N° d'ordre: 4298

THÈSE

présentée

devant l'Université de Rennes 1

pour obtenir

le grade de : Docteur de l'Université de Rennes 1
Mention Traitement du Signal et Telecommunications

par

Meenakshi Mani

Équipe d'accueil : Visages - IRISA
École Doctorale : Matisse
Composante universitaire : SPM

Titre de la thèse :

Quantitative Analysis of Open Curves in Brain Imaging: Applications to White Matter Fibers and Sulci

À soutenir le 31 janvier 2011 devant la commission d'examen

M. :	Patrick	Perez	Président
MM. :	Daniel	Rueckert	Rapporteurs
	Bertrand	Thirion	
MM. :	Anuj	Srivastava	Examinateurs
	Christian	Barillot	

To Saraswati
the gossamer touch of inspiration manifest

Acknowledgements

I would first like to thank the members of my thesis committee: Patrick Perez played the role of president, Daniel Rueckert and Bertrand Thirion were the examiners and Christian Barillot and Anuj Srivastava constituted the remainder of the jury. I thank them all for the time they took from their busy schedules and for their attention and interest during the defense. I am also grateful to Bertrand Thirion for his detailed notes and suggestions to improve the text.

The successful completion of this thesis was tied in part to my association with Christian Barillot and Anuj Srivastava. I would like to thank Christian, my thesis director, for inviting me to do a thesis in France and for giving me the opportunity to work independently. It has to be mentioned that Christian has been very fair to me. This is not my usual experience or even expectation. I had also visited Anuj Srivastava's lab during the course of my research. The fruitful conversations we had during that period have helped me develop the thematic outline for this thesis.

I would be remiss at this point if I did not also acknowledge the resources I used on the internet. Indeed it is fair to say that this thesis would not have been possible without this online interaction.

Finally, I gratefully acknowledge the help I received for the following:
Neuroanatomy: I benefitted a great deal from Xavier Morandi's neurosurgical expertise. He spent many hours helping me identify sulci.
Software: In addition to the Visages in-house software for segmenting sulci and for DTI tractography I used Matlab and R code developed by Michael Trosset, Anuj Srivastava and his student Sebastian Kurtek. A very special thanks to Alexandre Abadie for help with the Visages tools which helped me get started during my first year.
Computer related support: Phillipe Aoustin for hardware issues and Thierry Fauconnier for setting up the videconference during my thesis defense.
French translations: Olivier Luong, Camille Maumet, Franck Michel, Aymeric Stamm and Pierre Maurel helped translate the French sections in this thesis. Olivier Luong and Benoit Combès also helped me with all manner of administrative issues and I am very very grateful.
Financial support: The INRIA CORDIS grant provided full support. The UEB grant gave additional support during my visit to Anuj Srivastava's lab.
Administrative assistance: Céline, Odile, Angelique, Aline, Catherine Tannoux and other INRIA staff who handled paperwork and logistics.

Rennes, February 2011.

Contents

Chapter 1

Overview

Il y a dans le cerveau humain environ 100 sillons corticaux, et plus de 100 milliards de faisceaux de matière blanche. Si le nombre, la configuration et la fonction de ces deux structures anatomiques diffèrent, elles possèdent toutefois une propriété géométrique commune: ce sont des courbes ouvertes continues. Cette thèse propose d'étudier comment les caractéristiques des courbes ouvertes peuvent être exploitées afin d'analyser quantitativement les sillons corticaux et les faisceaux de matière blanche.

Les quatre caractéristiques d'une courbe ouverte–forme, taille, orientation et position–ont des propriétés différentes, si bien que l'approche usuelle est de traiter chacune séparément á l'aide d'une métrique ad hoc. Nous introduisons un cadre riemannien adapté dans lequel il est possible de fusionner les espaces de caractéristiques afin d'analyser conjointement plusieurs caractéristiques. Cette approche permet d'apparier et de comparer des courbes suivant des distances géodésiques. Les correspondances entre courbes sont établies automatiquement en utilisant une métrique élastique.

Dans cette thèse, nous validerons les métriques introduites et nous montrerons leurs applications pratiques, entre autres dans le cadre de plusieurs problèmes cliniques importants.

Nous commençons par considérer une courbe ouverte. Dans cette fenêtre, la vue pose des difficultés particulières, fait apparaître des problèmes et des possibilités.

Le choix de la représentation restreint notre analyse aux sillons et aux fibres de la matière blanche, les deux structures anatomiques modélisables par des courbes ouvertes en 3D. La figure 2.1 illustre la relation entre l'anatomie et un ensemble des courbes. Une courbe ouverte est une des moyens les plus simples de représenter mathématiquement un sillon ou une fibre de matière blanche. On pourrait également utiliser une surface (ouverte), mais celle-ci est plus difficile à modéliser et implanter.

Une courbe ouverte peut être représentée dans un espace euclidien ou riemannien. Nous utilisons les deux dans notre analyse. Dans un espace euclidien, un spline (une courbe paramétrisée définie par morceaux par des polynômes) peut être une bonne approximation de formes de courbes complexes. La distance $\mathbb{L}^2$ que nous calculons entre deux splines peut être utilisée pour le regroupement ou tout autre classification. L'avantage principal de cette approche est sa facilité d'implantation, et le fait que des

méthodes existantes provenant de l'apprentissage automatique et des statistiques peuvent être intégrées telles quelles dans l'analyse. Dans l'analyse riemannienne, la courbe ouverte est représentée par un point dans l'espace des fonctions continues. Cet espace de fonctions est de dimension infinie et non linéaire, et des méthodes avancées comme celles provenant de l'analyse fonctionnelle, la géométrie différentielle ou la théorie des groupes est nécessaire pour représenter et analyser les courbes. L'avantage de la représentation riemannienne est que nous pouvons nous appuyer la géométrie intrinsèque (riemannienne) de l'espace pour calculer la distance entre deux points (courbes). Les courbes ouvertes sont essentiellement des structures non linéaires. L'espace des courbes ouvertes est également non linéaire. Dans de tels espaces, les mesures linéaires ne sont pas bien définies et les résultats pourraient être incohérents.

1.1 Organisation de cette thèse

Cette thèse est composée de deux parties : la première est consacrée à l'analyse des sillons; la seconde, à l'analyse d'une fibre de matière blanche. Ces deux structures anatomiques sont traitées séparément car elles posent des problèmes particuliers : les sillons ont une grande variabilité , et la tractographie des fibres de matière blanche produit de grandes quantités de données. Chaque structure nécessite une approche adaptée.

De plus des applications associées à ces deux structures anatomiques font apparaître des résultats différents. Un ensemble classique d'outils mathématiques sont utilisés à la fois dans les parties I et II. Le chapitre 3 est un aperçu d'un framework riemannien complet pour l'analyse de courbes ouvertes.

1.1.1 Partie I

Les discussions dans la partie I se rapportent à la tâche difficile de l' etiquetage des sillons. Le chapitre 4 présente de l'étiquetage en reliant chacun de ces deux mots par des caractéristiques innée aux sillons. L'étiquetage ou l'identification est l'application la plus importante liée aux sillons. Le chapitre 5 propose une solution partielle à ce problème difficile.

1.1.2 Partie II

Dans notre travail, nous considérons des métriques basées sur la forme, l'échelle, l'orientation et la position: des caractéristiques physiques associées à une fibre. Ces caractéristiques ont des propriétés différentes, de sorte que la méthode habituelle a été de concevoir différentes métriques et des espaces pour les traiter individuellement. Dans le chapitre 3, nous avons décrit un cadre Riemannien global où les espaces de caractéristiques jointes permettent une analyse de combinaisons de caractéristiques. Nous pouvons comparer les courbes et les recaler en utilisant les formules pour les métriques et géodésiques données. Les correspondances entre les courbes sont automatiquement estimées grâce à l'utilisation d'une métrique élastique.

Nous commençons par une discussion sur le regroupement (clustering), une étape importante et nécessaire dans l'analyse de la matière blanche. Le clustering nous aide à organiser le grand volume de fibres obtenu par IRM de diffusion (DTI) et dans la première phase de notre travail, nous avons étudié les méthodes et les métriques existantes afin de développer des distances pour nos propres études. Le chapitre 7 décrit ces travaux.

Dans les trois chapitres suivants, nous montrons comment les outils et les métriques fournis par le cadre riemannien peuvent être efficaces dans le contexte de l'analyse de la matière blanche. Dans le chapitre 8, nous utilisons les configurations spécifiques des fibres dans le corps calleux pour montrer comment les résultats de clustering varient en fonction des différentes métriques. Cela plaide pour le choix judicieux de métriques en fonction des applications.

Ce cadre fournit également des outils pour le calcul de statistiques sur des courbes. Ces statistiques sont un premier pas vers une analyse statistique rendant possible des tâches d'inférence telles que l'estimation de paramètres et les tests d'hypothèses pour les comparaisons de groupes. Pour l'analyse des fibres de la matière blanche, en raison du nombre important de fibres concernées, il est également pratique de commencer par représenter un ensemble de fibres par une moyenne et une variance qui décrivent les caractéristiques essentielles du faisceau de fibres. Le chapitre 9 traite de ce sujet plus en détail.

L'étude de maladies de la matière blanche est l'un des principaux objectifs de l'analyse de la matière blanche. Dans le chapitre 10, nous montrons comment les distances caractéristiques décrites dans le chapitre 3 peuvent être utilisées pour détecter des changements morphologiques dûs à une maladie dégénérative. Puisque c'est la première fois que de telles mesures de distance sont utilisées pour traiter ce problème, notre approche peut être considérée comme nouvelle.

1.1.3 L'étiquetage des sillons utilisant une échelle multidimensionnelle

Au cours des 15 dernières années, deux idées ont émergé pour traiter le probléme de la variabilité des sillons pour leur étiquetage et leur identification. La première est l'utilisation d'un large ensemble de descripteurs. Nous fondons cette étude sur la seconde idée, celle de filtrage relationnel que nous mettons en œuvre dans un cadre d'apprentissage supervisé. Nous utilisons une mise à l'échelle multi-dimensionnelle (MDS) afin de recriéer les relations structurelles dans notre sous-espace. Les sillons non étiquetés sont projetés dans ce même sous-espace, et on leur attribue ensuite une classe par la méthode des plus proches voisins (NN). Nous adoptons une approche heuristique pour le choix de la mesure de distance, qui est au centre de toutes les méthodes à base de graphe. Dans un travail préliminaire, nous avons constaté qu'une distance basée uniquement sur la forme ne permettait pas de différencier les sillons de manière adéquate. Les distances basées sur des caractéristiques spatiales ont toutefois donné des résultats encourageants. Nous avons donc évalué trois distances spatiales conçues pour permettre de comparer différentes configurations de sillons. Dans des expériences

utilisant une stratégie leave-one-out, 90% des 180 sillons provenant de 10 classes de sillons sont classés avec succès comme le sont ceux de 6 patients atteints de tumeurs corticales.

Avec ce système d'apprentissage supervisé, nous offrons une solution simple et intuitive à un problème difficile. L'utilisation de l'appariement relationnel nous donne la souplesse nécessaire pour tenir compte des variations normales dans les sillons. Nous pouvons aussi, dans une certaine mesure, classer des sillons issus de données pathologiques ou partiellement segmentés, qui sont différents de l'ensemble d'apprentissage.

1.2 Contributions de cette thèse

Nous avons etudié l'analyse des courbes ouvertes dans deux contextes très différents , les sillons et les fibres de matière blanche. Les sillons sont peu nombreux et variables. Les fibres, d'autre part, sont nombreuses, mais homogènes au sein d'un faisceau de fibres.

Le problème clé dans l'analyse des sillons est l'étiquetage. Les experts en neuroanatomie eux mémes ont parfois des difficultés pour réaliser cette tâche. Leur niveau de performance constituera donc la limite supérieur à atteindre pour les methodes d'étiquetage automatique.

Nous abordons le problème de l'étiquetage automatique en limitant tout d'abord la portée du problème aux sillons primaires. Ce probléme est alors bien défini et nous lui apportons une solution simple et facile à implémenter. Nous proposons ensuite de formuler des solutions en considérant ce problème comme un problème d'appariement de graphes.

Le coeur du problème dans l'analyze des fibres de matière blanche (obtenues en DTI ou HARDI) est de gérer de façon efficace l'importante quantité de fibres. Ceci peut être réalisé par regroupement de fibres, par regroupement en utilisant des approximations comme la méthode Nyström , par échantillonnage des données du faisceau de fibres ou en calculant une moyenne représentative comme nous le faisons dans le chapitre 10. Les méthodes basées sur le volume éludent cette question en traitant un structure de la matière blanche comme une seule unité anatomique. En tractographie cependant analyser les fibres au niveau individuel est généralement indispensable. Tractographie et algorithmes de classification ont tendu ceci possible. Des améliorations sont nécessaires, mais je suis d'avis que l'optimisation du pipeline de pré-traitement est un travail d'ingénierie.

Bien que la lourdeur de tailements soit un inconvénient, méthodes basée sur la tractographie offrent la possibilité d' étudier des paramètres locaux le long des fibres. Ceci est important dans l'étude de maladies de la substance blanche et, à ce stade, c'est ici que les contributions les plus importantes peuvent être aportées à la communauté médicale.

Dans cet esprit, nous introduisons une nouvelle série de méthodes géométriques pour l'analyse des faisceaux de fibres. Ces outils étendent les méthodes d'analyse de forme des fibres en proposant une analyse plus globale combinant les quatre caractéristiques physiques des fibres–forme, orientation, échelle et position– quand elles sont , modélisées

sous forme du courbes. Nous motivons l'utilisation de ces outils avec deux exemples: Dans le premier cas, nous montrons comment nous pouvons améliorer les résultats de clustering en sélectionnant les métriques de distance appropriées en fonction des problèmes rencontrés. Nous pouvons personnaliser les solutions de cette façon car nous travailler avec un plus grand choix d'outils.

Nous avons aussi montré du façon convaincante, à travers un exemple clinique, que ces mesures utilisées de concert avec permettent de détecter des changements physiques dus à une maladie (sclérose en plaque (SEP) dans ce cas). Jusqu'ici l'analyse de ce problème a été limitée à l'utilisation d'ACP ou de variantes de celle-ci. Les paramètres de distance que nous utilisons offrent différentes façons de mesurer les changements et les expériences que nous menons à partir de cette méthode permettent de regarder le problème sous un nouveau jour. En fait, cette étude est sans doute le premier exemple d'analyse géométrique de ce type dans le contexte de l'analyse d'images médicales. Avec des études plus poursées cette méthode pourrait être utilisée pour calibrer les stades de la maladie. Ceci permettra aux médecins de cibler le traitement à administrer de plus efficace.

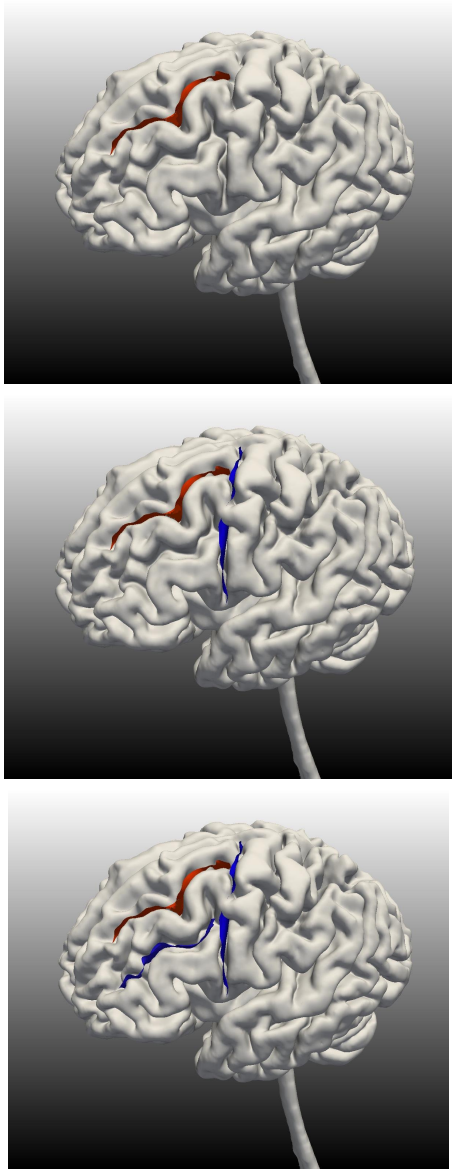


Figure 1.1: Les courbes des sillons sont labellisées en se basant sur leurs dependances spatiales avec les autres sillons.

The relative placement of sulcal curves is used to label them. We can identify the superior frontal sulcus (top panel, in red) by locating the precentral sulcus (center panel, blue), which is perpendicular to it, and the inferior frontal sulcus (bottom panel, blue) which runs parallel to it.

Chapter 2

Introduction

There are about a hundred sulci in the human brain and over a hundred billion white matter fibers. These two anatomical structures differ in many respects but they share a common geometric description: they are both open continuous curves. This thesis is a study of how the physical attributes of open curves can be used to advantage in the many varied quantitative applications of white matter fibers and sulci.

We begin by considering the open curve. The brain imaging landscape from this window has its own particular challenges, open problems and possibilities. The view offers the following:

Related to the representation

The choice of representation restricts our analysis to sulci, white matter fibers and the veins and arteries of the vascular system in the brain. These three anatomical structures can be modeled by open curves in 3D but we will limit the discussions in this thesis to sulci and white matter fibers. Figure 2.1 illustrates the relationship between the anatomy and a set of curves.

An open curve is one of the simplest ways to mathematically represent an individual sulcus or white mater fiber tract. An (open) surface, which may also be used, is more difficult to model and implement. The closed curve is a third option, but it is once again a mathematically more complex representation. Anatomical structures are not easy abstractions of geometrical primitives. Were this the case, a closed curve representation of a square or circle would be the right choice.

The details of the mathematical modeling aside, for the purposes of analysis (of sulci or white matter fibers), an open curve is the more useful representation. In the case of sulci, the fact that they are highly variable in shape and structure (see Chapter 4), also means that there is no additional information in a surface representation. The bottom curve (the section of the sulcus closest to the inner cortex– see Figure 2.1b) is the most stable part of a sulcus and this is what we use in our analysis. For white matter fiber analysis too, a curve model enables us to analyze individual tracts and, as we discuss in Chapter 6, there are benefits to this analysis over a volume-based one.

An open curve may be represented in a Euclidean or Riemannian space. We use both

in our work. In a Euclidean space, a spline (a parameterized curve defined piecewise by polynomials) can easily approximate shapes of complex curves. The $\mathbb{L}^2$ distance we compute between two such curves can be used for clustering or other classification tasks. The main advantage with this approach is the ease of implementation and the fact that existing methods from machine learning and statistics can be readily incorporated into the analysis.
In the Riemannian analysis, the open curve is represented as a point in the space of continuous functions. This space of functions is infinite dimensional and nonlinear and advanced methods such as those from functional analysis, differential geometry and group theory are needed to represent and analyze curves. The advantage of a Riemannian representation is that we can use the intrinsic (Riemannian) geometry of the space to compute distances between two points (curves). Open curves are essentially nonlinear structures. The space of open curves is also nonlinear. Linear measurements in such non-Euclidean spaces are not well-defined and the results may not be consistent.

Related to the geometrical analysis

The open curve is a geometric structure and we use it in analysis in this context. Geometrical analysis in medical imaging is most commonly a shape study although the term itself covers other possibilities. We use a mathematical model for open curves that allows us to incorporate shape, orientation, scale and position, the physical features associated with the curve. We expect to be able to apply this mathematical model to new kinds of problems or to reinterpret old ones.

One such new study design is the investigation of the morphological changes in the corpus callosum in Chapter 10. For the purposes of studying shape changes in its midsagittal profile, the corpus callosum is usually represented as a volume (comprised of thousands of individual curves) or is *skeletonized*. By measuring shape and orientation distances between individual curves, we have been able to detect changes in the callosal profile due to the progression of multiple sclerosis. These results have not been shown before for multiple sclerosis and potentially links atrophy, the mechanism believed to cause the changes, to the local white matter of the corpus callosum.

Sulci, on the other hand, are not easily interpreted in a geometrical study. This is due to the variability, which then translates to a problem of selecting a set of features that is unique to a class of sulci. The possibility of shape analysis has been explored before and in Chapter 5 and Appendix A we present some of our findings. Sulcal variability, as it relates to identification and labeling is, as we show in Chapter 5, better addressed by a graph pattern matching paradigm.

Organization of the Thesis

This thesis is divided into two parts: the first part is devoted to the analysis of sulci; the second, to white matter fiber analysis. We discuss these two anatomical structures separately because we encounter different problems with them: Sulci are highly variable; the white matter fiber tractography output, which is how these fibers are made available

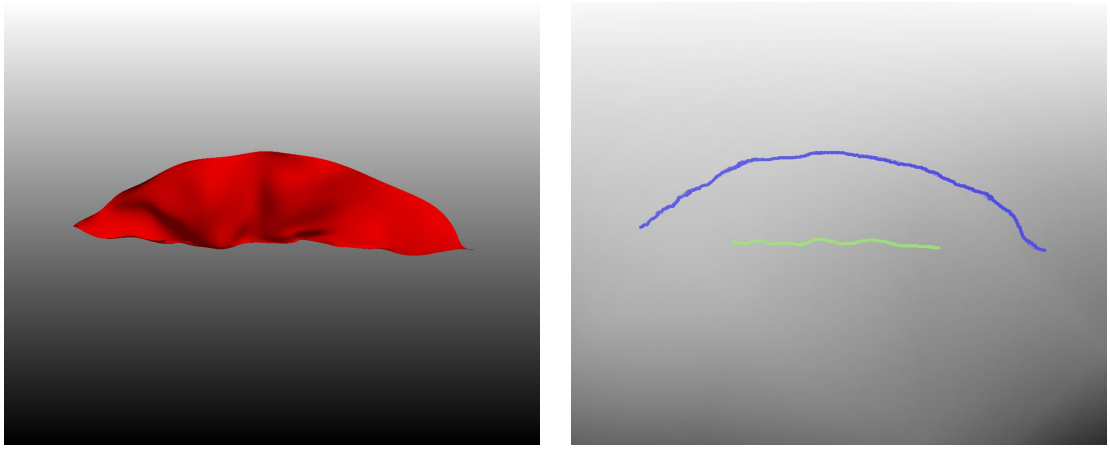
for image analysis, constitute a very large data set. Each structure requires a different approach. Furthermore, useful applications associated with these two anatomies emphasize different results.

A common set of mathematical tools are used for analysis in both Part I and II. Chapter 3 is an overview of a comprehensive Riemannian framework for the analysis of open curves [1]. This is a flexible mathematical framework: different combinations of features define joint Riemannian manifolds; the distances between curves in these manifolds quantify their differences. We have constructed this thesis around the idea that analysis outcomes and possibilities can be improved by using the full set of geometrical features associated with the open curve. This framework facilitates this and in doing so broadens the scope of geometrical analysis.

The discussions in Part I relate to the difficult task of labeling sulci. Chapter 4 introduces the *labeling problem* by linking each of these two words with characteristics that are innate to sulci. Labeling or identification is the most important application related to sulci. Chapter 5 offers a partial solution to this hard problem.

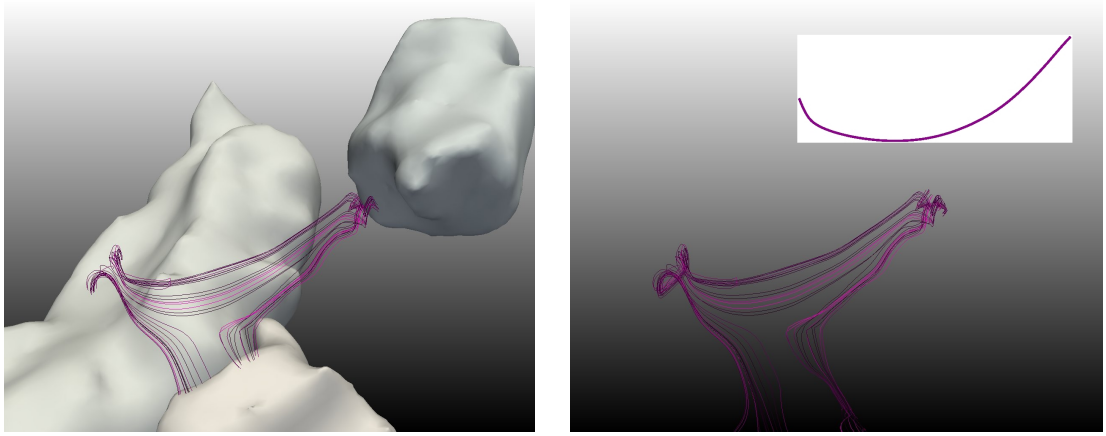
In Part II we focus on white matter fibers. Chapter 6 traces the historical development of quantitative white matter fiber analysis and supplies an important context for tract-based analysis. Tract-based studies offer the potential to analyze the length profiles of fiber bundles but there are few studies of this kind. A tract-based geometrical analysis of fiber tracts is a good application of the tools we use and in Chapter 10 we demonstrate how such studies are designed. A synopsis of the individual chapters of Part II can also be found at the end of Chapter 6.

In Chapter 11 we conclude by listing some of the contributions of the work presented in this document.



(a) reconstructed sulcal surface

(b) top (blue) and bottom (green) curves



(c) DTI fibers of the arcuate fasciculus

(d) DTI fiber modeled as a curve

Figure 2.1: Sulcal curves or white matter fiber tracts (from DTI or HARDI sources) can be modeled as 3D open, continuous curves.

Chapter 3

Comprehensive Riemannian Framework for Open Curves

The underlying theme in this thesis is the use of the full set of physical features – shape, scale, orientation, and position – in open curve analysis. This information can significantly aid the clustering, classifying, labeling and data analysis of open curves. In this chapter, we describe a flexible Riemannian framework for open curves which defines joint feature spaces or manifolds to study combinations of these features in a consistent way.

3.1 Introduction

White matter fibers reconstructed from DT-MRI images can be described as 3-dimensional open, continuous curves. Sulci can also be reconstructed to give a similar geometric description (Figure 2.1). A physical description of these structures would involve shape, scale, orientation, and position, the physical features associated with the curves. Of these features, shape is most commonly used in medical image analysis.

Several recent papers have proposed the use of a formal Riemannian framework for shape analysis of continuous curves [2, 3]. This type of framework has many advantages: (i) It provides techniques for comparing, matching, and deforming shapes of curves under the chosen metric. The correspondences for these tasks are established automatically. (ii) It also provides tools for defining and computing statistical summaries of sample shapes for different shape classes [4].

The Riemannian framework for open curves described in this chapter uses the same core ideas that were developed for elastic shape analysis of continuous closed curves. With this framework, we can compare and quantify differences between open curves in a coherent way. These comparisons are based on different feature combinations and each such combination constitutes a manifold. The five feature combinations we consider are:

1. Shape, orientation, scale and position: $\mathcal{S}_1$

2. Shape, orientation, scale: $\mathcal{S}_2$
3. Shape and scale: $\mathcal{S}_3$
4. Shape and orientation: $\mathcal{S}_4$
5. Shape: $\mathcal{S}_5$

These manifolds are united by a set of considerations and mathematical techniques; we describe this common methodology in Section 3.2. We then give mathematical descriptions for each of these manifolds (Sections 3.3.1–3.3.5).

This chapter serves as a mathematical background for the discussions in later chapters, particularly those in Part II of this thesis. The material presented is based directly on Mani et al. [1]. The methodology was developed by Anuj Srivastava and his collaborators; their publications, in particular those on Riemannian analysis of elastic curves, are other reference sources [3, 5, 6, 7, 8]. Concepts from differential geometry and group theory are discussed informally. For a more formal and complete treatment, the reader is referred to do Carmo [9, 10] for differential geometry and Rotman [11] for group theory.

3.2 General Methodology

The mathematical framework we describe draws from ideas in differential geometry, algebra and functional analysis and is an extension of the mathematical techniques developed for elastic shape analysis of continuous closed curves. Joshi et al. [3, 5, 12] introduced the square-root velocity function (SRVF) for the analysis of closed curves and Balov et al. [6] describe the analysis of open curves using the square-root function (SRF). The same general principles presented in these papers, i.e. the use of an elastic metric for shape analysis and a path-straightening approach to construct geodesics, apply here. The elastic metric, which is a Riemannian metric originally proposed by Younes [13], allows a curve to stretch and bend as it deforms along a geodesic. The path-straightening method [14] uses an arbitrary path to initialize the geodesic between two curves; the geodesic is iteratively computed using variational methods.

The mathematical tools we use enable us to define a representation space for open curves, ensure that the curves are invariant to certain shape-preserving transformations and compute geodesic distances between two curves. Each space defined uses a unique combination of features and, by extension, invariances but they share common procedural steps. These are outlined below:

1. Each fiber or sulcal curve is represented as an open continuous parameterized curve, β, as seen in Figure 3.1.
2. In order to compare curves we use elastic curve matching and for this the parameterized curve β is represented by a function. With the choice of function representation we can make the curve invariant to translation. We can also, at this stage, make the curve invariant to scale by manually scaling it.

3. The representation space described above (2) is a preshape space, $\mathcal{C}$, because many curves that are rotated and reparameterized versions of each other can actually be different elements. To unify these different representations of the same curve, one defines an equivalence class of functions. The set of all these equivalence classes is called the shape space $\mathcal{S}$.
   (*Note that the terms preshape space and shape space are used in a general sense for the representation spaces in manifolds $\mathcal{S}_1 - \mathcal{S}_5$. The methodology we use is an extension of work done in shape analysis where these terms were first applied and for consistency we use the same nomenclature for all the feature spaces.*)

4. A Riemannian structure is imposed on $\mathcal{S}$ and curves are compared by computing geodesics between their representations in $\mathcal{S}$. The geodesic length is a quantification of the difference between two curves in the joint feature space under consideration. The geodesics are computed using numerical algorithms.

We detail these steps describing the representations, the invariances and the computation of distances in the sections below. Figure 3.3 illustrates these steps and Figure 3.4 the relationship between the manifolds.

3.2.1 Representation Space

The curve is represented in an $\mathbb{L}^2$ space of square integrable functions first in a preshape space and then in a shape space which is a quotient space of the preshape space. There are thus two levels of representation. In this section, we will discuss the representation space under four headings: the curve representation, the function representation, the preshape space and the shape space.

3.2.1.1 Curve Representation

Let $\beta : [0,1] \to \mathbb{R}^3$ be an open continuous parameterized curve such that its speed $\|\dot{\beta}(t)\|$ is non-zero everywhere. The parameterization determines the rate at which the curve is traversed. The norm $\|\beta\| = \sqrt{\int_0^1 \|\beta(t)\|^2 dt}$ is a square-integrable function so we can refer to the space of all such curves as $\mathbb{L}^2([0,1], \mathbb{R}^3)$ or more simply $\mathbb{L}^2$.

An advantage to using continuous curves instead of landmarks which are reference points along the curve is that we avoid the problems associated with the selection of these points which many of the existing approaches to shape analysis have to contend with. As a trade-off, tools from functional analysis are needed.

3.2.1.2 Function Representation

The curve β is represented either by the *square-root velocity function* (SRVF) or the *square-root function* (SRF). These two function representations are important for the following reasons: First, the use of the $\mathbb{L}^2$ metric on the space of SRVFs and SRFs generates an elastic metric and a framework for elastic shape comparisons; Second, these functions preserve the $\mathbb{L}^2$ metric under reparameterization (see Section 3.2.3);

Lastly, a versatile framework can be built around these representations–the shape space can be easily modified and different combinations of features included.

