Universidade Estadual de Campinas

Faculdade de Engenharia Elétrica e de Computação

ANDRÉ LUÍS DA COSTA

Framework for the analysis of the corpus callosum micro structure along its extension

Framework para a análise da micro estrutura do corpo caloso ao longo de sua extensão

Campinas

2018

ANDRÉ LUÍS DA COSTA

FRAMEWORK FOR THE ANALYSIS OF THE
CORPUS CALLOSUM MICRO STRUCTURE ALONG ITS EXTENSION

FRAMEWORK PARA A ANÁLISE DA
MICRO ESTRUTURA DO CORPO CALOSO AO LONGO DE SUA EXTENSÃO

<div align="right">

Thesis presented to the School of Electrical and Computer Engineering of the University of Campinas as partial requirement fulfillment for the degree of Doctor in Electrical Engineering. Concentration area: Computer Engineering.

Tese apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Engenharia Elétrica. Área de concentração: Engenharia de Computação.

</div>

Supervisor: PH. D. ROBERTO DE ALENCAR LOTUFO

Co–supervisor: PH. D. LETÍCIA RITTNER

ESTE EXEMPLAR CORRESPONDE À VERSÃO
FINAL DA TESE DEFENDIDA PELO ALUNO
ANDRÉ LUÍS DA COSTA, E ORIENTADA PELO
PROF. DR. ROBERTO DE ALENCAR LOTUFO

———————————————————————

Campinas

2018

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Rose Meire da Silva - CRB 8/5974

Informações para Biblioteca Digital

Título em outro idioma: Framework para a análise da microestrutura do corpo caloso ao longo de sua extensão
Palavras-chave em inglês:
Corpus callosum
White matter
Microstructure
Signatures
Área de concentração: Engenharia de Computação
Titulação: Doutor em Engenharia Elétrica
Banca examinadora:
Roberto de Alencar Lotufo [Orientador]
Hélio Pedrini
Shin - Ting Wu
Ricardo José Ferrari
Carlos Ernesto Garrido Salmon
Data de defesa: 06-12-2018
Programa de Pós-Graduação: Engenharia Elétrica

COMISSÃO JULGADORA – TESE DE DOUTORADO

Candidato: André Luís da Costa RA: 058954

Data da Defesa: 06 de dezembro de 2018

Título da Tese: "Framework para a análise da micro estrutura do corpo caloso ao longo de sua extensão".

Prof. Dr. Roberto de Alencar Lotufo
Prof. Dr. Hélio Pedrini
Prof. Dra. Shin - Ting Wu
Prof. Dr. Ricardo José Ferrari
Prof. Dr. Carlos Ernesto Garrido Salmon

A ata de defesa, com as respectivas assinaturas dos membros da Comissão Julgadora, encontra-se no SIGA (Sistema de Fluxo de Dissertação/Tese) e na Secretaria de PósGraduação da Faculdade de Engenharia Elétrica e de Computação.

Acknowledgment

Abstract

The corpus callosum is of great interest for the medical and research community, and its characteristics have been associated with many psychological disorders and brain diseases. Localized analysis of its features is a usual procedure, particularly for the diagnosis of multiple sclerosis and other inflammatory diseases. In this work, we propose a framework for extracting microstructure features along the corpus callosum extent into a signature function, allowing global and localized analyses to be performed in the 1–D domain of the signature, instead of the 3–D domain of the original image. Our solution is a succession of several specialized methods, which were designed to solve specific parts of the signature generation pipeline, including defining a plane of local symmetry for the corpus callosum internal fibers, perform the corpus callosum segmentation, trace the structure median axis, and extract the features along the median axis. A dataset with images from 80 distinct acquisitions from healthy subjects was used to evaluate both, the fiber's symmetry plane, and the generated signatures. Results show that the plane predicted by our method is significantly distinct from the planes predicted by traditional mid–sagittal plane estimation methods, with a larger difference on the inclination relative to the axial plane, of about 2 degrees on average. The signatures present a similar pattern in most cases but retain individual characteristics. The signatures generated by our proposed framework provide an unprecedented way to perform the analysis of the corpus callosum microstructure features, which is inherently localized, and independent from the structure morphology. Our solution open new possibilities for future related research and development in the field.

Key-words: Corpus callosum. White matter. Microstructure. DTI. Signature.

Resumo

O corpo caloso é de grande interesse para a comunidade médica e de pesquisa, e suas características têm sido associadas a muitos distúrbios psicológicos e doenças cerebrais. A análise localizada de suas características é um procedimento usual, particularmente para o diagnóstico de esclerose múltipla e outras doenças inflamatórias. Neste trabalho, propomos um framework para extrair características da microestrutura ao longo da extensão do corpo caloso em uma função de assinatura, permitindo que análises globais e localizadas sejam realizadas no domínio 1–D da assinatura, em vez do domínio 3–D de a imagem original. Nossa solução é uma sucessão de vários métodos especializados, que foram projetados para resolver partes específicas do pipeline de geração de assinatura, incluindo a definição de um plano de simetria local para as fibras internas do corpo caloso, realizar a segmentação do corpo caloso, traçar o eixo médio da estrutura e extrair as características ao longo do eixo médio. Um dataset com imagens de 80 aquisições distintas de indivíduos saudáveis foi usado para avaliar tanto o plano de simetria da fibra quanto as assinaturas geradas. Os resultados mostram que o plano predito pelo nosso método é significativamente distinto dos planos preditos pelos métodos tradicionais de estimativa do plano sagital médio, com uma diferença maior na inclinação em relação ao plano axial, de cerca de 2 graus em média. As assinaturas apresentam um padrão similar na maioria dos casos, mas retêm características individuais. As assinaturas geradas pelo nosso framework proposto fornecem uma maneira inédita de realizar a análise das características da microestrutura do corpo caloso, que é inerentemente localizada e independente da morfologia da estrutura. Nossa solução abre novas possibilidades no campo para pesquisa e desenvolvimento futuros relacionados.

Palavras-chave: Corpo caloso. Substância branca. Microestrutura. DTI. Assinatura.

List of Figures

List of Tables

Contents

1. Introduction

The *corpus callosum* (CC) is the largest white matter structure found in mammalian brain [1, 20]. It is a structure basically composed by axons that interconnect cortical areas from both brain hemispheres, integrating motor, sensory and cognitive functions [24]. As illustrated in Fig. 1.1, the fibers that cross the corpus callosum connect to a large portion of the cortex, in both brain hemispheres, and present a symmetrical arrangement.
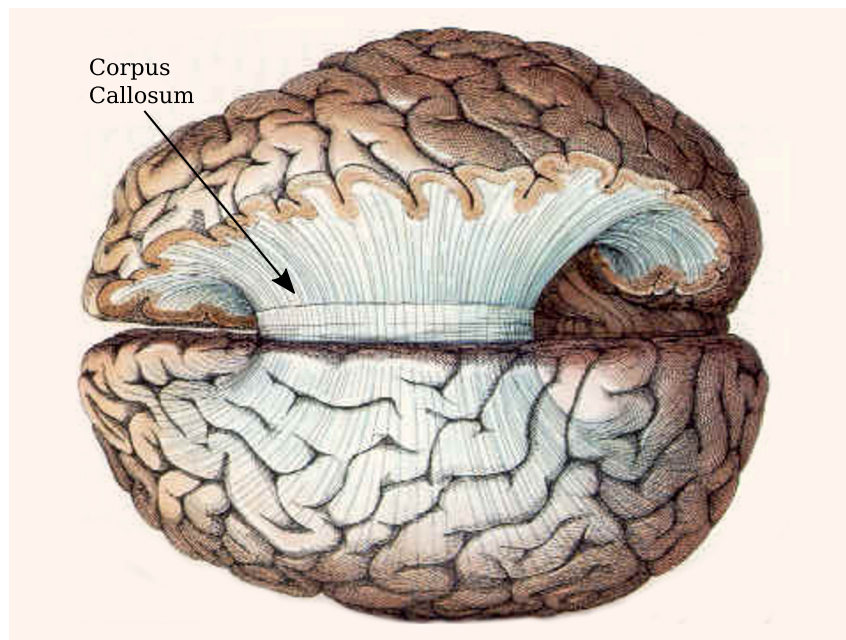
Figure 1.1: Corpus callosum illustration showing how its fibers have an symmetrical arrangement, and spread to a large portion of the cortex. Adapted from website brainmind.com.

For decades the CC has been of interest in medical research and with the advent of *magnetic resonance imaging* (MRI) new non–invasive approaches became possible, expanding the horizons in this field. There are many recent publications in the literature from studies that investigate the association of CC characteristics from MRI to subject attributes, like sex and age [29]; or to patient conditions, such as brain diseases [4, 10] and psychological disorders [37, 16]. Because the corpus callosum is associated with a large portion of the cortex, its volume and composition have been associated with neuronal loss [4], and with intelligence [28, 19]. In a recent study, Garg et al. [15] found that localized lesions in the corpus callosum provide clues for the diagnosis of multiple sclerosis and other inflammatory diseases.

While in early MRI studies the T1-weighted modality was widely employed to assess the CC shape and thickness, more recent studies include assessment of white matter microscopic characteristics [8] by using *diffusion tensor imaging* (DTI). Although

the analysis of microstructure features is completely different from the usual thickness evaluation, the most commonly used microstructure analysis method is very similar to the one used for the thickness, which is based on a 2–D image that corresponds to the *mid–sagittal plane* (MSP). This method makes sense for thickness evaluation, as the CC form is well defined in the sagittal plane, and is considered a suitable representation for the whole structure if the correct MSP is employed [29]. However, for assessing the CC microstructure, the MSP method imposes severe limitations to the sampling of CC internal features. In fact, the values derived from DTI at the middle of the corpus callosum may actually be inadequate to represent the whole structure, as appointed by Mollink et al. [31]. Therefore, it seems that the currently employed method has been inherited from early MRI anatomical research without proper considerations on the implications it would have in face of the new analysis requirements relative to the assessment of the white matter microstructure.

It is usual for microstructure analyses to split the corpus callosum total area into a few subregions, called partitions, that are associated with specific cortical areas [44, 2, 17, 14, 42]. There are many proposed partition schemata, each with a different number and sizes of partitions. One such partition scheme, shown in Fig. 1.2, separate the corpus callosum area into five partitions, which were defined based on a histological study [44, 17]. The justification for the usage of this approach is that the distinct cortical areas have its axons crossing specific areas of the corpus callosum. Indeed, the tissue in different portions of the corpus callosum has distinct properties, such as axon diameter and density of fibers. Therefore, it is clear that localized analysis of the microstructure features along the corpus callosum is quite important and meaningful.
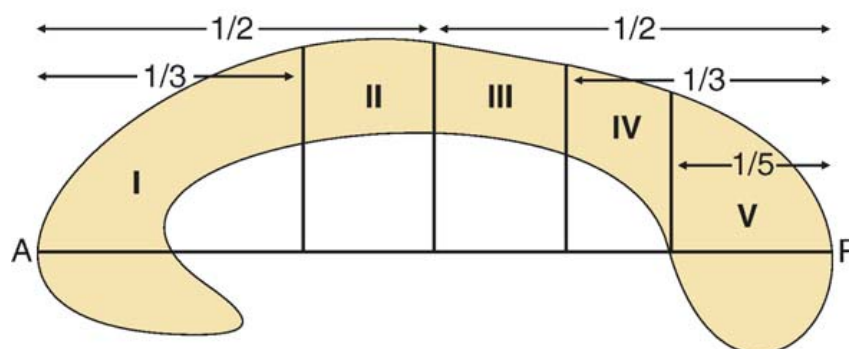
Figure 1.2: Witelson's partitioning scheme. Adapted from [17].

Another approach for the analysis of localized microstructure features along the corpus callosum is to perform many readings regularly spaced through the structure extent. Park et al. [36] have employed such technique as a support analysis in their work, and later Rittner et al.[42] used an improved version of their method to perform automated adaptive parcellation of the corpus callosum area. Although this is not exactly

the same as the partitioning analysis, both approaches share the fundamental concept of localized feature analysis, that is relative to the structure extent. Thus, the approach used by Park et al. [36] can be thought off as a generic partitioning analysis method, with a high number of small regular partitions. Interestingly, any of the partitioning schemata can be represented by grouping these regular small "partitions". In fact, a useful property of Park's approach is that it results in a profile of the corpus callosum features, in the form of a 1–D function, with the domain directly associated to the structure extent.

In this work, we propose a framework for extracting the corpus callosum features across its structure extension into a profile function that we refer to as a signature. Our design was inspired on the profile functions used by Park et al. [36], and includes solutions for several issues we have identified, such the definition of an explicit median axis that represents the corpus callosum extent, making the readings regularly spaced in relation to the median axis, and reading feature values in a 3–D neighborhood around each reference point, instead of choosing only one feature value, as in the previously proposed methods [36, 42]. We call our solution a framework because it is a succession of methods that form a pipeline for extracting the microstructure information from a 3–D space into 1–D space. The signatures are intended to be self-sufficient for the microstructure analysis, without the need to perform other procedures in the original 3–D image space. The main goals in this work are: to show how the signatures can represent the features along the corpus callosum structure; and to provide a basic set of methods, including implementations, that are able to generate such signatures. Secondary goals include promoting the discussion about corpus callosum characterization techniques, and exposing how inadequate are the methods commonly employed nowadays.

1.1 Challenges and issues

The issues we address in this work are complex and multidisciplinary in nature. In addition, there is a lack of standards for the type of analysis we are proposing. Hence, it is difficult to establish formal statements and proof for correctness in many parts of our work. Because of that, we trust our intuition and do our best to provide suitable solutions for each of the tasks required to achieve the goal of generating a reliable signature function. We carefully designed each step of our method to minimize the influence from initial conditions and other issues that could introduce variations to the geometrical references used to extract the signature, which ultimately consists of a median axis ranging from the anterior to the posterior extremes of the corpus callosum. The proposed solution is fully automated and directly applied in 3–D space.

One of the biggest challenges we faced was the low spatial resolution and quality of DTI. However, this is also a challenge for every solution based on that imaging modality. As an example, it is common for the corpus callosum thickness to have only a few (around three) voxels in certain parts. In special for the partitioning analysis methods, at least one partition will have too few voxels to represent the tissue composition in that region. Our solution is more robust to this problem because we acquire the samples from 3–D neighborhood, thus considerably increasing the number of voxels that contribute for the definition of each sample distribution. The low spatial resolution issue was more challenging for us while estimating a smooth contour for the corpus callosum shape, as described in Section 4.1.1.

