, Y. Chen, and T. H. Mareci. A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex DWI. *IEEE TMI*, 23(8):930–939, 2004.

[119] D. Wassermann and R. Deriche. Simultaneous manifold learning and clustering: Grouping white matter fiber tracts using a volumetric white matter atlas. In *MICCAI 2008 Workshop - Manifolds in Medical Imaging: Metrics, Learning and Beyond.*, 2008.

[120] D. Wassermann, M. Descoteaux, and R. Deriche. Diffusion maps clustering for magnetic resonance q-ball imaging segmentation. *International Journal on Biomedical Imaging*, 2008.

[121] J. Weickert, C. Feddern, M. Welk, B. Burgeth, and T. Brox. PDEs for tensor image processing. In J. Weickert and H. Hagen, editors, *Visualization and Processing of Tensor Fields*, pages 399–414. Springer, January 2006.

[122] Y. T. Weldeselassie and G. Hamarneh. DT-MRI segmentation using graph cuts. SPIE Medical Imaging, 2007.

[123] C.-F. Westin, S. E. Maier, H. Mamata, A. Nabavi, F. A. Jolesz, and R. Kikinis. Processing and visualization of diffusion tensor MRI. *Medical Image Analysis*, 6(2):93–108, 2002.

[124] Y. Xia, U. Turken, S. L. Whitfield-Gabrieli, and J. D. Gabrieli. Knowledge-based classification of neuronal fibers in entire brain. In *MICCAI*, 2005.

[125] D. Xu, S. Mori, D. Shen, P. C. van Zijl, and C. Davatzikos. Spatial normalization of diffusion tensor fields. *Magnetic Resonance in Medicine*, 50(1):175–182, July 2003.

[126] J. Yang, D. Shen, C. Davatzikos, and R. Verma. Diffusion tensor image registration using tensor geometry and orientation features. In *MICCAI*, 2008.

[127] P.-T. Yap, G. Wu, H. Zhu, W. Lin, and D. Shen. Timer: Tensor image morphing for elastic registration. *Neuroimage*, 47(2):549–63, 2009.

[128] B. Yeo, T. Vercauteren, P. Fillard, X. Pennec, P. Golland, N. Ayache, and O. Clatz. DTI registration with exact finite-strain differential. In *ISBI*, 2008.

[129] A. Yershova and S. M. LaValle. Deterministic sampling methods for spheres and SO(3). In *ICRA*, 2004.

[130] T. Zaraiskaya, D. Kumbhare, and M. D. Noseworthy. Diffusion tensor imaging in evaluation of human skeletal muscle injury. *Journal of Magnetic Resonance Imaging*, 24(2):402–408, 2006.

[131] F. Zhang, E. Hancock, C. Goodlett, and G. Gerig. Probabilistic white matter fiber tracking using particle filtering and von mises-fisher sampling. *Medical Image Analysis*, 13(1):5 – 18, 2009.

[132] F. Zhang and E. R. Hancock. Tensor mri regularization via graph diffusion. In *BMVC*, 2006.

[133] H. Zhang, P. A. Yushkevich, D. C. Alexander, and J. C. Gee. Deformable registration of diffusion tensor MR images with explicit orientation optimization. *Medical Image Analysis*, 10(5):764 – 785, 2006.

[134] J. Zhang, G. Zhang, B. Morrison, S. Mori, and K. A. Sheikh. Magnetic resonance imaging of mouse skeletal muscle to measure denervation atrophy. *Experimental Neurology*, 212(2):448 – 457, 2008.

[135] S. Zhang and D. H. Laidlaw. DTI fiber clustering in the whole brain. In *IEEE Visualization*, page 28, 2004.

[136] S. Zhang and D. H. Laidlaw. DTI fiber clustering and cross-subject cluster analysis. In *ISMRM*, 2005.

[137] U. Ziyan, M. R. Sabuncu, W. E. L. Grimson, and C.-F. Westin. Consistency clustering: A robust algorithm for group-wise registration, segmentation and automatic atlas construction indiffusion mri. *International Journal of Computer Vision*, 85:279–290, 2009.

[138] U. Ziyan, M. R. Sabuncu, L. J. ODonnell, and C.-F. Westin. Nonlinear registration of diffusion mr images based on fiber bundles. In *MICCAI*, 2007.

[139] U. Ziyan, D. Tuch, and C.-F. Westin. Segmentation of thalamic nuclei from DTI using spectral clustering. In *MICCAI*, 2006.

[140] O. Zvitia, A. Mayer, and H. Greenspan. Adaptive mean-shift registration of white matter tractographies. In *ISBI*, 2008.