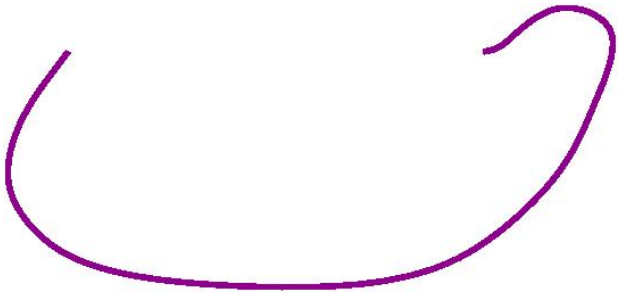


Figure 3.1: $\beta : [0,1] \rightarrow \mathbb{R}^3$ is an open continuous parameterized curve. A sulcal curve or white matter fiber can be described as an open curve β.

SRVF The square-root velocity function of the curve β is

$$q(t) = \frac{\dot{\beta}(t)}{\sqrt{\|\dot{\beta}(t)\|}} \ , \quad q : [0,1] \rightarrow \mathbb{R}^3 \ . \tag{3.1}$$

From Eqn. 3.1 we see that q is only dependant on the velocity term $\dot{\beta}$. It is therefore called the ***square root velocity function*** and the norm $||q||$ is the square root of the instantaneous speed along the curve β. It is possible to recover the original curve β, within a translation, using $\beta(t) = \int_0^t ||q(s)||q(s)ds$. The function q is invariant to translation in $\mathbb{R}^3$. By manually scaling the curves to the same length we can also remove the scale information at this stage.

SRF The square-root function is defined thus:

$$h(t) = \sqrt{\|\dot{\beta}(t)\|}\beta(t) \ , \quad h : [0,1] \rightarrow \mathbb{R}^3 \ . \tag{3.2}$$

Since h is dependant on β, the full set of attributes of the open curve, including global translation, are incorporated in the expression. The $\sqrt{\|\dot{\beta}(t)\|}$ term is a scalar that ensures invariance to reparameterization.

It is not easy to recover the curve β from the SRF because this involves solving a higher-order ordinary differential equation. Consequently, we cannot draw geodesic paths between two curves or compute sample statistics, though we can still compute distances between the curves. Because of this limitation, we only apply the SRF to cases where we want to include position as a feature; for all other cases, we use the SRVF.

3.2.1.3 Preshape Space

The space of all square-root or square-root velocity representations of curves $\mathcal{C}$ is an infinite dimensional vector space of all functions in $\mathbb{L}^2([0,1],\mathbb{R}^3)$. It has the form:

$$\mathcal{C} = \{f : [0,1] \to \mathbb{L}^2([0,1],\mathbb{R}^3)\},$$

where f is the function representation of the curve.

An inner product can be defined on $T_f(\mathcal{C})$, the tangent space of $\mathcal{C}$ at the point f:

$$\langle w_1, w_2 \rangle = \int_0^1 \langle w_1(t), w_2(t) \rangle \, dt, \qquad w_1, w_2 \in T_f(\mathcal{C}).$$

This metric, defined infinitesimally using elements of the tangent space at a point, is the ***Riemannian metric***. The use of the $\mathbb{L}^2$ metric on the space of square-root velocity representations generates an elastic metric; the Riemannian metric is therefore an elastic metric in this space.

The differentiable manifold $\mathcal{C}$ with this Riemannian metric is a ***Riemannian manifold***. With this Riemannian structure we can derive the following: (i) geodesic paths between curves; (ii) the exponential map; (iii) the inverse exponential map.

The elements of $\mathcal{C}$ do not represent the shape of a curve uniquely. A reparametrization of β, using an element $\gamma \in \Gamma$, where Γ is the group of ***diffeomorphisms*** (a smooth bijective map with smooth inverses) from $[0,1]$ to itself, results in a different square-root velocity function while preserving its shape. Similarly, any rigid rotation of β changes q but not its shape. Since $\mathcal{C}$ has many elements of the same shape, we call it the ***preshape*** space of open curves. (As noted before, we use the term preshape space in a general sense; it applies to the preshape space of $\mathcal{S}_1 - \mathcal{S}_4$, as well as to $\mathcal{S}_5$.)

In this framework, the three different preshape spaces we use are:

1. $\mathcal{C}_1 = \{h : \mathbb{R}^3 \to \mathbb{R}^3\}$. This is the preshape space for the $\mathcal{S}_1$ manifold. The h-function (SRF) is used here.

2. $\mathcal{C}_2 = \{q : \mathbb{R}^3 \to \mathbb{R}^3\}$. This is the preshape space for the $\mathcal{S}_2$ and $\mathcal{S}_3$ manifolds. The q-function (SRVF) is used since the space is invariant to translation.

3. $\mathcal{C}_3 = \{q : [0,1] \to \mathbb{R}^3 | \int_0^1 \|q(t)\|^2 dt = 1\}$. This is the space of all unit-length, open, elastic curves and is the preshape space for $\mathcal{S}_4$ and $\mathcal{S}_5$. The q-function is used and the curves are scaled to remove variability due to scale. The space is thus invariant to translation and uniform scaling.

3.2.1.4 Shape Space

To unify all the different representations of the same curve, we define an ***equivalence class*** or ***orbit*** of functions. In cases where invariance to reparameterization and orientation is sought, the orbit has the form

$$[f] = \{(\gamma, Of) | \gamma \in \Gamma, O \in SO(3)\}.$$

Here f is the function representation, i.e., the h-function or the q-function; Γ is the reparameterization group; $SO(3)$ is the 3-dimensional rotation group and (γ, Of) is the reparameterized and rotated curve. Equivalence relations are a by-product of group operations. Group actions with which we achieve invariance to reparameterization or orientation are more fully described in Section 3.2.2.

The elements of the orbit are considered equivalent, but the orbits themselves are distinct and do not intersect; collectively they define a disjoint set which is the ***shape*** space: $\mathcal{S} = \mathcal{C}/(\Gamma \times SO(3))$. This is now a ***quotient*** space of the preshape space $\mathcal{C}$ and has unique elements.
(*As with preshape space, the term shape space is generic to all five feature spaces, $\mathcal{S}_1 - \mathcal{S}_5$. These manifolds, discussed in Sections 3.3.1–3.3.5, are differentiated by the preshape spaces and equivalence classes that define them; only $\mathcal{S}_5$ is an actual shape space*).

There is a second type of equivalence class we are interested in. For manifolds $\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{S}_4$, invariance to reparameterization but not to orientation is desired. The orbit in this case has the form

$$[f] = \{(\gamma, f) | \gamma \in \Gamma\}$$

and the corresponding quotient space is $\mathcal{S} = \mathcal{C}/\Gamma$.

Since the shape space $\mathcal{S}$ is a quotient space of the preshape space $\mathcal{C}$, it inherits the Riemannian metric from $\mathcal{C}$. With this structure, we can compute a geodesic distance in $\mathcal{S}$ between two orbits $[f_1]$ and $[f_2]$. More details on geodesic distance computations are provided in Section 3.2.3.

3.2.2 Invariances, Equivalence Relationships

Since the shape of an object does not change when it is translated, scaled or rotated, an important requirement in shape analysis is that metrics be invariant to these transformations. Parameterized curves require an additional invariance. A reparameterization only changes the speed with which a curve is traversed, not its shape. Reparameterization is thus another shape preserving transformation. Metrics should be invariant to it as well.

There are two ways in which invariances to transformations are achieved. The first comes directly from the function representation. With the choice of function we can remove translation or scale information. In the $\mathcal{S}_2$ and $\mathcal{S}_3$ manifolds, for instance, we remove translation and in the $\mathcal{S}_4$ and $\mathcal{S}_5$ manifolds we remove both translation and scale. The second way to achieve invariance to transformations is by establishing equivalence classes. We elaborate on this in Sections 3.2.2.3 and 3.2.2.4.

We apply the four transformations, *translation, nonrigid uniform scaling, rigid rotation* and *reparameterization* to the shape manifold, $\mathcal{S}_5$. These transformations are discussed below. The invariance requirements for the four other manifolds, $\mathcal{S}_1 - \mathcal{S}_4$, are different but they are dealt with in a similar manner.

3.2.2.1 Translation

The *shape+orientation+scale+position* manifold, $\mathcal{S}_1$, includes position information and so we want the function representation we use to reflect this. For this manifold, the value of the function is designed to change with the translation or the change in position of the curve.
The four other manifolds, $\mathcal{S}_2 - \mathcal{S}_5$, do not include position as a feature and are invariant to translation. The function representation, q, used in those cases, is expressed exclusively in terms of the time derivative of the curve, β, and does not have a position component.

3.2.2.2 Scale

In order to remove the influence of the scales of curves in the quantitative analysis, we can rescale them to be of the same length. The rescaling, which is done with the preshape representation, leaves the function representation, q, and the Riemannian metric unchanged. The preshape space, however, is reduced. It is computationally convenient to scale the curve, β, to unit length; $\int_0^1 ||\dot{\beta}(t)||dt = \int_0^1 ||f(t)||^2 dt = 1$ then holds. Since they have unit norm, the set of all functions associated with curves of length one are elements of a hypersphere in $\mathbb{L}^2$. The differential geometry of a sphere is well-known and because of this, analysis and computations in subsequent steps is greatly simplified.

Two of the manifolds we consider, *shape*, $\mathcal{S}_5$, and *shape+orientation*, $\mathcal{S}_4$, are invariant to scale and receive this treatment. The shape attribute is by definition invariant to scale. In the case of the *shape+orientation* space, orientation, not scale, is added to the shape representation. Thus, this manifold retains the scale invariance of the shape space.

3.2.2.3 Reparameterization

The curve β is parameterized. This parameterization introduces an additional source of variability since arbitrary parameterizations are included in the representation. For any two curves, β_1 and β_2, different parameterizations, in general, result in different distances between them. We account for this variability by applying ideas from group theory. The shape of the curve does not change due to reparameterization so we would like to treat this mapping as we do rigid rotations and other shape-preserving transformations. We define a reparameterization group (i.e. the set of all reparameterized curves) and the action of this group on the preshape space results in reparameterization orbits. The elements of these orbits are equivalent. This enables us to compare orbits of curves, a comparison that is now independent of parameterization.

Γ, as noted before, is the set of all orientation-preserving diffeomorphisms (i.e., the direction at different points along the curve do not change due to the diffeomorphism). For the SRVF, i.e. the q-function, the map, $\mathcal{C} \times \Gamma \to \mathcal{C}$, is a group action defined by:

$$(q, \gamma) \to \sqrt{\dot{\gamma}}(q o \gamma).$$

The new function (q, γ) is the reparameterized curve and its distance from the original function q, is, in general, non-zero. We define two elements q_1 and q_2 as equivalent (i.e. $q_1 \sim q_2$) if for some $\gamma \in \Gamma$, $q_2 = (q_1, \gamma)$. A set of equivalent elements constitute an equivalence class:

$$[q] = \{(q, \gamma) | \gamma \in \Gamma\}.$$

We see that the equivalence class, which is obtained as a result of the group action Γ on $\mathcal{C}$, enables us to achieve invariance to reparameterization. The equivalence classes partition $\mathcal{C}$ into disjoint sets. The quotient space that results is $\mathcal{S} = \mathcal{C}/\Gamma$.

For the SRF, i.e. the h-function, the group action $\mathcal{C} \times \Gamma \to \mathcal{C}$ is defined by:

$$(h, \gamma) \to \sqrt{\dot{\gamma}}(h o \gamma).$$

Since the parameterized curve, β, is the starting point for all the manifolds, $\mathcal{S}_1 - \mathcal{S}_5$, we need to ensure invariance to reparameterization for each one of them.

3.2.2.4 Orientation

Let $O \in SO(3)$ be a rotation matrix and let $SO(3)$, the set of all possible rotations in $\mathbb{R}^3$, act isometrically (i.e., it is a distance-preserving map) on $\mathcal{C}$ as follows: $SO(3) \times \mathcal{C} \to \mathcal{C}$. $(O, q) = \{Oq(t) | t \in [0, 1]\}$ is the rotated curve that results from this group action. In cases where we want the analysis to be invariant to this group action we define the orbit of q under $SO(3)$ as

$$[q]_o = \{(Oq) | O \in SO(3)\} \subset \mathcal{C}.$$

The elements of $[q]_o$ are equivalent; they are rotated versions of each other.

The action of $SO(3)$ is combined with the action of Γ since, as noted in Section 3.2.2.4, the quotient spaces $\mathcal{S}_1 - \mathcal{S}_5$, need to also be invariant to reparameterization. The joint action of Γ and $SO(3)$ gives the larger orbit

$$[q] = \{(\gamma, Oq) | \gamma \in \Gamma, O \in SO(3)\} \subset \mathcal{C}. \tag{3.3}$$

The corresponding quotient space is $\mathcal{S} = \mathcal{C}/(\Gamma \times SO(3))$. We will consider how the two group actions, Γ and $SO(3)$, interact next.

3.2.2.5 The Product Group $(\Gamma \times SO(3))$

There are two important properties associated with the product group $(\Gamma \times SO(3))$:

1. The actions of $SO(3)$ and Γ on $\mathcal{C}$ commute. It is due to this that we can form an equivalence class (Eqn. 3.3) and define the action of the product group. We also make use of this property when we find distances and need to optimize iteratively (see Section 3.2.3.1).

2. The joint action of $(\Gamma \times SO(3))$ on $\mathcal{C}$ is by isometries with respect to the distance metric. This implies that the inner product on $T_{[q]}$, the tangent space to $\mathcal{S}$, is independent of the choice of $\tilde{q} \in [q]$. The geodesic distance between two points in $\mathcal{S}$ is given by:

$$d_{\mathcal{S}}([q]_0, [q]_1) = \min_{\tilde{q}_1 \in [q_1]} d_{\mathcal{C}}(q_0, \tilde{q}_1). \tag{3.4}$$

3.2.3 Geodesics and Distances

We can compare curves by quantifying their differences using a distance function. The standard $\mathbb{L}^2$ metric, often used in quantitative analysis of fibers, is given by

$$\|\beta_1 - \beta_2\|_2 = \sqrt{\int_0^1 \|\beta_1(t) - \beta_2(t)\|^2 dt}.$$

This metric is not, in general, invariant to reparameterizations, i.e.

$$\|\beta_1 \circ \gamma - \beta_2 \circ \gamma\| \neq \|\beta_1 - \beta_2\|.$$

Since we require invariance to reparameterization, we solve this issue by introducing the SRF and the SRVF representations. These functions preserve the $\mathbb{L}^2$ distance under reparameterization so that for any two curves β_1, β_2, with the corresponding functions f_1 and f_2, and any $\gamma \in \Gamma$, we have that $\|(f_1, \gamma) - (f_2, \gamma)\|_2 = \|f_1 - f_2\|_2$. Because of this equality, we are able to define the distance between the two curves as

$$d(\beta_1, \beta_2) = \gamma^* = \underset{\gamma \in \Gamma}{\operatorname{argmin}} \left(\|f_1 - (f_2, \gamma)\|_2\right) \ . \tag{3.5}$$

The minimization is performed by taking the path-straightening approach and using the standard dynamic programming algorithm for the computations. This is described in Section 3.2.3.1.

For the $\mathcal{S}_3$ and $\mathcal{S}_5$ manifolds, invariance to orientation is sought. Since we need to take the variability due to reparameterization also into account, the distance function now becomes

$$d(\beta_1, \beta_2) = \gamma^* = \underset{\gamma \in \Gamma, O \in SO(3)}{\operatorname{argmin}} \|f_1 - O(f_2, \gamma)\|_2 \ . \tag{3.6}$$

We solve this with a joint optimization described in Section 3.2.3.1.

In our open curve analysis, we are interested in the distance function in $\mathcal{S}$, the shape space. This distance function is inherited from $\mathcal{C}$. The distance between two curves β_1 and β_2 is the distance between their orbits $[\beta_1]$ and $[\beta_2]$ and is the pairwise shortest distance between elements in these two orbits. Once the optimal reparametrization and/or orientation of f_2 are obtained, we can compute the geodesic path between the orbits $[f_1]$ and $[f_2]$. Figure 3.2 shows the geodesic path between two curves in the $\mathcal{S}_2 - \mathcal{S}_5$ manifolds.

3.2.3.1 Optimization

The optimal distance is obtained by minimizing the cost function. For optimization over reparameterization, Γ, we use dynamic programming or gradient descent. When we need to optimize over both orientation and reparameterization, we perform a *joint optimization*; the optimal Γ is obtained by dynamic programming and the optimal orientation over $SO(3)$ is computed using Procrustes alignment. We briefly describe these procedures in this section.

Dynamic Programming Dynamic programming (DP) is a numerical optimization algorithm where we obtain an optimal path by solving the problem sequentially. At each stage we select an optimal trajectory from all the possible trajectories by optimizing (minimizing) the cost function. For this analysis, the cost function is the distance function (Eqns. 3.5 and 3.6) that matches the point $f_2(\gamma(t))$ with the point $f_1(t)$. Since the cost function is defined by the $\mathbb{L}^2$ norm, it is additive over the path $(t, \gamma(t))$ and can be cast as a DP problem. The algorithm forms a graph from $(0,0)$ to $(1,1)$ in $\mathbb{R}^2$ and searches over all the paths on that grid, such that the slope of the graph is strictly between 0^o and 90^o. This constraint is placed so that $0 < \dot{\gamma}^* < \infty$. The cumulative cost over the entire grid gives us an approximation to γ^*. The DP algorithm gives an exact solution.

A gradient-based optimization is an alternative to DP. The gradient-descent gives a local solution. Its estimate for γ^* is less accurate than DP but, as a trade-off, it is computationally less expensive.

Joint Optimization Consider the optimization problem:

$$(\gamma^*, O^*) = \underset{\gamma \in \Gamma, O \in SO(3)}{\operatorname{argmin}} \|f_1 - \sqrt{\dot{\gamma}} O f_2(\gamma)\|^2 \ . \tag{3.7}$$

We encounter this problem with the $\mathcal{S}_3$ and $\mathcal{S}_5$ manifolds where the curves are invariant to both orientation and reparameterization and it requires a joint optimization solution. The individual solutions for optimal orientation and registration are given below:

1. **Optimal Rotation**: For a fixed $\gamma \in \Gamma$, the optimization problem over $SO(3)$ in Eqn. 3.7 is solved by Procrustes alignment. We find $O^* = UV^T$ where O^* is the optimal rotation for aligning two curves and USV^T is the singular valued decomposition of $A = \int_0^1 f_1(t)(\sqrt{\dot{\gamma}} f_2(\gamma(t)))^T dt$. (In cases where the determinant of A is negative, one needs to modify V by multiplying the last column by -1. We have $O^* = U\tilde{V}^T$ in this case. This is a known result from rigid alignment of objects when the points across objects are already registered.)

2. **Optimal Registration**: For a fixed O, the optimization problem in Eqn. 3.7 over Γ can be solved using the DP algorithm described above. The cost function is defined by the $\mathbb{L}^2$ norm and, thus, is additive over the path $(t, \gamma(t))$. The algorithm forms a finite-dimensional grid in $[0,1]^2$ and searches over all the paths on that grid, satisfying the required constraints, to obtain an approximation to γ^*.

Since we have algorithms for optimizing over the two components individually, we can go back and forth between the two steps till convergence is reached.

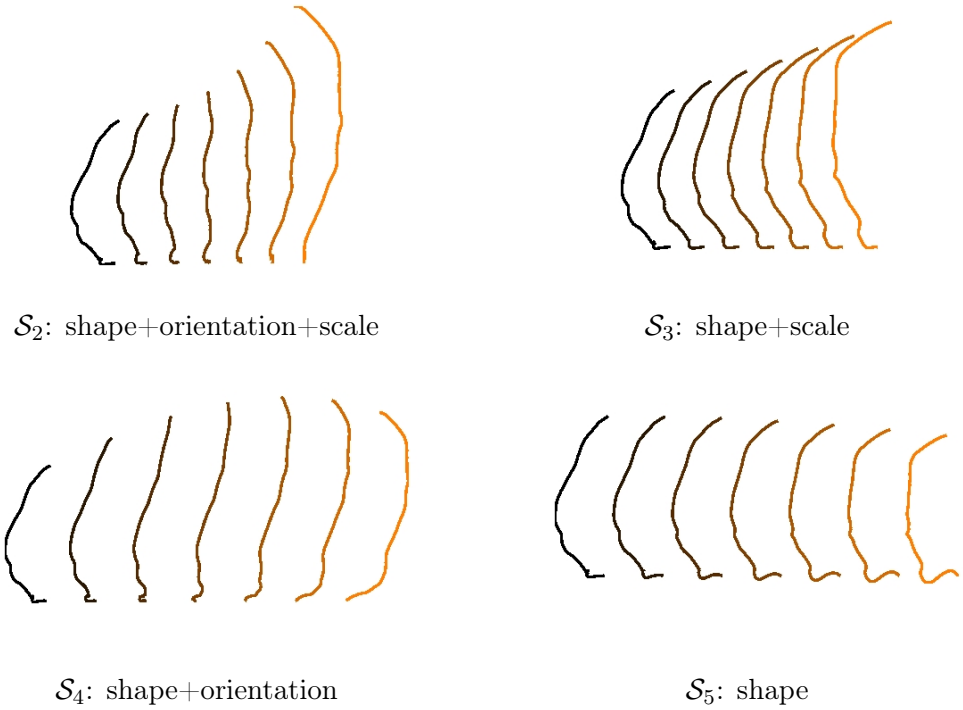


Figure 3.2: **Evolution of one curve into another along the geodesic path.** The two curves are DTI fibers that vary in shape, scale and orientation as seen in the $\mathcal{S}_2$ manifold. The reparameterization allows for bending and stretching of curves. This elastic matching results for a smooth and natural transition between curves. In $\mathcal{S}_3$, the two curves are oriented similarly so we see the evolution of the shape and scale. In $\mathcal{S}_4$, they have the same scale so we see the transformation of shape and orientation. The two curves have the same orientation and scale in $\mathcal{S}_5$, and so we see one shape transform into the other. Figure credit: Anuj Srivastava/FSU.

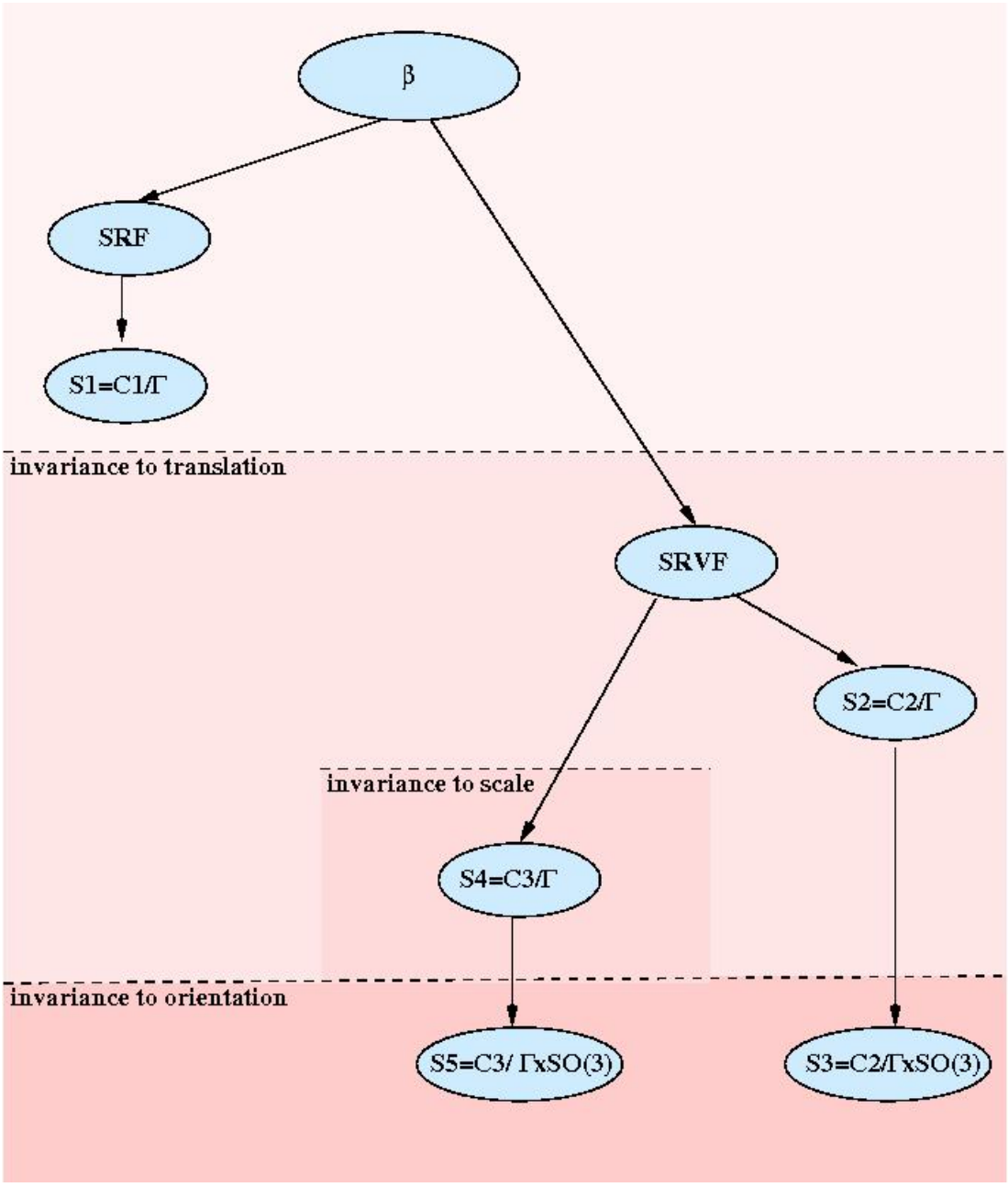


Figure 3.3: **Relationship between manifolds $\mathcal{S}_1 - \mathcal{S}_5$.** There is a hierarchy of steps that takes us from the curve β to its representation in the shape space. $\mathcal{S}_4$ and $\mathcal{S}_5$, for instance, follow the same mathematical procedures but the $\mathcal{S}_5$ space is in addition also invariant to orientation.

3.3 Description of the Riemannian Manifolds, $\mathcal{S}_1 - \mathcal{S}_5$

In this section, we present five Riemannian manifolds which capture five different combinations of the physical features of interest. We also provide metrics for comparing

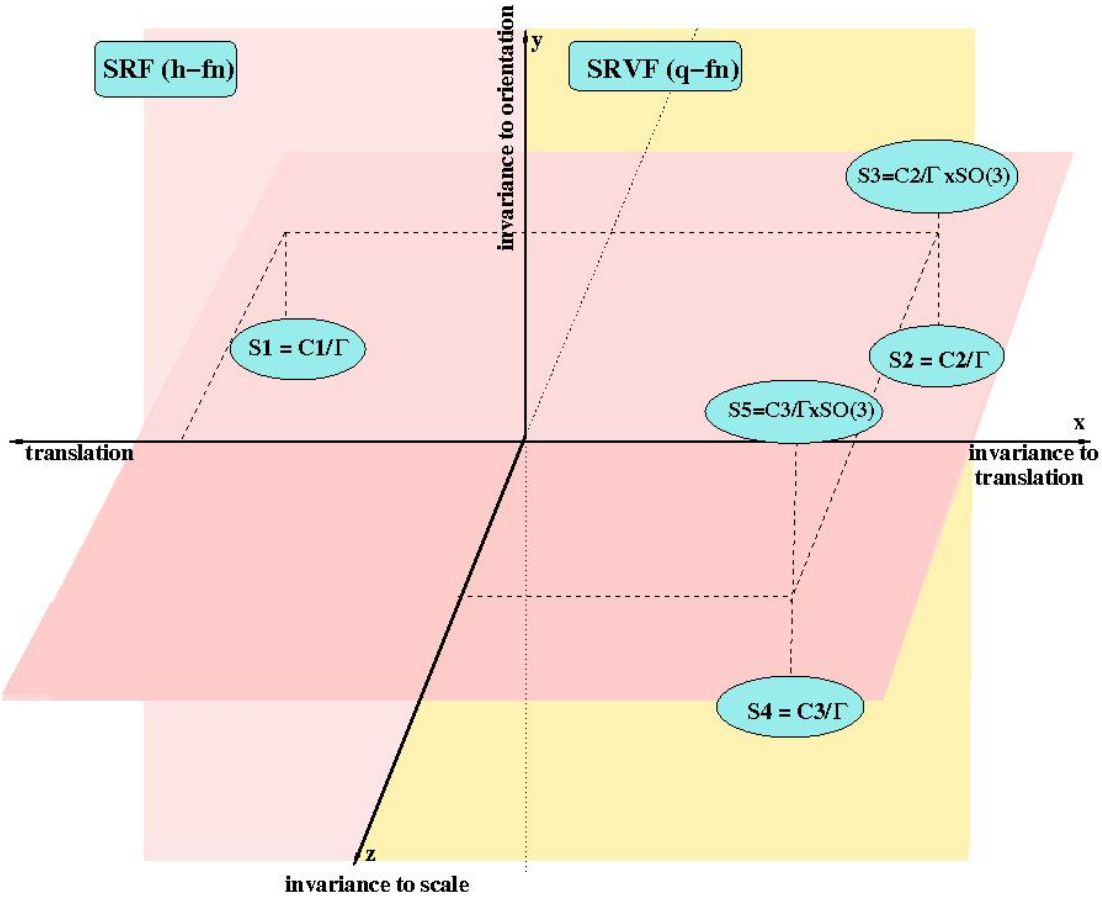


Figure 3.4: **Manifolds $\mathcal{S}_1 - \mathcal{S}_5$ and the invariances associated with them.** The comprehensive Riemannian framework allows for flexible feature combinations. Each of these constitute a Riemannian manifold and is associated with a subset of the four shape-preserving invariances: translation, orientation, scale and reparameterization. To visualize this, translation, orientation and scale are superimposed on an xyz-coordinate system. In the figure, which is multiperspective, since $\mathcal{S}_1$ includes position (translation), orientation and scale as features along with a SRF representation, it is located on the $-ve$ side of the $x-$, $y-$ and $z-$axes. $\mathcal{S}_2$, which is invariant to translation and uses the SRVF representation, is on the $+ve$ side of the $x-$axis. $\mathcal{S}_3$, invariant to translation and orientation, is on the $+ve$ $x-y$ quadrant; $\mathcal{S}_4$, invariant to translation and scale, is on the $+ve$ $x-z$ quadrant. Finally $\mathcal{S}_5$, the *shape* manifold, is invariant to translation, orientation and scale and has $+ve$ x, y, z values.

curves based on these features. In particular, for each manifold we provide: (i) a geodesic distance between curves that depends only on selected features (and is independent of the parameterization of curves), and (ii) a geodesic path between the two curves. We begin with $\mathcal{S}_1$, a space where all the features are utilized. Each subsequent manifold

is restricted to fewer features and $\mathcal{S}_5$, the last of these, uses only shape. Table 3.1 is a helpful summary of the manifolds and their preshape spaces and shape spaces.

3.3.1 Shape, orientation, scale and position, $\mathcal{S}_1$

We start by considering a situation where we are interested in comparing curves using all the four physical features – shape, scale, position and orientation. This might be useful when one of the four feature vectors (say position) predominates and the other three features are used to fine-tune the classification of curves. We shall see an example of this in Appendix A where we attempt to cluster and label sulci.

Preshape space: In this feature space, we use the square-root function (SRF) to represent the curve β:

$$h(t) = \sqrt{\|\dot{\beta}(t)\|}\beta(t) \ , \quad h : [0,1] \to \mathbb{R}^3 \ .$$

The SRFs are elements of the full $\mathbb{L}^2$ space so the preshape space is:

$$\mathcal{C}_1 = \{h \in \mathbb{L}^2([0,1], \mathbb{R}^3)\}.$$

Shape space: Since orientation, scale and translation are included in the representation, the only computation we need consider is invariance to reparametrization. The shape space is thus, $\mathcal{S}_1 = \mathcal{C}_1/\Gamma$.