Another issue we have observed is that there is a lack of consensus for which MSP estimation plane to use, and in some cases, the MSP is estimated from precarious or invalid criteria [27, 29], which can potentially have impacted the overall results in many studies. Further, even without any mistakes done during implementation, some variations on the MSP estimation are to be expected, because there are many valid MSP estimation methods available that will lead to slightly different estimated versions. In fact, every MSP estimation method will lead to an estimated plane that will be a slightly different version of the ideal MSP, which is unknown and impossible to precisely determine in regards to the current broad definition of what an MSP is. Not surprisingly, it is common to find contradicting results among different studies [29], which is also attributed to the lack of standardization. Therefore, selecting an MSP method for our framework was not an easy task, especially because practically all MSP estimation methods are not directly related to the corpus callosum structure. Fortunately, we decided to investigate an alternative way to defined such reference plane and found that it is possible to estimate a plane of local symmetric for the corpus callosum internal fibers. Surprisingly, this fiber symmetry plane is significantly different from other MSP, and this result represents an important finding in this work.

The analysis and visualization of the signatures we have proposed is also very challenging, mostly because there are no references about how to do this. Our solution of storing the whole empirical distribution function for each reference reading point along the median axis is unprecedented. In addition, we designed the framework to be generic, allowing the generation of signatures for any microstructure feature map and, therefore, we cannot make any assumptions about the distributions. Visualization is a great tool that can certainly be applied to the signatures, to enable experts to perform quick visual inspections and analyses. However, this subject requires further specific research that could not be conducted in this work.

1.2   Chapter closing

Chapter 2 present an overview about the fundamental concepts associated with our work, and our proposed methods, together with detailed descriptions of other methods, datasets, e other resources used through this study. Chapter 3 describe our proposed method for estimating the corpus callosum symmetry plane, named *mid–callosal plane* (MCP), from the symmetrical organization of its internal fibers. The MCP is the intermediary reference for the subsequent steps needed to estimate the corpus callosum median axis, which is the ultimate reference for generating the signatures. These steps and the signature extraction procedure are described in Chapter 4, along with experimental results. Finally, in Chapter 5 we discuss our findings related to the application of our proposed framework and present perspectives for future works and improvements. Associated publications and submissions are also listed in Chapter 5.

2. Materials and methods

What will you find in this chapter? In this chapter we provide: an overview for the methods evaluated in this work, including our proposed one; information about the dataset used in the experiments, and about evaluation issues; coding languages, libraries, and third–party software; equipment configuration; descriptions about the object of interest – which in our case is the corpus callosum – and its representations; and fundamental concepts related to this work.

2.1 Fundamental concepts

In this Section, we provide a brief review of fundamental topics associated with our work, in order to make it easier to be understood. Although we make use of several mathematical tools in our method's implementation, those concepts are not going to be covered in this Section, as they are already explained along the method description through this document. Instead, this Section will cover basic concepts about *magnetic resonance imaging* (MRI), DTI, and relevant properties from the brain and the corpus callosum such as anatomy and organization.

2.1.1 Magnetic resonance imaging

The MRI is a technology that allows seeing inside the human body from outside without causing any harm to the living tissue. Currently, the MRI technology enables several distinct imaging modalities that target specific properties of the living tissue, such as structural anatomy, chemical spectrography, diffusion of water molecules, and others [40, 26, 32]. This review is restricted to the basics of MRI, followed by an explanation about diffusion imaging, and how diffusion data is fitted to the DTI model.

The principle for the MRI technology comes from a property many atomic nuclei have, which is an angular momentum due to their inherent rotation, or spin. Because the nuclei are electrically charged, the current flowing about the spin axis generates a small magnetic field, giving a non–zero spin atomic nucleus an associated magnetic dipole. In natural conditions, the nuclei magnetic dipoles are pointing at random directions but will align themselves with the lines of induction of a strong enough magnetic field [40]. The magnetic behavior from a nuclei population can be estimated, defining a bulk magnetization vector $\vec{M}$, which represents the net effect of all magnetic moments

combined at the material being examined. In natural conditions, the bulk magnetization is zero, but when a magnetic field is applied on the sample the nuclei dipoles become oriented, and a non–zero bulk magnetization vector will point at a direction parallel to the magnetic field. By convention, this external strong magnetic field is called B_0, and its direction defines the z axis.

The second element to MRI is resonance, which is used to cause a precession movement to the atomic nuclei from the desired species. Precession is the denomination given to the conical motion the axis from a gyroscope will perform about the vertical axis, when the gyroscope axis is tipped away from the vertical. Figure 2.1(a) illustrate the precession movement as an analogy to a top. The spinning atomic nuclei behave the same way, and if tipped away from the magnetic field lines of induction will start a precession movement. In order to do the nuclei tipping, a second weaker magnetic field is applied, oscillating at the same frequency as the natural nuclei species precession frequency, which is given by the Larmor equation [40, 26]

$$\omega_0 = \gamma B_0, \tag{2.1}$$

where γ is the natural unique gyroscopic ratio from the nuclear species, and B_0 is the strength of the external magnetic field. Therefore, the MRI is a selective process and can be used to assess the composition of a material by targeting different nuclei species. Nevertheless, most medical imaging modalities are restricted to atoms of hydrogen, which, very fortunately, are the most sensitive spinning nuclei to resonance, in addition to being abundant in the human body [40, 26].

The oscillating secondary magnetic field is applied to the material through a *radio frequency* (RF) coil. This magnetic field rotates in the xy plane with appropriate angles relative to the main magnetic field direction in order to cause the precession movement to an ensemble of nuclei, causing the bulk magnetization vector $\vec{M}$ to rotate about the z axis. An RF pulse long and strong enough can make $\vec{M}$ rotate in the plane xy, or even invert its direction in relation to the main external magnetic field. After the RF excitation is turned off, $\vec{M}$ will return to its original position relative to z through a process called *free precession*, as illustrated in Fig. 2.1(b). The energy released during the free precession is detected by an RF receiver coil, which is used to form the image.

Distinct image modalities can be generated by observing different aspects of the free precession signal from specific RF pulse sequences. The most straight forward aspect is intensity, which reveals the density of nuclei from the targeted species in a given portion of space. Another aspect is the time it takes for the signal to decay, i. e., how long the bulk magnetization vector $\vec{M}$ takes to go back to the original state precedent
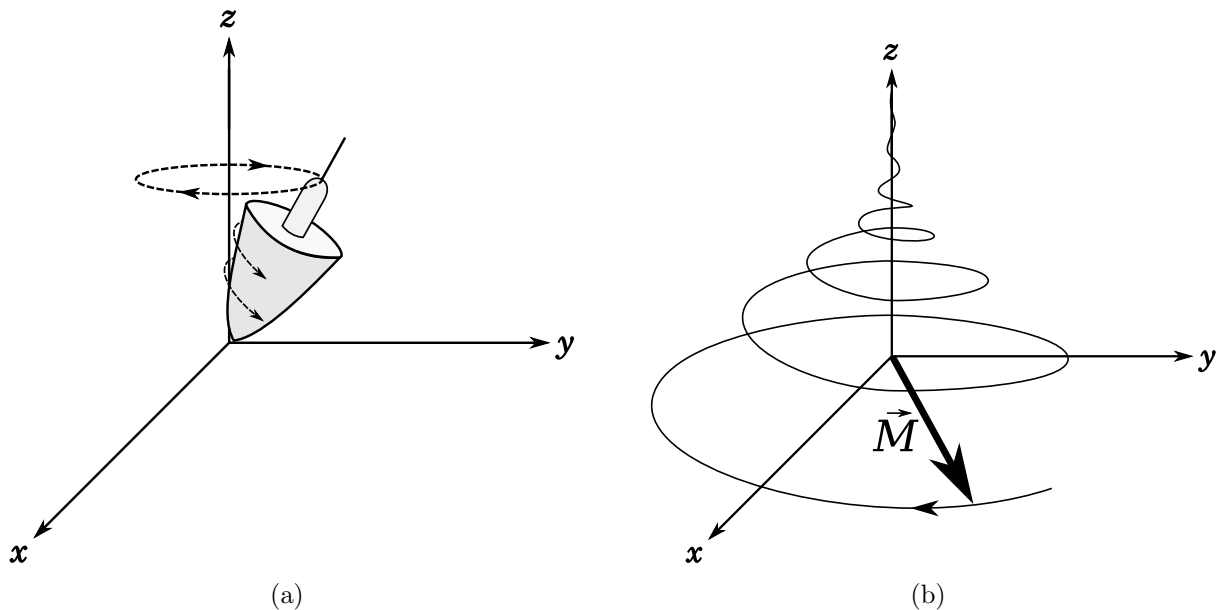
Figure 2.1: Precession: (a) illustration of the precession movement as an analogy to a spinning top that was tipped away from the vertical (z axis); (b) the pathway the tip of the bulk magnetization vector $\vec{M}$ makes during the free precession, after the RF pulse. Adapted from [26].

to the RF pulse. Given that the time required for the relaxation process to complete is influenced by physical properties of the tissue such as state and temperature, it is possible to generate contrast between liquids, and hard and soft tissues. There are two relaxation times that can be observed: the so–called $T1$, or longitudinal relaxation time, the actual time it takes for the magnetized spin system to reach thermal equilibrium after the RF pulse, causing $\vec{M}$ to return to its original state; the other time is referred to as $T2$, or transverse relaxation time, which is the time it takes for the vector $\vec{M}$ to reorient with the external main magnetic field. While in liquids $T1$ and $T2$ are similar and can last for up to a few seconds, in solids or in low–temperature conditions they are very different, being $T2$ very fast, in the order of microseconds, and $T1$ very slow, possibly lasting for hours [40]. Thus, $T2$–weighted images will have brighter voxels in areas with liquids, while in $T1$–weighted images the brighter voxels will represent harder tissues, such as bones.

2.1.2 Diffusion tensor imaging

This imaging modality was designed to measure the diffusion of water molecules inside a voxel. Unlike flow and bulk motion, which are directional, diffusion is characterized by the displacement of molecules at random directions, also known as Brownian motion [32]. In a *diffusion weighted image* (DWI), the signal intensity during relaxation after an RF pulse is sensitized to the amount of water diffusion in a given direction. For

this purpose, a special device is used to introduce a linear inhomogeneity (gradient) to the main external magnetic field. The inhomogeneity can be applied at only one direction at a time, but it can be any direction. In addition, the strength of the gradient and its polarity can be controlled by this device.

In order to understand how to signal intensity can be diffusion sensitized, let us recall that according to 2.1 the precession frequency of an atomic nucleus has a direct relation to the external magnetic field strength. This relationship is explored by applying a pair of dephasing and rephasing gradients, as illustrated in Fig. 2.2, after the RF excitation. Note that in the first time frame t_1 the magnetic field is homogeneous, and the atoms from three water molecules (red, green and blue circles) at different locations are processing on phase. In t_2 a dephasing gradient is applied and now each water molecule is experiencing a distinct strength from the external magnetic field, making the individual signals from each molecule to go out of phase (change in the precession frequency), which affects the combined signal. Despite the homogeneity of the magnetic field being reestablished in t_3, the individual signals are still out of phase. Finally, in t_4 a rephasing gradient is applied to the sample, which has the same strength as the gradient that was applied in t_2, but with inverted polarity. The rephasing gradient makes the individual signals to regain phase, recovering the combined signal to what it should be if the field was always homogeneous. Although in the illustration from Fig. 2.2 the locations of the water molecules are fixed, and thus nothing seems to have happened, the signal after t_4 is diffusion sensitized. If there are water molecules randomly moving in the gradient direction a portion of the molecules are going to change position, causing them to not be properly rephased in t_4, which will affect the combined signal, turning it into something different than what it was expected to be. Finally, the diffusion constant is obtained by comparing the diffusion–weighted signal to another signal with no diffusion weighting; or one weighted with a distinct gradient strength.

The diffusion constant is specific for the applied gradient direction, and although diffusion by itself is inherently isotropic, i. e., is the same in all directions, a gradient applied to a different direction may yield distinct diffusion constants for the same voxels when scanning biological tissue. This happens because in organized tissue the diffusion in some directions are restricted, causing the diffusion to be anisotropic. In fibrous tissue, for instance, as are white matter axonal tracts, the diffusion is going to be preferential along the fibers. Therefore, several DWI acquired from distinct gradient directions are required for a good estimation of the tissue properties by diffusion constants. This implies a high time cost for acquiring a set of DWI, which associated with the lower signal to noise ratio limits the image spatial resolution in relation to other MRI modalities, such as T_1–weighted. In addition, due to the longer time the patient has to be in the scanner, the DWI is more susceptible to bulk motion, thus requiring heavy post–processing.
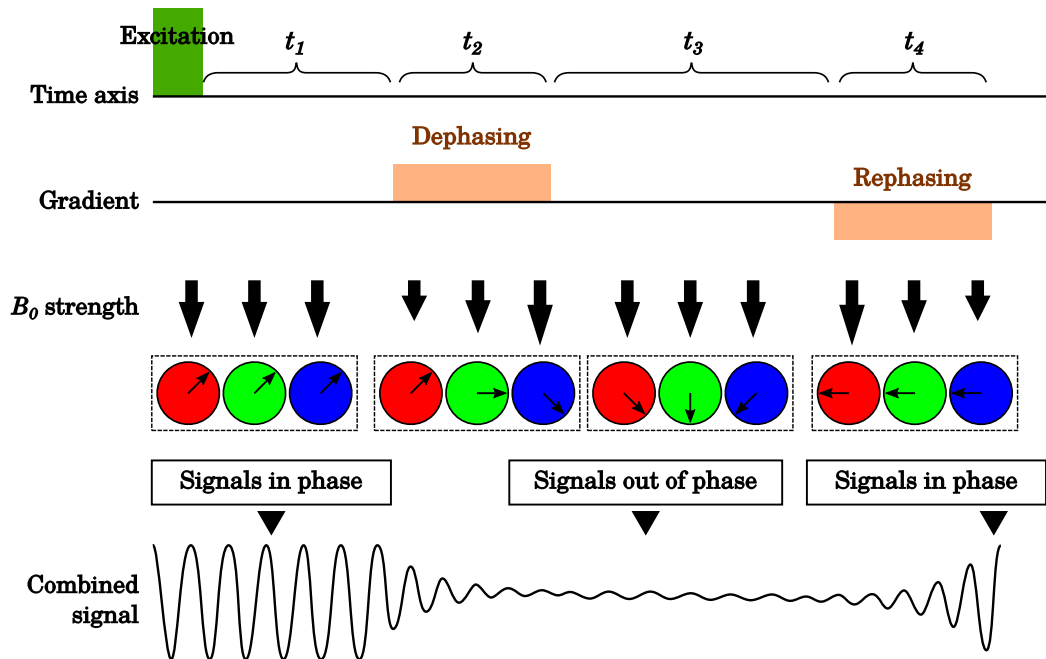
Figure 2.2: Illustration about the main concepts involved in a dephasing and rephasing pulse sequence, which is used to sensitize the signal to diffusion. Adapted from [32].

Nonetheless, it reveals unique valuable information about the tissue characteristics that cannot be obtained from anatomical images.

While in 3–D space the isotropic diffusion can be represented by a sphere with a radius determined by the diffusion constant, the anisotropic diffusion requires an ellipsoid to be properly represented (assuming that diffusion property is elliptic). A 3×3 tensor is used to estimate the anisotropic diffusion, hence DTI. The tensor is a symmetric matrix, which implicates six independent parameters to be estimated. At least seven DWI are required to define the tensor, being six from independent gradient directions, and one with the least diffusion weighting, which is used to calculate the diffusion constant from the other six DWI. However, usually more gradient directions are used in order to account for measurement errors. Once the tensor has been estimated, an eigenvalue system can be solved to obtain the eigenvalues λ_1, λ_2, and λ_3, and their respectively associated eigenvectors ν_1, ν_2, and ν_3. The eigenvectors define the orientations of the three orthogonal ellipsoid axes, and the eigenvectors their length.

Information can be extracted from a DTI

Although it is possible to render the tensors as 3–D ellipsoids for visualization, this is not a practical approach because the size becomes too small to see, unless magnified, in addition to other issues such as view angle, light source, and shading. Therefore, it is

important to reduce the information to fit a grayscale or RBG image, as well for specific analysis types. For instance, the *mean diffusivity* (MD), defined as [32]

$$MD = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3},\qquad(2.2)$$

provide an estimation of the diffusion amount that is independent from gradient direction. Another widely used DTI derived scalar is the *fractional anisotropy* (FA), defined as [32]

$$FA = \sqrt{\frac{1}{2}}\frac{\sqrt{(\lambda_1 - \lambda_2)^2 + (\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_3)^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}},\qquad(2.3)$$

which is an index normalized in the range $[0, 1]$ for how anisotropic the diffusion is, being $FA = 0$ an isotropic diffusion. Although these scalars can be used to generate contrast in grayscale images, their are particularly useful and commonly employed on analyses about the underling tissue microstructure [35, 32, 8, 10, 37].

The preferential diffusion direction, defined by the eigenvector ν_1, is also a piece of very important information that can be extracted from the DTI tensors, which reveal the tissue microstructure organization. It can be plotted as a color image for visualization purposes, with the vector's x, y, and z components respectively mapped to the RBG channels, and weighted by FA to reduce the influence of eigenvectors representing not organized tissue [32]. Besides visualization, the preferential diffusion directions are useful for several applications, such as tractography [7, 17, 25], and segmentation of white matter structures [34, 33]. Particularly for our work, the preferential diffusion directions play a fundamental role in the estimation of the corpus callosum symmetry plane, which is employed as the basic reference for the subsequent processes. We also use the preferential diffusion directions combined with FA values to perform reliable segmentation of the corpus callosum.

2.1.3 Corpus callosum anatomy

The corpus callosum is a large white matter structure composed by commissural fibers, which interconnect areas from the contralateral brain hemispheres, allowing learning and memory in one hemisphere to be shared with the other [1, 12, 20]. As shown in Fig. 2.3, parts of the corpus callosum along its extension receive the denominations *rostrum*, *genu*, *body*, and *splenium*, although there are no clear anatomical markers between them. The bulk of the frontal, parietal, occipital, and temporal lobes are interconnected by fibers crossing through the corpus callosum, being this structure essentially a dense bundle of organized fibers, which results in high FA values for the structure, as can be

seen in Fig. 2.3(b) where the corpus callosum body in an axial slice take an x–like shape that is symmetrical in relation to the mid–callosal plane.

(a)   (b)
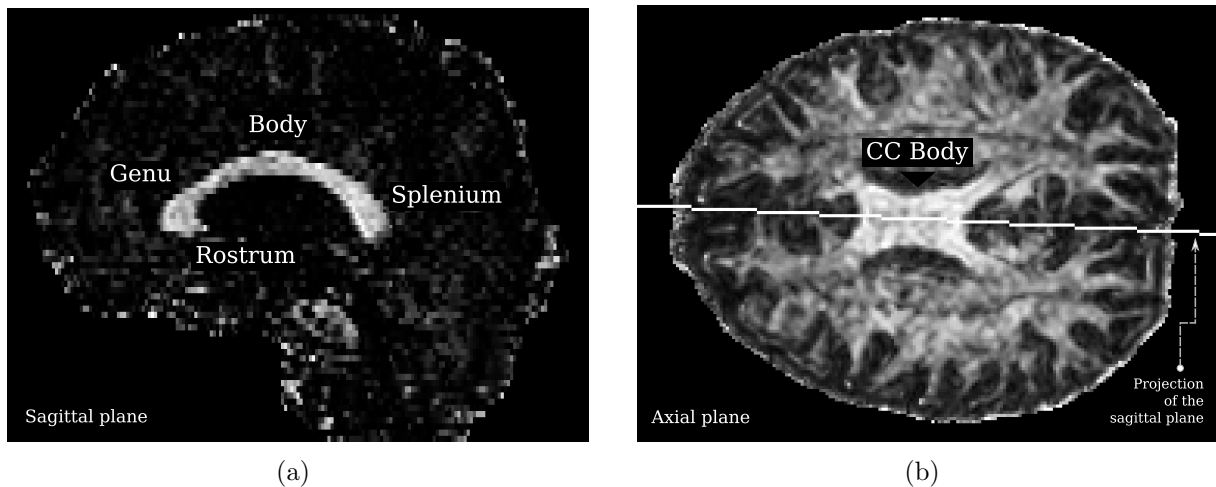
Figure 2.3: Visualization of the corpus callosum in DTI derived scalar maps. Figure 2.3(a) is a FA map weighted by the length of the x component from the eigenvectors ν_1, and is showing the most known corpus callosum form. Figure 2.3(b) is an axial slice passing through the body of the corpus callosum with contrast due to FA values. In this axial plane the corpus callosum appear as an x–like shape, and not all of its form is well define as in the sagittal plane.

An outstanding characteristic from the corpus callosum fibers is that they have a high degree of organization, being the overall orientation of fibers near the corpus callosum middle perpendicular to its own symmetry plane, as shown in Fig. 2.4. This property is useful for distinguishing the corpus callosum in the sagittal plane from other white matter structures that also has high FA values, such as the fornix. Figure 2.3(a) demonstrate how distinct the corpus callosum appear in the sagittal plane when the FA values are weighted by the projection of the eigenvector ν_1 over the plane's normal vector, which makes segmentation a much simpler task. Furthermore, the corpus callosum internal fibers have a consistent symmetrical organization, which is a piece of useful information for establishing a robust reference directly associated with the structure itself. Although the corpus callosum symmetry plane is expected to be near the MSP estimated from other information, it would still be a more appropriate reference because this symmetry plane is a reflection of the own structure organization.

2.2   Our proposed method

The main goal of our proposed method is to extract microstructure features along the corpus callosum extension, generating a 1–D function that we refer to as a

Figure 2.4: Corpus callosum internal microstructure organization in an axial slice crossing its body. Dashed lines represent the estimated corpus callosum boundaries, and the small arrows represent the orientat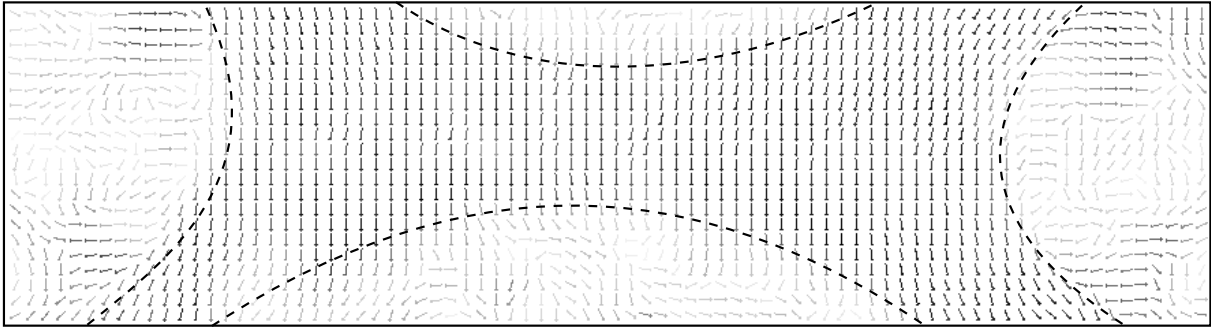ions of the eigenvectors ν_1. The grayscale of arrows represent the associated FA values, being black and white, respectively, 1 and 0. Adapted from [9].

signature. The corpus callosum extension is defined by its median axis, which must be estimated as a first step for the whole process, as shown in Fig. 2.5. The median axis goes from the anterior to the posterior corpus callosum extremes and can be seen as a backbone for the structure. Further details about the median axis, refer to Section 4.1.3. Once the median axis is estimated, it serves as a path across the tridimensional space that guides the feature extraction procedure. Equidistant points relative to the median axis define local weighted neighborhoods, from where the features are extracted. The distributions of values correspondent to the local neighborhoods are stored as estimates of the respective empirical distribution functions. A case of a FA signature is shown in Fig. 2.5, as a plot of percentiles computed from the estimated distributions. It is important to note that multiple signatures can be generated for one single DTI, as there are multiple features that can be derived from the diffusion images, and the tensors. For further details about the signature extraction process, refer to Section 4.1.4.

Each main step in the whole pipeline presented in Fig. 2.5 deal with very distinct problems. While in the last step the concerns are strictly related with how to establish the neighborhoods, and to represent the extracted features, the first step is concerned only with the issues related with defining a reliable geometrical reference for the corpus callosum extension, i. e., the median axis. Although in this overview the first step seen to be simpler than the last step, it is actually more complex, requiring a few specialized substeps to achieve the final goal. One very important substep, and the first one is to predict the plane of symmetry for the corpus callosum, where lies the final median axis. Although any conventional MSP could be used as a reference plane in this substep, we proposed a new method to estimate this plane from the symmetry of fibers from inside the own corpus callosum structure. Thus, in Chapter 3 we describe our proposed method to estimate the corpus callosum symmetry plane. The other substeps needed to estimate the final median axis, together with the procedures involved in the
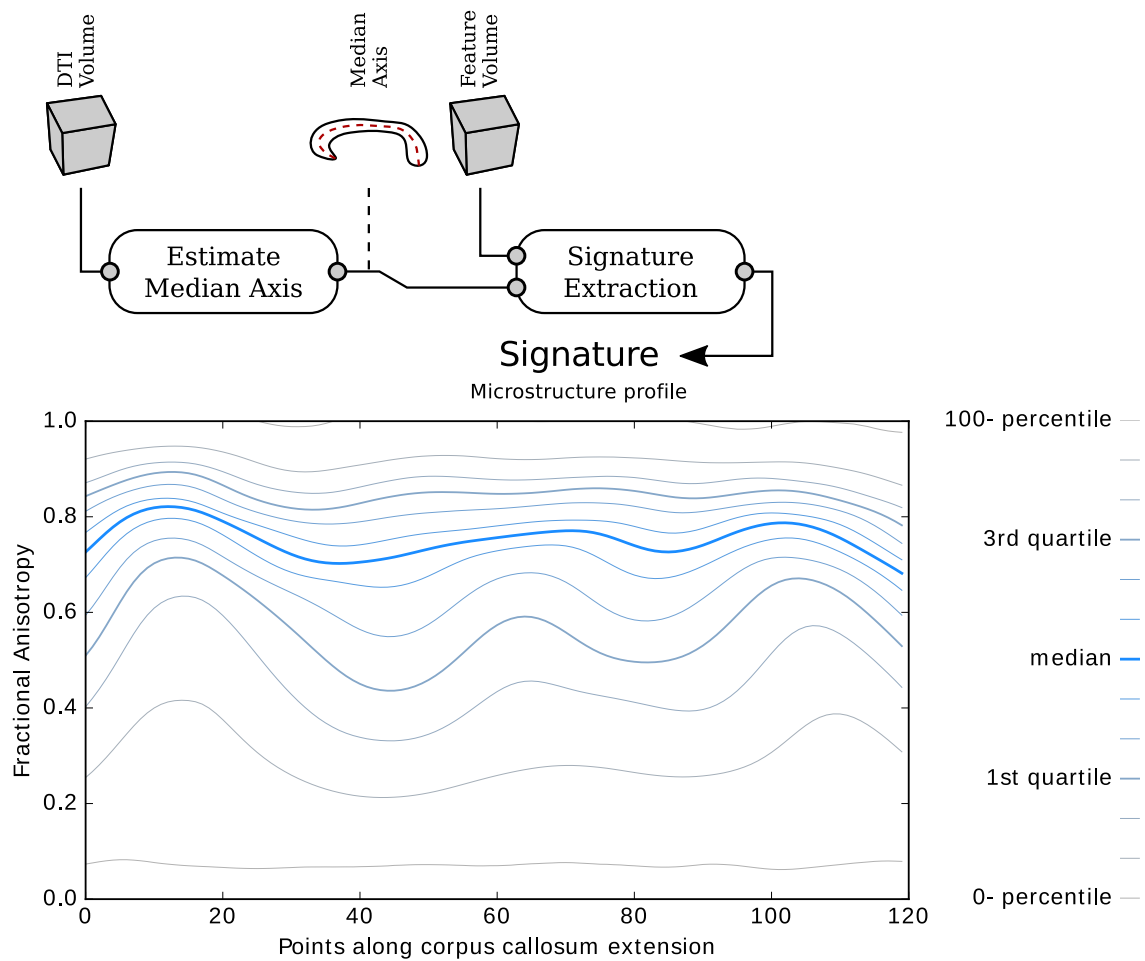
Figure 2.5: Overview for the signature extraction method with two main steps: first, estimate the corpus callosum median axis, which is a continuous curve in 3–D space from the anterior to the posterior extremes of the structure; second, given the median axis and a volume of microstructure features with same domain as the DTI volume, extract the features and generate a corresponding signature. The signature represents the variation of feature distributions across the structure, as illustrated in the plot of percentiles computed from the signature distributions.

signature generation step, are described in Chapter 4.

It is important to note that our proposed method assumes the DTI were already generated from the DWI, with the proper procedures to enhance image quality, such as motion correction and noise reduction. Although DTI quality enhancement is an issue that must not be neglected in any study, it is out of the scope of our method, and, therefore, is not included in the general pipeline we have depicted in Fig. 2.5. In this work, we used third–party software to generate the DTI from a DWI dataset, as explained in details at Section 2.4.