When comparing two curves, we look at the difference between their SRFs in the shape space. For a curve β, the curve $\tilde{\beta}(t) \equiv \beta(\gamma(t))$ is simply the old curve with a new parameterization. For this reparameterized curve $\tilde{\beta}$, the SRF is given by $\tilde{h}(t) = \sqrt{\dot{\gamma}(t)}h(\gamma(t))$. We use (h, γ) to denote this reparameterized SRF. Now, it can be shown that for any two curves, β_1, β_2, with the corresponding SRFs, h_1 and h_2, and any $\gamma \in \Gamma$, we have that $\|(h_1, \gamma) - (h_2, \gamma)\|_2 = \|h_1 - h_2\|_2$. Because of this equality, we can define a distance between the two curves as:

$$d_1(\beta_1, \beta_2) = \min_{\gamma \in \Gamma} \left(\|h_1 - (h_2, \gamma)\|_2 \right) \ . \tag{3.8}$$

This minimization is performed using the standard dynamic programming (DP) algorithm, and it results in a quantification of differences in curves that is associated with the aforementioned four features. The geodesic path between the curves is the straight line:

$$\psi(\tau) = (1 - \tau)h_1 + \tau(h_2, \gamma^*) \ , \tag{3.9}$$

where γ^* is the optimal reparameterization obtained earlier in minimization using DP.

3.3.2 Shape, orientation and scale, $\mathcal{S}_2$

Here we consider the case where we compare curves using all the feature information except position. The need for such computations which use shape, scale and orientation distances may arise when the focus of our study is a localized section of brain anatomy.

Preshape space: For this feature space we use the q-function to represent β:

$$q(t) = \frac{\dot{\beta}(t)}{\sqrt{\|\dot{\beta}(t)\|}} \ , \quad q : [0,1] \to \mathbb{R}^3 \ .$$

This function is different from the square-root function used in Section 3.3.1 in that this definition is based only on the velocity function $\dot{\beta}$. Thus we call it the *square-root velocity function* (SRVF) [3]. Since this function is invariant to a global translation of β, analysis based on it will not depend on the global coordinates of the curves. The SRVFs are elements of the full space but since we use q instead of h we define a new preshape space:

$$\mathcal{C}_2 = \{q \in \mathbb{L}^2([0,1], \mathbb{R}^3)\}.$$

Shape space: Since we want to include orientation and scale in the computations, the only invariance we need to consider is reparameterization. The shape space is thus, $\mathcal{S}_2 = \mathcal{C}_2/\Gamma$.

The SRVF of the reparameterized curve is given by $(q, \gamma) \equiv \sqrt{\dot{\gamma}(t)}q(\gamma(t))$, where q is the SRVF of the original curve. As before, it can be shown that for any two curves β_1, β_2, with the corresponding SRVFs q_1 and q_2, and for any $\gamma \in \Gamma$, we have that $\|(q_1, \gamma) - (q_2, \gamma)\|_2 = \|q_1 - q_2\|_2$. Once again, we can define a distance between the two curves as:

$$d_2(\beta_1, \beta_2) = \min_{\gamma \in \Gamma} \|q_1 - (q_2, \gamma)\|_2 \ . \tag{3.10}$$

This minimization is performed using the DP algorithm, and it results in a quantification in differences in curves based on the three features–shape, orientation, and scale. The geodesic path between the two curves is a straight line given by:

$$\psi(\tau) = (1 - \tau)q_1 + \tau(q_2, \gamma^*) \ , \tag{3.11}$$

where γ^* is the optimal reparameterization obtained by DP.

3.3.3 Shape and scale, $\mathcal{S}_3$

This combination of features allows us to compare curves based on their shapes and scales. For WM analysis, we note that the lengths of WM tracts are determined by the brain regions they connect. Two prominent fiber bundles, the inferior longitudinal fasciculus (ilf) and inferior fronto-occipital fasciculus (ifo) are both long. They connect the occipital lobe to the temporal lobe in the first case and to the frontal lobe in the second case and can be separated from each other by their shape. Short association fibers connecting adjacent gyri intermingle with these fiber bundles but can be separated out because of their length. Sulcal curves also come in different lengths; the central sulcus, which is a primary sulcus, is long, the tertiary sulci are much shorter. Being able to discriminate sulci and fibers on the basis of their shape and length is important.

Preshape space: For the *shape+scale* manifold, $\mathcal{S}_3$, we use the same SRVF representation as the $\mathcal{S}_2$ (*shape+orientation+scale*) manifold. Since the SRVFs are elements of the full space the preshape space is once again $\mathcal{C}_2 = \{q \in \mathbb{L}^2([0,1], \mathbb{R}^3)\}$.

Shape space: Curves in the $\mathcal{S}_3$ feature space are invariant to orientation. To remove the rigid motions from the $\mathcal{S}_2$ representation, we rotate the curve β by a rotation matrix $O \in SO(3)$, where $SO(3)$ the set of all possible rotations in $\mathbb{R}^3$. The shape space in this case is $\mathcal{S}_3 = \mathcal{C}_2/(\Gamma \times SO(3))$.

The SRVF of the rotated curve is given by Oq where q is the SRVF of the original curve. We also need to ensure invariance to reparameterization. Consequently, the SRVF of a rotated and reparameterized curve is given by $\sqrt{\dot{\gamma}(t)}Oq(\gamma(t))$. The distance function is a joint optimization over $SO(3)$ and the group of orientation preserving diffeomorphisms, Γ. It is given by:

$$d_3(\beta_1, \beta_2) = \min_{\gamma \in \Gamma, O \in SO(3)} \|q_1 - O(q_2, \gamma)\|_2 \ . \tag{3.12}$$

Let γ^* and O^* be the reparameterization and the rotation that minimize the right side in this equation. Then, the geodesic path between any two curves, which once again is a straight line, is:

$$\psi(\tau) = (1 - \tau)q_1 + \tau(O^* q_2, \gamma^*) \ . \tag{3.13}$$

3.3.4 Shape and orientation, $\mathcal{S}_4$

The short U-fibers–white matter that connect adjacent gyri–are similar in shape but have different orientation. There are other structures that are identical in most respects but because of the bilateral symmetry in the brain, are oppositely oriented. Orientation is thus an important feature in brain image analysis.

We describe *shape+orientation* distances in this section. In Chapter 10, we use these distances along with shape distances to detect changes in the curvature of the corpus callosum along the median plane.

Preshape space: For the *shape+orientation* manifold, $\mathcal{S}_4$, we begin with the SRVF of the $\mathcal{S}_2$ (*shape+orientation+scale*) manifold and remove the scale component. To do this, we manually rescale the curves to be of the same length as described in Section 3.2.2.2. The mathematical representation of the SRVF remains the same as in $\mathcal{S}_2$ but the space of SRVFs, reduces from the full $\mathbb{L}^2$ space to a hypersphere as a result. This preshape space, the space of all unit-length (the curves are usually scaled to unit length) elastic curves, is defined as $\mathcal{C}_3 = \{q : [0,1] \to \mathbb{R}^3 \mid \int_0^1 ||q(t)||^2 dt = 1\}$.

The advantage of using a sphere is that the differential geometry is well-known and analytical expressions can be found for exponential maps, inverse exponential maps and geodesic paths. For example, if q_1 and q_2 are two elements of a unit hypersphere, the geodesic distance between them is given by the length of the shortest arc connecting them on the sphere. This length is $d_{\mathcal{C}_3} = \cos^{-1}(\int_0^1 \langle q_1(t), q_2(t) \rangle \, dt)$.

Shape space: As with the $\mathcal{S}_1$, $\mathcal{S}_2$, and $\mathcal{S}_3$ manifolds, invariance to reparameterization is achieved by the action of the group of diffeomorphisms, Γ, on $\mathcal{C}_3$. We shall denote this by $(q, \gamma) = \sqrt{\dot{\gamma}(t)}q(\gamma(t))$ and define the equivalence class of q as $[q] = \{(q, \gamma) \mid q \in S_\infty, \gamma \in \Gamma\}$. The shape space is $\mathcal{S}_4 = \mathcal{C}_3/(\Gamma)$. Note that we do not remove the $SO(3)$ group action on $\mathcal{C}_3$.

The distance between two curves does not depend on the reparameterization of the curves, i.e., for any q_1, q_2 and γ,

$$\cos^{-1}(\int_0^1 \langle q_1(t), q_2(t)\rangle \, dt) = \cos^{-1}(\int_0^1 \langle (q_1,\gamma)(t), (q_2,\gamma)(t)\rangle \, dt).$$

This leads to the definition of a distance between two curves which depends only on their shapes and orientations. The geodesic distance is calculated by minimizing the following:

$$d_4(\beta_1,\beta_2) = \min_{\gamma\in\Gamma} \left(\cos^{-1}(\int_0^1 \langle (q_1,\gamma)(t), (q_2,\gamma)(t)\rangle \, dt) \right) . \tag{3.14}$$

$\mathcal{S}_4$ is a sphere so the geodesic or shortest path between the two curves is a great circle. It can be specified analytically by:

$$\psi(\tau) = \frac{1}{\sin(\theta)}[\sin(\theta - \tau\theta)q_1 + \sin(\tau\theta)(q_2, \gamma^*)] \, , \tag{3.15}$$

where $\theta = d_4(\beta_1, \beta_2)$.

3.3.5 Shape manifold, $\mathcal{S}_5$

White matter fiber bundles have a well-defined shape and structure which is determined by the regions they connect and the constraints of the surrounding anatomy. Shape analysis of white matter fibers is an active area of study and is a starting point for our investigations in geometric modeling.

Preshape space: If we are interested in analyzing only the shape of 3-dimensional open curves, we must achieve invariance to translation, scaling, rotation and reparametrization. Since the q-function is defined in its entirety by the derivative of β, translational invariance is automatically removed. In order to achieve scale invariance, we manually scale all curves to be of unit length. This allows us to define the preshape space, $\mathcal{C}_3 = \{q : [0,1] \to \mathbb{R}^3 | \int_0^1 ||q(t)||^2 dt = 1\}$. This is the space of all unit length, elastic curves. We define a Riemannian metric and the tangent space on this manifold.

Shape space: We define the shape space to be:

$$\mathcal{S}_5 = \mathcal{C}_3/(\Gamma \times SO(3)).$$

The elements of $\mathcal{S}_5$ are the orbits of the type:

$$[q] = \{\sqrt{\dot{\gamma}} O q(\gamma) | q \in S_\infty, \gamma \in \Gamma, O \in SO(3)\} \, .$$

The shape space $\mathcal{S}_5$ inherits a Riemannian structure from the preshape space $\mathcal{C}_3$. The geodesic distance between any two orbits $[q_1]$ and $[q_2]$ is given by:

$$d_5(\beta_1,\beta_2) = d_{\mathcal{S}_5}([q_1],[q_2]) = \min_{\gamma\in\Gamma, O\in SO(3)} d_{\mathcal{C}_3}(q_1, \sqrt{\dot{\gamma}} O q_2(\gamma)) \, , \tag{3.16}$$

Table 3.1: Summary of the Comprehensive Riemannian Framework Manifolds

Manifold	function representation	preshape space	quotient space
shape + orientation + scale + position	$h(t) = \sqrt{\|\dot{\beta}(t)\|}\beta(t)$	$\mathcal{C}_1$ $\mathbb{L}^2$ space	$\mathcal{S}_1 = \mathcal{C}_1/(\Gamma)$
shape + orientation + scale	$q(t) = \frac{\dot{\beta}(t)}{\sqrt{\|\dot{\beta}(t)\|}}$	$\mathcal{C}_2$ $\mathbb{L}^2$ space	$\mathcal{S}_2 = \mathcal{C}_2/(\Gamma)$
shape + scale	$q(t) = \frac{\dot{\beta}(t)}{\sqrt{\|\dot{\beta}(t)\|}}$	$\mathcal{C}_2$ $\mathbb{L}^2$ space	$\mathcal{S}_3 = \mathcal{C}_2/(\Gamma \times SO(3))$
shape + orientation	$q(t) = \frac{\dot{\beta}(t)}{\sqrt{\|\dot{\beta}(t)\|}}$	$\mathcal{C}_3$ hypersphere	$\mathcal{S}_4 = \mathcal{C}_3/(\Gamma)$
shape	$q(t) = \frac{\dot{\beta}(t)}{\sqrt{\|\dot{\beta}(t)\|}}$	$\mathcal{C}_3$ hypersphere	$\mathcal{S}_5 = \mathcal{C}_3/(\Gamma \times SO(3))$

The manifolds, $\mathcal{S}_1$–$\mathcal{S}_5$, are quotient spaces that result when the reparameterization group, Γ, or the product group, $(\Gamma \times SO(3))$, are removed from the respective preshape spaces. The optimizations are done using DP and Procrustes alignment. $\mathcal{C}_1$ and $\mathcal{C}_2$ are different $\mathbb{L}^2$ spaces since the curves have different function representations. $\mathcal{C}_3$ is a unit hypersphere because the curves have been scaled to unit length to remove the effects of scale.

where $d_{\mathcal{C}_3}$ is the distance in the preshape space. A closer look at that distance function reveals the following:

$$\underset{\gamma\in\Gamma, O\in SO(3)}{\operatorname{argmin}} \cos^{-1}\left\langle q_1, \sqrt{\dot{\gamma}}Oq_2(\gamma)\right\rangle = \underset{\gamma\in\Gamma, O\in SO(3)}{\operatorname{argmin}} \|q_1 - \sqrt{\dot{\gamma}}Oq_2(\gamma)\|^2 \ , \tag{3.17}$$

where the $||.||$ is simply the $\mathbb{L}^2$ norm on the representation space. This equality implies that minimizing the arc length on a unit sphere is the same as minimizing the chord length. If one is minimized then so is the other. We can therefore use the $\mathbb{L}^2$ norm since it is computationally more efficient.
The actual geodesic between $[q_1]$ and $[q_2]$ in $\mathcal{S}_5$ is given by $[\psi_t]$, where ψ_t is the geodesic in $\mathcal{C}_3$ between q_1 and $\sqrt{\dot{\gamma^*}}O^*q_2(\gamma^*)$. Here (O^*, γ^*) are the optimal transformations of q_2 that minimize the right side in Eqn. 3.16. For $\theta = d_5(\beta_1, \beta_2)$, the geodesic path is a great circle given by:

$$\psi(\tau) = \frac{1}{\sin(\theta)}[\sin(\theta - \tau\theta)q_1 + \sin(\tau\theta)(O^*q_2, \gamma^*)] \ .$$

Part I

Analysis of Sulcal Curves

Chapter 4

The Sulcal Labeling Problem

The cortical surface is characterized by alternating ridges and furrows. The sulci (*singular* sulcus), as the Latin suggests, are the fissures or grooves; they serve as counterpoint to the raised gyri (see Figure 4.1). The sulci, in a sense, exist because they do not exist. Their utility derives from this fact as demonstrated by the following:

- In neurosurgery they function as channels which give a surgeon access to parts of the brain even deep within the subcortex. As the noted neurosurgeon M.G. Yaşargil writes in his foreword to the Ono atlas [15]: "any point within the cranium can be reached by following the corridors of the sulci". Tissue damage from incisions is thus minimized.

- They also serve as orienting landmarks in neurosurgery. The major sulci, in addition, partition important functional areas of the brain. This information reinforces their usefulness as landmarks. The central sulcus, for instance, demarcates the sensory-motor cortex. The sylvian fissure, one of the most identifiable cortical features, is the locus of language cortex. Both these sulci are important reference points in a variety of contexts and applications.

- The sulcal grooves are filled with cerebrospinal fluid (CSF) which make them easy to identify in T1-weighted images (where CSF is dark in contrast to the brighter gray/white matter).

To be useful in the neurosurgical applications described, we first need to identify and label the sulci. There are other applications that would also benefit from a reliable labeling scheme. Internal changes in the brain, either due to aging or pathology, for instance, alter the cortical surface. Labeling is the first step in a systematic study that allows us to quantify these changes for the differential diagnosis of disease.

Just as there is a need to have a scheme to identify sulci, it is equally desirable that it be implemented as an automatic labeling scheme. The reasons for this are obvious: manual labeling is tedious and the assessments are subjective. The problems associated with a manual procedure are brought to the fore when we work with large databases.

The labeling of sulci is a challenging problem. This is because, cortical sulci are highly variable. Sulci vary not just across individuals but even between the hemispheres

of a single brain [15]. They vary in shape, scale, placement and branching morphology. They may be continuous (present as one uninterrupted segment) in some individuals, fragmented (exist as multiple segments) in others and altogether absent in yet others. This poses a problem for feature selection and classification. Figure 4.3 illustrates how the variability can make feature selection difficult.

19th century illustrations such as those from Horsley [16], trace the wide variations along a sulcal fold. A whole nomenclature has developed since then to account for the branch variations possible along a single sulcus. (An example from the Ono atlas is illustrative–see Figure 4.2). For this reason, there is no gold standard in sulcal labeling; one neuroanatomist may disagree with another.

While accurate sulcal identification can be a challenge even for expert neuroanatomists, there are sulci that are to some degree more consistent, and for which anatomical correspondence can be established across subjects. These are the larger primary sulci (Figure 4.1(a)). The localization of these sulci allows us to generate a probabilistic map which can be used to label candidate sulci. A graph can also be constructed and unlabeled sulci (or the more variable secondary and tertiary sulci) can be identified against this reference. These two ideas have been incorporated into automated and semi-automated labeling methods in various ways.

A straightforward implementation of the probabilistic atlas paradigm can be seen in Le Goualher et al. [17, 18]. Statistical probabilistic anatomical maps (SPAM) [19] give the probability for each sulcal class so that at any given location, unlabeled sulci are assigned the most probable label for that location. A different spatial distribution model is used by Lohmann et al. [20]. A point distribution model [21] computes the shape of sulcal basins across a training set. Any unlabeled sulcus can be expressed as a linear combination of the eigenvalues generated from the PCA of this shape covariance matrix; an optimization over the linear function would give the best label.

Spatial distribution models give spatial bounds but this is not adequate to discriminate between the sulci in a local region. They are usually combined with graphs which model connections between sulci thus giving local context. In the combined strategy, the spatial information is used to supply spatial priors [22, 23], localization constraints or to narrow the search space in an optimization or matching process [24].

An example of a graph approach, where a model graph is matched to a random graph of sulci from a test subject, is presented in Rivière et al. [25]. Sulcal features, represented as the node attributes and neighborhood relations modeled as the edges of a graph are matched by training a set of neural networks to optimize the potential function at each graph node. Yang et al. [24] also apply graph matching to labeling. A similarity function that uses statistical information about spatial extent, orientation, shape and neighborhood structure is optimized using a genetic algorithm.

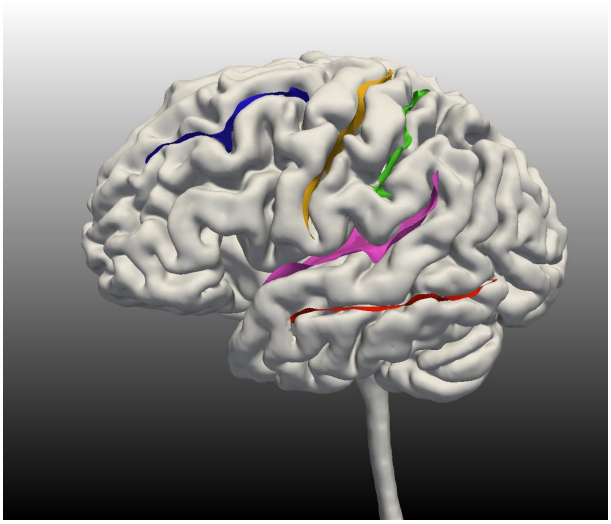
Mechouche et al. [26] use graph-matching in a very different way. An atlas represented via a semantic ontology is matched to pre-segmented candidate brains using descriptive logic. The main engine for the matching is a constraint solver (CSP), software that uses advanced numerical techniques that are computationally expensive.

Distinct from these matching approaches, deformable atlas matching [27], a method widely used in brain imaging to deform a template to a candidate brain, has been

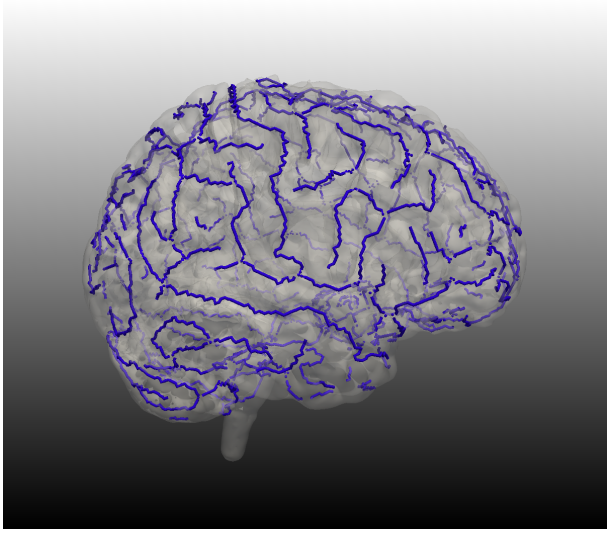
applied to sulcal labeling as well [28, 29]. The automatic surface registration presented in Vaillant et al. [28] is a hierarchical procedure guided by sulcal and gyral spatial probability constraints. In Jaume et al. [29], a multi-resolution strategy is used to match the lengths of sulci at different scales. The labels of the final match are successfully propagated to patient brains, but, it should be noted, to corresponding sulcal regions rather than to individual sulci. Elastic or fluid deformation algorithms [30] in general ensure a close match to local regions, but, with complex cortical folding patterns that vary from subject to subject, the matching process may give many local optima [25]. The application of these algorithms to this problem has been questioned in many other papers as well [17, 20, 31].

Graph and other learning-based matching approaches typically use large feature sets or computationally expensive algorithms. In [32], 8000 features at different scales are used to learn likelihoods to generate a probability map. Under a Bayesian formulation, the best curve is computed by dynamic programming using the probability map and shape priors. In [33], a large sample space of candidate curves is generated to learn shape priors which are used to define the potential functions of a graphical model. In [25], once again, a large number of training features are generated by oversampling the sulcal segments to be identified. Yang et al. [24] set up graph matching as a computationally expensive combinatorial optimization problem.

In Chapter 5 we present a simpler approach that retains the idea of a graph, but as a global representation of sulcal class data. Structural motifs of a distance graph act as a reference against which unlabeled sulci can be matched.



(a)



(b)

Figure 4.1: (a) Sulci are the grooves on the cortical surface. Five primary sulci of the left hemisphere (in color) are shown. (b) Sulcal curves (in blue) form a complex and widely varying pattern on the cortex.

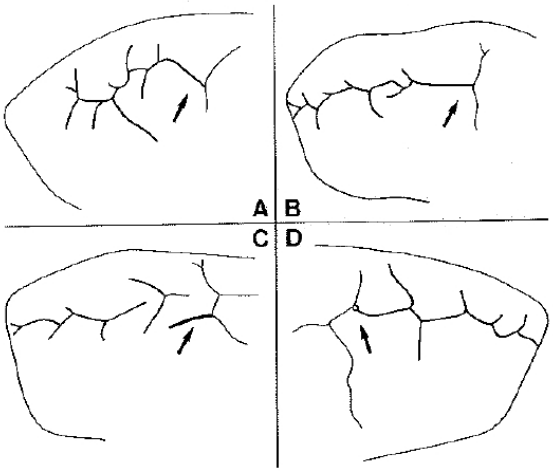
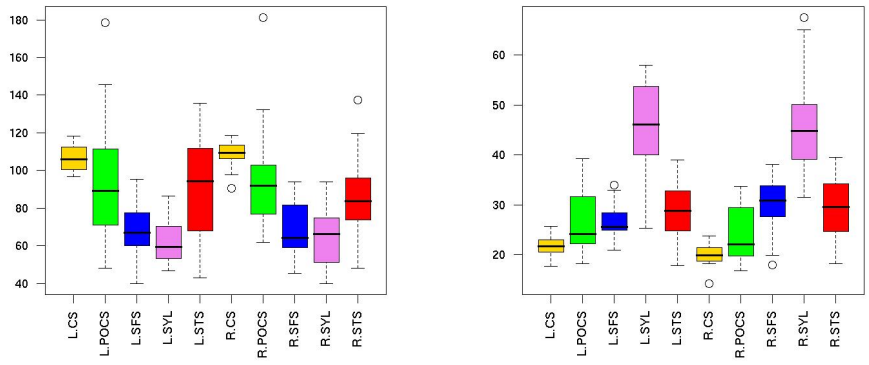


Figure 4.2: **Sulcal patterns** for the posterior end of the superior frontal sulcus. Types A, B, C, D, are possible variations for this sulcus (for the 25 postmortem brains examined, 4 variations were found). The pattern in the two hemispheres of a single subject may differ; the left may be Type B and the right may be Type C. The lengths of the small segments and the connections made to other sulci also vary. Reproduced from Ono et al. [15].



(a) length boxplot (b) depth boxplot

Figure 4.3: **The variability in sulcal features** is seen here in the length and depth boxplots for 18 subjects. Of the 10 types of sulci shown, only the sylvian fissure (in violet, see (b)) can be identified by a single feature, its depth measurement. This is not surprising since the deep sylvian fissure is one of the most easily identifiable parts of the cortical surface. The other classes of sulci cannot be separated out solely on a length or depth measurement.

Chapter 5

Labeling Sulci: The Relational Pattern Matching Solution

In the last 15 years, two ideas have emerged to address the variability of sulci in labeling and identification problems. The first is the use of a spatial distribution model of sulci to match candidate brains. The second is the explicit or implicit use of graphs to match the linkages between sulci. We base the work presented in this chapter on the second idea, which, we have implemented in a supervised learning framework. We use multidimensional scaling (MDS) to recreate the structural relationship in an embedded space. Unlabeled sulci, projected into the same space using an out-of-sample procedure, are given class designations based on nearest neighbor (NN) search. We take a heuristic approach in selecting the distance measure, which is central to all graph-based methods. In preliminary work, we found that a pure shape feature distance could not discriminate sulci adequately. Distances based on spatial features, however, gave encouraging results. Based on this ad hoc assessment, we evaluate 3 spatial distances designed to measure a (variable) sulcal graph structure. In experiments using a leave-one-out strategy, 90% of the 180 sulci drawn from 10 sulcal classes are successfully classified as are data from 5 out of 6 patients with cortical tumors.

With this supervised learning scheme, we offer a simple and intuitive solution to a challenging problem. The use of relational matching gives us the flexibility to accommodate normal variation in the sulci. We can also, to some extent, classify sulci such as pathological or partially segmented data, that are different from the training set.

This chapter is organized as follows: In Section 5.1 we explain our classification method which uses MDS and NN for labeling the major sulci. In Section 5.2, we describe the distance measures used. Section 5.3 has the experimental details and results. In the discussion that follows, Section 5.4, we consider the performance on both healthy and pathological data and offer insights on the success of our classification scheme.

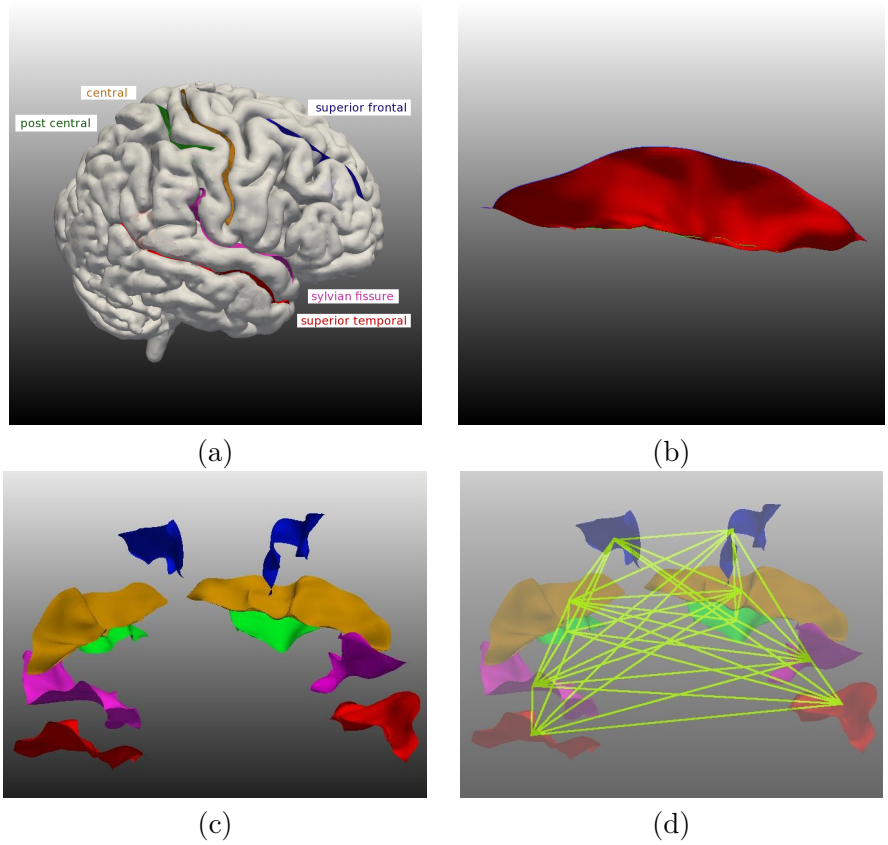


Figure 5.1: (a) The five major sulci of the right hemisphere used in this study. (b) An individual sulcus. The top or superficial curve, which points in the direction of the outer cortex, is in blue. The bottom curve is in green. (c) The ten sulci used in this study and their spatial relation (anterior coronal view). (d) Graph representation where the distances are the edges and the sulci are the nodes.

5.1 Methods

In Chapter 4 we discussed some of methods proposed to label sulci. Here we present a simpler approach that retains the idea of a graph, but as a global representation of sulcal class data. Structural patterns of a distance graph act as a reference against which unlabeled sulci can be matched. An example of a distance graph is shown in Figure 5.1(d). Multidimensional scaling (MDS) is a natural choice for this implementation. It gives a geometric interpretation which can be used to reproduce the structural relationships between sulci in a low dimensional space. We can thus build a reference atlas against which unlabeled sulci, using the same relational constraints, can be identified using nearest neighbor criterion (NN). The approach we take to classify unlabeled sulci, falls under the ambit of supervised learning. The steps are as follows:

1. First, using multidimensional scaling, we embed a map using dissimilarity data computed from a database of labeled sulci.
2. Next, once again using MDS, unlabeled sulci are introduced into the embedded space. Since this is an *out-of-sample* procedure, the original embedding remains unchanged.
3. Finally, the unlabeled points are assigned using nearest neighbor classification.

In describing this combination of steps as supervised learning, we would like to emphasize two points: First, each of the labeled nodes in the embedded space represents a class average rather than the individual data points of a class. MDS is not used as a precursor to clustering, but only to provide a common space to place labeled and unlabeled points. Second, the reference MDS map is generated solely from labeled sulci and not from a combination of labeled and unlabeled sulci as is done in semi-supervised learning. Each of the components of the labeling scheme is discussed below.