2.3   Related works

Analysis of the corpus callosum characteristics across the extension of its structure is neither new nor restricted to the microstructure assessment. For instance, by using a median corpus callosum axis drawn from the anterior to the posterior extremes of the structure, Downhill Jr. et al. [13] established evenly spaced perpendicular divisions to break the whole corpus callosum region into 30 consecutive smaller sections. Then, the areas from the sections define a 1–D profile function that can be used to compare correspondent sections from another subject. Just like in other works, including ours, the profile can also be plotted for better visualization of the localized attributes. Although Park et al. [36] have used the profile functions to assess the conditions of the microstructure, the fundamental concept is the same: to produce detailed measurements associated with relative locations along the corpus callosum extension. The correspondence on the profile relative positions from distinct cases is more clearly seen in the work of Downhill Jr. et al. [13]. Because there are fewer divisions that are thick and numbered, it is more evidence that a section from one image corresponds to the section from another image that happens to have the same associated number, i. e., is at an equivalent position relative to the corpus callosum extension.

Although Park et al. [36] do not use a median corpus callosum axis in their method, the points used to break the shape edge into two segments are equivalent to the points used to trace the median axis by Downhill Jr. et al. [13], which makes the implicit median axis in Park's method equivalent. However, the absence of an explicit median axis force Park et al. [36] to define the spacing between readings on another reference, which was the top shape segment, resulting in an uneven spacing relative to the median axis. Rittner et al. [42] have adapted Park's method to estimate the corpus callosum median axis, but due to the way the reading reference points were determined, they are also uneven in relation to the median axis. Defining a median axis may seem like a minor design detail at first glance, but it is actually the most important definition for extracting a profile for the corpus callosum extension because it eliminates ambiguity in the structure representation. In our work, the median axis is the ultimate reference for defining the reading points that generate the signatures, which are evenly spaced in relation to the median axis.

One very interesting property the profile functions have is that traditional corpus callosum partition schemata are as simple to implement as defining breaking points, as shown in Fig. 2.6. At the image domain instead, the partitioning process is complex and subject to mistakes that would potentially introduce ambiguity to the resulting partitions. However, there is a key difference in the sampling between the two methods that would

make the analyses in both domains distinct. In order for the analysis in the profile domain to be the same as the one done in the image domain, each reading point should be associated to a region in a similar fashion as in the work of Downhill Jr. et al. [13], with appropriate representation. Nonetheless, that should not be the goal, because the traditional partitioning is also a flawed process, as already discussed. Therefore, the real goal in our work was to improve the way the corpus callosum microstructure features can be analyzed. For that, we extract for each reference point across the median axis a complete representation for the distribution of surrounding values in all three dimensions. Such representation allows for any kind of comparison one intends to perform without recurring to the original image space. This makes our method very distinct from the ones proposed by other authors, besides the many other implementation particularities that we address with different approaches. In fact, the ways Park et al. [36] and Rittner et al. [42] sample the microstructure values associated with the reading points are inadequate, due to the extreme undersampling. In both methods, only one value is selected to represent each reading point, which is far from being sufficient for correctly represent the distributions associated with the reading locations.
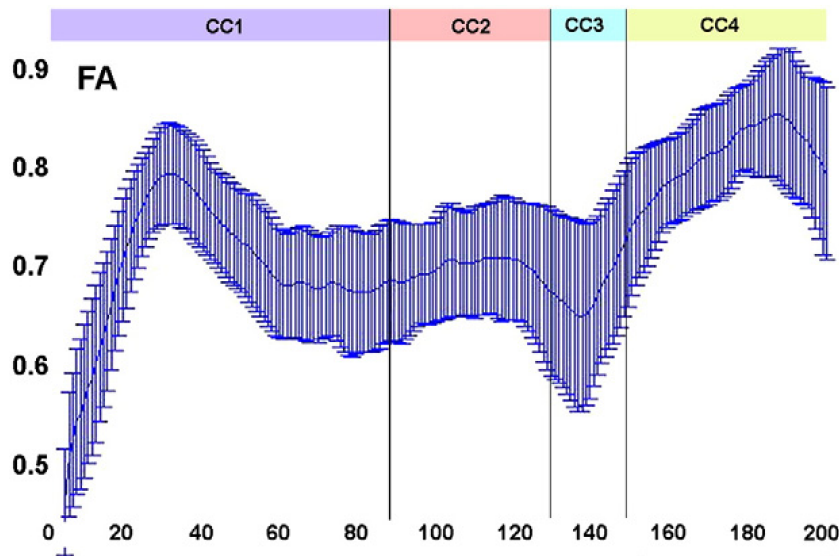
Figure 2.6: Profile of FA features computed by Park's method for a set of images, represented as average values ± standard deviation. In this plot there are four partitions, which corresponds to the topographical division of the corpus callosum based on Aboitiz's diagrams. Adapted from [36].

2.3.1 Mid–sagittal plane estimation methods

Even though the median axis is the geometrical reference in our method, the MSP still plays an important role, since it is the foundation for the median axis estimation. Choosing a proper MSP estimation method is not a trivial task because there are several

distinct implementations available. Most methods, however, are based on two distinct approaches, as appointed by Bergo et al. [6], one of which search for the plane that better define a bilateral symmetry of the head or brain; the other, aim at fitting a plane to the interhemispheric fissure. While the symmetry approach commonly considers the whole head instead of only the brain, some authors consider it as a better option because in some cases the interhemispheric fissure is not well defined, or it is so large that a degree of freedom is added to the plane estimation process [43, 39]. Thus, some methods are designed as hybrids from both approaches, like the one proposed by Stegmann et al. [43], where the plane is estimated as a function of the brain symmetry, but considering only a local portion closest to the interhemispheric fissure. It is clear that Stegmann et al. [43] were concerned about miss–alignment between the head symmetry plane and the interhemispheric fissure, which seems to be a reasonable concern as biological tissue can present several types of variations.

Independent from the approach an MSP estimation method uses, almost all solutions currently available were designed to work over anatomical images such as T1– weighted or CT, which has considerably higher spatial resolution than DTI. This implicates an intermodality registration for analyses that involve the corpus callosum microstructure assessment in DTI. Such a procedure will increase the complexity of the analysis pipeline and can potentially introduce errors. One way to avoid the intermodality registration would be to adapt one of the traditional methods to run over scalar maps derived from DTI or over the B_0 data volume from the DWI set. This solution, however, will be conditioned to a much lower spatial resolution and will rely on data that are not best representations for the brain's anatomical structures, which can potentially compromise the effectiveness of the plane estimation. One interesting solution proposed by Prima et al. [39] was to estimate the MSP as a function of the symmetry of DTI tensors. While Prima's method will be subject to the same spatial resolution problem, the tensors encode much richer information about the brain structure, the white matter fiber's organization. In practice, what their method does is to find the plane of symmetry relative to the brain fibers. The approach Prima et al. [39] employed in their method is especially valuable for the corpus callosum, as it is a white matter structure and composed essentially from fibers. In our work, we employed a very similar approach to estimate a reference plane exclusive for the corpus callosum, where we optimize a function of the symmetry of the tensor's preferential directions in a region of interest around the corpus callosum.

To understand the properties of our proposed plane estimation method, we compare results with the planes predicted by two traditional MSP estimation methods. One of them was proposed by Bergo et al. [6] and is based on the interhemispheric fissure, where the method goal is to maximize the cerebrospinal fluid area intercepted by the plane. The second method, proposed by Liu et al. [27], estimate the head and brain

global anatomical symmetry using the combination of several symmetry axes, which are computed independently for each of the axial slices in the image volume. These two methods were arbitrarily chosen, but following the criteria of being fully automated methods, and each one belonging to one of the two main categories of MSP methods. Our goal is to compare how is the alignment of the corpus callosum local fiber symmetry plane to the interhemispheric fissure, and to the brain/head global symmetry plane.

2.4 Dataset

The base dataset used in our experiments is composed of 80 pairs of images, each composed of a T1-weighted image and a DTI, both from the same subject and acquired during the same scan. The images were obtained between the years 2010 and 2012 using a 3T Philips Achieva MRI scanner at Unicamp from a sample of healthy subjects with ages between 8 and 60 years old, and an average age of 35 years (± 13). Of the images acquired, 23 were from males and 57 from females. All the patients were informed in advance and signed to indicate free and informed consent, and it was approved by the Research Ethics Committees of FCM–UNICAMP (number CEP 920/2007; CAAE: 0669.0.146.000-07).

We acquired the T1-weighted images using a SENSE protocol with spatial resolution $1\ mm \times 1\ mm \times 1\ mm$. The DTI were generated from *diffusion weighted images* (DWI), which were acquired using a SENSE protocol as $2\ mm$ axial slices with spatial resolution $1\ mm \times 1\ mm$ in 32 gradient directions with a b-value of 1000. We employed the software FSL5.0 [21] to fit the DTI from the acquired DWI.

2.4.1 Augmented dataset

The augmented dataset was used exclusively for an experiment designed to test the sensitivity of the MSP plane estimation methods. We created this new dataset with images generated from the original dataset by applying a rotation transform with the origin in the middle of the image grid, and adding Gaussian noise. The rotation transforms were defined as a combination of two elemental rotations around the y and z axes. The angle for each elemental rotation was defined randomly with a zero mean normal distribution, and $\sigma = 5°$. A total of five new images were generated from each of the original ones, with distinct rotations, being the augmented dataset composed of 480 pairs of images. Different levels of Gaussian noise were added to each of the new images, and the final signal–to–noise ratio was between 14 and 35 decibels relative to the signal

of the whole brain. For the DTI, we applied the noise to each of the eigenvalues and for the first eigenvector coordinates, simulating a random rotation. We normalized the eigenvectors after applying the noise.

2.4.2 Software and system configuration

All experiments were executed in a computer with an I7 processor, with four nuclei of 2 GHz each, and with 16 GiB of memory. All code was written in Python language, without the usage of multi–thread, except when this resource was implemented by the employed libraries. The libraries we have used were Numpy, Scipy, Nibabel for loading the Nifti formatted images, and PIL and matplotlib for generating images and plots for visualization purposes. All pre–processing for the images used in this work was done using the FSL 5.0 [21] software.

3. Mid–callosal plane estimation

Because the corpus callosum seamlessly connects to other white matter structures its boundaries are not enough defined for a complete 3–D segmentation. However, the corpus callosum does have a well-defined symmetry, which is similar to the whole brain symmetry. Therefore, as happens with the whole brain, in normal conditions, the reference for this symmetry is a plane, which is assumed to coincide with the whole brain symmetry plane, and is known as the *mid–sagittal plane* (MSP). Several computational methods were proposed to define the MSP in MR and CT images, but their methodology in most cases fall into two categories [6]: optimizing a plane of symmetry from the whole brain in structural images, almost always including the skull; or estimating the middle of the interhemispheric fissure. One very interesting method for defining the MSP was proposed by Prima et al. [39], which optimize the symmetry of tensors from a DTI in the whole brain. This method is very distinct from the others because it is a function of the microstructure of the brain instead of the macro (anatomical) structure. Nevertheless, this method also searches for the symmetry of the whole brain as many of the other methods.

Although it is evident that the corpus callosum is aligned with the interhemispheric fissure, we consider a mistake to assume it will always be well aligned enough to use the MSP as the reference. Therefore, in our method, we aim at finding a reference plane specific for the corpus callosum structure, which is estimated from the symmetry of its internal fibers. Provided we are no longer interested into finding the MSP, but instead a reference plane specific for the corpus callosum, it will be helpful to define a distinct name for this new plane to avoid future misunderstandings. Therefore, we call this new corpus callosum specific reference plane the *mid–callosal plane* (MCP). The MCP is a sagittal–oriented plane that split the corpus callosum into two symmetrical halves. Since the corpus callosum is a fibrous tissue, in essence, we expect that defining the MCP as a function of the symmetry of the corpus callosum internal fibers will lead to a more robust reference. Thus, in this work, we introduce a fully automated method to estimate the MCP using the preferential directions from tensors in a DTI. The basis for this new method comes from our previous exploratory study [9], which observed the symmetrical arrangement of the corpus callosum internal fibers. In addition, our method has a similar function to assess the fibers symmetry as in the method proposed by Prima et al. [39], which was a coincidence since we gained knowledge of this work only after our method was already implemented. Nonetheless, our methods have several distinctions, where the major ones are:

- our objective function uses only the tensor's preferential direction;

- our method is restricted to a region of interest around the corpus callosum instead of using the whole brain;

- and, our method has high computational efficiency, requiring just a few seconds to run in a conventional modern computer.

To validate our proposed method, we conducted experiments using 80 images from healthy subjects, and we compared the estimated MCP to the MSP predicted by two other fully automated methods: one was proposed by Liu et al. [27], and optimizes the head and brain anatomical symmetry; the other was proposed by Bergo et al. [6], and searches for a plane with a better fit to the inter-hemispheric fissure. Those two methods are representatives of the two most usual categories of automatic MSP estimation methods [6].

This Chapter is organized as follows: Section 3.1 describe our proposed method, detailing the procedures involved in estimating the MCP; the experimental procedures, results and discussion are described in Section 3.2; and, finally, Section 3.3 provides a conclusion for this Chapter.

3.1   Method description

Let $V : \mathcal{S} \rightarrow \mathbb{R}^3$, be a vector field given by the eigenvectors associated to the major eigenvalues from tensors of a DTI, which represents the preferential diffusion orientations. The subset $\mathcal{S} \in \mathbb{R}^3$ is the volume image in a standard space with basis given by the canonical vectors $\vec{e}_x = (1, 0, 0)$, $\vec{e}_y = (0, 1, 0)$, and $\vec{e}_z = (0, 0, 1)$, respectively pre–aligned with the patient axes left–right, anterior–posterior, and superior–inferior, as illustrated in Fig. 3.1. The method goal is to find a plane $\Pi_{(\bar{\theta},\bar{\phi},\bar{\rho})}$ that defines an optimal symmetric arrangement of the CC internal fibers, as

$$\Pi_{(\bar{\theta},\bar{\phi},\bar{\rho})} = \underset{\Pi_{(\theta,\phi,\rho)}}{\arg\min} \, \mathrm{Sym} \left(\Pi_{(\theta,\phi,\rho)} \right), \tag{3.1}$$

where $\Pi_{(\theta,\phi,\rho)}$ is a candidate plane with inclination defined by a rotation transform $R(\theta, \phi)$, as in 3.3, and with displacement defined by a point $\rho \in \mathcal{S}$.

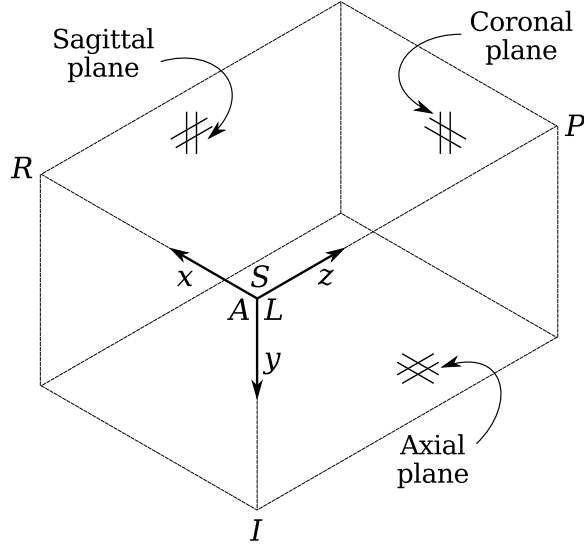The inclination of a candidate plane $\Pi_{(\theta,\phi,\rho)} : ax + by + cz + d = 0$ is defined

Figure 3.1: Representation of the volume orientation employed in this work.

by the normal vector

$$\vec{n}_\Pi = \begin{bmatrix} a \\ b \\ c \end{bmatrix} = R\,\vec{e}_x = R \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \tag{3.2}$$

where R is a 3×3 rotation matrix given by any combination of other rotation matrices. Now, let $R = R(\theta, \phi)$ be a rotation transform defined by the combination of a rotation $R_z(\theta)$ about the z axis, and a rotation $R_y(\phi)$ about the y axis, then

$$\begin{aligned} R(\theta, \phi) &= R_y(\phi)\; R_z(\theta) \\ &= \begin{bmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \cos\phi\cos\theta & \cos\phi-\sin\theta & \sin\phi \\ \sin\theta & \cos\theta & 0 \\ -\sin\phi\cos\theta & -\sin\phi-\sin\theta & \cos\phi \end{bmatrix}. \end{aligned} \tag{3.3}$$

Replacing 3.3 in 3.2, we have

$$\begin{aligned} \vec{n}_\Pi &= \begin{bmatrix} \cos\phi\cos\theta & \cos\phi-\sin\theta & \sin\phi \\ \sin\theta & \cos\theta & 0 \\ -\sin\phi\cos\theta & -\sin\phi-\sin\theta & \cos\phi \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \cos\phi\cos\theta \\ \sin\theta \\ -\sin\phi\cos\theta \end{bmatrix}. \end{aligned} \tag{3.4}$$

Therefore, the inclination of a candidate plane $\Pi_{(\theta,\phi,\rho)}$ can be expressed as two elemental rotations about the y and z axes. This will be handful to specify the search space for our method.

In a similar way, we can rotate all three canonical vectors to define a new basis $B(\theta, \phi) = \{\vec{v}_x, \vec{v}_y, \vec{v}_z\}$ of $\mathbb{R}^3$, obtaining the change of basis matrix

$$
\begin{aligned}
C_{(\theta,\phi)} &= \begin{bmatrix} \vec{v}_x & \vec{v}_y & \vec{v}_z \end{bmatrix} \\
&= R(\theta, \phi) \begin{bmatrix} \vec{e}_x & \vec{e}_y & \vec{e}_z \end{bmatrix} \\
&= \begin{bmatrix} \cos\phi\cos\theta & \cos\phi - \sin\theta & \sin\phi \\ \sin\theta & \cos\theta & 0 \\ -\sin\phi\cos\theta & -\sin\phi - \sin\theta & \cos\phi \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \cos\phi\cos\theta & \cos\phi - \sin\theta & \sin\phi \\ \sin\theta & \cos\theta & 0 \\ -\sin\phi\cos\theta & -\sin\phi - \sin\theta & \cos\phi \end{bmatrix}.
\end{aligned}
\tag{3.5}
$$

This is essential for evaluating the symmetry between pairs of vectors in V, as explained in the next Section.

3.1.1   Vector symmetric difference

For the purposes of the MCP method the definition of symmetry between vectors is directly related to the plane inclination, i. e., the same pair of vectors can be either symmetric in relation to a given plane Π', and asymmetric in relation to another plane Π'', where $\Pi' \parallel\!\!\!/ \; \Pi''$. Let $\vec{v}_a \in V$ be a vector with coordinates in the standard basis, and $[\vec{v}_a]_{B(\theta,\phi)} = (v^x, v^y, v^z)$ be the coordinates of $\vec{v}_a$ in relation to the basis $B(\theta, \phi)$ defined by the inclination of a candidate plane $\Pi_{(\theta,\phi,\rho)}$. Then, a vector $\vec{v}_b \in V$ is considered symmetric to the vector $\vec{v}_a$ if $[\vec{v}_b]_{B(\theta,\phi)} = (-v^x, v^y, v^z)$. Figure 3.2 illustrate two symmetric vectors $\vec{v}_a$ and $\vec{v}_b$ in relation to a basis $B(\theta, \phi) = \{\vec{v}_x, \vec{v}_y, \vec{v}_z\}$ defined by a plane $\Pi = \Pi_{(\theta,\phi,\rho)}$. By this definition the plane act like a mirror, where two symmetric vectors are the reflection of each other. However, given we want to measure the symmetry of orientation, the vector $\vec{v}_c = -\vec{v}_b$ can also be defined as symmetric to $\vec{v}_a$, where $[\vec{v}_c]_{B(\theta,\phi)} = (v^x, -v^y, -v^z)$.

While the definition of symmetry between vectors here established clearly depends on the inclination of the candidate plane, we can define the *vector symmetric difference* operator in terms of the vectors components, assuming the proper change of basis for the vectors to be already done. Thus, let $\vec{u}_1 = (x_1, y_1, z_1)$ and $\vec{u}_2 = (x_2, y_2, z_2)$
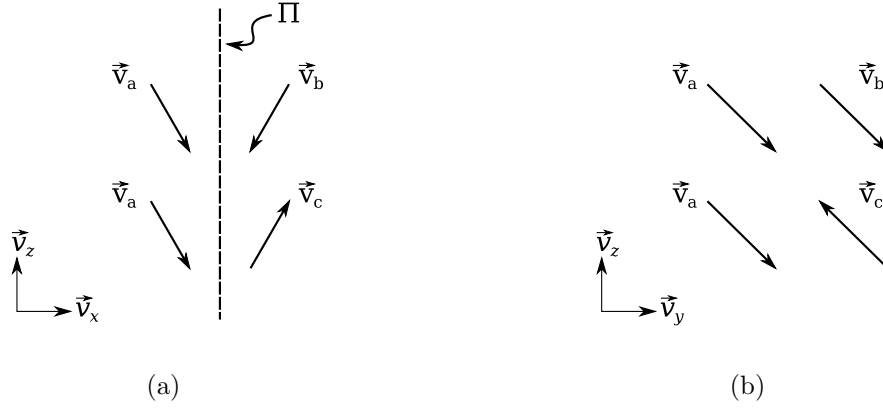
Figure 3.2: Illustration of three orientation symmetric vectors in relation to a plane $\Pi = \Pi_{(\theta,\phi,\rho)}$.

be two vectors of same length in any appropriate basis, the symmetric difference is given by

$$\mathrm{SD}(\vec{u}_1, \vec{u}_2) = |x_1 + x_2| + |y_1 - y_2| + |z_1 - z_2|. \tag{3.6}$$

Now, let $\vec{v}_1$ and $\vec{v}_2 \in V$ be two vectors in the standard basis, and let $B = B(\theta, \phi)$ be a basis defined by the inclination of a candidate plane $\Pi_{(\theta,\phi,\rho)}$. The symmetric difference between $\vec{v}_1$ and $\vec{v}_2$ is

$$\mathrm{SD}_B(\vec{v}_1, \vec{v}_2) =$$
$$\min\left(\mathrm{SD}\left(\left[\frac{\vec{v}_1}{||\vec{v}_1||}\right]_B, \left[\frac{\vec{v}_2}{||\vec{v}_2||}\right]_B \right), \mathrm{SD}\left(\left[\frac{\vec{v}_1}{||\vec{v}_1||}\right]_B, -\left[\frac{\vec{v}_2}{||\vec{v}_2||}\right]_B \right) \right), \tag{3.7}$$

where $[\vec{v}]_B = C_{(\theta,\phi)}^{-1}\vec{v}$ are the coordinates in relation to the basis $B = B(\theta, \phi)$ of the vector $\vec{v}$.

3.1.2 Selection of eigenvector pairs

Let $\vec{v}_1 = V(p_1)$ and $\vec{v}_2 = V(p_2)$ be two eigenvectors respectively associated to the points $p_1 = (\rho_1^x, \rho_1^y, \rho_1^z)$ and $p_2 = (\rho_2^x, \rho_2^y, \rho_2^z) \in \mathcal{S}$. Recall the mirror analogy, the points p_1 and p_2 must be equidistant to the candidate plane $\Pi_{(\theta,\phi,\rho)}$, and lie in a line with same orientation as the normal vector $\vec{n}_\Pi$ of $\Pi_{(\theta,\phi,\rho)}$. Hence, the symmetric difference between $\vec{v}_1$ and $\vec{v}_2$ is meaningful only in relation to the plane $\Pi_{(\theta,\phi,\rho)}$ with

$$\vec{n}_\Pi = p_2 - p_1 = \begin{bmatrix} \rho_2^x - \rho_1^x \\ \rho_2^y - \rho_1^y \\ \rho_2^z - \rho_1^z \end{bmatrix} \quad \text{or} \quad \vec{n}_\Pi = p_1 - p_2 = \begin{bmatrix} \rho_1^x - \rho_2^x \\ \rho_1^y - \rho_2^y \\ \rho_1^z - \rho_2^z \end{bmatrix}, \tag{3.8}$$

that intercept the point $p = \frac{p_1 + p_2}{2}$. Therefore, the choice of p_1 and p_2 fully determine the candidate plane $\Pi_{(\theta,\phi,\rho)}$.

Inversely, the determination of p_1 and p_2 can be made by first choosing the candidate plane $\Pi_{(\theta,\phi,\rho)}$. Let $p \in \Pi_{(\theta,\phi,\rho)}$ be a pivot point in the candidate plane, and $\vec{n}_\Pi$ be the normal vector for $\Pi_{(\theta,\phi,\rho)}$, with $||\vec{n}_\Pi|| = 1$. Then, $p_1 = p + \delta\vec{n}_\Pi$ and $p_2 = p - \delta\vec{n}_\Pi$. This definition is suitable for defining a set of vector pairs to measure the degree of symmetry of a given candidate plane. Therefore, let $\mathcal{P} = \{\text{SD}_B(\vec{v}_1, \vec{u}_1) \ldots \text{SD}_B(\vec{v}_n, \vec{u}_n)\}$ be a set of symmetric differences between n pairs of vectors determined by a candidate plane $\Pi_{(\theta,\phi,\rho)}$. Then, the degree of symmetry associated with $\Pi_{(\theta,\phi,\rho)}$ is given by

$$\text{Sym}(\Pi_{(\theta,\phi,\rho)}) = \text{median}(\mathcal{P}). \tag{3.9}$$

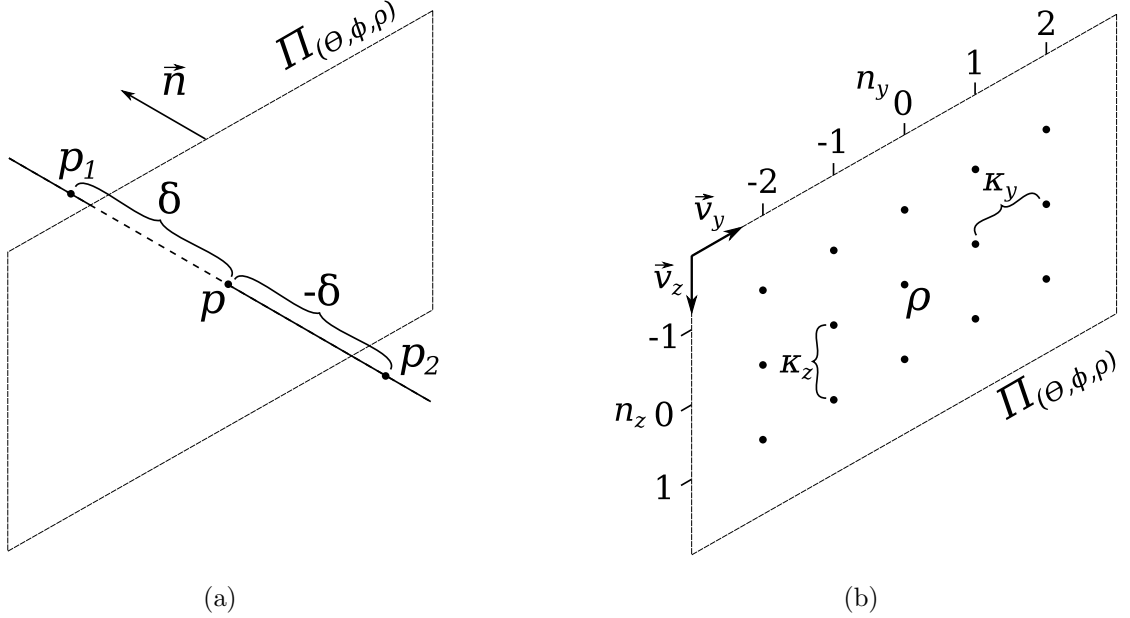(a)   (b)

Figure 3.3: Selection of eigenvector pairs: (a) how a pair of eigenvector locations are determined from a pivot point $p \in \Pi_{(\theta,\phi,\rho)}$; (b) illustrate how pivot points are defined.

Although the vector pairs can be defined by randomly varying the pivot point p and the distance δ from the plane, it is better to use a regular approach to maximize coverage of the possible vector pairs, and minimize repetition. Therefore, each pivot point is determined according to a position (n_y, n_z) in a bi–dimensional grid as

$$p_{(n_y, n_z)} = \rho + n_y \kappa_y \vec{v}_y + n_z \kappa_z \vec{v}_z, \tag{3.10}$$

where $\rho \in \Pi_{(\theta,\phi,\rho)}$ is a origin point that defines the candidate plane $\Pi_{(\theta,\phi,\rho)}$, n_y and $n_z \in \mathbb{Z}$ define the position in the grid, and κ_y and $\kappa_z \in \mathbb{R}$ are the steps respectively in the directions of the vectors $\vec{v}_y$ and $\vec{v}_z$ from the basis $B(\theta, \phi)$ defined in 3.5. The steps κ_y and

κ_z are defined according to the spatial resolution of the DTI. For each pivot point $p_{(n_y,n_z)}$ a number of different distances δ from the candidate plane can be defined.

3.1.3 Corpus callosum segmentation

The corpus callosum is a brain structure composed essentially by neural fibers, being the water diffusion inside it expected to be anisotropic. In sagittal–like planes, the corpus callosum is highly distinguishable in *fractional anisotropy* (FA) images, as can be seen in the sample shown in Fig. 3.4(a). But there are other fibrous structures adjacent to the corpus callosum that would require the segmentation method to be more complex that it should be. However, the fiber orientations from these other structures are very different from the orientations inside the corpus callosum. While in the corpus callosum the fibers tend to be perpendicular to the sagittal plane, in the other adjacent structures the fibers are prone to be parallel to the plane. Therefore, using the fiber orientation information to filter the FA image, we can obtain a more suitable image to segment the corpus callosum, like the one shown in Fig. 3.4(b).