5.1.1 Classical Multidimensional Scaling (CMDS)

Multidimensional scaling is a method that allows us to compute an optimal configuration, X, for a set of n observations using only their interpoint distances. If $\Delta = [\delta_{ij}]$ is a non-negative, symmetric matrix of dissimilarities computed from a set of (hidden) coordinates, X^d of rank d, with MDS, we can represent X^d as X^p, $p \leq d$, such that the distance matrix D computed from it is close to Δ. An intermediate step would be to derive the scalar product matrix, $B = XX^T$, from the squared dissimilarities by double centering. An eigen decomposition, $B = \Gamma\Lambda\Gamma^T$, gives us Λ, a diagonal matrix of eigenvalues, and Γ, a matrix of orthonormal eigenvectors. X, the square root of XX^T, is then $\Gamma\Lambda^{\frac{1}{2}}$. Since B is of rank p, X is also of rank p. Thus, we see that in CMDS, the coordinates, X, of the embedded configuration are implicit and need to be extracted from the Euclidean inner product matrix B. The eigenvalue problem solves the optimization:

$$\min_{x_i \in \mathbb{R}^p} \sum_{i=2}^{n} \sum_{j=1}^{i-1} (\|x_i - x_j\|^2 - \delta_{ij}^2). \tag{5.1}$$

5.1.2 Out-of-sample Embedding

For classification, our objective is to introduce k unlabeled sulci, $y_1,\ldots,y_k$, without disturbing an existing configuration. Using CMDS to embed $n+k$ points, however, gives an entirely new map computed from inner products which are centered with respect to the centroid of the $n+k$ points. The optimization we seek instead is:

$$\min_{y \in \mathbb{R}^d} 2 \sum_{i=1}^{n} \sum_{j=1}^{k} (\|x_i - y_j\| - a_{i(n+j)})^2 + \sum_{i=1}^{k} \sum_{j=1}^{k} (\|y_i - y_j\| - a_{(n+i)(n+j)})^2, \tag{5.2}$$

where $A = [a_{ij}]$ is the augmented $(n + k) \times (n + k)$ dissimilarity matrix. We can get y_i, the coordinates for the k points with this while the original labeled points, x_i, which were previously computed, remain fixed. This is an *exact solution* for the out-of-sample problem for CMDS but it is a nonlinear optimization problem. If, though, the second term of the objective function above is dropped, the resulting function is convex. Solutions to this give an approximate out-of-sample embedding. The treatment is detailed in Trosset et al. [34].
We use Equation 5.2 later, in the context of a leave-one-out strategy to assess the utility of different data features in terms of generalisation error (as assessed with cross validation). In this instance the x correspond to the training set and the y the test set.

5.1.3 Nearest Neighbor Classification

For unclassified points, embedded in the out-of-sample step described above, class membership is determined based on their proximity to the class nodes. A minimum euclidean distance criterion is used.

In the nearest neighbor framework, we can also compute, as a confidence measure, the relative probability of an unlabeled sulcus being assigned to a class. Let us consider the entire embedded space as a Voronoi tesselation with a training node at the center of each cell. A *softmin* function over distances in this space will give a continuous probability that decreases the further a point is from a node. If d is the euclidean distance between an unlabeled point i and a node j, for n classes, the relative probability with which an unlabeled point selects a labeled node as its nearest neighbor is:

$$p_{ij} = \frac{e^{\frac{-d_{ij}^2}{\sigma^2}}}{\sum_{k=1}^{n} e^{\frac{-d_{ik}^2}{\sigma^2}}}. \tag{5.3}$$

The variance, σ, can be heuristically determined from training set data. Such computations cannot be made through a spectral embedding alone.

5.2 Feature Set

An important consideration with graph-based methods such as MDS, is the distance measure. In preliminary work [35] we used a geometric shape representation for open curves. The results indicate that *shape* distance, i.e. the geodesic distance between sulcal curves, is not a good feature for our classification scheme. A pure shape feature is obtained after removing *rotation, translation* and *scaling*. As part of our heuristic approach, we looked at each of these three components separately to determine their effects. We found that the position-based feature, which complements the spatial structure inherent in our pattern matching approach, gave very promising results. Here, we extend this study to three distance measures based on the position distance.

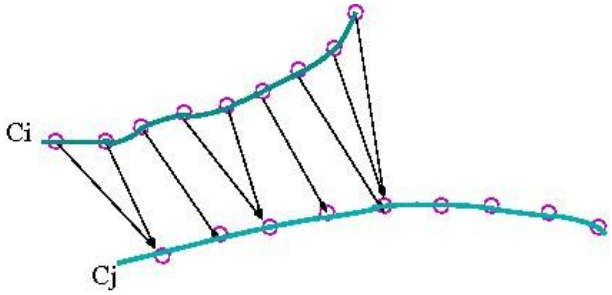


Figure 5.2: The set of n minimum distances from C_i to C_j used in the mean and median closest point computation. This is a directed distance; the symmetric distance is used in the distance matrix computations.

5.2.1 Distance Measures Evaluated

We evaluate three pairwise spatial distance measures between individual sulci. Since there are many different configurations to consider, with curves at various angles and of varying length, a measure that is robust to this variability will give the best results.

1. **Distance from midpoint**: The distance was computed from the mid-point of each sulcal curve.

2. **Mean closest point (MCP)**:

$$\begin{aligned} d_{MCP} &= \text{mean}(d_m(C_i, C_j), d_m(C_j, C_i)) \\ \text{where } d_m(C_i, C_j) &= \textbf{mean}_{i \in C_i} \min_{j \in C_j} d(i, j) \end{aligned}$$

   This distance (d_{MCP}) is described in [36, 37]. For each of n equally spaced points along curve C_i, the minimum distance to a point in curve C_j is measured. d_m is the mean of the set of these minimum distances. Since this is a directed distance, $d_m(C_i, C_j) \neq d_m(C_j, C_i)$ in general. The symmetric distance is found by computing the average of $d_m(C_i, C_j)$ and $d_m(C_j, C_i)$. See Figure 5.2.

3. **Median closest point (NCP)**: This is similar to the mean closest point, but we compute a median for the set of minimum distances instead of a mean.

5.2.2 Preliminary Features Evaluated

In our preliminary evaluation [35], in addition to the position feature, we used four other descriptors, pure shape, mean depth, length and orientation. These contain physical information not captured by the relative position of a 3D sulcal curve alone. The features, their class averages and the distances between the class averages are summarized below. It is these distances that we used as the distance feature in our characterization.

1. **Elastic Shape of a Curve:** We use the square-root velocity description for shapes of 3D sulcal curves: $q(t) = \frac{\dot{\beta}(t)}{\sqrt{\|\dot{\beta}(t)\|}}$ where $\beta : [0, 1] \to \mathbb{R}^3$ is a unit-length, parameterized curve such that its speed $\|\dot{\beta}(t)\|$ is non-zero everywhere.

2. **Mean Depth:** The depth profile for each sulcus was first obtained by resampling 100 points along the length of the sulcus and computing the distance between points on the top sulcal curve with corresponding points on the bottom curve. The mean depth was then calculated by averaging these. The top and bottom curve are shown in Figure 5.1b.

3. **Length:** The length of a sulcus was computed using $\int_0^1 \|\dot{\beta}(t)\|dt$, the length of the parameterized open curve β. β is defined in Chapter 3.

4. **Orientation:** Using singular value decomposition (SVD), the relative orientations, O, between sulcal curves were computed for each subject.

Class Averages
The class average for each of these descriptors was calculated by averaging across the subjects. For the relative orientation, the average was obtained by computing a SVD of $\sum_i O_i$. The Karcher mean gave us the mean shape.

Distances
The distance matrix was computed using the feature distances between two sulci. For position we computed a euclidean distance. For depth and length, we used an absolute distance. For orientation, we used $d_o = \frac{\|\log(O_1 O_2^T)\|}{\sqrt{2}}$, where $\|.\|$ is the Frobenius norm. For shapes, the geodesic distance was used.
The composite distance between any two sulci was also computed by using a combination of the scaled distances, $d = w_1 d_s + w_2 d_p + w_3 d_d + w_4 d_l + w_5 d_o$, where the w_i could be chosen by trial and error.

Table 5.1: Feature set used in preliminary analysis

	Descriptors	**Class Averages**	**Distance features**
1	Spatial Position		euclidean distance
2	Shape	Karcher mean	geodesic distance
3	Orientation	SVD($\sum_i O_i$)	$\frac{\|\log(O_1 O_2^T)\|}{\sqrt{(2)}}$
4	Length		$\|l_1 - l_2\|$
5	Mean Depth		$\|\bar{d}_1 - \bar{d}_2\|$

(1) Training set represented by class averages
(2) Individual unlabeled sulci constitute test sulci
(3) Distance matrix computed using feature distances between two sulci

5.3 Data Analysis

To assess the capability, i.e. the performance and limitations of the method we propose, we evaluated the following:
1) Preliminary analysis of spatial position, shape, orientation, length and depth
2) Evaluation of spatial position distances
3) Evaluation of spatial position distances with a larger data set
4) Partial sulci, i.e., sulci that are split into segments
5) Tumor data
Clustering, an alternative method, has also been used to group the data (see Appendix A). The results of the two approaches are discussed in Section 5.4.

5.3.1 Data and Initial Processing

Subjects: We used a database of 18 T1-MR 3D SPGR brain images of healthy subjects, matched for sex (male), handedness (right), and age (35 ± 10 years).
Sulci used: Five major sulci, the central, postcentral, superior frontal, sylvian fissure and superior temporal, (shown in Figure 5.1a), were extracted from each hemisphere. This gave us $18 \times 10 = 180$ sulci for classification.
Segmentation: The sulci were extracted using the *active ribbon* method [18, 38]. A 1D curve, skeletonized on the superior surface of the sulcal groove, evolves through the depth of the groove tracing out a 2D median surface in the process. This algorithm has been adapted from the active contour model [39].
Final form of the data: Performance was evaluated on two separate classifiers, one constructed with the top sulcal line and the other with the bottom line of a sulcus (see Figure 5.1b).

5.3.2 Evaluating Dimensionality, Assessing Fit

MDS reproduces a high dimensional input in a lower dimensional space. In Figure 5.3, an *elbow* in the scree plot at $d = 3$, represents the optimal dimension beyond which only minor reductions in stress are achieved. The Shepard diagram, which plots the scaled MDS distances against the original dissimilarity data, was then used to assess fit for 2 and 3D data. A good fit is indicated by a minimal spread in the scatter plot and evidence of a monotonically ascending relation. A comparison between the two Shepard plots, shows that the output distances are highly correlated to the input dissimilarities for the 3D data. As a final test, the MDS map we obtain, reflects the actual arrangement of sulci in the brain.

5.3.3 Classification

Our goal was to label pre-segmented sulci without making assumptions about predefined relationships or using semantic information. For this we designed a 10-class MDS classifier. We evaluated the classification using leave-one-out (LOO) cross-validation.

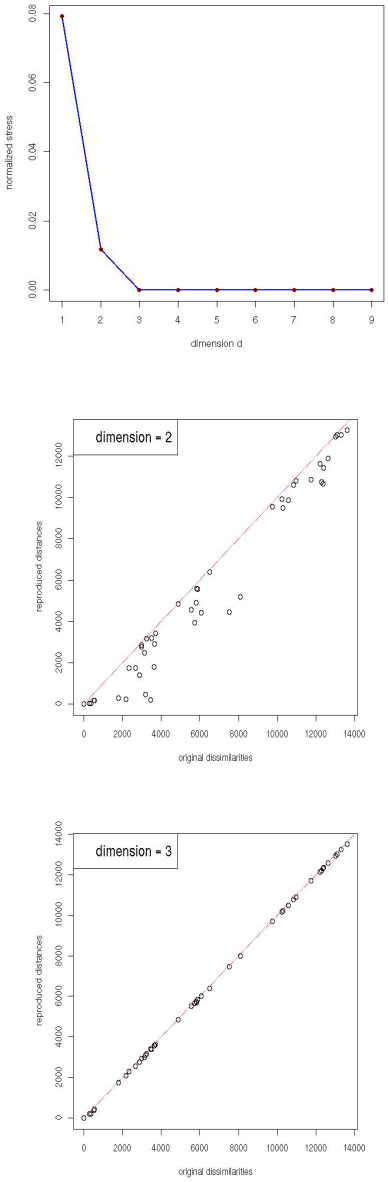


Figure 5.3: Evaluating the number of dimensions required to represent the data. The scree plot (left), shows that 3 dimensions is adequate for the MDS map. The Shepard plot for 2D data (middle) shows more spread than the corresponding 3D plot (right).

The 18 sulci for each of the 10 sulcal classes were split 17:1 into a training set and a testing set.

5.3.3.1 Preliminary Characterization

The best results for the 18 LOO iterations were obtained when spatial distances were used for the dissimilarities. Classification was also done for shape, using geodesic paths for distances, and for scaled combinations of shape, position, length, depth and orientation. The shape classifier labeled all the central sulci, which are very regular, but was less successful with the other classes. Combining the features did not improve the success rate. Rather, the results achieved by the spatial distance classifier were compromised (Figure 5.4). The depth and length classifiers could not discriminate right and left brain sulci, and while the orientation information improved results in cases where the orientations of the sulci differed, these were offset by the errors introduced using an *average* orientation as a feature.

5.3.3.2 Spatial Distances

For the distance measured from the mid-point of the sulcal curve, we got an overall true positive rate of 90% and 90.6% for the top and bottom sulcal classifiers respectively (Table 5.2), and were able to correctly label all the superior frontal sulci (100%). The central and postcentral sulci run parallel to each other as do the sylvian fissure and superior temporal sulcus. These were sometimes misclassified, one for the other.

The results for the mean (d_{MCP}) and median (d_{NCP}) closest point distance classifiers were 86% and 84% respectively (Figure 5.5). The MCP and NCP calculations underestimate distances when two sulci are perpendicular. It is perhaps because of this uneven application of distances that we got poorer results.

5.3.3.3 Partial Sulci

While the central sulcus and sylvian fissure are highly continuous [15], some of the other large sulci are interrupted by buried gyri or split into 2 or 3 pieces as an artifact of the segmentation. To simulate this, we split the test sulci into two at the mid-point. Of the $180 \times 2 = 360$ pieces tested using LOO cross-validation with the mid-point distance classifier, 77% were correctly labeled. This is in line with published results [25]. The MCP and NCP classifiers performed poorly here.

5.3.3.4 Tumor Data

The algorithm was tested on data from 6 patients with medium to large tumors (see Figure 5.7). The lesions had displaced sulci or had otherwise altered the arrangement of the surrounding cortical tissue. The sulci in the unaffected regions were successfully labeled, i.e. they had the same error rate as for the healthy subjects. The value in our method lies in the fact that we were also able to classify sulci in the region surrounding the tumor.

5.4 Discussion

Our supervised MDS spatial distance classifier offers several advantages.

First, it is a fast and lightweight implementation of CMDS and NN search. CMDS is an $O(N^3)$ time algorithm. For NN, we need to search through the entire training set to classify each new point. Both of these operations are computationally intensive. By computing class averages in the feature space, we get a compact representation of the training data, thereby significantly reducing the computational load.
Also, because we use an out-of-sample strategy, we further limit the number of computations. For the training distance matrix, we need only make $\frac{n(n-1)}{2}$ one-time measurements, where n is the number of classes. These, along with the MDS reference map are computed off-line. The only runtime computations are those that involve unlabeled sulci.

Second, our results suggest that we can very quickly assign a given sulcus to a small region of the left or right hemisphere. 179 out of 180 sulci were correctly identified when we relaxed the classification criteria to include the second nearest sulcus. For most of the misclassifications, the first and second nearest sulci had almost the same probability of being assigned to the correct class (see Section 5.1.3).

As an alternative to our method, we consider clustering, a method which is also often combined with MDS. When we tried to group the sulci using spectral clustering, 121/180 (67%) sulci were correctly classified (see Figure A.1). Clustering may not be the best choice for sulcal data since distance-based clustering algorithms do not use the structural pattern information from the graphs. The post- and precentral sulci, for instance, both equidistant and parallel to the central sulcus but on opposite sides, group together. Also, relying solely on distances when the data is inherently noisy could also lead to incorrect grouping. Some of the results from spectral clustering are presented in Appendix A.

Spectral clustering has been successfully applied to DTI data [37], for which, the large number of fibers warrant the use of an unsupervised or semi-supervised approach. Sulci, by contrast, are orders of magnitude fewer in number.

Third, the classifier is robust to normal population variation. The results we obtained, were for data that had not been spatially normalized and thus representative of the full sample variation. Also, the results for both the variable top and stable bottom sulci, each independently classified, were comparable (Table 5.2).

The relational paradigm allows us to identify anatomical regions for which, because of inherent variability, a feature set is difficult to select. We were even able to correctly identify displaced and distended sulci, taken from tumor patients. The preliminary success with labeling tumor datasets strongly suggests that we can design a general-purpose tool, i.e. one that can be used to identify both normal and pathological sulcal data. In general, diagnostic tools seek to differentiate healthy tissue from disease–tumor detection being an example. Here, we have demonstrated a method which can *unify* class data for applications where it is important to do so.

The classifier works best when only a small number of well separated sulcal classes are used in the training set. We saw the performance degrade from 90.6% to 84.3%

when two additional classes, the left and right precentral sulcus were added (Figure 5.6). This is a limitation of the current design. There are ways to extend the classification however. We can, for instance, train with different sets of small, well-separated sulcal classes. There are many ways in which these individual results can be combined to influence the final solution.

5.5 Concluding Remarks

In this chapter, we describe a novel scheme to label sulci which uses MDS in a supervised learning framework. Our method is conceptually simple and easy to implement, using for the most part standard, off-the-shelf algorithms. It is computationally fast because we use a compact representation of data and an out-of-sample procedure. It also requires no registration allowing us to use standalone sulci from a database. We considered several distance measures in the preliminary evaluation; some, like the scale distance, have not been widely used in medical imaging applications. Our method has the flexibility to be used both for healthy and diseased data. While there is a need for such an application, the conventional approach has been to design tools to differentiate rather than unify anatomy on the basis of group differences. Finally, our method gives impressive results, accurately labeling 91% of the 180 sulci we tested.

By training with different sets of sulci, we can extend the classification and this will be the direction for future work. Another direction for future work is to apply this classification tool in different contexts using it to label white matter fibers bundles and other anatomical structures.

Table 5.2: **Confusion Table for Spatial Distance (midpoint) LOOCV.** The true positive rate (blue), gives the % of the sulci correctly identified in 18 LOO tests. The off-diagonal terms (red) give the % of false negatives.

Top Sulcal Curve

Test Label: True Label		1	2	3	4	5	6	7	8	9	10
Left central	1	94.5	16								
L. postcentral	2	5.5	84								
L. sup. frontal	3			100							
L. sylvian fissure	4				78	11					
L. sup. temporal	5				22	89					
Right central	6						94.5	22			
R. postcentral	7						5.5	78			
R. sup. frontal	8								100		
R. sylvian fissure	9									84	16
R. sup. temporal	10									16	84

Bottom Sulcal Curve

Test Label: True Label		1	2	3	4	5	6	7	8	9	10
Left central	1	89	16								
L. postcentral	2	11	84								
L. sup. frontal	3			100							
L. sylvian fissure	4				84	11					
L. sup. temporal	5				16	89					
Right central	6						94.5	22			
R. postcentral	7						5.5	78			
R. sup. frontal	8								100		
R. sylvian fissure	9									94.5	5.5
R. sup. temporal	10									5.5	94.5

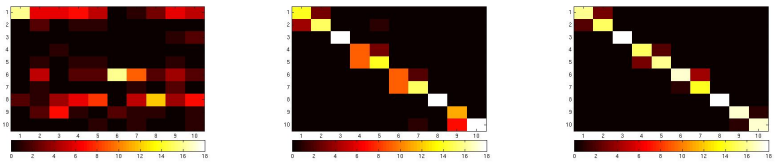


Figure 5.4: Confusion map for classification: the shape distance results (left) improve considerably when the shape and spatial distances (middle) are combined; the spatial distance alone (right) gives the best perfomance.

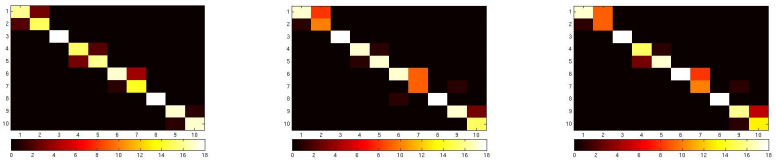
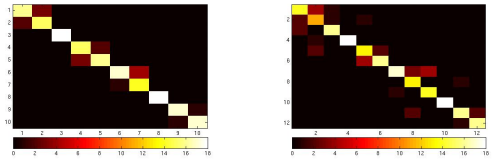


Figure 5.5: Confusion map for spatial distance classification: the midpoint distance classifier (left) gave the best results. The MCP (middle) and NCP (right) classifiers gave roughly similar results.



(a) 10 sulcal classes (b) 12 sulcal classes

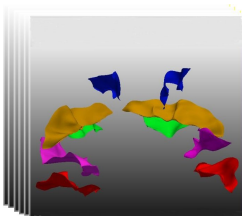
Figure 5.6: Labeling 10 vs 12 classes: The results for the spatial distance deteriorated when the two precentral sulci were added, increasing the number of classes to 12. The precentral sulcus was sometimes miscategorized as the adjacent sylvian fissure, the central or superior frontal sulcus.

Supervised Labeling of Brain Sulci using Multidimensional Scaling

Problem Statement

To label a database of pre-segmented sulci. These share a common referential system but are not spatially normalized.

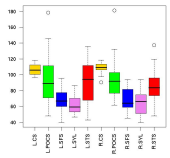
We make no assumptions about:
- predefined relationships between the sulci
- semantic information
- (*desideratum*) the kinds of subjects they come from



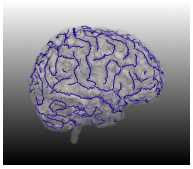
The Sulcal Labeling Problem

1. Large feature variation in the population (feature selection problem)
2. Convolutions vary from subject to subject and across hemispheres (template matching problem)

Variability in Length

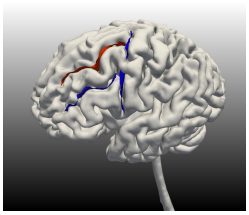


Complexity of Patterns



Expert Annotation–The Graph Approach

In manual annotation, neuroanatomists identify sulci by looking at their relation to other sulci



Classical Multidimensional Scaling

MDS gives a physical map with which we can reproduce the structural relationships between sulci. The eigenvalue problem solves the optimization:

$$\min_{x_i \in \mathbb{R}^p} \sum_{i=2}^{n} \sum_{j=1}^{i-1} (\|x_i - x_j\| - \delta_{ij})^2$$

For the Out-of-Sample Extension†:
The objective is to introduce k unlabeled sulci, $y_1, \ldots, y_k$, without disturbing an existing configuration. The exact solution is:

$$\min_{x_c \in \mathbb{R}^d} (2 \sum_{i=1}^{n} \sum_{j=1}^{k} (\|x_i - y_j\| - a_{i(n+j)})^2 + \sum_{i=1}^{k} \sum_{j=1}^{k} (\|y_i - y_j\| - a_{(n+i)(n+j)})^2)$$

If we drop the second term, the resulting convex expression can be solved numerically to give an approximate out-of-sample embedding.

†Trosset, M.W., Priebe, C.E., "The Out-of-Sample Problem for Classical Multidimensional Scaling," *Comp. Stat. & Data Anal.* **52**(10), (200 8): 4635-4642.

Classification with MDS

1. Compute a distance matrix for the class averages of a set of *training sulci*
2. Generate a reference MDS map
3. Introduce unlabeled sulci to the existing map by expanding the distance matrix and obtain an out-of-sample embedding
4. Membership is assigned using the nearest-neighbor criterion
5. The classification can be evaluated using leave-one-out crossvalidation (LOOCV)

Feature Set

	Descriptors	Class Averages	Distance features
1	Spatial Position†		euclidean distance
2	Shape‡	Karcher mean	geodesic distance
3	Orientation	SVD($\sum_i O_i$)	$\frac{\|\log(O_1 O_2^T)\|}{\sqrt{(2)}}$
4	Length		$\|l_1 - l_2\|$
5	Mean Depth		$\|d_1 - d_2\|$

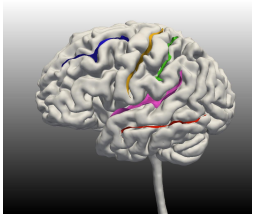
⋆*Training set represented by class averages*
⋆*Individual unlabeled sulci constitute test sulci*
⋆*Distance matrix computed using feature distances between two sulci*

†3 different position distances evaluated-midpoint, mean closest point, median closest point
‡joint work with Anuj Srivastava, Department of Statistics, Florida State University

Results

- 18 T1-MR 3D SPGR images of healthy subjects
  - sex (male)
  - handedness (right)
  - age (35 ± 10 years)
- 10 major sulci/subject
  - left, right central sulcus
  - left, right postcentral sulcus
  - left, right superior frontal sulcus
  - left, right sylvian fissure
  - left, right superior temporal sulcus

18x10 = 180 sulci

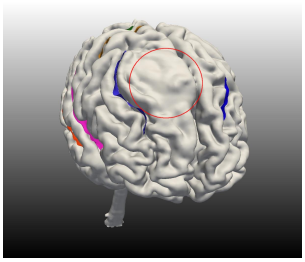


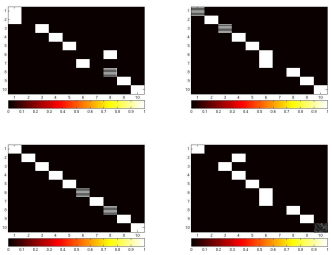
MDS Spatial Position Distance LOOCV

	1	2	3	4	5	6	7	8	9	10
L.central	89	16								
L.postcentral	11	84								
L.sup. frontal			100							
L.sylvian fiss.				84	11					
L.sup. temp.				16	89					
R.central						94.5	22			
R.postcentral						5.5	78			
R.sup. front.								100		
R.sylvian fiss.									94.5	5.5
R.sup. temp.									5.5	94.5

The % of the sulci correctly identified in 18 LOO tests (blue).
The off-diagonal terms (red) give the % of false negatives.
10 classes, 18 tests/class = 180 tests. **Success rate = 163/180 = 90.6%**

Labeling Tumor Data

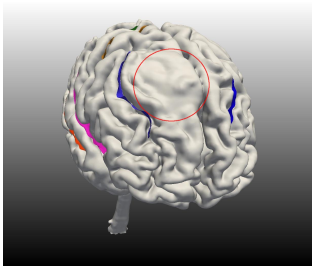




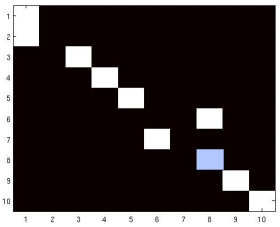
The labeling error for 4 patients in the unaffected cortical regions is the same as for the healthy subjects. In the affected regions(shaded boxes), the sulci are correctly labeled in 3 cases. The sulcus does not exist in the last case.

Advantages of the MDS spatial distance classifier

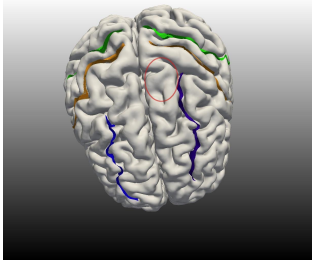
1. Quickly assign a sulcus to a small region of the left or right hemisphere
   - $\frac{180}{180}$ sulci correctly identified when classification criteria is relaxed to nearest or next-nearest neighbor
2. Classifier robust to normal population variation
   - both top and bottom sulci gave comparable results
   - even identify sulci displaced by large tumors (use as a general purpose tool)
3. Fast, lightweight
   - CMDS and NN search performed on group averaged data
   - Small feature set–only need $\frac{n(n-1)}{2}$ (i.e. 45 for $n = 10$) one-time distance measurements to construct a reference map
4. Easy to implement using off-the-shelf algorithms



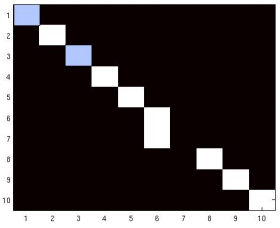
(a)right superior frontal sulcus



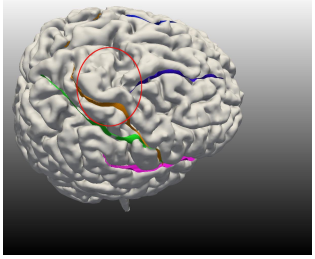
(b) displaced sulcus identified



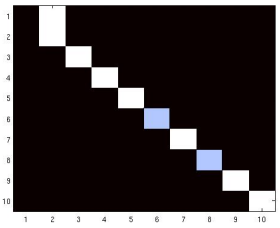
(c)left superior frontal & central



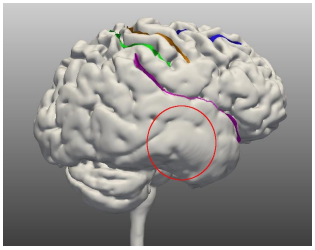
(d) displaced sulci identified



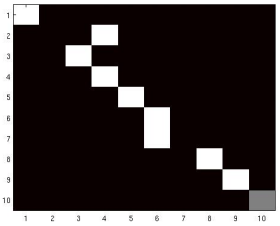
(e)right central & superior frontal



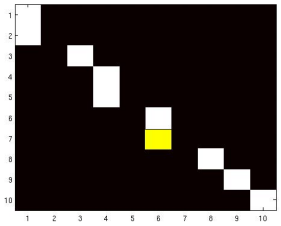
(f) displaced sulcus identified



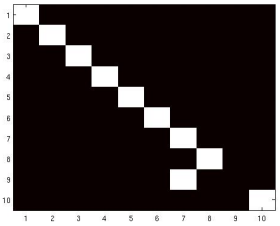
(g)right superior temporal



(h) no sulcus due to tumor



(i)displaced sulcus misidentified



(j) deep tumor, sulcus unaffected

Figure 5.7: **Labeling sulci displaced by tumors for 6 subjects.** The affected sulci are shown on the left; the labeling results, on the right and bottom. The labeling error rate in the unaffected cortical regions is the same as for the healthy subjects. In the affected regions (shaded boxes), the sulci are correctly labeled (mauve) or do not exist (grey) in 5/6 cases. It is misidentified in only one case, (i) (yellow).

Part II

White Matter Fiber Analysis

Chapter 6

Quantitative Analysis of White Matter Fibers

White matter (WM) is a part of the central nervous system. It is comprised primarily of myelinated axons which give it its whitish appearance. The axons are long extensions of nerve cells, the nerve cell bodies themselves being embedded in the gray matter of the brain. This architecture equips the WM to serve as a communication link between the gray matter and the rest of the body.

A WM fiber can be modeled as an open curve, the geometrical structure which is the focus of the investigations in this thesis. Part II of this document is devoted to the quantitative analysis of WM fibers. To frame the discussions which follow in the next four chapters, we begin this chapter with an abbreviated history.

6.1 A short history

The modern scientific study of white matter has its roots in the 19th century when links were being established between mental dysfunction and neuroanatomy. Correlations made between postmortem abnormalities in the brains of mental patients and clinical evaluations while they were living led to important discoveries. The identification of the arcuate fasciculus (AF) as a language pathway that connected the two language centers, the Broca and Wernicke regions, is one famous example. In that case, Carl Wernicke, who was developing language network models, made the association between lesions in the AF and the various aphasias he had observed.

The impetus from these findings crossed over to other developments. Theodor Meynert, a reputed neuroanatomist of that period, had classified prominent white matter tracts–or fasciculi as they were then called–based on the kinds of connections they made. Burdach and Déjérine published postmortem atlases and both prominently featured white matter dissections. New techniques for histopathological analysis were introduced. Notable among these was Golgi's staining method and Santiago Ramón y Cajal's use of it in his histolgical studies of nerve fibers. Quantitative investigations in white matter fiber analysis benefited from these cumulative efforts which made studies in

fiber thinning, demyelination and microstuctural damage possible. Today, postmortem dissections still give the most precise quantitative assessments.

With the advent of soft tissue imaging technology–computed tomography (CT) in 1972 and magnetic resonance imaging (MRI) in 1977–it was possible to examine living brains. In 1982, physicians observed multiple sclerosis (MS) lesions (see Figure 6.1 for lesions) in a living patient for the first time. Since then, clinicians have increasingly relied on brain scans for diagnosis and treatment. With this *in vivo* technology, disease progression in patients could be tracked by physicians and researchers, either individually, or as part of a longitudinal study of cohorts. This has led to a better understanding of white matter degenerative disease and has improved treatment options.