(a)   (b)
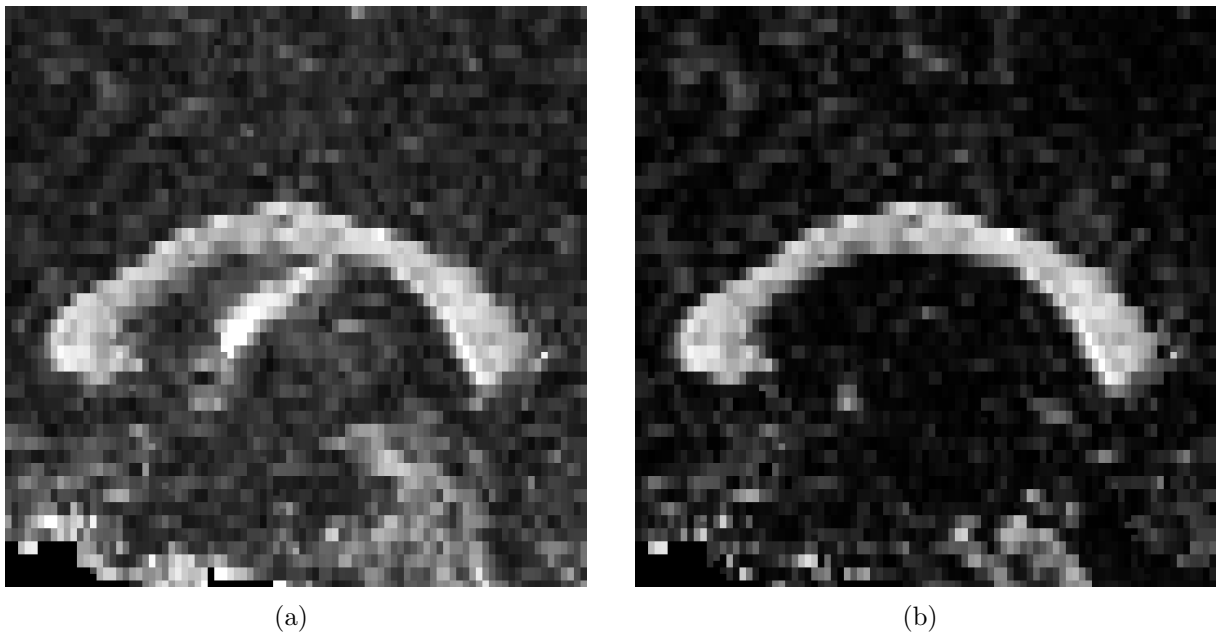
Figure 3.4: The corpus callosum in a candidate plane $\Pi_{(\theta,\phi,\rho)}$: (a) FA values only; (b) FA values weighted according to the orientation of the first eigenvectors.

For the MCP method purpose, the segmentation of the corpus callosum $\mathcal{S}g^{\Pi}$ is always defined in a plane $\Pi = \Pi_{(\theta,\phi,\rho)}$ as a bi–dimensional segmentation. Let $FA : \mathcal{S} \to \mathbb{R}$ be a volume with the FA values for the voxels in $\mathcal{S}$; and β be a binary 2–D image defined as

$$\beta(p) = \begin{cases} 1 & \text{if} \quad FA(p) \, \|\operatorname{proj}_{\vec{n}_\Pi} V(p)\| \geq t \\ 0 & \text{otherwise} \end{cases}, \tag{3.11}$$

where $p \in \Pi = \Pi_{(\theta,\phi,\rho)}$, $\| \text{proj}_{\vec{n}_\Pi} V(p)\|$ is the length of the orthogonal projection of the eigenvector $V(p)$ over the normal vector $\vec{n}_\Pi$, and $t \in \mathbb{R}$ is a threshold. Then, the segmentation $\mathcal{S}g^\Pi = \{p_1, p_2, \ldots, p_n\}$, with $p_n \in \Pi$, is defined by the largest connected component from β, including the holes, if any.

3.1.4 Optimal plane search

The first step to find the optimal plane $\Pi_{(\theta,\phi,\rho)}$ is to define an initial origin point $\rho \in \mathcal{S}$. Let $\Pi_d : x = d$ be the sagittal plane closest to the interhemispheric fissure, defined using the method proposed by Freitas et al. [14]. Then, the origin ρ is defined by the mass center of the corpus callosum segmentation $\mathcal{S}g^{\Pi_d}$ in the plane Π_d. The initial origin and the plane Π_d represent a starting guess to the search for the optimal symmetry plane $\Pi_{(\bar{\theta},\bar{\phi},\bar{\rho})}$. Whenever a candidate plane $\Pi_{(\theta,\phi,\rho)}$ is tested for having a better degree of symmetry than all the other previously tested, the origin point ρ is updated to be the corpus callosum mass center for that candidate plane $\Pi_{(\theta,\phi,\rho)}$.

The next step is to predict the inclination of the optimal plane $\Pi_{(\theta,\phi,\rho)}$. Let $\Pi_{(\theta,\phi,p_i)}$ be a candidate plane with inclination determined by the rotation matrix $R(\theta,\phi)$, and displacement given by a point $p_i \in P_{(\theta,\phi)}$, defined as

$$p_i = \rho + \Delta_i \vec{n}_\Pi, \tag{3.12}$$

where $\Delta_i \in \mathbb{R}$ is the distance from the plane $\Pi_{(\theta,\phi,p_i)}$ to the origin ρ previously defined. Then, the optimal degree of symmetry for the inclination $R(\theta,\phi)$ is

$$\text{Sym}^R(\theta,\phi) = \min_{p_i \in P_{(\theta,\phi)}} \text{Sym}(\Pi_{(\theta,\phi,p_i)}). \tag{3.13}$$

Figure 3.5(a) illustrate a bi–dimensional sample generated by the cost function $\text{Sym}^R(\theta,\phi)$, with both θ and ϕ varying in the range $[-12, 12]$ degrees, every $1°$. The sample show a radial pattern for the cost function, with a well defined global optimum. Thus, the optimal inclination $R(\bar{\theta},\bar{\phi})$ is given by the global minimum from the interpolated cost function $\hat{\text{Sym}}^R(\theta,\phi)$, illustrated in Fig. 3.5(b).

Because the optimal inclination $R(\bar{\theta},\bar{\phi})$ was estimated from an interpolated sample, it is most likely that none of the tested candidate planes $\Pi_{(\theta,\phi,\rho)}$ had the optimal inclination. Therefore, one final step is needed to find the displacement of the optimal plane $\Pi_{(\bar{\theta},\bar{\phi},\bar{\rho})}$, given by

$$\bar{\rho} = \arg\min_{p_i \in P_{(\bar{\theta},\bar{\phi})}} \text{Sym}(\Pi_{(\bar{\theta},\bar{\phi},p_i)}). \tag{3.14}$$

$$Sym^R\,(\theta,\phi) \qquad\qquad \hat{Sym}^R\,(\theta,\phi)$$
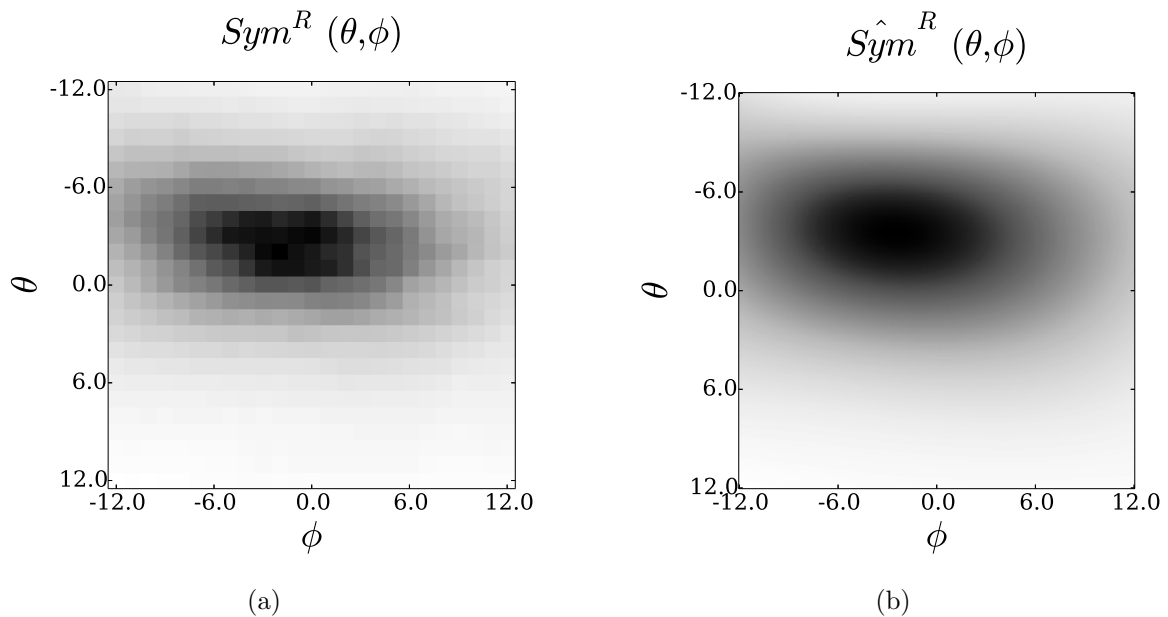
(a)   (b)

Figure 3.5: Sample space generated by the function Sym^R, which measure the degree of symmetry associated with a given plane inclination: (a) raw samples from a test; (b) interpolated samples used to predict the optimal plane inclination.

3.2   Experiments

The three methods we have used in the experiments were implemented by us in Python language, using the Numpy framework, along with several libraries, including Scipy, nibabel and skimage. The software FSL5.0 was used to perform several tasks involving the dataset, including fitting of the DTI model, skull–stripping, and multimodality registration.

Our method was implemented as described in Section 3.1, using for the selection of eigenvector pairs a grid of pivot points spaced by $1mm$ x $1mm$. Each pivot point defines three pairs of eigenvectors, with $\delta \in \{0.5mm, 1.5mm, 2.5mm\}$. A pair is discarded if one of its eigenvectors is associated with a fractional anisotropy value below 0.4, increasing the likelihood of the symmetrical difference between a valid pair to be meaningful. The pivot point grid is large enough to contain the entire corpus callosum. The search range was $\theta = [-12°, +12°]$, and $\phi = [-12°, +12°]$ for the inclination, with displacement given by $\Delta_i = [-5mm, +5mm]$. These parameters were defined arbitrarily based on the problem characteristics, and also according to results from preliminary tests.

The implementations of Liu's and Bergo's methods were made following the descriptions found respectively in [27] and [6], with some adaptations. For Liu's method, we employed a robust linear regression using the RANSAC approach, instead of using the least–median of squares as did the method authors. For Bergo's method, we have

significantly modified the plane location stage, introducing the concept of an origin point that is a reference for all transformation applied to the starting plane. The origin is the mass center point computed from the final brain mask found in the preprocessing stage. Also, in our implementation, there is no rotation around the axis perpendicular to the sagittal plane. The rotations were in range $[-5°, +5°]$, and the translations in range $[-4mm, +4mm]$.

3.2.1 Inter method comparison

The purpose of this experiment is to assess the difference between the planes estimated by each of the three methods, i. e.., the methods proposed by Liu and Bergo as well as our proposed method. The difference between planes is being expressed as a combination of two rotations θ and ϕ about the y and z axes, respectively, and a displacement computed as the distance from a point to a plane. The point is the center of mass of the corpus callosum segmentation at a reference plane, which is the one estimated by our method, except for the comparison of Bergo's method with Liu's method, where the plane predicted by Bergo's method is the reference. Figure 3.6 shows the differences between all computed planes.
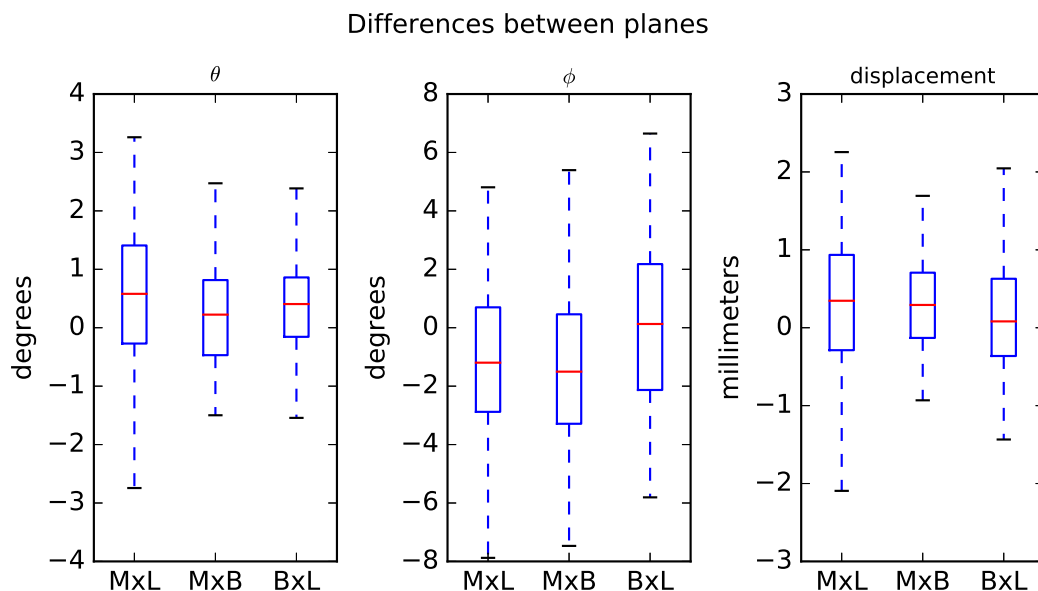
Figure 3.6: Differences between the planes estimated by our proposed method vs the planes given by the methods of Liu (MxL) and Bergo (MxB); and the differences between the methods of Bergo and Liu (BxL).

The results reveal that the planes estimated by our method were very close to the planes estimated by the methods of Liu and Bergo. The differences between our method and Bergo's method were slightly lower than those between our method and Liu's

method with respect to the dispersion of the angle θ and the displacement. The dispersion of differences between the methods of Bergo and Liu were also similar.