The MR signal can be assessed in different ways and the 1990s saw the emergence of two important MRI modalities. The first, Seiji Ogawa's 1990 proposal to use contrasts in blood oxygen response to map changes in brain activity [40], led to the development of functional magnetic resonance imaging (fMRI). The ability to view brain function in real time was a big step forward and is responsible for the widespread use of fMRI among clinical neurologists, behavioral scientists, neuroscientists and others.

Diffusion tensor magnetic resonance imaging (DTI) was the second important MRI modality introduced. Water constitutes a big part of living tissue–WM is 72% water [41]–and the physical flow of fluid is described by a diffusion process. In 1994, Peter Basser, James Matiello and Denis Le Bihan, in a landmark paper, proposed a tensor model for diffusion where the flow of water was described by the magnitude and direction of the principal eigenvector [42] at each image voxel. White matter fibers are inherently anisotropic and the first applications of DTI were studies of neural connectivity where fibers were tracked from end to end. Since the resolution of DTI is at the cellular level, it was also possible to detect disease–through indices such as fractional anisotropy (FA)–before it appeared in conventional MRI scans. Normal appearing white matter (NAWM) in MS is one example where compromised integrity manifests through lower average FA values.

A second-order tensor model is adequate for DTI reproductions of coherent fiber tracks but in cases where fibers meet or cross, only one of these directions is retained. DTI tractography of callosal fibers, where the lateral projections are attenuated, is illustrative of this limitation. To overcome this shortcoming, high angular resolution diffusion imaging (HARDI) images acquired in several spatially uniform directions are used. An orientation distribution function (ODF) that can model multiple maxima representing the different fiber directions replaces the simple tensor model at each voxel. HARDI datasets offer better resolution for important DTI applications such as connectivity studies and preoperative investigations.

6.2 Quantitative white matter analysis

It has been three decades since the introduction of the first MRI scans and there is large body of literature on quantitative white matter studies. It is useful to classify these efforts as volume-based or tract-based. The preprocessing path flows and possibilities

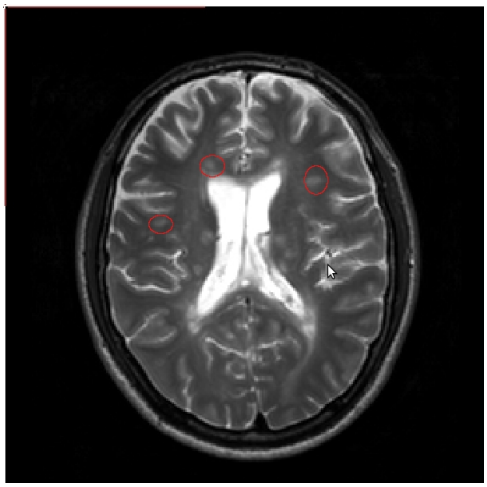


Figure 6.1: MRI scan shows MS lesions in a patient (circled in red).

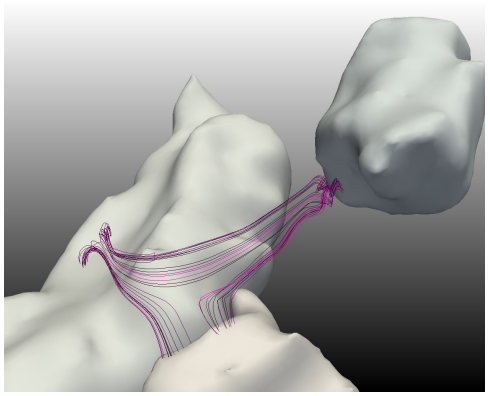


Figure 6.2: Tracts of the arcuate fasciculus that connect the Broca, Wernicke and Geschwind areas visualized through DTI. Tractography can aid connectivity studies.

for data interpretation differ in these two approaches.

Volume-based Anatomical structures or regions of interest are treated as volumes and quantitative information is smoothed or averaged in such a way that the local variation in individual tracts is not preserved. The white matter structures are usually segmented by thresholding FA maps though fiber tracts have also been used. Group-wise registration for population studies may be done at the voxel level or across individual volumes. A voxel-based coordinate system is used in the first case and a structure-based coordinate system in the second.

Tract-based The emphasis is on fiber tracts and parameters that vary along a fiber or anatomically defined bundle. Diffusion indices such as FA and physical descriptors such as shape are typically studied to assess fiber integrity or changes due to disease.

Tract-based image analysis was made possible after the first tractography algorithms were introduced in the late 90s. We devote the next section to developments in this area.

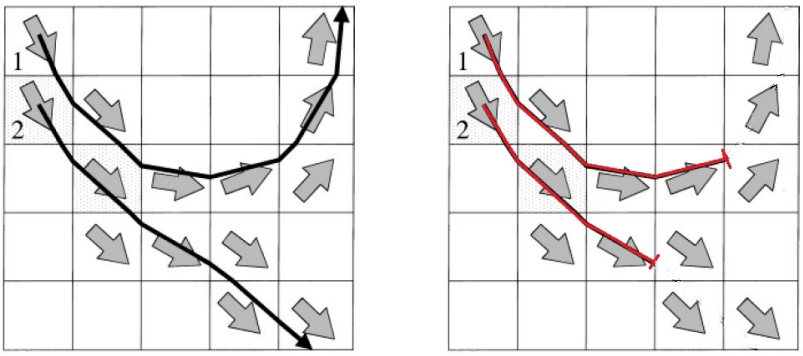
6.3 Tract-based white matter analysis

In 1999, Mori et al. ushered in tract-based image analysis by reconstructing fiber pathways in a rat brain [43]. Improvements to the basic tractography algorithm and work in clustering paved the way for data analysis. These three stages of what is a developing field are summarized below.

6.3.1 Tractography

Tractography or fiber tracking, as the name suggests, is a way to follow the direction of the local white matter diffusion from voxel to voxel. For DTI, the simplest algorithms follow the direction of the principal diffusion tensor eigenvector in a deterministic fashion. The reconstruction process, which includes curvature thresholds and other termination criteria, generates a tract or streamline (see Figure 6.3). More sophisticated approaches include interpolations for smoother pathways, the use of anatomical and topological constraints to guide the tracking and ways to deal with the uncertainty at each voxel due to noise and registration errors.

Streamline tractography is also used in conjunction with high angular resolution methods. HARDI acquisitions, which measure diffusion along several uniform directions in a sphere, are used to estimate orientation distribution functions. ODFs give the probability of diffusion for complex fiber architectures where there are multiple directions in a voxel. Q-ball imaging or spherical harmonic basis functions may be used to approximate the ODF. Higher order tensors may also be used to fit HARDI signals.



(a) (b)

Figure 6.3: **Fiber tracking** using the FACT algorithm. (a) The arrows at each voxel represent the direction of the principal eigenvector. A streamline connects the arrows. Reproduced from Xue et al. [44]. (b) FA thresholds and curvature thresholds serve as termination criteria. (Figure modified from the original in (a).)

6.3.2 Clustering

The mass of DTI fibers rendered is not immediately available for analysis. To organize and pare them down into meaningful fiber tracts, clustering is used and this has led to a study of these methods.

6.3.3 Mathematical frameworks

There are three steps in data analysis: pre-processing, modeling and statistical analysis. The pre-processing may involve registration, tractography segmentation or clustering and the information contained in these results is often incorporated into the data analysis model. FA is one such artifact and the analysis of FA profiles along tracts is usually included in most studies. The quantitative frameworks use both simple ideas such as the alignment of fibers for statistical analysis and more advanced mathematical concepts such as currents or techniques in Riemannian geometry. The methods listed below supply a context for the work presented in this thesis.

Fiber-tract oriented statistics This model by Corouge et al. [45, 46] is a prototype for computing statistics along a fiber bundle. The core idea was to give a compact description of the geometry and diffusion properties of a fiber bundle along its length profile. Corresponding points for a set of fibers are aligned and averages of DTI indices along the cross-section are computed. This can be done with linear or nonlinear Riemannian spaces. This basic idea has also been extended to a population atlas where information contained in tractography results [47] or registration results [48] are evaluated along a tract-based coordinate system.

Statistical modeling and EM clustering of WM fiber bundles Individual fibers are parameterized and aligned with the goal of performing tract-oriented analy-

sis over a population. A statistical model of pre-segmented fiber bundles is then calculated which serves as a prior for mixture model clustering. The Expectation-Maximization algorithm is used to infer membership probabilities and cluster parameters. Atlas guided clustering gives anatomically meaningful bundles. This is work by Maddah et al. [49].

Statistical model of WM fiber bundles based on currents This is a flexible framework where fiber bundles are modeled as currents. The space of currents is a vector space equipped with an inner product and norm which defines the distance between two bundles. This is a global distance that does not use point-wise correspondences [50]. Durrleman et al. [51] give a statistical model for a fiber bundle atlas and it variability in a population.

Comprehensive Riemannian framework In this framework, Mani et al. [1] use various combinations of shape, scale, orientation and position, the physical features associated with white matter fibers, to define Riemannian feature spaces. This is useful since applications have different objectives and often need different sets of tools and metrics for optimal results. For each joint manifold, a (geodesic) distance metric that quantifies differences between fibers as well as tools for computing statistical summaries of samples are defined. Correspondences between fibers are implicitly established during these pairwise comparisons. The framework may be extended to accomodate features such as scalar diffusion indices [7].

Statistical mapping of medial models of WM tracts Yushkevich et al. [52] use a different geometric model by giving major fiber tracts a medial representation. While this is strictly not a tract-based model, it is included here because local statistics of tensor-based features can be computed.

6.3.4 Tract-based analysis in this thesis

This thesis is motivated by the premise that the many varied applications of white matter fiber analysis, differential diagnosis of white matter diseases and neurosurgery being two of them, would benefit from the availability of metrics and methods that optimize the results sought.

6.4 Organization of Part II

In our work, we consider metrics based on shape, scale, orientation and position, the physical features associated with a fiber. These features have different properties so the usual approach has been to design different metrics and spaces to treat them individually. In Chapter 3, we described a comprehensive Riemannian framework where joint feature spaces allow for analysis of combinations of features. We can match and compare curves using the formulas for distance metrics and geodesics given. Correspondences between curves are automatically established through the use of an elastic metric.

We begin with a discussion of clustering, an important and necessary step in WM analysis. Clustering helps us organize the large volume of fibers obtained through DTI tractography and in the early phase, we reviewed existing methods and distance metrics in order to develop metrics for our own studies. Chapter 7 describes these efforts.

In the three chapters that follow, we demonstrate how the tools and metrics provided by the Riemannian framework can be effective in the context of WM analysis. In Chapter 8, we use specific tract configurations in the corpus callosum to show how clustering results vary with the different distance metrics. This nicely argues for the judicious selection of metrics in various applications.

The framework also provides tools for computing statistical summaries of curves. These summaries are a first step in statistical analysis enabling tasks of inference such as parameter estimation and hypothesis testing for group comparisons. For white matter fiber analysis, because of the large number of fibers involved, it is also convenient to first represent a set of fibers with a mean and variance that describes the essential characteristics of the fiber bundle. Chapter 9 covers this topic in more detail.

The study of WM disease is one of the primary objectives in WM analysis. In Chapter 10, we show how the feature distances described in Chapter 3 can be used to detect changes in morphology due to degenerative disease. Since this is the first time such distance measures have been used in addressing this problem, our approach may be considered novel.

Chapter 7

Clustering of White Matter Fibers

Clustering is an unsupervised method that assigns a set of data to subsets called clusters. The elements of a cluster are similar in some sense and the similarity is determined by similarity distances. For white matter fibers, the clusters are fiber bundles and the similarity criteria is used to ensure that the bundles correspond to anatomical structures.

Clustering is an important technique in white matter fiber analysis. Tractography algorithms help us visualize fibers and were naturally the first methods to be developed. Clustering followed soon after. A method was needed to organize the tractography output, which, for whole brain tractography can be several thousands of fibers. In this sense, tractography and clustering together constitute the pre-processing stages of quantitative analysis. The combined process has been referred to as *tractography segmentation* by O'Donnell et al. [37].

The primary goal of clustering is to partition the fibers into bundles that correspond to anatomical structures. Connectivity studies can be most directly related to both clustering and tractography but there are other applications. In preoperative surgical planning, a knowledge of tract locations, some of which may have been displaced due to lesions or injury, is needed. This is used to guide the surgery.

Quantitative frameworks built around a clustering algorithm can also be used to achieve a broader set of objectives. Maddah et al. [49, 53], for instance, obtain statistical parameters, and use a labeled atlas as a prior to label new bundles. O'Donnell et al. [37] also construct an atlas. By leveraging the Nyström method described in Section 7.1.2, only a fraction of the pairwise distance computations required for multiple subjects are made. They then apply an out-of-sample scheme to label new subjects.

Ding et al. [54] published some of the earliest work on DTI fiber clustering. There has been substantial progress since then. Both hierarchical [55] and spectral clustering [37, 56, 57] methods have been tried. The distance measures used can be broadly classified as *feature-based* or *point-correspondence based*. An example of a feature-based implementation is by Brun et al. [56] where a 9-dimensional vector that captures features such as shape, position and the start and termination points of a fiber, was used. We have also evaluated feature distances which we discuss in Section 7.2.3. Point-correspondence distances like the Hausdorff-style Mean Closest Point distance match

points along fibers [36, 37, 54]

In this chapter we report on our own efforts to group DTI tracts. We regard this procedure as an important preprocessing step for subsequent quantitative analysis. We choose spectral clustering because of its ease of implementation and the fact that it is robust. We carry out experiments with the algorithm parameters and distance metrics. Spectral clustering is reviewed in Section 7.1.1.

7.1 Spectral Clustering

Spectral clustering is relatively new but popular algorithm to cluster data. The method, as the name suggests, uses eigenvectors of data matrices to cluster points. The first spectral clustering algorithm to appear was by Shi and Malik [58] in 1997. They used normalized cuts (NCut) for graph partitioning and applied it to an image segmentation problem. Spectral clustering methods are now used in diverse domains. Brun et al. [56] were the first to adapt the method to cluster DTI fibers.

7.1.1 The Normalized Cut Algorithm

Spectral graph partitioning methods can be applied to clustering. For this we construct a weighted graph, $G(V, E)$, in which the data points are the vertices, V, and the distances between them, the edges, E.

A graph can be partitioned into two or more sets by removing the edges. This is called a *cut* (see Figure 7.1). Optimal partitions are obtained by minimizing the cut. The *normalized cut* is a kind of graph cut that takes the number of connections at each vertex into account. Outliers, which are poorly connected, are less likely to form unique partitions. The normalized cut is given by:

$$NCut(A, B) = \frac{cut(A, B)}{Vol(A)} + \frac{cut(B, A)}{Vol(B)}.$$

A and B are two disjoint subsets ($A \cup B = V$, $A \cap B = 0$) and $Vol(X)$ is the total connection from nodes in X to all the nodes in the graph (here X is A or B). Since the minimization of the NCut is NP-hard, we relax the criteria. The solution to this semi-optimal cut is the Fiedler vector, which is the second eigenvector of the graph's Laplacian.

There are two basic approaches to partitioning the graph with NCut: The first is a *bipartition* [58] where only a single eigenvector is used. The algorithm is applied recursively to obtain k smaller clusters. The *multiway* algorithms [59, 60] provide a second option. They use k eigenvectors to directly split the points into k clusters. The method by Ng et al. [60], which uses k eigenvectors simultaneously, is described next.

Let $\mathcal{S} = \{s_1, s_2, \ldots, s_n\}$, $\mathcal{S} \in \mathbb{R}^l$. To group these n points into k clusters:

1. Compute pairwise distances, d, between the elements of $\mathcal{S}$. This is an $n \times n$ matrix.
2. Construct the affinity matrix $K \in \mathbb{R}^{n \times n}$ such that $K_{ij} = exp(-d_{ij}^2/2\sigma^2)$ if $i \neq j$; $K_{ii} = 0$.
3. Construct the matrix $L = D^{-1/2}KD^{-1/2}$. This is a form of the Laplacian. D is the diagonal matrix (the $(i,i)^{th}$ element is $\sum_i A_{ij}$).
4. After spectral decomposition, find the k *largest* eigenvectors of L. The resulting matrix is $X = [x_1, x_2, ...x_k] \in \mathbb{R}^{n \times k}$.
5. Renormalize X so that each row has unit length. Each row is a point in $\mathbb{R}^k$.
6. Apply the *K-means* algorithm to obtain k clusters.
7. Assign the original data points s_i to a cluster point.

This algorithm is versatile and can be applied to a range of situations. The distance, d, can be a Euclidean distance between points, some other similarity measure or even a nonlinear function; K is a gaussian kernel though another radial basis function can be used; σ is the scaling parameter which controls the affinity between points–the larger the distance between two points, the smaller the affinity.

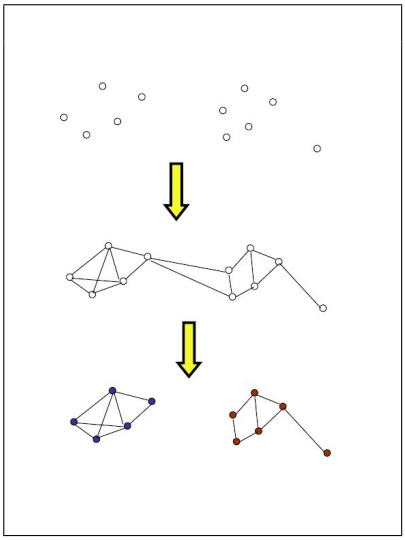


Figure 7.1: The graph repesentation of data. The graph can be partitioned with a *cut*. Reproduced from Arik Azran/Spectral clustering talk (www.videolectures.net).

7.1.2 Nyström Approximation

The singular value decomposition (SVD) of K, which is at the heart of every spectral method, requires $\mathcal{O}(n^3)$ operations. When n is large, the time required to factor K makes it unfeasible and the matrix itself becomes too large to be stored in memory. This presents a practical problem when clustering a large data set such as the DTI tractography output.

One solution is to sample from the data, using a small subset to approximate the entire affinity matrix K. The Nyström method, which is a numerical approximation to an eigenfunction problem, was introduced into the machine learning community for this purpose. It can be incorporated into the NCut algorithm to give an approximate solution for the eigenvectors associated with the data points [61].

The method was originally proposed by E.J. Nyström in the late 1920s, as an approximation to the integral equation [62]:

$$\int_a^b K(x,y)\phi(y)dy = \lambda\phi(x). \tag{7.1}$$

Using the same notation as in [61], K is a symmetric positive definite (SPD) kernel and (ϕ, λ) are eigenfunction-eigenvalue pairs. We can discretize this equation by applying a quadrature rule over the interval $[a, b]$. The original problem can now be reinterpreted as a discrete matrix eigenvalue problem:

$$K\hat{\Phi} = n\hat{\Phi}\Lambda,$$

where K is the Gram matrix, $\hat{\Phi}$ are the approximate eigenvectors and Λ, the corresponding eigenvalues. We extend these eigenvectors by getting estimates of eigenfunctions. We have solved a simpler eigen problem due to the discretization step. We then use this solution to extend the eigenvectors, this time using the full kernel.

The matrix form of the Nyström estimation is to write K as a Gram matrix $K = Z^T Z$ for $z_1 ... z_p$. We can then use a subset of the vectors z_i and their correlations with all the other vectors to find a low rank approximation $\hat{K}$.

We partition the Gram matrix $K \in \mathbb{R}^{p\times p}$ into submatrices A, B and C:

$$K = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}. \tag{7.2}$$

$A \in \mathbb{R}^{n\times n}$ is a square, symmetric matrixr; $B \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{m\times m}$ (see Figure 7.2). Noting that the matrix K is positive definite, we can write: $K = Z^T Z$. We then partition Z so that $Z = [X\ Y]$, $X \in \mathbb{R}^{p\times n}$ and $Y \in \mathbb{R}^{p\times m}$. This expansion of K gives:

$$K = Z^T Z = \begin{bmatrix} X^T X & X^T Y \\ Y^T X & Y^T Y \end{bmatrix}.$$

A can be diagonalized ($A = U\Lambda U^T$) and with some matrix manipulation we get:

$$\hat{K} = \begin{bmatrix} A & B \\ B^T & B^T A^{-1} B \end{bmatrix}. \tag{7.3}$$

Comparing K (Eqn. 7.2) and $\hat{K}$ (Eqn. 7.3), we see that the submatrix C is approximated by $B^T A^{-1} B$. This implies that it is only necessary to compute A and B. The error associated with the approximation is given by $||C - B^T A^{-1} B||$, which is the norm of the Schur complement.

7.1.2.1 Practical Application of the NCut-Nyström Method

In order to cluster data points (the DTI fibers) using the NCut-Nyström algorithm, we partition the affinity matrix K into submatrices A, B and C. For A, we compute affinities of a *random* sample of n data points. B represents the affinities between those n points and the rest of the data matrix. The submatrix C, which is much larger, is approximated by the Nyström algorithm. The partitioning of the matrix K is shown in Figure 7.2.

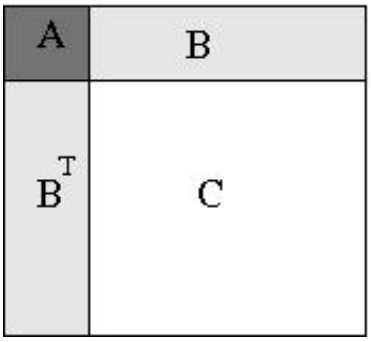


Figure 7.2: **The affinity matrix** K can be partitioned into submatrices A, B and C. $A \in \mathbb{R}^{n \times n} < B \in \mathbb{R}^{n \times m} \ll C \in \mathbb{R}^{m \times m}$.

There are two issues to be considered for the practical implementation:
(i) The calculation of A and B, depending on the size and distance metric used, could be computationally intensive. The diagonalizations that are required could also be expensive.
(ii) There is a tradeoff between the size of A and the spectral reconstruction error. For small A, there are fewer computations; for larger A, the reconstruction accuracy is better. The improvement in accuracy as A increases, however, reaches a plateau and so we can find an optimal partition for A and B.

7.2 Clustering the Corpus Callosum

The corpus callosum (CC) is both easy to identify in MRI scans and easy to segment. This may be one reason why it is the brain structure of choice in the published literature on clustering algorithms. The extensive use of the CC makes it a baseline against which to measure results. We use the CC in our own experiments, in part, for this reason. The main motivation for our choice is the fact that in our projects, the applications

are specific to the CC (in Chapters 8, 9, 10). To facilitate this work, we first establish optimal clustering parameters and distances. In this section, we present these results.

To evaluate the distance metrics, we look at: (i) the Mean Closest Point distance (many authors, e.g. [36, 37], have used this to cluster DTI data); (ii) a barycenter distance; and (iii) the Riemannian distance metrics for the open curve described in Chapter 3. For each of these, tests over an extensive matrix where σ and k are varied, have been performed. σ, the scaling parameter, and k, the input of the *k-means* algorithm specifying the number of clusters desired, are the two inputs to the basic spectral clustering algorithm.

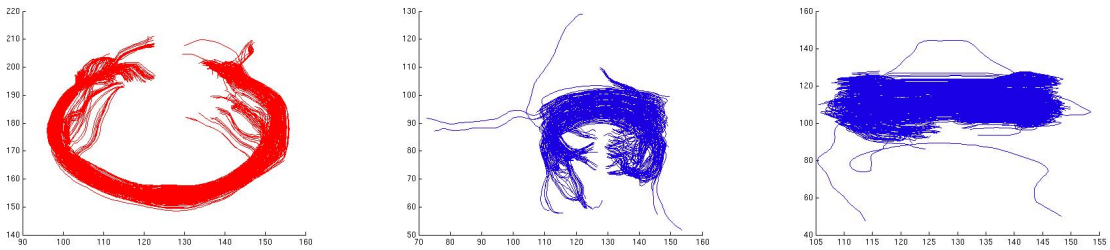


Figure 7.3: **Clusters of the corpus callosum.** The three clusters correspond to three fiber bundles: the splenium (left); the rostrum (center); and corpus (right). The MCP distance was used.

7.2.1 Mean Closest Point (MCP) Distance

The mean closest point distance is described in Chapter 5. We evaluated the following test matrix for 12 same-sex normal controls:

	$k=4$	$k=6$	$k=8$	$k=10$	$k=12$	$k=14$
$\sigma = 20$ mm				x	x	x
$\sigma = 30$ mm	x	x	x	**x**	**x**	**x**
$\sigma = 60$ mm	x	x	x	**x**	**x**	**x**

† **x** *(in bold text) represents the most frequent clustering parameters tested.*

A histogram of the distance matrix which shows the distribution of pairwise distances guided the initial selection of the σ values. The DTI tractography output is large (>10,000 fibers) so a random sample of 1000 fibers from the CC was used instead. The distance matrix was computed in two blocks: A = 250×250 and $B = 250 \times 750$ fibers since the Nyström version of the NCut algorithm was used.

Good clusters that corresponded to anatomical structures, in particular, the splenium, genu, tapetum and cingulam (which is not part of the CC but is often segmented along with it), were obtained for $\sigma = 30$ mm and $k = 10, 12$ and 14. With these input parameters, we were able to reproduce the results across 12 subjects, getting similar-looking clusters in each case. The clusters for two subjects are shown in Figure 7.4.

7.2.2 Barycenter Distance

The *barycenter* distance is the distance between the barycenters (centroid) of two fibers. The barycenter is computed using:

$$\mathcal{C} = \frac{\sum_{i=1}^{n} x_i}{n}$$

where $x_1, x_2, \ldots, x_n$ are the points used to represent the discretized version of the continuous curve.

Clustering experiments, over a test matrix, were performed in this case too. Since the barycenter metric has a different distance distribution, new σ values were selected from the distance matrix histogram. The number of points on each fiber used for the barycenter distance computation was $n = 15$. Affinity matrices A and B were computed for data from two normal subjects.

For both these subjects, the clusters obtained were cleaner compared to the results for the MCP distance on the same two subjects (see Figure 7.5). This assessment was based on external visual judgment. A sagittal view of the clustering (Figure 8.1) shows the clean divisions which correspond to a ground truth of anatomically defined bundles.

These results we see may be explained by the physical arrangement of fibers in the CC. For a successful clustering, fibers adjacent to each other should be grouped together. The barycenter distance calculates small distances for fibers that are very similar to each other along their entire length so that adjacent fibers are grouped together into a cohesive cluster. At the same time, the barycenter computation penalizes fibers that vary somewhere along their length. The MCP distance accommodates more variability. A fiber may have approximately the same pairwise distance for different fibers with different configurations.

Data from the distance matrix we compute for the two distance metrics also support the explanation put forth here. The σ values for the MCP were an order of magnitude larger than the σ values for the barycenter distance matrix.

7.2.3 Riemannian Distance Metrics

The Riemannian distance metrics defined in Chapter 3 were also evaluated over an extensive test matrix. The distances were:
1. Shape + orientation + scale + position distance: d_1
2. Shape + orientation + scale distance: d_2
3. Shape + scale distance: d_3
4. Shape + orientation distance: d_4
5. Shape distance: d_5

Experiment 1: For each of these 5 distances, we constructed a distance matrix. The effect of the σ and k was tested over several runs. These feature distances, in particular the shape distance, were not able to separate the CC into the clusters we desired. Figure 7.6 shows the results for the *shape* distance.
To explain the clustering results on the CC, we consider the following:

i) Most of CC fiber tracts rendered by DTI tractography have the same U-shape. A shape feature distance will not be able to separate tracts which have the same shape.
ii)The tractography is terminated after a specified length so the length of fibers is not a very useful feature.
iii) Fibers in the corpus, the long middle section of the CC, have approximately the same orientation. The orientation changes only for fibers at the anterior and posterior ends of the CC.
iv) Position is the only feature that changes across the length of the CC.

Experiment 2: A second set of runs were evaluated by combining each of these 5 distances with the MCP and the *barycenter* distance (e.g. shape + MCP distance, shape + barycenter distance). With the addition of a position distance, the clustering results were almost as good as the results with the MCP or *barycenter* distance alone.

7.3 Conclusion and Future Direction

We are interested in clustering the corpus callosum into anatomically defined bundles which will be used for quantitative analysis. We looked at different distance metrics in addition to the Mean Closest Point distance used by O'Donnell et al. [37] and some other authors. Through extensive experiments, we have determined that the distance best suited for partitioning the corpus callosum into smaller fiber bundles is the *barycenter* distance. The clusters we obtained corresponded with anatomical tracts. The splenium bundles, the genu, the corpus and the cingulum are easier to separate and we got good clusters. We were also able to separate the isthmus which is harder to separate from the posterior corpus (see Fig. 8.1).

In addition to the CC, we have successfully used this distance to cluster the arcuate fasciculi, the genu and splenium of the CC and the corticospinal tract.

The Riemannian feature distances performed poorly because, as discussed in Section 7.2.3, the CC is not the best anatomical structure to apply *shape* or scale distances. In Chapter 8, we apply these same distances to the CC. We target sections of the CC (rather than cluster the entire structure) and demonstrate the utility of these metrics in a specific context. There is a high computational cost associated with the Riemannian distance metrics so their use should be limited to problems where the number of fibers is small.

One aspect of the Nyström-NCut algorithm that would be useful to study is the optimal partition of the A and B matrices. Experiments by various authors in the machine-learning literature suggest that the error associated with the spectral reconstruction (i.e. the accuracy of the estimated matrix C) is influenced by the size of A. There is an optimal size for A beyond which the accuracy does not improve [61, 63]. In these experiments we choose $A = 250$ which is 25% of the full set of 1000 fibers used in the experiments. By finding the optimal A (which is smaller) we would be able to leverage the full efficiency of the Nyström procedure.

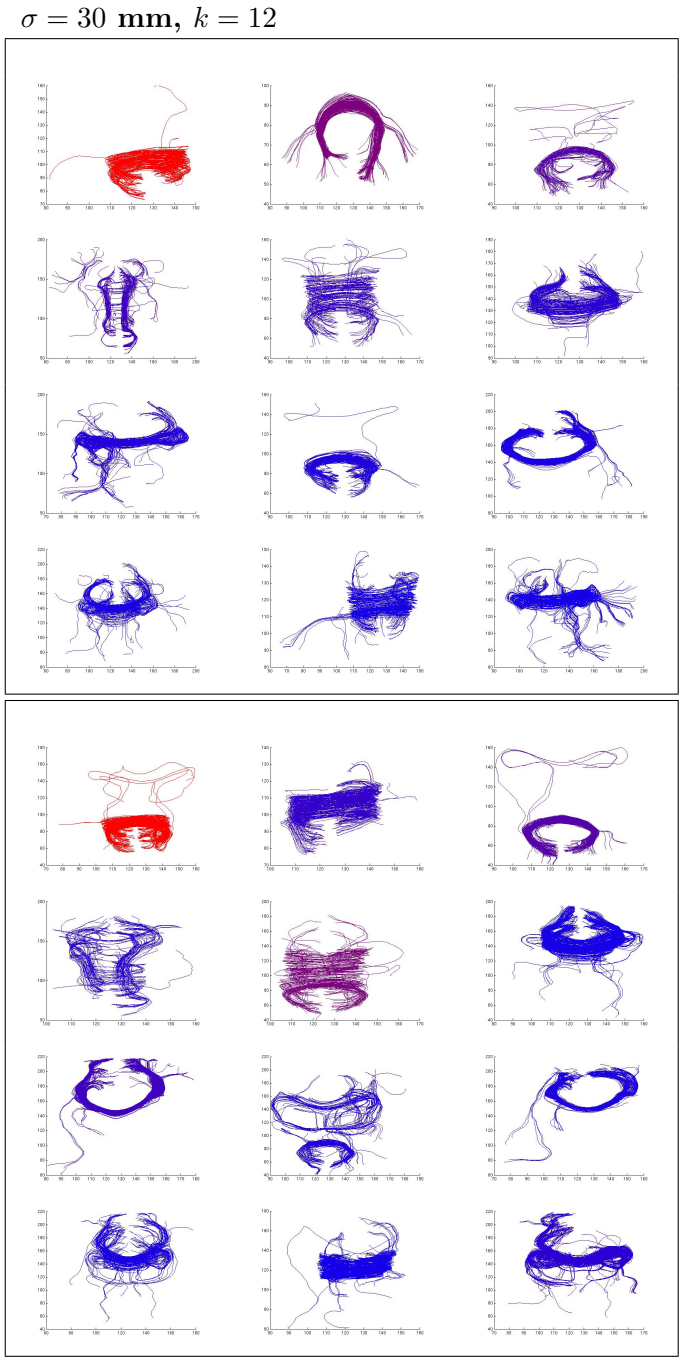


Figure 7.4: Clustering the CC with the **MCP distance** ($\sigma = 30$mm, $k = 12$). The top panel shows the CC of one subject partitioned into $k = 12$ clusters. The bottom panel shows the clustering results for a second subject. There was a correspondence between the bundles for the two subjects shown here, as well as for the data from the 10 other subjects.

$\sigma = 1,\ k = 12$

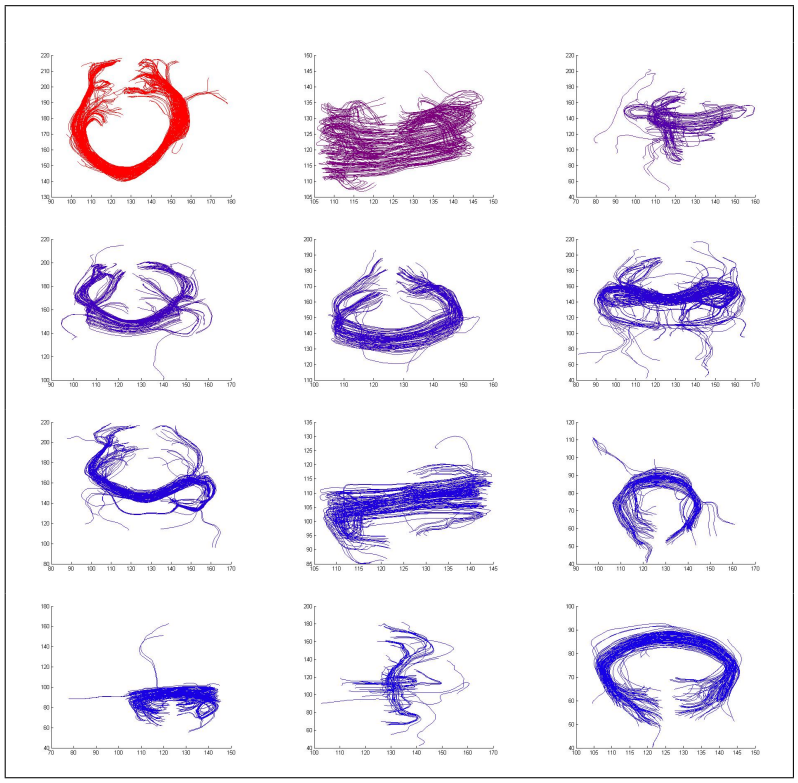


Figure 7.5: Clustering using the **barycenter distance** ($\sigma = 1,\ k = 12$). The CC of one subject is partitioned into $k = 12$ clusters.

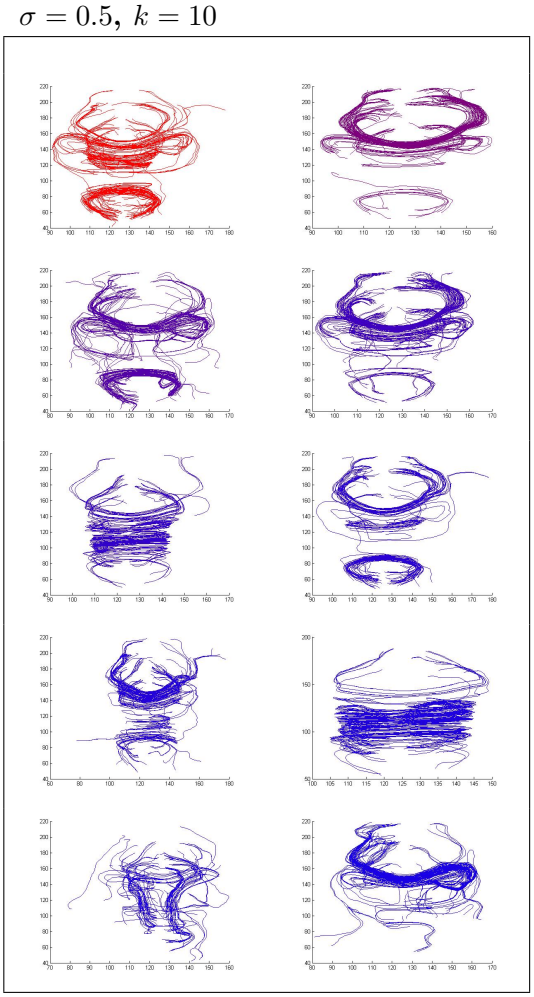
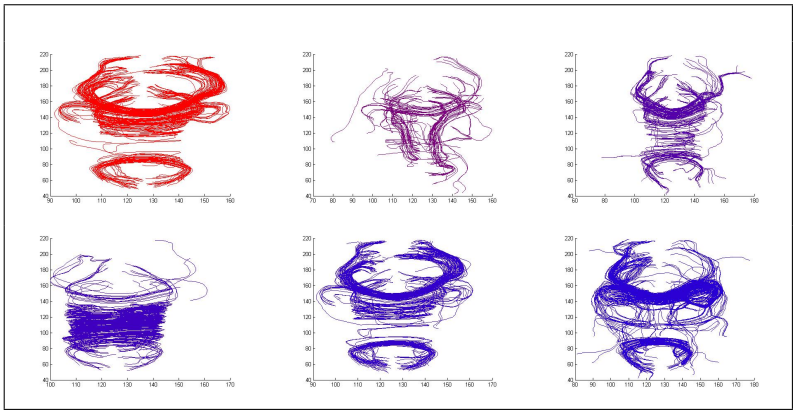


Figure 7.6: Clustering using the **Riemannian shape distance**. The CC of one subject is partitioned into $k = 10$ clusters. CC fiber tracts rendered by DTI tractography have a similar U-shape, so some of the clusters are indistinguishable. Clusters in the first two rows, for instance, are examples of oversegmentation. Only the cingulam bundle (bottom left) has a distinct shape and forms a separate cluster.

$\sigma = 0.5,\ k = 6$



$\sigma = 0.5,\ k = 12$

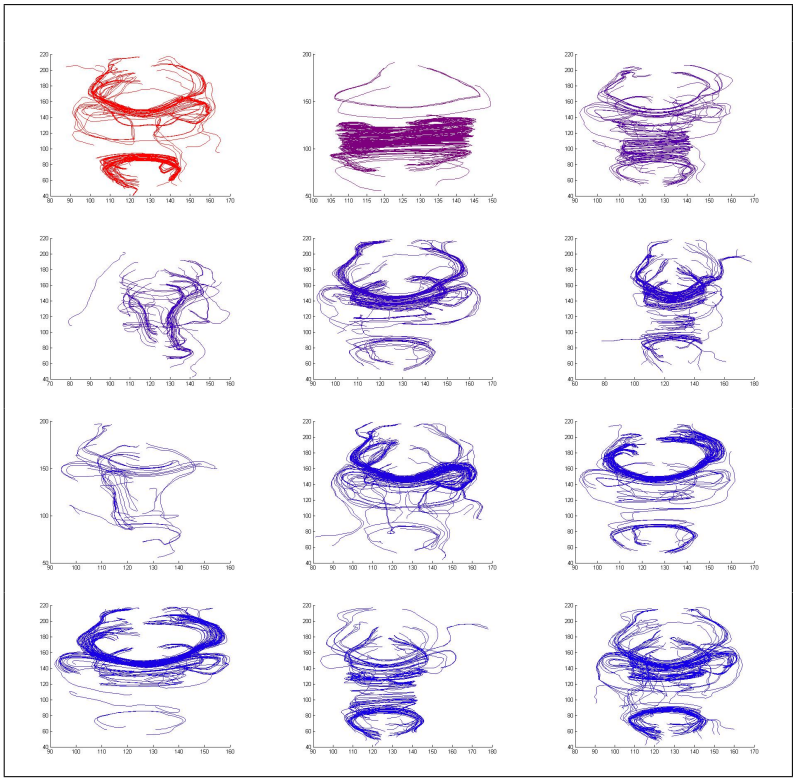


Figure 7.7: More examples of oversegmentation with the **shape distance** for $k = 6$ and $k = 12$.

Chapter 8

Clustering Corpus Callosal Fibers

Each of the distances derived in Chapter 3 are based on a different combination of features and so we expect different outcomes in applications such as clustering. In this chapter, we demonstrate this with fibers of the corpus callosum (CC henceforth), showing how some distances are better suited for clustering a given configuration than others.

8.1 Clustering the Corpus Callosum

The CC is the largest white matter fiber bundle in the brain (Figure 8.1). It is centered along the midsagittal plane with radiations that extend to the prefrontal and frontal cortex in the anterior brain, the sensory-motor cortex in the middle and the parietal and occipital lobes in the posterior brain. The fibers vary not just in topography but also in composition. Thin, lightly myelinated, densely packed fibers predominate in the genu at the anterior end, and to a lesser extent, in the splenium, at posterior end. The mid-corpus, on the other hand, has thicker, longer fibers [64]. Due to this heterogeneity, pathologies such as multiple sclerosis [65] and schizophrenia [66] selectively affect specific sections of the CC and so there is interest in parcellating the different regions for study.

Clustering CC fibers is made difficult due to the lack of obvious landmarks and also because a limitation in DT-MRI tractography [67] renders fibers that are similar in shape and scale along the length of the CC.

8.1.1 Method used in the Experiments

We used spectral clustering for our experiments. The distance functions were Riemannian metrics defined in Chapter 3. For the benefit of the reader, we list these distances below:

1. d_1: the shape, orientation, scale and position distance defined in $\mathcal{S}_1$
2. d_2: the shape, orientation and scale distance defined in $\mathcal{S}_2$
3. d_3: the shape and scale distance defined in $\mathcal{S}_3$

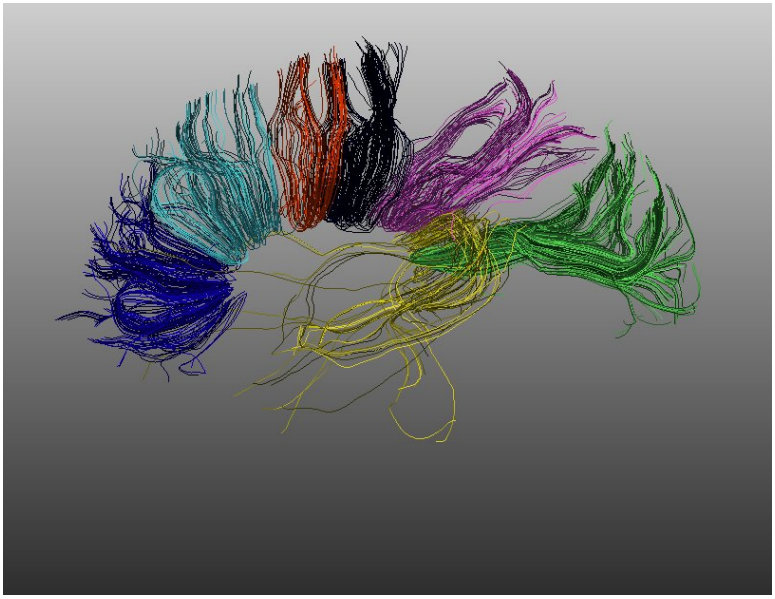


Figure 8.1: A midsagittal view of the corpus callosum. The rostrum and genu (blue), rostral body (cyan), anterior corpus (red), posterior corpus (black), isthmus (magenta), tapetum (yellow) and splenium (green) subdivisions are based on the Witelson [68] classification.

4. d_4: the shape and orientation distance defined in $\mathcal{S}_4$

5. d_5: the shape distance defined in $\mathcal{S}_5$

The CC was rendered by DTI tractography using a second-order Runge-Kutta algorithm. The data came from a healthy male subject in the 20-30 year old age bracket. The partitions were made *a priori* using guidelines from the Wittelson [68] classification (see Figure 8.1).

8.1.2 Clustering Results

It is usually the case that the combined shape, orientation and scale information is richer and gives better discrimination than shape information alone. In the case of the CC, the genu (blue) and splenium (green) are at opposite ends and are easily grouped using the joint shape, orientation and scale metric, d_2 (Figure 8.4). They differ in orientation and scale but have similar shape (see Figures 8.2 and 8.3), so using shape distances, d_5, alone results in poor clustering. The results for shape clustering (using d_5) are shown in Figure 8.4 for comparison.

There are instances, however, where a combination of shape, orientation and scale do not yield the best results. When the arrangement is altered slightly to include a third bundle, the anterior section of the corpus (red), the joint shape and orientation distance, d_4, consistently gives better results than the d_2 metric (Figure 8.5). This can

be explained by the fact that the genu (blue) and corpus (red) have roughly the same scale and tend to cluster together with a d_2 metric. The three bundles, however, are oriented in different directions (approximately $0°, 90°, 180°$ in relation to each other) and so the d_4 metric, which does not include scale, provides better discrimination.

A second example is the clustering of the isthmus (magenta) and the splenium (green) at the posterior end of the CC (Figure 8.6). The scale parameter, similar for the two groups, dominates the clustering and using either the shape and scale, d_3, or the d_2 distance gives poor results. The d_4 metric, which does not include scale information, gives the best results (Figure 8.6).

From these examples, we see that the choice of distance metric can affect clustering results even when small changes in configuration are made. A broader range of problems in white matter fiber analysis can similarly benefit from the judicious selection of distance metrics.

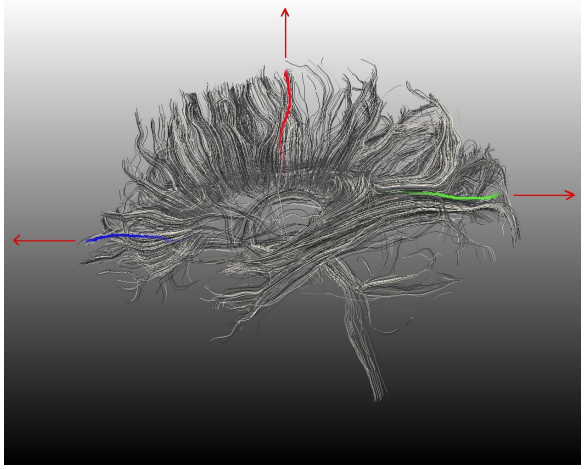


Figure 8.2: The genu (blue), corpus (red) and splenium (green) are oriented in different directions.

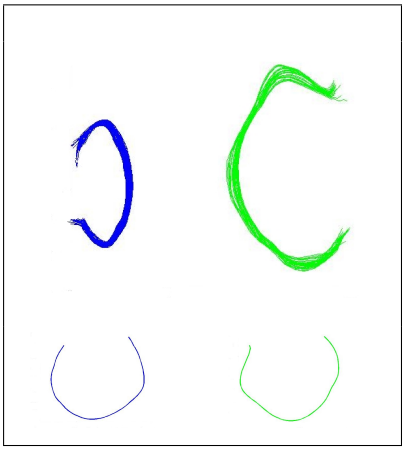
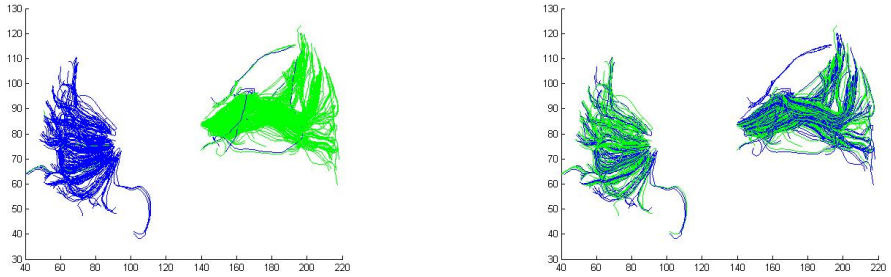
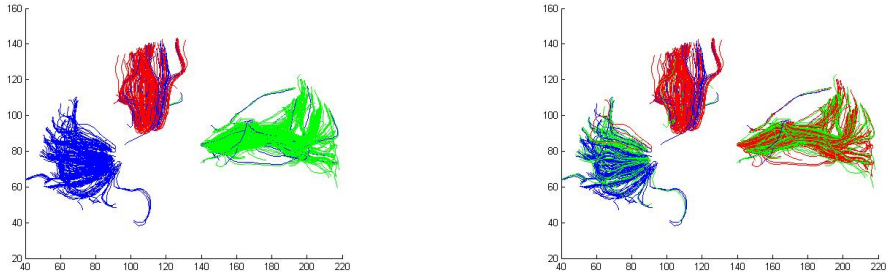


Figure 8.3: The genu (blue) and splenium (green) have different scale (top) but are similar in shape (bottom).



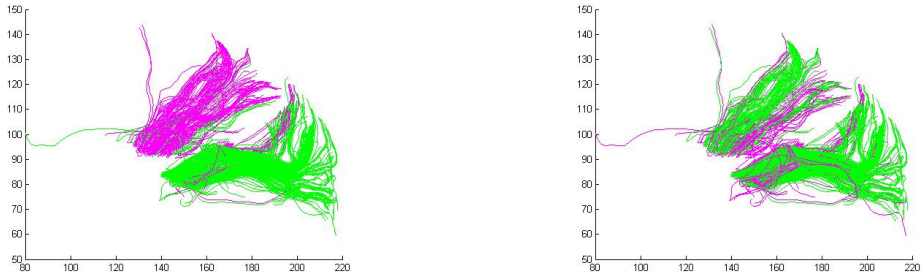
(a) shape+orientation+scale (d_2) (b) shape (d_5)

Figure 8.4: Clustering the genu and splenium at the anterior and posterior of the CC respectively. Here, shape information alone (b) is not adequate for clustering.



(a) shape+orientation (d_4) (b) shape+orientation+scale (d_2)

Figure 8.5: Clustering of the genu, corpus and splenium located in the anterior, middle and posterior sections of the CC. Including the scale information results in poorer clustering (b).



(a) shape+orientation (d_4) (b) shape+scale (d_3)

Figure 8.6: Clustering the isthmus and splenium, both at the posterior end of the CC.

Chapter 9

Statistical Summaries of White Matter Fibers

A key concept in medical image analysis is the idea of a mean template, i.e., a statistical average around which deviations can be assessed. In the context of white matter fibers, we seek to represent an anatomically defined fiber bundle with a mean and variance that describes its essential characteristics. The mean is of interest in this analysis for the following reasons:

1. There are a large number of fibers involved in white matter fiber analysis. The CC has 300 million fibers alone; the whole brain, over 100 billion [67]. The tractography output which is some fraction of this can still be several thousand fibers. Due to this large volume, a practical first step in any study, and one that is advocated by the author, is to compute a representative mean of a fiber bundle.

2. The tractography output is subject to error. Noise, imperfections in the image and the presence of regions of low anisotropy due to fiber crossings all contribute to this. In order to make the streamline output more robust we can average over the fiber bundle. This strategy is also useful when a representative bundle is sought and there are discontinuities and other fiber damage due to disease.

3. To facilitate statistical analysis for population studies where the underlying problem is one of assigning membership to a group.

Means may be computed for the following situations:

(i) for a group of fibers within a fiber bundle

(ii) for an intra- or inter-subject collection of fibers from many bundles

(iii) for an intra- or inter-subject collection of means of fiber bundles

The computation of statistical summaries is usually included in any framework for white matter fiber analysis. These mathematical frameworks were briefly summarized in Chapter 6.3.3. Maddah et al. [49, 53], for instance, compute clusters and their mean

trajectories as follows: A fiber trajectory is treated as an ordered set of points, each of which is a cluster center. Points from other fiber trajectories, similarly sampled, are assigned to these cluster centers. The mean curve can be reconstructed from these updated cluster centers. This is one example of a mean computation.

In the case of the manifolds defined in Chapter 3, Riemannian structures on these spaces enable us to compute means and variances and thus construct statistical models that capture the variability of the white matter fiber configurations under study. In this chapter we show how means are computed in a nonlinear Riemannian manifold. A procedure for computing the means of fiber bundles is described in Section 9.1.1 The mean for a given fiber bundle in the shape manifold, $\mathcal{S}_5$, is the mean shape. The means computed in $\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{S}_3$, and $\mathcal{S}_4$ can similarly be described by the combination of features used. We compute means for some of these manifolds. These are then used in analyses such as the study of morphological changes in the CC in Chapter 10.

9.1 Means in a Nonlinear Manifold

In a Euclidean space, for a set of k points, $x_1, x_2 \dots x_k$, the sample mean is:

$$\bar{x}_k = \frac{1}{k} \sum_{i=1}^{k} x_i, \qquad x_i \in \mathbb{R}^n. \tag{9.1}$$

In a nonlinear manifold, a simple summation is no longer possible. We can, however, make an *extrinsic* computation by embedding the manifold in a vector space, computing the Euclidean mean and projecting the result back into the manifold. A disadvantage is that the mean computed depends on the choice of embedding.

A second possibility is an *intrinsic* computation, i.e., one where we use *intrinsic* manifold computations to compute the mean.

9.1.1 Karcher Means of Fiber Bundles

To compute an intrinsic mean within a manifold, $\mathcal{S}$, we use the concept of the mean as the *centroid of a density*. This idea was put forward by Fréchet [69] to calculate means in a Riemanniann manifold. The computation involved a minimization but the existence and uniqueness of the resulting mean could not be guaranteed (see [70] for details). Karcher's proposal that a local instead of a global mean be used [71], led to a practical implementation. We shall henceforth refer to this local mean as the Karcher Mean.

For a given collection of fibers $\{\beta_1, \beta_2, \dots, \beta_n\}$, with function representations $\{f_1, f_2, \dots, f_n\}$, the Karcher mean is defined as:

$$\bar{\mu}_n = \underset{f \in \mathcal{S}}{\operatorname{argmin}} \sum_{i=1}^{n} d_s(f, f_i)^2 \ . \tag{9.2}$$

A gradient approach such as the gradient descent algorithm we employ, is used to find the local minimum of the cost function. The steps for the shape manifold, $\mathcal{S}_5$, are outlined below.

1. **Initial Condition**: In the shape manifold, $\mathcal{S}_5$, we represent the fibers with the SRVF (i.e the q-function) (see Chapter 3, Eqn. 3.1). Since the q_i's are elements of a hypersphere, we can use their *extrinsic mean* in $(\mathbb{L}^2[0,1])^3$ to initialize the gradient algorithm. We use the classical average:

$$\mu(t) = \frac{1}{n}\sum_{i=1}^{n} q_i(t), \quad t \in [0,1] \ ,$$

   followed by the normalization $\mu \mapsto \mu/\|\mu\|$, as the initial value for the gradient search.

2. **Iterative Update**: Let μ be the current estimate of the Karcher mean. We find the tangent vectors, $v_i \in T_\mu(\mathcal{S})$, $i = 1, 2, \ldots, n$, such that the geodesic starting from μ with the initial velocity v_i, reaches q_i in unit time. (In practice, we first find the optimal rotations and reparameterizations of q_i to best match μ, and then use the inverse exponential map to compute v_is.) Next, the average direction, $\bar{v} = \frac{1}{n}\sum_{i=1}^{n} v_i$, is computed. The average vector $\bar{v}$ is actually the negative gradient of the cost function given in Eqn. 9.2. We update μ with the Riemannian exponential map, $\exp_\mu(\epsilon\bar{v})$, which provides a means of travel on the manifold. Here, ϵ is a small positive number. Repeated application of these steps leads to a local minimizer of cost function. A local minimum is found by repeating these steps.

Since the gradient method does not guarantee a globally optimal solution, the final result is dependent on the starting condition. A heterogeneous set of fibers will give different means for different starting points. The shapes of white matter fiber bundles, however, are anatomically determined by the regions they connect and occupy. They are thus homogeneous and computing the Karcher mean of a fiber bundle does not usually give inconsistent results.

9.1.2 Karcher Means for the $\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{S}_3$, and $\mathcal{S}_4$ Manifolds

The explicit exponential and inverse exponential maps in the $\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{S}_3$, and $\mathcal{S}_4$ manifolds are different, but the Karcher means can otherwise be handled similarly. In Figures 9.1 and 9.2, we show the Karcher means for a genu and splenium bundle respectively. These computations were done in the *shape*, $\mathcal{S}_5$, and *shape+orientation*, $\mathcal{S}_4$, spaces. (The mean curves in these two manifolds are similar because the individual fibers do not vary greatly in orientation in this small sample set.)

9.1.3 Statistical Modeling of Fibers

We can construct parametric and nonparametric statistical models for different classes of fiber bundles. We first need to adapt the probability models to a nonlinear mani-

fold. Using hypothesis tests such as the likelihood ratio test, we can then answer such questions as whether a fiber bundle belongs to a population.

9.2 Measures of Dispersion

The Karcher mean is the local minimum of the variance function [70, 69] and so computing the mean of a fiber bundle also gives us a value for the variance. The variance function $V : \mathcal{S} \to \mathbb{R}$ is defined as

$$V(\mu) = \frac{1}{n} \sum_{i=1}^{n} d_s(\mu, q_i)^2 \ . \tag{9.3}$$

The iterative gradient procedure converges to the Karcher mean which can then be used to calculate the geodesic lengths in the variance computation.

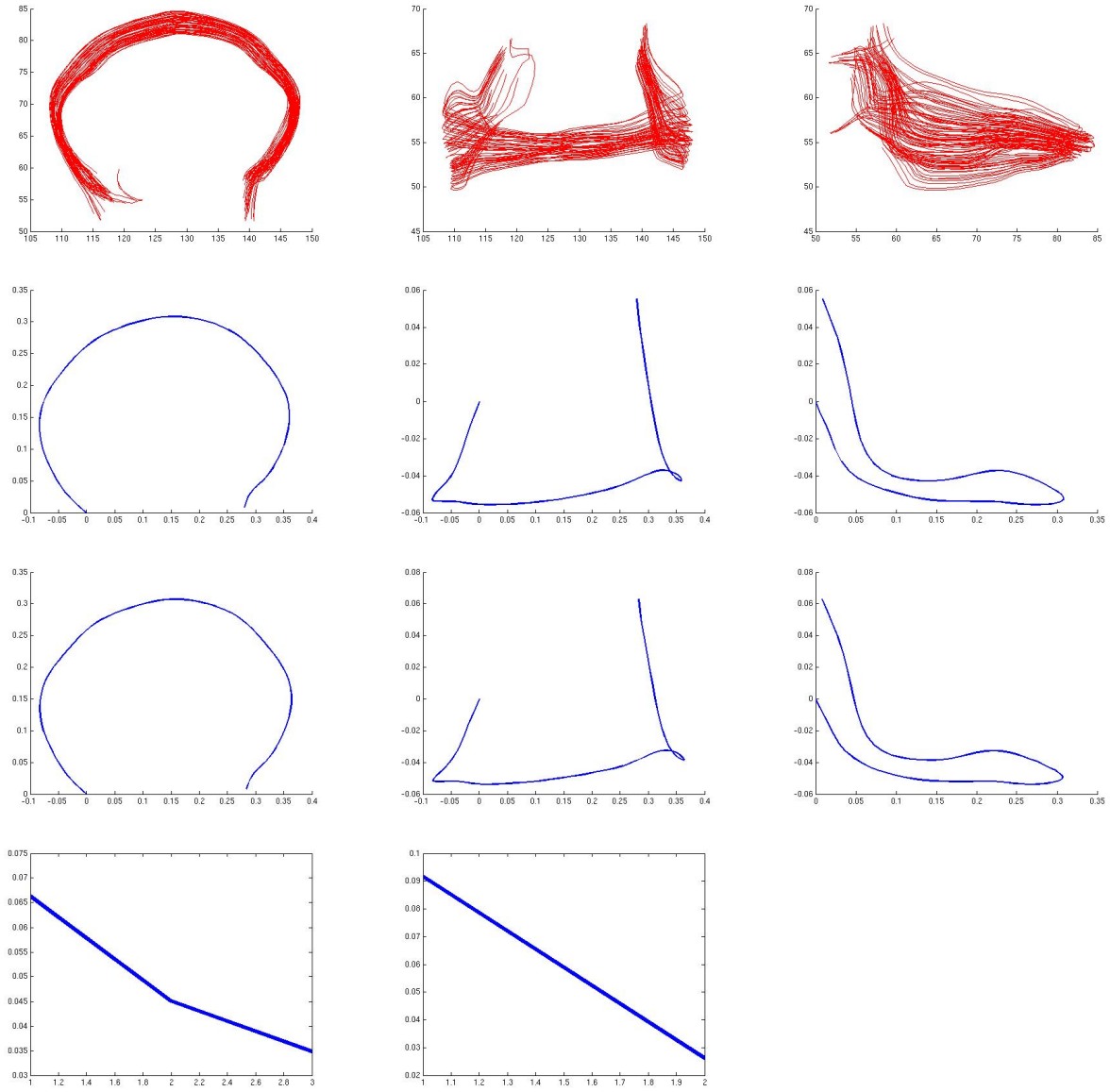


Figure 9.1: Top row: axial, coronal and sagittal views of a fiber bundle extracted from the **genu**; second row: mean curve in the *shape* manifold; third row: mean curve in the *shape+orientation* manifold; bottom row: decrease in cost function using gradient iterations for the *shape*(left) and *shape+orientation*(right) manifolds.

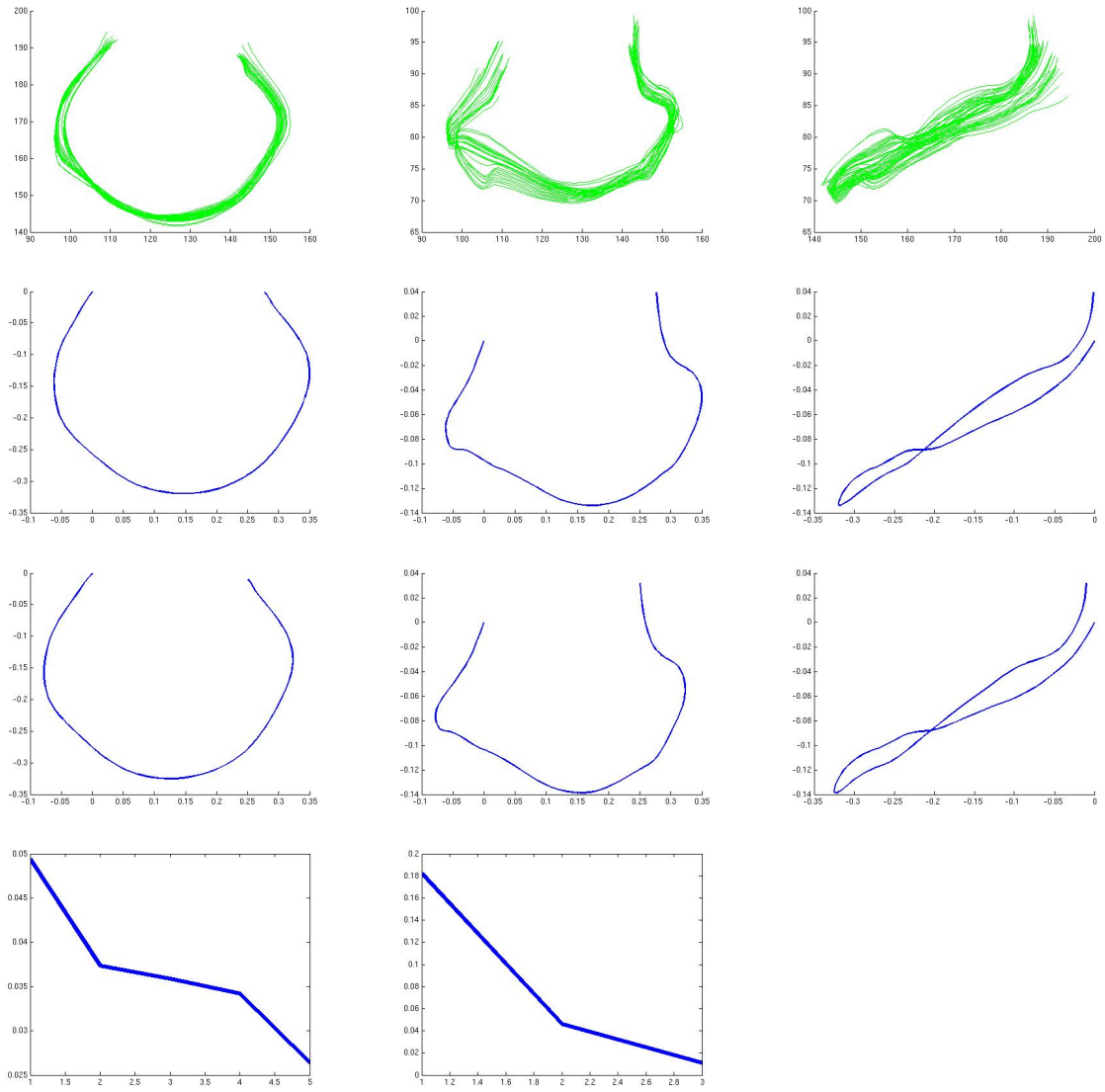


Figure 9.2: Top row: axial, coronal and sagittal views of a fiber bundle extracted from the **splenium**; second row: mean curve in the *shape* manifold; third row: mean curve in the *shape+orientation* manifold; bottom row: decrease in cost function using gradient iterations for the *shape*(left) and *shape+orientation*(right) manifolds.