However, by analyzing the centrality of the distributions of differences, as indicated by the median, we observe that there are some biases in the results. If the differences between the methods were just random, the center of the distributions should be near zero. The most significant bias is associated with the angle ϕ, indicating that the planes predicted by our method tend to have a distinct inclination relative to the coronal plane as opposed to the planes estimated by the methods of Liu and Bergo. The bias associated with the displacement follows the pattern of the bias related to the angle ϕ, indicating that the bias in the displacement is caused by the bias in the angle ϕ. Finally, the bias associated with the angle θ has a different pattern, which indicates that the planes predicted by our method and the method of Bergo had more similar inclinations relative to the axial plane compared to Liu's planes.

Despite the small differences between estimated planes, there are cases where the differences can be significant, highlighting the misalignment between the corpus callosum symmetric median plane and the other planes, which are predicted according to other brain structures. In some cases (Fig. 3.7), it is clear that the planes estimated by our method were better positioned with respect to the corpus callosum anatomical symmetry.

## 3.2.2   Sensitivity experiment

This experiment was designed to assess whether a method is able to estimate an equivalent plane on images from the same subject but from different acquisitions. This is an essential property for longitudinal studies and is an indicator that planes estimated for distinct subjects are also equivalent, which is harder to assess. However, it is known that distinct acquisitions from the same subject remain susceptible to variations due mainly to different positioning and noise. Therefore, we do not expect a method to give an exact estimate of the same plane, but a very close one. Here, we employ the same metric used in the inter–method difference experiment described in Section 3.2.1.

The procedure was to test the planes estimated for the generated images compared with the ones predicted for the original images, and this was done by measuring the differences between them.

According to the results (Fig. 3.8), our method and the method proposed by Bergo had low sensitivity to the initial conditions and the noise applied to the images,
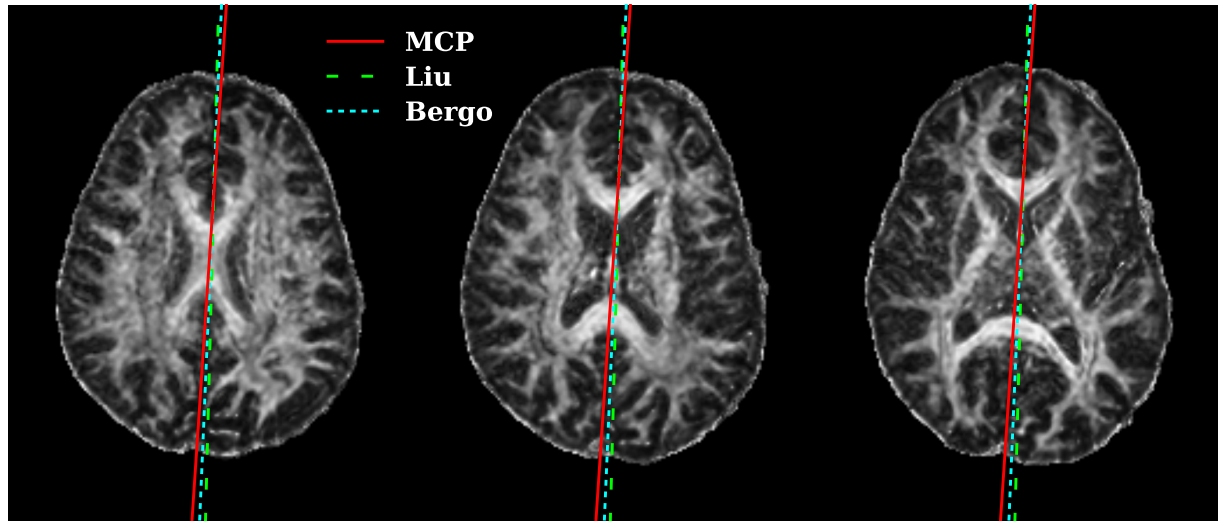
(a) female, 25yo
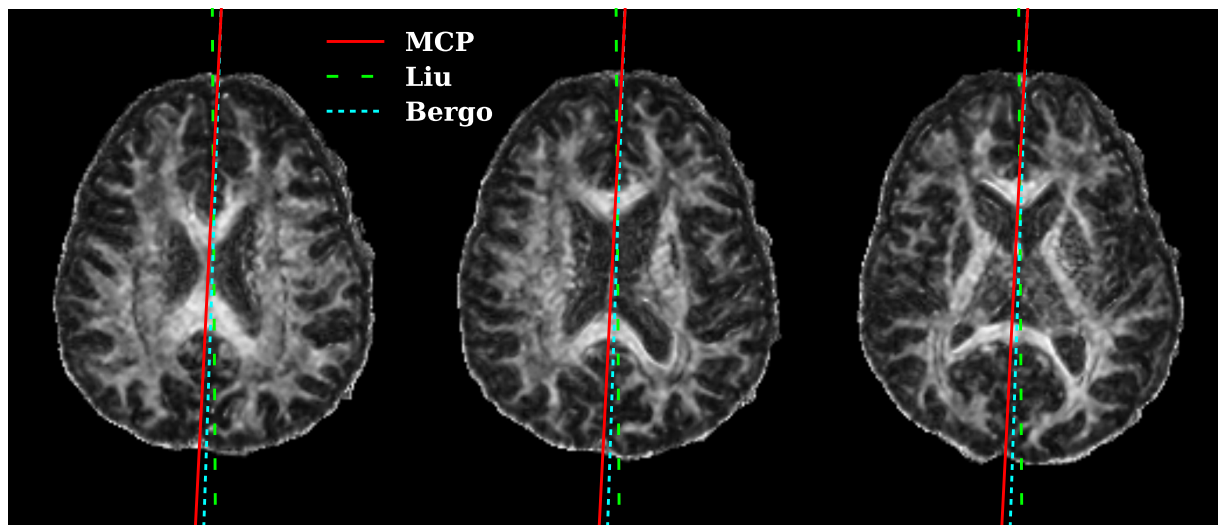
(b) female, 26yo

Figure 3.7: Visualization of the estimated planes for two sample images. Each visualization is composed of three axial slices taken perpendicular to the MCP estimated by our method, and they are parallel to each other. From left to right, the slices pass through the body of the corpus callosum, the middle of the corpus callosum just below the body, and the anterior and posterior extremes.
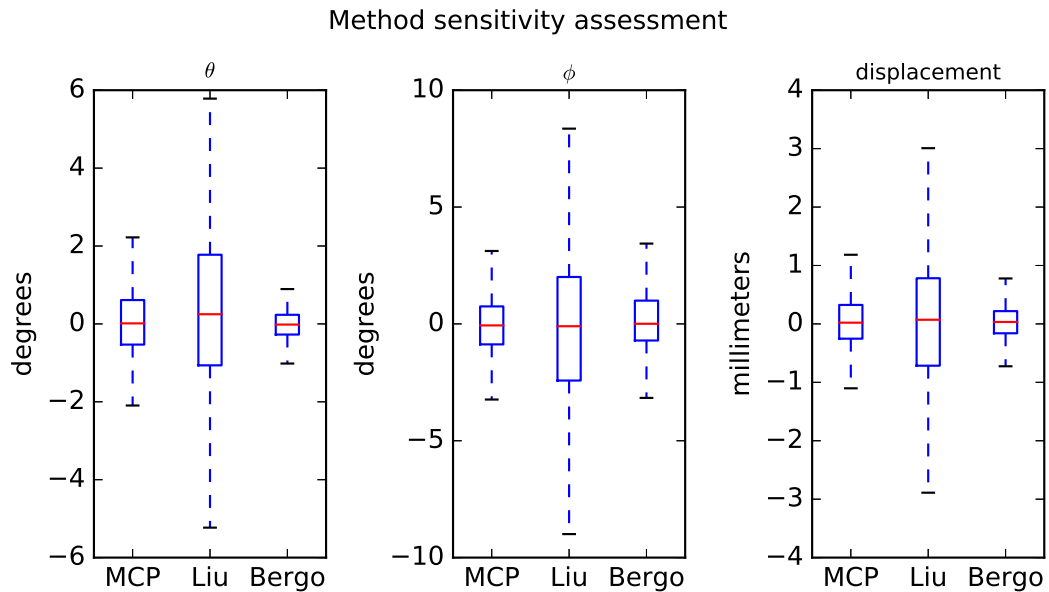
Figure 3.8: Sensitivity test results: difference between the planes estimated for the original image and the planes estimated for the corresponding images generated by each of the methods used in this study.

while the method proposed by Liu had a significantly higher sensitivity. In addition, because the medians were very near to zero for our method and the method of Bergo, the differences were random, i.e., these methods are unbiased. However, for Liu's method, we observe a small bias in the angle θ. This bias may influence the results associated with the angle θ shown in Fig. 3.6 in the inter-method difference experiment. If that is the case, removing this influence would make the pattern of the bias associated with the angle θ be similar to the ones found for the angle ϕ and the displacement. Because this scenario is likely to be true, we can confirm that the biases found in the inter-method experiment indicate that the planes estimated by our method were distinct from the planes estimated by the other methods. We consider two main potential sources for the bias: one is the intermodality registration procedure, and the other is the natural characteristics of the brain structures present on the dataset images.

3.2.3 Impact analysis

The experiments described in this Section were conducted to provide insights into the relevance of the distance between estimated planes in regards to the analysis of the corpus callosum characteristics.

Overlap experiment

It is known that in continuous space the intersection between two nonparallel planes is a straight line. However, in the discrete space of digital images, the intersection of two nonparallel planes is a set of voxels defined by the inclination and displacement of the planes and the image spatial resolution. Thus, this experiment was designed to map the angular distance and displacement between two planes to an expected percentage of the corpus callosum voxels that would be common to both planes.

Let Π_o be the estimated MCP, and $o \in \mathbb{R}^3$ be a point given by the mass center of the corpus callosum segmentation S_o relative to the plane Π_o. Then, another plane $\Pi_{(\theta,\phi,d)}$ generated by rotation of Π_o around the point o is used to define another corpus callosum segmentation $S_{(\theta,\phi,d)}$. The rotation transform is the combination of two elemental rotations respectively with θ and ϕ angles. Also, a displacement d if applied to this new plane. Thus, the percentage $\mathcal{P}_{(\theta,\phi,d)}$ of corpus callosum voxels common to both planes Π_o and $\Pi_{(\theta,\phi,d)}$ is computed as

$$\mathcal{P}_{(\theta,\phi,d)} = \frac{2 \, |S_o \cap S_{(\theta,\phi,d)}|}{|S_o| + |S_{(\theta,\phi,d)}|}, \tag{3.15}$$

which is a measure widely known as DICE similarity. The map shown in Fig. 3.9 was generated by averaging the values of 3.15 for several planes with $\theta = [-7°, +7°]$, $\phi = [-7°, +7°]$, and $d = [-5mm, +5mm]$. It gives an expected overlap percentage of the corpus callosum voxels for a given angular distance and displacement for images with spatial resolution of $1mm$ x $1mm$ x $1mm$.

It is noticeable that there is a relationship regarding the influence of angular distance and displacement over the intersection of corpus callosum voxels. For angular distances under 1°, the displacement has a stronger influence, being the expected overlap percentage down to zero as the displacement goes over $1mm$. That happens because the planes are almost parallel, and as the displacement exceeds the spatial resolution the intersection of the planes moves out of the working space subset. For larger angular distances, however, there is greater tolerance for the displacements between planes that will result in the intersection of corpus callosum voxels.

Using the map presented in Fig 3.9, it is possible to estimate the expected percentage of common corpus callosum voxels for the results obtained in the previous experiments described in Sections 3.2.1 and 3.2.2. Figure 3.10 show the correspondent expected overlap percentages computed for those previous experiment results.

It is now evident that the planes predicted by our method are closer to the
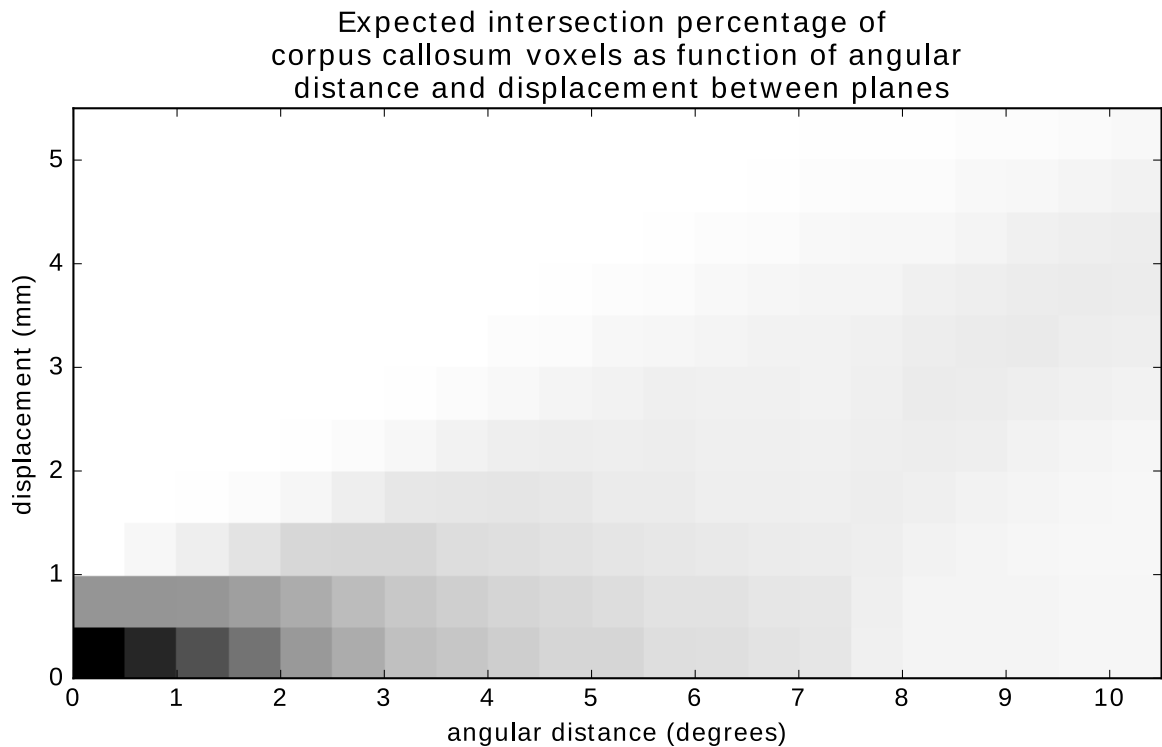
Figure 3.9: Overlap map computed from experimental data obtained by applying random rotation and translation transforms to MCP estimated for one arbitrarily selected image. Each position in the map represents the expected percentage of voxels inside the corpus callosum that are intercepted by both the original and the transformed planes, being black and white colors respectively correspondent to 100% and 0%.

planes predicted by Bergo's method than the ones by Liu's method, in regards to the image discrete space. In addition, the divergences found in the sensitivity experiment from Section 3.2.2 exert significant influence over the expected overlap ratio, which scores on average 60% at the best scenario.