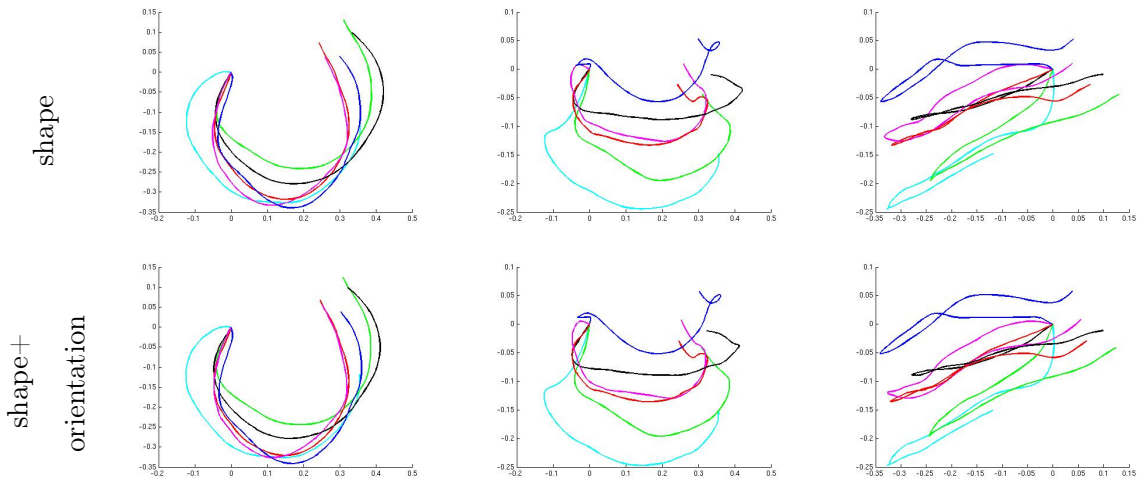


Figure 9.3: **Mean curves of the mid-splenium for 6 control subjects.** These are used to construct an atlas. Top row: *shape* manifold (axial, coronal, sagittal views); bottom row: *shape+orientation* manifold. In both manifolds, the curves are scaled to unit length so registration beyond the initial alignment to a common reference frame is not required.

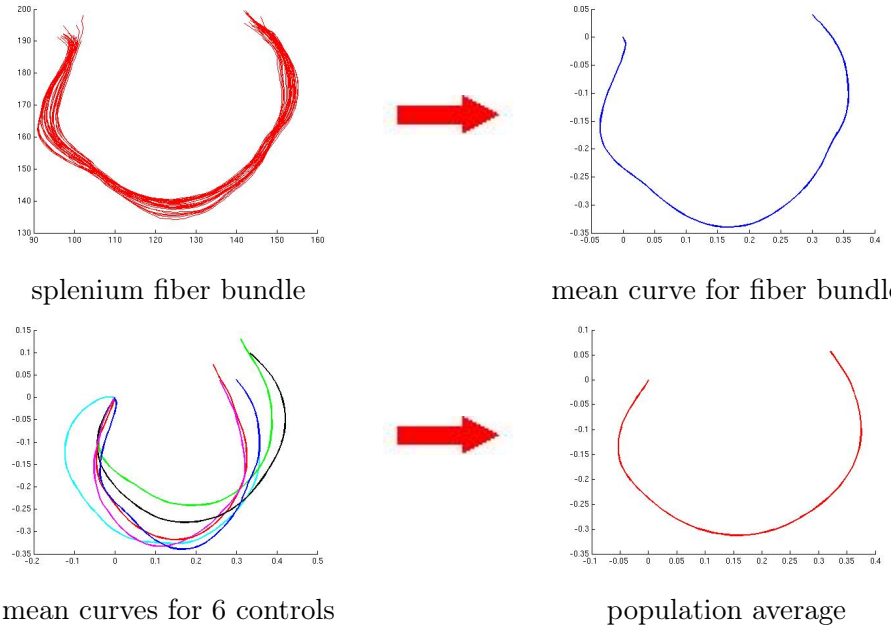


Figure 9.4: **Computing population means for fiber bundles.** Top row: the mean curve for a single splenium fiber bundle. Bottom row: Mean curves for 6 subjects are computed. The mean of these is the population average curve for the splenium.

Chapter 10

Morphological Changes in the Corpus Callosum

The corpus callosum (CC) connects the two brain hemispheres and thus plays a vital role in interhemispheric communication. Consequently, changes to its shape or structure, which occur because of aging [72, 73] or degenerative disease [74, 75, 76], is the subject of active study. There is an interest in linking these physical changes to neurological impairment [77]. There is also an interest in linking the different stages of a degenerative disease such as multiple sclerosis (MS), schizophrenia or Alzheimer, with physical alterations [78, 79].

One of the underlying mechanisms for these topographical changes is atrophy. In the case of MS, a degenerative disease associated with atrophy, whole brain grey matter volume loss occurs even in the earliest stages of the disease. What is not as well understood is the nature and extent of white matter volume loss [80]. The CC is an ideal brain structure to study atrophy in MS subjects. It has the largest concentration of white matter tracts and a high prevalence of compacted fibers and we expect it to be disproportionately affected by the mechanisms that result in atrophy.

The study of morphological changes in the CC, which falls under the purview of shape analysis, can be categorized as volume-based or tract-based. With volume-based analysis, two broad approaches have been followed in the literature. The first approach has been to present the CC as an abstraction using a geometric representation such as a fixed topology skeleton or a medial axis description. Classification using a linear discriminant method such as support vector machine (SVM) is then performed [81, 82]. In other studies, different principal component analysis (PCA) techniques have been used to study the modes of local and global variation [83]. The second set of approaches, more common in clinical studies, has been to divide the CC into well-defined partitions and to analyze quantitive parameters such as the width (dorsal-to-ventral) and thickness (anterior-to-posterior) [78, 84] of these sections.

With tract-based analysis, the research emphasis has been on the evaluation of microstructure functions such as fractional anisotropy (FA) and mean diffusivity (MD) along the lengths of tracts. Correspondence and averaging methods have been applied

to preserve the FA function for bundles and atlases of prominent CC tracts like the genu and splenium [45, 48]. Classification based on the geometrical shapes of fiber tracts has received less attention. Adluru et al. [85], in a recent paper, did a classification based on the shape histograms of fiber tracts but this is the only example of a tract-based shape study we know of to date.

In this chapter, we present new tract-based methods to assess fiber integrity. We use the Riemannian metrics described in Chapter 3 to look at changes in the shape and alignment of fibers or tracts and through a case-control study of MS subjects, demonstrate the feasibility of our methods. In particular, we look at the anterior CC which when viewed along the midsagittal plane forms a deep arc. We expect that atrophy would result in a straightening of this arc as the results in [83] suggest. We also expect that the distances computed in the *shape* and *shape+orientation* manifold, when considered together, are well suited to detect such changes.

10.0.1 Data

To study shape changes in the genu, DTI data from 20 subjects, 10 normal controls (NC) and 10 MS patients were used. The MS subjects had a mean age of 44 ± 12 years and the mean disease duration was 10 ± 10 years. Registration of the data set is not required (other than an alignment to a common coordinate system) but was performed in this case. To obtain the DTI fiber tracts, anatomical regions corresponding to the genu and rostrum were seeded and streamline tractography with fourth order Runge-Kutta integration was performed. A few fibers from the posterior-most section of the genu and another small set from the mid-section of the genu, were used to compute two average curves, one for each section. The Karcher mean algorithm, described in Chapter 9, was used for this. We shall refer to these two curves, shown in Figure 10.1, as the lower curve (LC) and the upper curve (UC) respectively. A total of 40 curves (2 curves/subject $\times$ 10 subjects/group $\times$ 2 groups) were used in this study. To maintain consistency, the entire dataset was post-processed by a single individual over a short period of time.

10.0.2 Distance Maps

Distance matrices of the NC and MS populations are used to look at LC-LC, UC-UC and LC-UC distances. The layout of such a map is shown in Figure 10.2. The first quadrant shows LC-LC distances of the 10 LC fibers in the group; the fourth quadrant shows the UC-UC distances of the 10 UC fibers; the second and third quadrants show the LC-UC distances between 20 fibers. We represent both *shape* and *shape+orientation* distances in this way. The distances have a common scale so comparisons can easily be made.

Distance maps are an efficient way to make intra- and interpopulation comparisons. We can study 3 distribution profiles–the LC-LC, UC-UC and LC-UC distances–within the NC population, the MS population and also between the NC-MS population. In principle, this gives us 9 ($3 \times 3 = 9$) distributions to analyze.

Since we can make the same 9 comparisons for each of the distance measures used, for the two distances (*shape* and *shape+orientation*) that we use in this chapter, there are 18 ($9 + 9 = 18$) distance distributions to study. We may also study intrapopulation differences for the LC-LC, UC-UC and LC-UC distances between the distance maps. In total, comparing the intra- and interpopulation differences between the *shape* and *shape+orientation* distances gives us 24 ($9 + 9 + 3 + 3 = 24$) distributions to study.

10.0.2.1 Shape Distances

The intrapopulation shape differences are small for the LC-LC and UC-UC comparisons. The LC-UC distances are only slightly greater. This pattern is observed in both the control and MS populations with the MS group having slightly larger differences. This suggests that the shapes of the two curves within a population are similar, with the MS population showing more variation. The variances in Table 10.1 computed for the shape manifold further support this visual summary.

10.0.2.2 Shape+orientation Distances

The *shape* distances are a baseline. With the additive orientation distances in the *shape+orientation* manifold, we expect the LC-UC distances to be greater since the two curves are differently oriented. This is what we observe in the control group with large distances in the second and third quadrants (see Figure 10.4). The striking visual showing the second quadrant differences between the two manifolds is further supported by the Wilcoxon signed-rank statistic. This is a nonparametric hypothesis test between paired dependent samples. The p-value of 0 that we get for both the NC and MS populations strongly indicates that the median distances are different for the two manifolds.

Figure 10.4 also shows that the second quadrant LC-UC distances in the *shape+orientation* manifold are both smaller and more variable for MS patients. (The LC-UC distances for each individual subject are plotted in Figure 10.5). The LC-UC variance for the MS subjects in Table 10.1 in fact shows that they are the most variable group. Because of the heterogeneity in variance between the NC and MS groups, a nonparametric rank-based test to compare populations was not performed.

Overall, the data suggests that the relative LC and UC orientation differences become smaller for the MS group as the disease progresses.

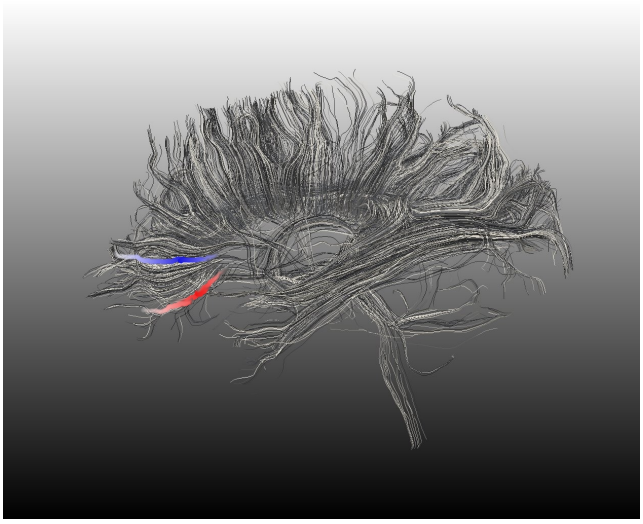


Figure 10.1: **Two curves, the LC in red and the UC in blue,** are used to determine changes in the curvature of the anterior CC. The LC is extracted from the mid-rostrum where the tractography is more reliable. The UC is extracted from the mid-genu.

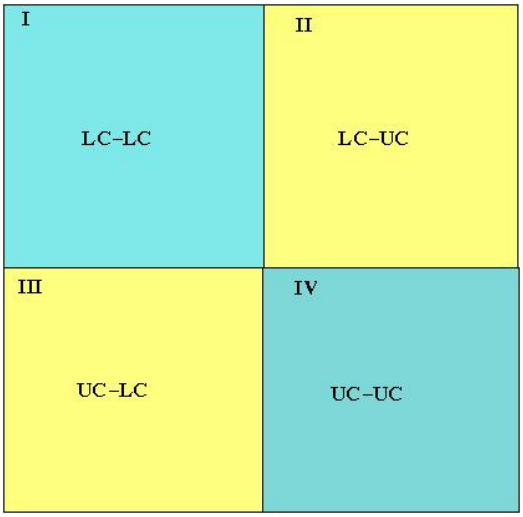
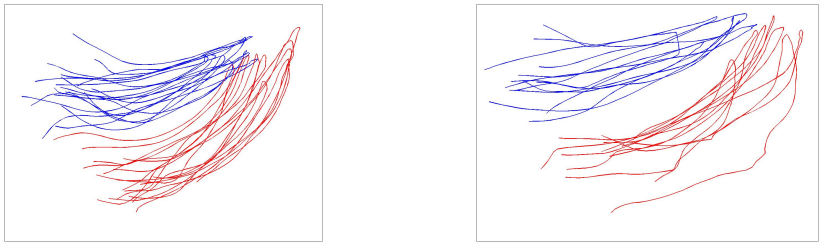


Figure 10.2: **Distance map layout.** Quadrant I represents the pair-wise distances of the 10 LCs. Same subject distances fall on the diagonal and are zero. The colormap variation shows intra-population variability. Quadrant IV, likewise, displays the UC distances. Quadrants II/III show the LC-UC/UC-LC distances. Each row displays the distance between the LC of one subject and the 10 UCs.



(a) NC (b) MS

Figure 10.3: **Raw curve data.** (a) Population of NC curves with the LC in red and UC in blue. (b) Curves from MS subjects.

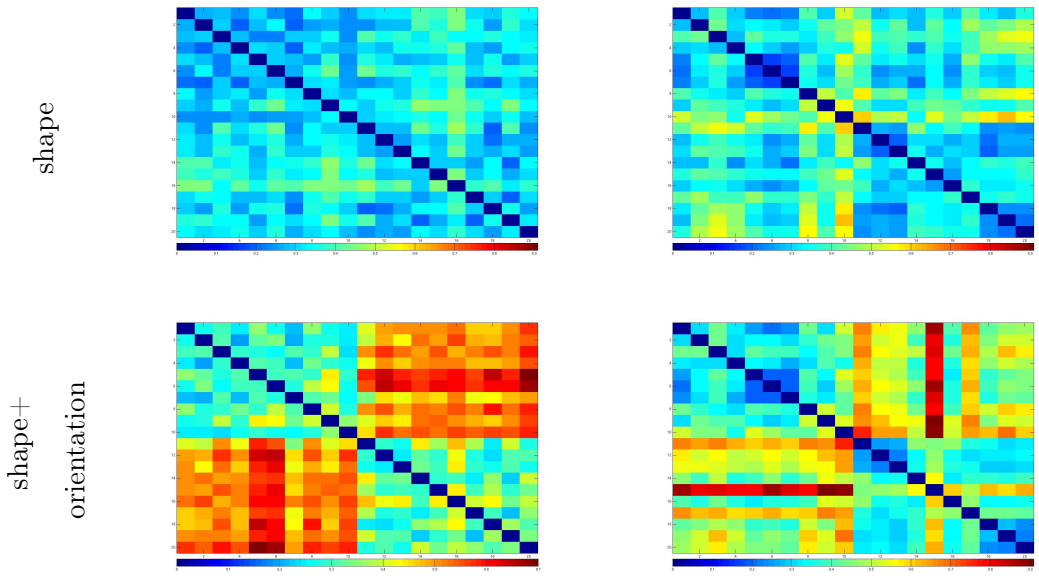


Figure 10.4: **Distance maps.** The LC-LC, UC-UC and LC-UC *shape* differences are small for the NC and MS populations. The LC-UC *shape+orientation* distances are large (quadrants II & III are orange) for the NC population. These distances are both smaller and more variable, by contrast, in the MS group. A common color scale enables us to compare groups.

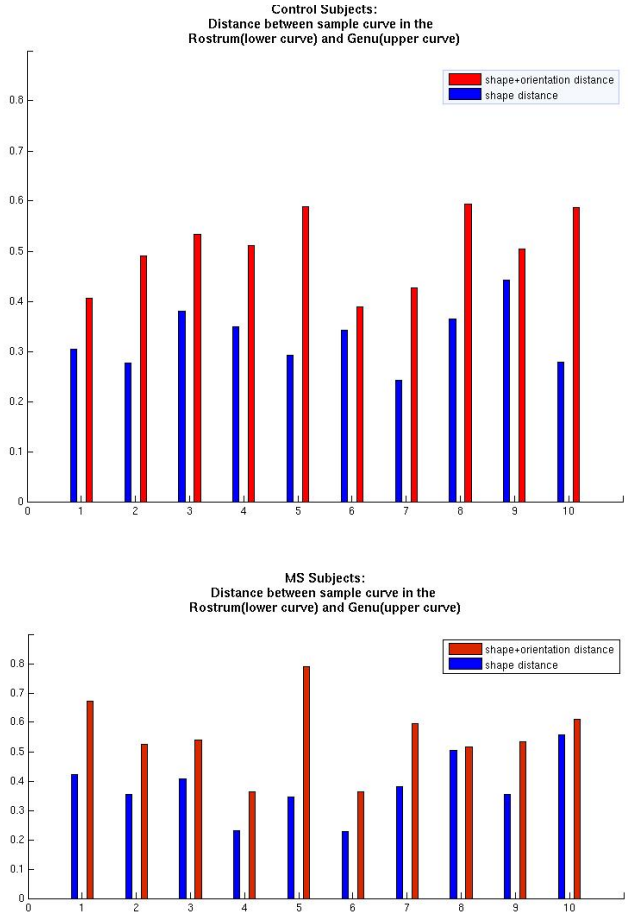


Figure 10.5: Distance between the LC and the corresponding UC in 10 subjects. The *shape+orientation* distance (red) is greater than the baseline *shape* distance (blue) as expected. The difference in these two distances is greater in the control population (top) and more consistent. The MS population (bottom) is more variable.

Table 10.1: Within-group variance for distance distributions

	NC			*MS*		
	LC-LC	**UC-UC**	**LC-UC**	**LC-LC**	**UC-UC**	**LC-UC**
shape	0.0026	0.0048	0.0030	0.0089	0.0039	0.0089
shape+orientation	0.0032	0.0043	0.0040	0.0080	0.0108	0.0176

† Distances between the same curve (d = 0) were not included in the variance computation

10.0.3 Discussion/Contribution

Distance measures betweeen two fibers, visually represented as distance matrices or maps, can be used for both intra- and interpopulation comparisons. In a comparative study of NC and MS subjects, the *shape+orientation* distance color maps have different signatures for the two populations with the MS group showing more variability.

This variability is suggestive of alterations to callosal shape that accompany illness progression. The genu curvature may either be getting more pronounced or straightening out. PCA models of shape changes in the CC indicate that the CC flattens out [83]. We can confirm this by designing an experiment with 3 curves from 3 different locations in the anterior CC. More expensive longitudinal studies that follow a patient over time, can also verify this.

The Riemannian framework can be extended to include additional features such as a fractional anisotropy (FA) or mean diffusivity (MD) function along the fiber [7]. Incorporating this information into our analysis will provide interesting insights into the links between shape changes, atrophy and FA.

Using only 10 MS subjects and 10 NC for comparison, we have demonstrated an effective new design that uses *shape* distances in tandem with *shape+orientation* distances to study shape, morphology and changes to these. Our feasibility study needs to be repeated on a larger sample matched for age, gender and handedness, factors that are known to affect callosal size and shape [68, 72]. The broader context of this work is to develop methods to identify and track progressive white matter disease. We envision that tools such as this will improve the clinical evaluation of patients suffering from MS and other similar kinds of degenerative disease.

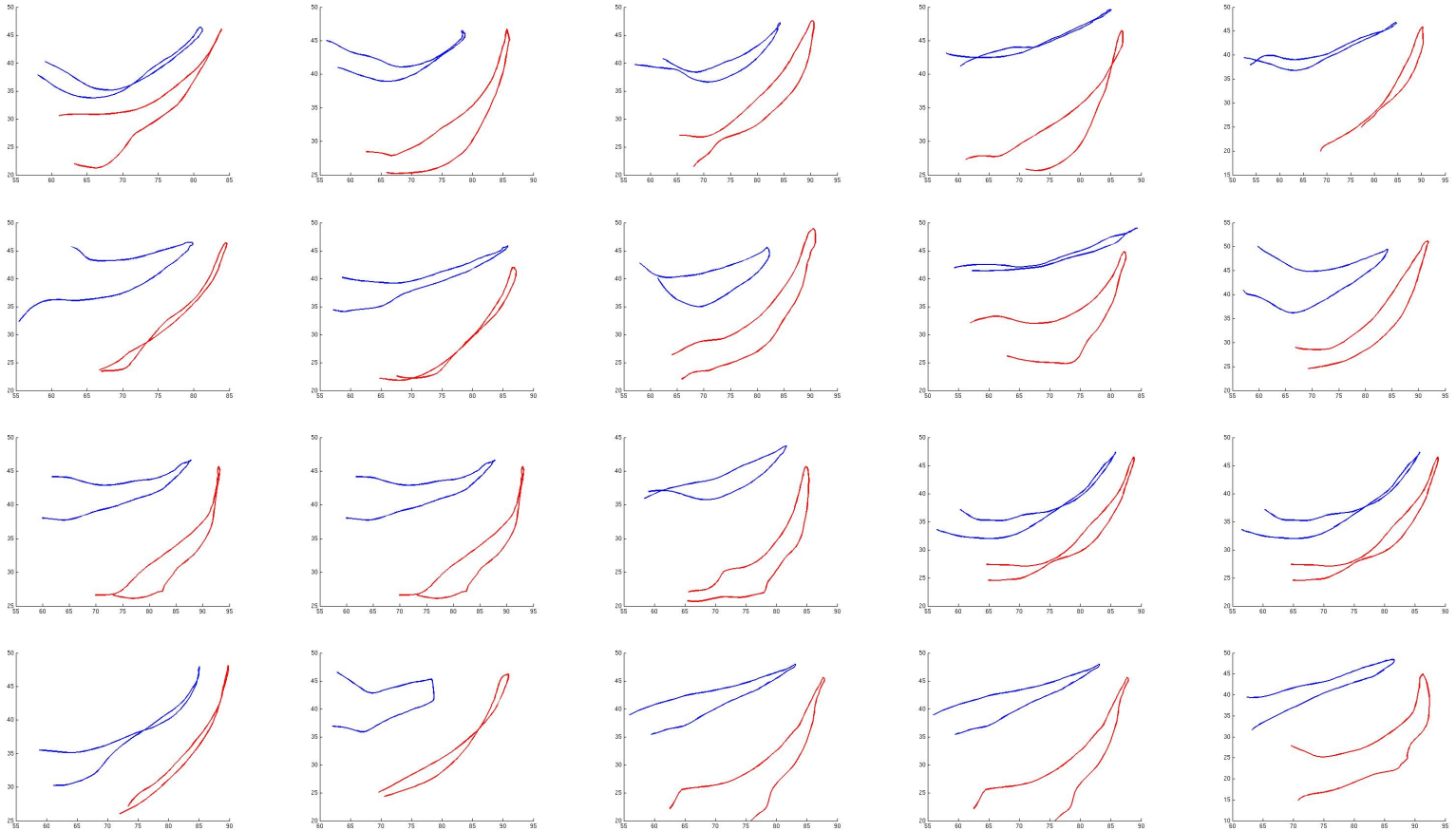


Figure 10.6: **Raw curve data for each subject.** The UC and LC for NC subjects (rows 1 & 2) and MS patients (rows 3 & 4).

Chapter 11

Conclusion

We have explored the analysis of open curves in two very different contexts, sulci and white matter fibers. Sulci are few and variable. DTI fibers, on the other hand, are numerous but somewhat homogeneous within a fiber bundle.

A hard problem in sulcal analysis is labeling. The difficulty expert neuroanatomists sometimes have with this task sets the upper parameters for the performance we can expect from automatic labeling. We address the automatic labeling problem by first limiting the scope of the problem to the primary sulci. This now becomes a well-defined problem for which we give a simple, lightweight, easy-to-implement solution. The insight we offer is that solutions can be found by treating this as a graph pattern matching problem.

One difficult problem in white matter fiber analysis is the practical matter of efficiently handling the large volume of fibers (obtained from DTI or HARDI). This may be achieved by clustering the fibers, by clustering using approximations such as the Nyström method, by sampling from the data in a fiber bundle or computing representative means as we do in Chapter 10. Volume-based methods side-step this issue by treating a white matter structure as a single anatomical unit. With tract-based analysis, however, processing individual fibers is usually a requirement. Tractography and clustering algorithms have made this possible. Improvements are needed but it is my view that making the pre-processing pipeline more efficient is an engineering effort.

While tedious processing is a disadvantage, tract-based methods offer the potential to study local parameters along a tract. This is important in the study of white matter disease and, at this juncture, it is here that the maximum contributions to the medical community can be made.

With this in mind, we introduce a new set of methods for the geometrical analysis of fiber tracts. These tools extend the shape analysis of fiber tracts to a more inclusive analysis: that of combinations of the four physical features–shape, orientation, scale and position– associated with a fiber when it is modeled as a curve. We motivate the use of these tools with two examples: In the first case, we show how we can improve clustering results by selecting appropriate distance metrics for the different problem configurations presented. We can customize solutions in this way because we work with

a wider selection of tools.

In the second example, we convincingly show, through a clinical example, that these metrics used in concert with each other, can detect physical changes due to disease (MS in this case). Analysis of this problem has been limited to PCA or some variant of it. The distance metrics we use offer different ways to measure changes and the experiments we design around them are a way to look at the problem anew. In fact, this may be the first example of geometrical analysis of this kind in medical image analysis. With extensive experiments this method could potentially be used to calibrate the stages of disease. This will enable physicians to target treatment more effectively.

In summary, the contributions of this thesis are as follows:

1. Graph pattern matching is an important strategy for labeling sulci. This insight takes its inspiration from manual labeling where the relative placement of sulci are cues for correct identification. The relational pattern design compensates for the variability in sulcal patterns across subjects. We can identify both normal sulci (tested for 10 classes of primary sulci) and displaced sulci to a high degree of accuracy. Furthermore, as a secondary benefit, atlases may be constructed without registering the sulci to a common template.

2. The graph pattern matching solution implemented as an out-of-sample MDS procedure, is fast, light-weight and easy to put-together.

3. We evaluate the use of combinations of the four physical features–shape, orientation, scale and position– for sulcal and DTI fiber applications.

4. Several distance metrics have been proposed for clustering white matter tracts. The Mean Closest Point distance has been mentioned in some papers as the distance of choice. We propose a *barycentric* distance for parcellating the CC. This gives cleaner clusters based on visual evaluation.

5. With the computation of statistical summaries, we demonstrate two things: (i) how to compute representative means of fiber bundles and (ii) how means might be used in the statistical analysis of a population. In the first case, the mean is a preliminary step which makes subsequent analysis easy. It can be used in different kinds of analyses such as the problem, in Chapter 10, where we compare two regions of the CC by looking at their mean curves.

6. The changes in shape that the corpus callosum undergoes as disease progresses is a problem that lends itself to geometrical analysis. We identify this problem, introduce new mathematical methods to study this, design a method–comparing curves from different sections of the CC–to detect these changes and finally present the results through effective visualization.

Appendix A

Clustering of Sulci

We have made extensive use of clustering to organize DTI fibers in our projects. In this appendix, we report on our efforts to cluster sulci.

A.1 Clustering Applied to the Sulcal Labeling Problem

Clustering algorithms are optimized to group *like* data. As applied to the problem of labeling sulci, we wish to partition a database of sulci into categories which can then be labeled. The specific task we are interested in is to cluster 180 sulci from 18 subjects into 10 classes (this is the same database used in Chapter 5). We use spectral clustering, described in Chapter 7, for this purpose. The primary motive for these experiments is to compare the performance of a simple clustering procedure to the labeling method we have devised.

A.1.1 Clustering with the Position Distance

As we saw in Chapter 5, the *position* distance is the most useful feature for labeling sulci using the out-of-sample MDS and nearest neighbor procedure. We compared clustering performance with this labeling algorithm. The *midpoint* distance, defined in Section 5.2 was used in both cases.

Figure A.1 displays the clustering results on the left (a). The columns of the confusion matrix correspond to known classes; the rows, the cluster results. This is the convention we follow in the rest of this chapter. 121/180 (67%) sulci were correctly grouped into clusters that corresponded with their known categories. The results of the labeling method are shown on the right (b). 163/180 (91%) sulci were correctly classified in that case.

Clustering uses distance information whereas our labeling algorithm uses both distance and structural pattern information. We see in the figure (a), that the left postcentral sulci (second column) were classified as the central suclus (first row) which is adjacent to it; some of the right central sulci (sixth column) were identified as the right postcentral sulcus (seventh row); the right postcentral sulcus (seventh column) was also classified as the left postcentral sulcus (second row). In another experiment, we were

unable to separate the post and precentral sulci, both of which are equidistant to the central sulcus but on opposite sides of it. In general, the clustering algorithm had greater difficulty in separating sulci that are oriented in the same direction.

A.1.2 Clustering with Riemannian distance metrics

In Chapter 5, we had evaluated the labeling performance for each of the five features– shape, mean depth, orientation, length/scale and position– individually. Position, as noted in Section A.1.1, gave the best results. To gain a better understanding of the interactions between shape, scale, orientation and position, we used the Riemannian distance metrics (defined in Chapter 3) to cluster the sulci. The distances are:

1. Shape + orientation + scale + position distance: d_1
2. Shape + orientation + scale distance: d_2
3. Shape + scale distance: d_3
4. Shape + orientation distance: d_4
5. Shape distance: d_5

For each of these five distances, we constructed a distance matrix. The effect of the scale parameter, σ, on the clustering results was tested over several runs, while k, the number of clusters, was fixed at 10. The clustering results for the best of these runs (shown in Figure A.2) are discussed below:

1. **Shape distance:** The shape of sulci are oftentimes indeterminate. The central sulcus is the most consistent of the sulci and we see marginally better results for this class (row 1; columns 1 and 6). The left and right central sulcus (columns 1 and 6 respectively) cluster in one group (row 1) since orientation is not preserved in the shape information.

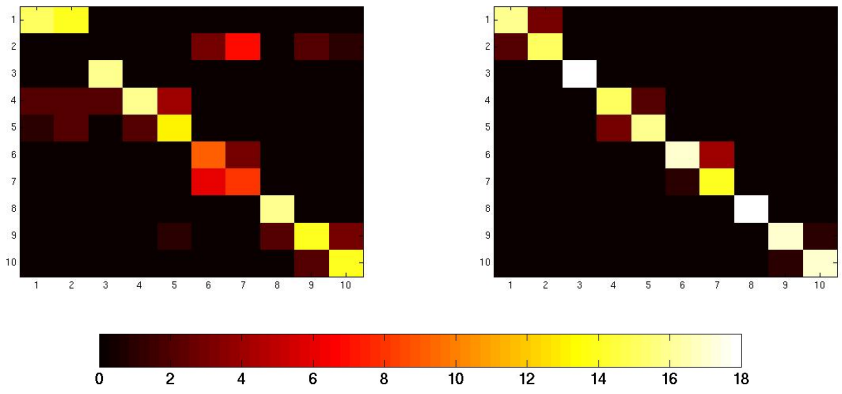
2. **Shape + scale distance:** The addition of scale information does not enhance the clustering results. The distribution of the lengths for each class of sulci for the unregistered, raw data, can be seen in Figure 4.3(a). We see that only the central sulcus, which has a narrow length distribution, can be reasonably identified by the combined shape and length parameter.

3. **Shape + orientation + scale distance:** When orientation is also included in the distance metric, the results improve. Once again, the results are better for the central sulci, which due to the orientation information, now separate into left and right clusters.

4. **Shape + orientation distance:** When the scale is removed, leaving us with just shape and orientation information, the results improve further. We have already seen (in 2 above) that adding scale to the distance metric does not help the cluster results.

5. **Shape + orientation + scale + position distance:** Adding position so that the full feature information of the sulcal curve is used, gives the best results.

In summary, results improve when orientation is added to the distance metric. Results improve further when position is also included. The scale distance does not contribute to better clustering and results improve when it is removed from the comprehensive distance metric. These clustering results support our findings for the feature evaluation of the labeling algorithm in Chapter 5. There, we saw that the shape distance results improve considerably when shape and position distances are combined; but that position distances alone give the best results.



(a) Clustering algorithm (b) Labeling algorithm

Figure A.1: **Clustering vs Labeling performance** evaluated using (the midpoint) position distance. There are 10 classes, each with 18 sulci. A total of $18 \times 10 = 180$ sulci are used in the experiments. The columns of the confusion matrix represent the ground truth/true labels (the column sum is 18); the rows are the clusters/test labels. The cluster labels were assigned based on the majority vote.

Table A.1: **Confusion Tables for the Clustering and Labeling algorithms.** The diagonal terms (in blue) give the number of sulci that were correctly identified (18 = 100%; this table is the raw data for the confusion matrix in Fig. A.1)

Clustering method

Cluster results: True Label		1	2	3	4	5	6	7	8	9	10
Left central	1	15	14								
L. postcentral	2						3	7		2	1
L. sup. frontal	3			16							
L. sylvian fissure	4	2	2	2	16	4					
L. sup. temporal	5	1	2		2	13					
Right central	6						9	3			
R. postcentral	7						6	8			
R. sup. frontal	8								16		
R. sylvian fissure	9					1			2	14	3
R. sup. temporal	10									2	14

Labeling method

Test Label: True Label		1	2	3	4	5	6	7	8	9	10
Left central	1	16	3								
L. postcentral	2	2	15								
L. sup. frontal	3			18							
L. sylvian fissure	4				15	2					
L. sup. temporal	5				3	16					
Right central	6						17	4			
R. postcentral	7						1	14			
R. sup. frontal	8								18		
R. sylvian fissure	9									17	1
R. sup. temporal	10									1	17

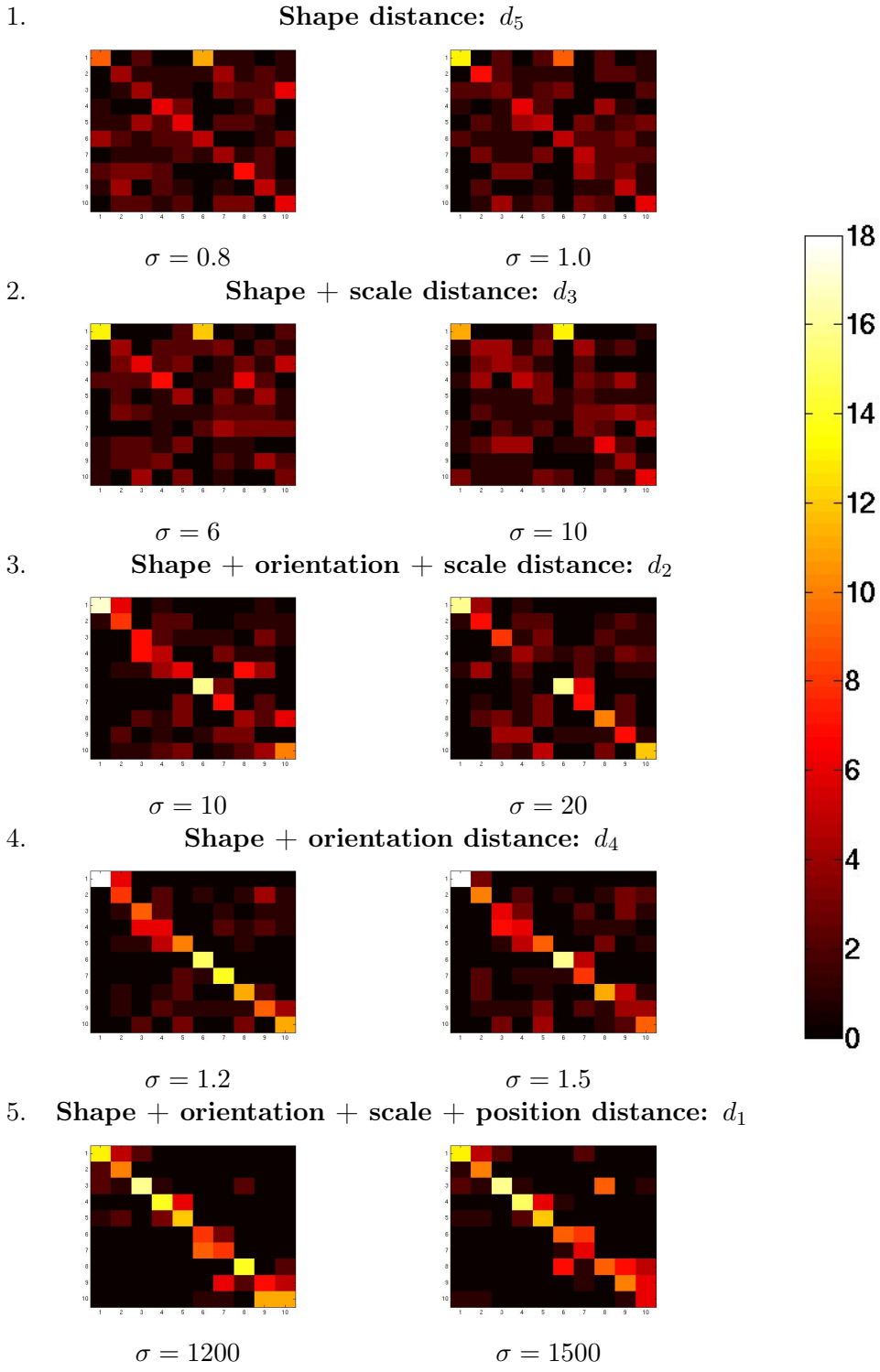


Figure A.2: **Clustering sulci using Riemannian feature distance metrics.** The shape distance (1 above) has the worst clustering results. The left and right central sulcus (columns 1 and 6 respectively) cluster in one group (row 1). Including orientation in the distance metric (3 above) rectifies this. With the addition of position (5 above), we get the best results.

Glossary

Comprehensive Riemannian Framework terms

Γ	the set of all orientation-preserving diffeomorphisms
$SO(3)$	the set of all possible rotations in $\mathbb{R}^3$
SRF	square root function (the h-function)
SRVF	square root velocity function (the q-function)
equivalence class	pg. 19
orbit	pg. 19
quotient space	pg. 20
Riemannian metric	pg. 19
Riemannian manifold	pg. 19
preshape space	pg. 19
shape space	pg. 19
DP	dynamic programming, pg. 24

White matter fiber analysis terms

AF	arcuate fasciculus
CC	corpus callosum
CT	computed tomography
DTI	diffusion tensor imaging
FA	fractional anisotropy
fMRI	functional magnetic resonance imaging
HARDI	high angular resolution diffusion imaging
MRI	magnetic resonance imaging
MD	mean diffusivity
MS	multiple sclerosis
NAWM	normal appearing white matter
ODF	orientation distribution function
PCA	principal component analysis
WM	white matter

Bibliography

[1] M. Mani, S. Kurtek, C. Barillot, and A. Srivastava, "A Comprehensive Riemannian Framework for the Analysis of White Matter Fiber Tracts," in *International Symposium on Biomedical Imaging (ISBI)*, (Rotterdam), pp. 1101–1104, IEEE, April 2010.

[2] L. Younes, P. W. Michor, J. Shah, and D. Mumford, "A metric on shape space with explicit geodesics," *Rend. Lincei Mat. Appl.*, vol. 9, pp. 25–57, 2008.

[3] S. H. Joshi, E. Klassen, A. Srivastava, and I. Jermyn, "A novel representation for Riemannian analysis of elastic curves in R^{n}," in *CVPR*, 2007.

[4] A. Srivastava, S. Joshi, W. Mio, and X. Liu, "Statistical shape analysis: Clustering, learning and testing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 4, pp. 590–602, 2005.

[5] S. Joshi, E. Klassen, A. Srivastava, and I. Jermyn, "Removing shape-preserving transformations in square-root elastic (SRE) framework for shape analysis of curves," in *Energy Minimization Methods in Computer Vision and Pattern Recognition*, vol. 4679 of *Lecture Notes in Computer Science*, pp. 387–398, 2007.

[6] N. Balov, A. Srivastava, C. Li, and Z. Ding, "Shape analysis of open curves in R^3 with applications to study of fiber tracts in DT-MRI data," in *EMMCVPR*, pp. 399–413, 2007.

[7] W. Liu, A. Srivastava, and E. Klassen, "Joint shape and texture analysis of objects boundaries in images using a Riemannian approach," in *Asilomar Conferences on Signals, Systems, and Computers*, 2008.

[8] E. Klassen, A. Srivastava, W. Mio, and S. Joshi, "Analysis of planar shapes using geodesic paths on shape spaces," *IEEE Patt. Analysis and Machine Intell.*, vol. 26, no. 3, pp. 372–383, 2004.

[9] M. P. do Carmo, *Differential Geometry of Curves and Surfaces*. Englewood Cliffs, NJ: Prentice-Hall, 1976.

[10] M. P. do Carmo, *Riemannian Geometry*. Boston, MA: Birkhäuser, 1992.

[11] J. Rotman, *An introduction to the theory of groups*. New York: Springer-Verlag, 1994.

[12] S. H. Joshi, *Inferences in Shape Space with Applications to Image Analysis and Computer Vision*. PhD thesis, Florida State University, Tallahassee, Florida, USA, 2007.

[13] L. Younes, "Computable elastic distances between shapes," *SIAM Journal on Applied Mathematics*, vol. 58, no. 2, pp. 565–586, 1998.

[14] E. Klassen and A. Srivastava, "Geodesics between 3D closed curves using path-straightening," in *ECCV 2006* (A. Leonardis, H. Bischof, and A. Pinz, eds.), vol. 3951 of *Lecture Notes in Computer Science*, pp. 95–106, Springer Berlin / Heidelberg, 2006.

[15] M. Ono, S. Kubik, and C. D. Abernathy, *Atlas of the Cerebral Sulci*. Thieme Medical Publisher, 1990.

[16] V. Horsley, "On the topographical relations of the cranium and surface of the cerebrum," in *Contribution to the surface anatomy of the cerebral hemispheres* (D. J. Cunningham, ed.), pp. 306–355, Royal Irish Academy, 1892.

[17] G. Le Goualher, D. L. Collins, C. Barillot, and A. C. Evans, "Automatic identification of cortical sulci using a 3D probabilistic atlas," in *MICCAI*, pp. 509–518, 1998.

[18] G. Le Goualher, E. Procyk, D. Collins, R. Venugopal, and C. Barillot, "Automated extraction and variability analysis of sulcal neuroanatomy," *IEEE Trans. Med. Imag.*, vol. 18, no. 3, pp. 206–217, 1999.

[19] A. C. Evans, D. L. Collins, P. Neelin, M. Kamber, and S. Marrett, "Three-dimensional correlative imaging: Applications in human brain mapping," in *Advances in Functional NeuroImaging: Technical Foundations* (R. Thatcher, M. Hallett, T. Zeffiro, E. John, and M. Huerta, eds.), pp. 145–162, Academic Press, 1994.

[20] G. Lohmann and Y. von Cramon, "Automatic labeling of the human cortical surface using sulcal basins," *IEEE Trans. Med. Imag.*, vol. 4, pp. 179–188, 2000.

[21] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[22] M. Perrot, D. Rivière, and J.-F. Mangin, "Identifying cortical sulci from localizations, shape and local organization," in *5th Proc. IEEE ISBI*, (Paris, France), pp. 420–423, May 2008.

[23] M. Perrot, D. Rivière, A. Tucholka, and J.-F. Mangin, "Joint Bayesian Cortical Sulci Recognition and Spatial Normalization," in *IPMI*, pp. 176–187, 2009.

[24] F. Yang and F. Kruggel, "A graph matching approach for labeling brain sulci using location, orientation, and shape," *Neurocomputing*, pp. 179–190, 2009.

[25] D. Rivière, J.-F. Mangin, D. Papadopoulos-Orfanos, J.-M. Martinez, V. Frouin, and J. Régis, "Automatic recognition of cortical sulci of the human brain using a congregation of neural networks," *Medical Image Analysis*, vol. 6, no. 2, pp. 77–92, 2002.

[26] A. Mechouche, X. Morandi, C. Golbreich, and B. Gibaud, "A Hybrid System Using Symbolic and Numeric Knowledge for the Semantic Annotation of Sulco-Gyral Anatomy in Brain MRI Images," *IEEE Transactions on Medical Imaging*, vol. 28, pp. 1165–1178, August 2009.

[27] R. Bajcsy and C. Broit, "Matching of deformed images," *IEEE Conf. on Pattern Recognition*, pp. 351–353, 1982.

[28] M. Vaillant and C. Davatzikos, "Hierarchical matching of cortical features for deformable brain image registration," in *IPMI*, pp. 182–195, Springer-Verlag, 1999.

[29] S. Jaume, B. M. Macq, and S. K. Warfield, "Labeling the Brain Surface using a Deformable Resolution Mesh," in *MICCAI*, pp. 451–458, 2002.

[30] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "Deformable templates using large deformation kinematics," *IEEE Trans. Imag. Proc.*, vol. 5, no. 10, pp. 1435–1447, 1996.

[31] P. Hellier and C. Barillot, "Cooperation between local and global approaches to register brain images," in *IPMI*, vol. 2082, pp. 315–328, Springer-Verlag, 2001.

[32] Z. Tu, S. Zheng, A. L. Yuille, A. L. Reiss, R. A. Dutton, A. D. Lee, A. M. Galaburda, I. Dinov, P. M. Thompson, and A. W. Toga, "Automated Extraction of the Cortical Sulci Based on a Supervised Learning Approach," *IEEE Trans. Med. Imag.*, vol. 26, no. 4, pp. 541–552, 2007.

[33] Y. Shi, Z. Tu, A. L. Reiss, R. A. Dutton, A. D. Lee, A. M. Galaburda, I. Dinov, P. M. Thompson, and A. W. Toga, "Joint Sulcal Detection on Cortical Surfaces with Graphical Models and Boosted Priors," *IEEE Trans. Med. Imag.*, vol. 28, no. 3, pp. 361–373, 2009.

[34] M. W. Trosset and C. E. Priebe, "The Out-of-Sample Problem for Classical Multidimensional Scaling," *Computational Statistics & Data Analysis*, vol. 52, no. 10, pp. 4635–4642, 2008.

[35] M. Mani, A. Srivastava, and C. Barillot, "The Labeling of Cortical Sulci using Multidimensional Scaling," *The Midas Journal*, 2008. `http://hdl.handle.net/10380/1502`.

[36] I. Corouge, G. Gerig, and S. Gouttard, "Towards a Shape Model of White Matter Fiber Bundles Using Diffusion Tensor MRI," in *ISBI*, pp. 344–347, 2004.

[37] L. J. O'Donnell and C. F. Westin, "Automatic tractography segmentation using a high-dimensional white matter atlas," *IEEE Trans. Med. Imag.*, vol. 26, no. 11, pp. 1562–1575, 2007.

[38] G. Le Goualher, C. Barillot, and Y. Bizais, "Modeling cortical sulci with active ribbons.," *IJPRAI*, vol. 11, no. 8, pp. 1295–1315, 1997.

[39] M. Kass, A. P. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.

[40] S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank, "Brain magnetic resonance imaging with contrast dependent on blood oxygenation," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 87, no. 24, pp. 9868–9872, 1990.

[41] G. J. Siegel, B. W. Agranoff, R. W. Albers, S. K. Fisher, and M. D. Uhler, eds., *Basic neurochemistry: molecular, cellular, and medical aspects.* Lipincott-Raven Press, sixth ed., 1999.

[42] P. J. Basser, J. Mattiello, and D. Le Bihan, "MR Diffusion Tensor Spectroscopy and Imaging," *Biophysical Journal*, vol. 66, pp. 259–267, 1994.

[43] S. Mori, B. J. Crain, V. P. Chacko, and P. C. M. Van Zijl, "Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging," *Annals of Neurology*, vol. 45, no. 2, pp. 265–269, 1999.

[44] R. Xue, P. C. van Zijl, , B. J. Crain, M. Solaiyappan, and S. Mori, "In vivo three-dimensional reconstruction of rat brain axonal projections by diffusion tensor imaging," *Magnetic Resonance in Medicine*, vol. 42, pp. 1123–1127, 1999.

[45] I. Corouge, P. T. Fletcher, S. Joshi, S. Gouttard, and G. Gerig, "Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis," in *Medical Image Analysis*, pp. 786–798, 2006.

[46] G. Gerig, S. Gouttard, and I. Corouge, "Analysis of brain white matter via fiber tract modeling," in *IEEE Engineering in Medicine and Biology Society, (EMBC)*, vol. 2, pp. 4421 –4424, Sep 2004.

[47] L. J. O'Donnell, C.-F. Westin, and A. J. Golby, "Tract-based morphometry for white matter group analysis," *NeuroImage*, vol. 45, no. 3, pp. 832–844, 2009.

[48] C. B. Goodlett, P. T. Fletcher, J. H. Gilmore, and G. Gerig, "Group analysis of DTI fiber tract statistics with application to neurodevelopment," *NeuroImage*, vol. 45, no. 1, Supplement 1, pp. S133 – S142, 2009. Mathematics in Brain Imaging.

[49] M. Maddah, W. E. L. Grimson, and S. K. Warfield, "Statistical Modeling And EM Clustering Of White Matter Fiber Tracts," in *ISBI*, pp. 53–56, 2006.

[50] M. Vaillant and J. Glaunès, "Surface matching via Currents," in *IPMI*, pp. 381–392, 2005.

[51] S. Durrleman, P. Fillard, X. Pennec, A. Trouvé, and N. Ayache, "A statistical model of white matter fiber bundles based on currents," in *IPMI '09*, (Berlin, Heidelberg), pp. 114–125, Springer-Verlag, 2009.

[52] P. A. Yushkevich, H. Zhang, T. J. Simon, and J. C. Gee, "Structure-specific statistical mapping of white matter tracts," *NeuroImage*, vol. 41, no. 2, pp. 448–461, 2008.

[53] M. Maddah, W. M. W. III, S. K. Warfield, C.-F. Westin, and W. E. L. Grimson, "Probabilistic clustering and quantitative analysis of white matter fiber tracts," in *IPMI*, vol. 20, pp. 372–383, July 2007.

[54] Z. Ding, A. W. Anderson, and J. C. Gore, "Classification and quantification of neuronal fiber pathways using diffusion tensor MRI," *Magn. Reson. Med.*, vol. 49, pp. 716–772, 2003.

[55] S. Zhang, S. Correia, and D. H. Laidlaw, "Identifying white-matter fiber bundles in DTI data using an automated proximity-based fiber-clustering method," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, pp. 1044–1053, 2008.

[56] A. Brun, H. Knutsson, H.-J. Park, M. E. Shenton, and C.-F. Westin, "Clustering Fiber Traces Using Normalized Cuts," in *MICCAI (1)*, pp. 368–375, 2004.

[57] L. Jonasson, P. Hagmann, J. P. Thiran, and V. J. Wedeen, "Fiber tracts of high angular resolution diffusion MRI are easily segmented with spectral clustering," in *ISMRM*, p. 1310, 2005.

[58] J. Shi and J. Malik, "Normalized Cuts and Image Segmentation," in *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97)*, CVPR '97, (Washington, DC, USA), pp. 731–737, IEEE Computer Society, 1997.

[59] M. Meila and J. Shi, "Learning Segmentation by Random Walks," in *In Advances in Neural Information Processing Systems*, pp. 873–879, MIT Press, 2001.

[60] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On Spectral Clustering: Analysis and an algorithm," in *Advances in Neural Information Processing Systems*, pp. 849–856, MIT Press, 2001.

[61] C. Fowlkes, S. Belongie, and J. Malik, "Efficient Spatiotemporal Grouping Using the Nyström Method," in *In Proc. IEEE Conf. Comput. Vision and Pattern Recognition*, pp. 231–238, 2001.

[62] E. Nyström, "Über Die Praktische Auflösung von Integralgleichungen mit Anwendungen auf Randwertaufgaben," *Acta Math.*, vol. 54, no. 1, pp. 185–204, 1930.

[63] S. Kumar, M. Mohri, and A. Talwalkar, "On sampling-based approximate spectral decomposition," in *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML '09, (New York, NY, USA), pp. 553–560, ACM, 2009.

[64] F. Aboitiz, A. B. Scheibel, R. S. Fisher, and E. Zaidel, "Fiber composition of the human corpus callosum," *Brain Research*, vol. 598, no. 1-2, pp. 143 – 153, 1992.

[65] N. Evangelou, D. Konz, M. M. Esiri, S. Smith, J. Palace, and P. M. Matthews, "Regional axonal loss in the corpus callosum correlates with cerebral white matter lesion volume and distribution in multiple sclerosis," *Brain*, vol. 123, no. 9, pp. 1845–1849, 2000.

[66] J. Foong, M. Maier, C. A. Clark, G. J. Barker, D. H. Miller, and M. A. Ron, "Neuropathological abnormalities of the corpus callosum in schizophrenia: a diffusion tensor imaging study," *Journal OF Neurology Neurosurgery and Psychiatry*, vol. 68, no. 2, pp. 242–244, 2000.

[67] S. Mori, S. Wakana, L. M. Nagae-Poetscher, and P. C. M. van Zijl, *MRI Atlas of Human White Matter*. Elsevier, 2005.

[68] S. F. Witelson, "Hand and Sex Differences in the Isthmus and Genu of the Human Corpus Callosum: A Postmortem Morphological Study," *Brain*, vol. 112, no. 3, pp. 799–835, 1989.

[69] M. Fréchet, "Les éléments aléatoires de nature quelconque dans un espace distancié," *Annales de l'institut Henri Poincaré*, vol. 10, no. 4, pp. 215–310, 1948.

[70] X. Pennec, "Probabilities and statistics on Riemannian manifolds: Basic tools for geometric measurements," in *NSIP*, pp. 194–198, 1999.

[71] H. Karcher, "Riemannian center of mass and mollifier smoothing," *Communications on Pure and Applied Mathematics*, vol. 30, no. 5, pp. 509–541, 1977.

[72] F. Aboitiz, E. Rodriguez, R. Olivares, and E. Zaidel, "Age-related changes in fibre composition of the human corpus callosum: sex differences," *NeuroReport*, vol. 7, no. 11, pp. 1761–1764, 1996.

[73] E. V. Sullivan, E. Adalsteinsson, and A. Pfefferbaum, "Selective age-related degradation of anterior callosal fiber bundles quantified in vivo with fiber tracking," *Cereb. Cortex*, vol. 16, no. 7, pp. 1030–1039, 2006.

[74] A. Y. Hardan, N. J. Minshew, and M. S. Keshavan, "Corpus callosum size in autism," *Neurology*, vol. 55, no. 7, pp. 1033–1036, 2000.

[75] P. W. Woodruff, I. C. McManus, and A. S. David, "Meta-analysis of corpus callosum size in schizophrenia," *Journal of Neurology, Neurosurgery and Psychiatry*, vol. 58, pp. 457–461, 1995.

[76] J. H. Simon, L. D. Jacobs, M. K. Campion, R. A. Rudick, D. L. Cookfair, R. M. Herndon, J. R. Richert, A. M. Salazar, J. S. Fischer, D. E. Goodkin, N. Simonian, M. Lajaunie, D. E. Miller, K. Wende, A. Martens-Davidson, R. P. Kinkel, I. Munschauer, F. E., and C. M. Brownscheidle, "A longitudinal study of brain atrophy in relapsing multiple sclerosis," *Neurology*, vol. 53, no. 1, pp. 139–148, 1999.

[77] H. Hampel, S. J. Teipel, G. E. Alexander, B. Horwitz, D. Teichberg, M. B. Schapiro, and S. I. Rapoport, "Corpus Callosum Atrophy Is a Possible Indicator of Region- and Cell Type-Specific Neuronal Degeneration in Alzheimer Disease: A Magnetic Resonance Imaging Analysis," *Arch Neurol*, vol. 55, no. 2, pp. 193–198, 1998.

[78] M. Walterfang, A. G. Wood, D. C. Reutens, S. J. Wood, J. Chen, D. Velakoulis, P. D. McGorry, and C. Pantelis, "Morphology of the corpus callosum at different stages of schizophrenia: cross-sectional study in first-episode and chronic illness," *The British Journal of Psychiatry*, vol. 192, no. 6, pp. 429–434, 2008.

[79] S. J. Teipel, W. Bayer, G. E. Alexander, Y. Zebuhr, D. Teichberg, L. Kulic, M. B. Schapiro, H.-J. Moller, S. I. Rapoport, and H. Hampel, "Progression of Corpus Callosum Atrophy in Alzheimer Disease," *Arch Neurol*, vol. 59, no. 2, pp. 243–248, 2002.

[80] R. A. Bermel and R. Bakshi, "The measurement and clinical relevance of brain atrophy in multiple sclerosis," *Lancet Neurol*, vol. 5, pp. 158–170, 2006.

[81] P. Yushkevich, S. M. Pizer, S. Joshi, and J. S. Marron, "Intuitive, localized analysis of shape variability," in *In International Conference on Information Processing in Medical Imaging*, pp. 402–408, Springer-Verlag, 2001.

[82] P. Golland, W. E. L. Grimson, and R. Kikinis, "Statistical shape analysis using fixed topology skeletons: Corpus callosum study," in *IPMI*, pp. 382–387, 1999.

[83] O. Ishaq, G. Hamarneh, R. Tam, and A. Traboulsee, "Longitudinal, regional and deformation-specific corpus callosum shape analysis for multiple sclerosis," in *International Conference of IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2110–2113, 2007.

[84] J. E. Downhill, M. S. Buchsbaum, T. Wei, J. Spiegel-Cohen, E. A. Hazlett, M. M. Haznedar, J. Silverman, and L. J. Siever, "Shape and size of the corpus callosum in schizophrenia and schizotypal personality disorder," *Schizophrenia Research*, vol. 42, no. 3, pp. 193 – 208, 2000.

[85] N. Adluru, C. Hinrichs, M. Chung, J.-E. Lee, V. Singh, E. Bigler, N. Lange, J. Lainhart, and A. Alexander, "Classification in DTI using shapes of white matter tracts," in *International Conference of IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2719 –2722, 2009.

List of Figures

List of Tables

Résumé

Il y a dans le cerveau humain environ 100 sillons corticaux, et plus de 100 milliards de faisceaux de matière blanche. Si le nombre, la configuration et la fonction de ces deux structures anatomiques diffèrent, elles possèdent toutefois une propriété géométrique commune: ce sont des courbes ouvertes continues. Cette thèse se propose d'étudier comment les caractéristiques des courbes ouvertes peuvent être exploitées afin d'analyser quantitativement les sillons corticaux et les faisceaux de matière blanche.

Les quatre caractéristiques d'une courbe ouverte–forme, taille, orientation et position–ont des propriétés différentes, si bien que l'approche usuelle est de traiter chacune séparément à l'aide d'une métrique ad hoc. Nous introduisons un cadre riemannien adapté dans lequel il est possible de fusionner les espaces de caractéristiques afin d'analyser conjointement plusieurs caractéristiques. Cette approche permet d'apparier et de comparer des courbes suivant des distances géodésiques. Les correspondances entre courbes sont établies automatiquement en utilisant une métrique élastique.

Dans cette thèse, nous validerons les métriques introduites et nous montrerons leurs applications pratiques, entre autres dans le cadre de plusieurs problèmes cliniques importants. Dans un premier temps, nous étudierons spécifiquement les fibres du corps calleux, afin de montrer comment le choix de la métrique influe sur le résultat du clustering. Nous proposons ensuite des outils permettant de calculer des statistiques sommaires sur les courbes, ce qui est un premier pas vers leur analyse statistique. Nous représentons les groupes de faisceaux par la moyenne et la variance de leurs principales caractéristiques, ce qui permet de réduire le volume des données dans l'analyse des faisceaux de matière blanche. Ensuite, nous présentons des méthodes permettant de détecter les changements morphologiques et les atteintes de la matière blanche.

Quant aux sillons corticaux, nous nous intéressons au problème de leur labellisation.

Mots clés: imagerie cérébrale, analyse des sillons, anomalie de la matière blanche, analyse de formes, regroupement, courbes ouvertes, variété riemanniene

Abstract

There are about a hundred sulci in the human brain and over a hundred billion white matter fibers. These two anatomical structures differ in their number, their physical arrangements and in the function they serve but they do share a common geometric description: they are both open continuous curves. This thesis is a study of how the physical attributes of open curves can be used to advantage in the many varied quantitative applications of sulci and white matter fibers.

Shape, scale, orientation and position, the four physical features associated with open curves, have different properties so the usual approach has been to design different metrics and spaces to treat them individually. We take an alternative approach using a comprehensive Riemannian framework where joint feature spaces allow for analysis of combinations of features. We can compare curves by using geodesic distances which quantify their differences.

We validate the metrics we use, demonstrate practical uses and apply the tools to important clinical problems. To begin, specific tract configurations in the corpus callosum are used to showcase clustering results that depend on the Riemannian distance metric used. This nicely argues for the judicious selection of metrics in various applications, a central premise in our work. The framework also provides tools for computing statistical summaries of curves. We represent fiber bundles with a mean and variance which describes their essential characteristics. This is both a convenient way to work with a large volume of fibers and is a first step towards statistical analysis. Next, we design and implement methods to detect morphological changes which can potentially track progressive white matter disease.

With sulci, we address the specific problem of labeling. An evaluation of physical features and methods such as clustering leads us to a pattern matching solution in which the sulcal configuration itself is the best feature.

Keywords: brain image analysis, white matter fiber analysis, sulcal analysis, white matter pathologies, shape analysis, clustering, open curves, Riemannian manifolds.