3.2.4 Performance issues

The computational cost is a very important factor in an analysis pipeline. Although not as important as solving the problem, a tool must be also efficient to help increase the pipeline output. In that regards, our method performed very well, taking about 8 seconds to estimate the MCP. The methods of Liu and Bergo took respectively about 22 seconds and 55 seconds. All methods were very consistent about execution time.

Further, in order to use Liu's or Bergo's methods, a multimodality registration is required, adding not only time to execute but also complexity to the analysis pipeline. This is especially impactful for analysis pipelines that rely exclusively on DTI information.
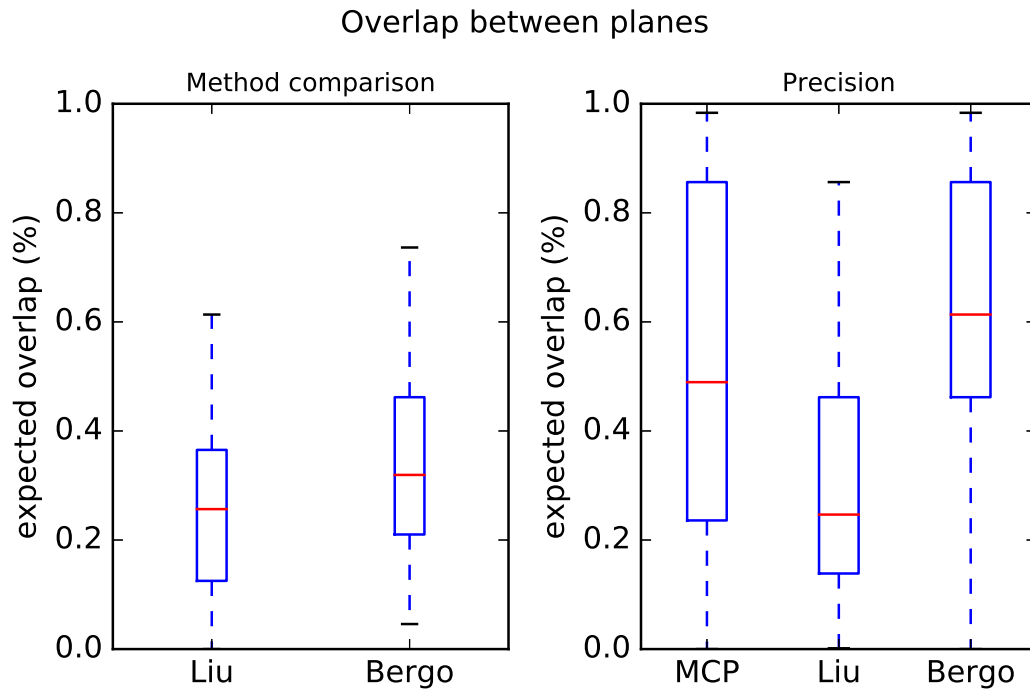
Figure 3.10: Expected overlap between planes estimated in previous experiments by crossing the plane divergences with the data in the overlap map from Fig. 3.9: (a) expected overlap ratios related to the divergences between planes estimated with Liu's and Bergo's methods against our proposed method; (b) expected overlap related to the divergences registered in the sensitivity experiment.

3.3   Chapter conclusion

In this Chapter, we introduced a method to estimate the corpus callosum reference plane by optimizing the symmetry of its fibers encoded in diffusion tensor images. According to experimental results, the new method is capable of estimating a valid and reliable reference for the corpus callosum analysis. The proposed method is also very efficient, being significantly faster than the other methods involved in this study. In addition, using the new method simplifies the analysis pipeline for studies that require only DTI data.

As expected, the overall results show the difference between MCP and MSP are subtle. Nonetheless, we could confirm that the middle of the corpus callosum, the interhemispheric fissure and the middle of the head are not always well aligned, despite in most cases the differences between planes did not produce a significant impact over the FA and MD distributions inside the corpus callosum. In regards to that matter, any of the tested methods would be suitable for the corpus callosum 2–D analysis.

The proposed method has the intrinsic advantage of performing the optimiza-

tion of features that are directly and exclusively related to the corpus callosum. This makes the plane it estimates the canonical corpus callosum reference. In addition, the fact that almost all the known methods for estimating the MSP require a multi–modality registration, which adds complexity and computational cost to the analysis pipeline, is also a strong factor for choosing the method proposed in this work instead. Therefore, we strongly recommend the proposed method as the standard for future studies involving the characterization of the corpus callosum.

4. Corpus callosum signatures

A corpus callosum signature, in the context of this work, is an array–like descriptor that is determined by a set of measurements of the corpus callosum characteristics along its extension. The signatures presented here were inspired by the work of Park et al. [36] and Rittner et al. [42], where functions of DTI scalar values were proposed and used for analysis. In this work, we focus on methodological issues involved in the process of generating such signatures in order to improve reliability. The pipeline diagram presented in Fig. 4.1 details the main step "Estimate Median Axis", first introduced in Fig. 2.5. It is worth noting that the intermediary results in the median axis estimation step are critical to ensuring the overall signature quality.
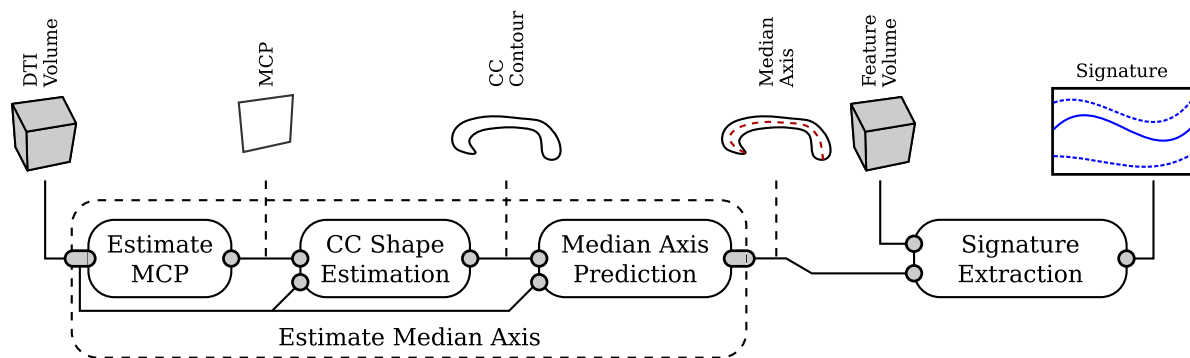

Figure 4.1: Pipeline diagram detailed with substeps of the main step "Estimate Median Axis". The substep "Estimate MCP" was described in Chapter 3, while the substeps "CC Shape Estimation" and "Median Axis Prediction" are being described in this Chapter.

The first substep is to estimate the mid-callosal plane, as described in Chapter 3. The second substep, described in Section 4.1, is to perform a reliable corpus callosum segmentation, resulting in a continuous representation of its shape. The third substep, described in Sections 4.1.2 and 4.1.3, is responsible for determining the median axis as a continuous spline. The final main step of the pipeline, responsible for reading the features along the corpus callosum structure extension, represented by the median axis, is described in Section 4.1.4. It should be noted that every step in the pipeline can influence the final result. Thus, the solutions proposed for each step where designed to produce high test-retest reliability.

4.1 Corpus callosum segmentation and shape representation

The corpus callosum segmentation at the MCP is very important since its shape defines the location of the median axis, which is the ultimate reference for the signature extraction. In addition, our solution to determine the median axis extremes relies on a continuous and smooth shape boundary, as described in Section 4.1.3. Therefore, the challenge in this step is two–fold:

1. perform a correct 2–D segmentation of the corpus callosum structure;

2. estimate a continuous and smooth spline function from the low–resolution discrete corpus callosum segmentation, which must preserve the shape characteristics.

The method we employed for the segmentation was the one proposed by Freitas et al. [14], modified to work with reference planes not aligned with the volume grid. The continuous representation solution was to explore the partial volume effect by combining multiple shape versions from rotation transforms applied to the MCP original image.

We choose to represent the corpus callosum shape by means of a parametric spline [5] that starts and ends at the same point [1], thus defining a "closed circuit". The spline interpolation functions *splrep()* and *splev()* from the scipy [22] library were used to implement our corpus callosum shape representation. The spline implementation used requires an ordered set of points in $\mathbb{R}^2$. Therefore, the initial task is to segment the corpus callosum in its original space, and subsequently identify and sort the points that belong to the segmentation edge. Once the ordered set of points that define the corpus callosum boundary is obtained, the generation of the continuous spline representation is straight forward.

For performing the segmentation we explore the fact that the corpus callosum is a white matter brain structure that is composed essentially of neural fibers. Therefore, the diffusion of water inside the corpus callosum is expected to be anisotropic, making the corpus callosum highly distinguishable in FA images, as can be observed in the sample shown in Fig. 4.2(a). However, there are other fibrous structures adjacent to the corpus callosum that would require the segmentation method to be complex, but the fiber orientations from these other structures are very different from the orientations inside the corpus callosum. While in the corpus callosum, the fibers tend to be perpendicular to the

---

[1]Due to a constraint of the spline implementation we used the start and end are not exactly the same point, but are close enough.

sagittal plane, in the other adjacent structures, the fibers tend to be parallel to the plane. Therefore, using the fiber orientation information to filter the FA image, we can obtain a more suitable image for the corpus callosum segmentation, as shown in Fig. 4.2(b).
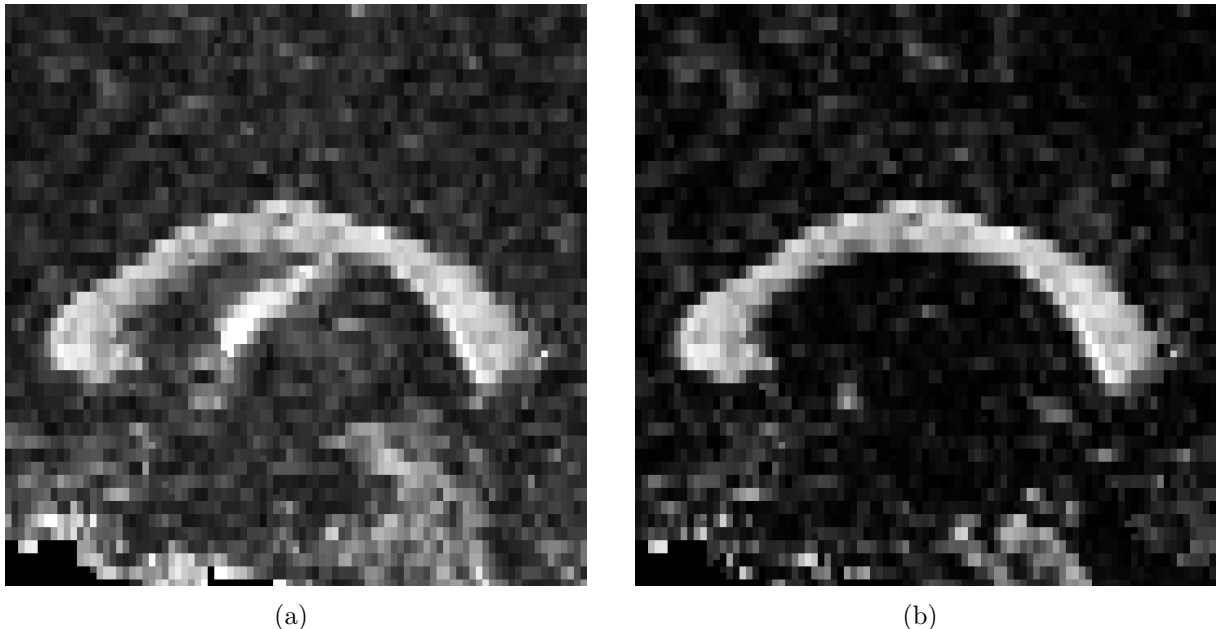

(a)   (b)

Figure 4.2: Corpus callosum in a mid–callosal plane sample: (a) FA values only; (b) FA values weighted according to the orientation of the first eigenvectors.

The segmentation of the corpus callosum $\mathcal{S}\mathrm{g}^{\Pi}$ is always defined in a plane Π as a bi-dimensional segmentation. Let $\mathrm{FA} : \mathcal{S} \to \mathbb{R}$ be a volume with the FA values for the voxels in $\mathcal{S}$, and let β be a binary 2-D image defined as

$$\beta(p) = \begin{cases} 1 & \text{if} \quad \mathrm{FA}(p) \, \| \operatorname{proj}_{\vec{n}_{\Pi}} V(p) \| \geq t \\ 0 & \text{otherwise} \end{cases}, \tag{4.1}$$

where $p \in \Pi$, $\| \operatorname{proj}_{\vec{n}_{\Pi}} V(p) \|$ is the length of the orthogonal projection of the eigenvector $V(p)$ over the normal vector $\vec{n}_{\Pi}$, and $t \in \mathbb{R}$ is a threshold. Then, the segmentation $\mathcal{S}\mathrm{g}^{\Pi} = \{p_1, p_2, \ldots, p_n\}$, with $p_n \in \Pi$, is defined by the largest connected component from β, including the holes, if any. For healthy subjects a fixed arbitrary t is enough to produce a good segmentation, which was empirically defined as $t = 0.4$.

4.1.1   Corpus callosum shape estimation

The shape of the corpus callosum is estimated from its segmentation using a parametric spline function

$$r(u) = (x(u), y(u)), \tag{4.2}$$

where $u \in (R)^{[0,1]}$ is the domain of the curve, which defines its natural parametrization. The function $r(u)$ represents the corpus callosum shape as a planar continuous curve [30]. This representation provides a framework for the shape analysis and is essential for some tasks we perform to obtain the corpus callosum median axis, such as identifying the axis extremes, as described in Section 4.1.2.

To improve the quality of the corpus callosum shape estimation we first apply a scale to the weighted–FA image before performing the segmentation, thus increasing the spatial resolution. This is important because some corpora callosa have very thin portions where the tracking of the segmentation edge could fail. In addition, the spline curve location would be misplaced because of the large pixel sizes of low spatial resolution. Therefore, we perform a symmetric scale transformation with a factor of 4 and use linear interpolation to provide some degree of smoothness to the corpus callosum segmentation.

The next procedure is to perform the corpus callosum segmentation and identify the boundary pixels that will be used to predict the continuous shape. The edge pixels $\mathcal{E} = \{e_1, e_2, \ldots, e_n\}$ are pixels $e_n \in \mathcal{S}g^{\Pi}$ that belong to the segmentation and have at least one neighbor that is a background pixel $p \notin \mathcal{S}g^{\Pi}$, regarding a 4-neighborhood schema. The Fig. 4.3(a) show the segmentation and the edge pixels for a sample image. The edge pixels $\mathcal{E}$ are used to estimate the spline that represents the corpus callosum continuous shape.

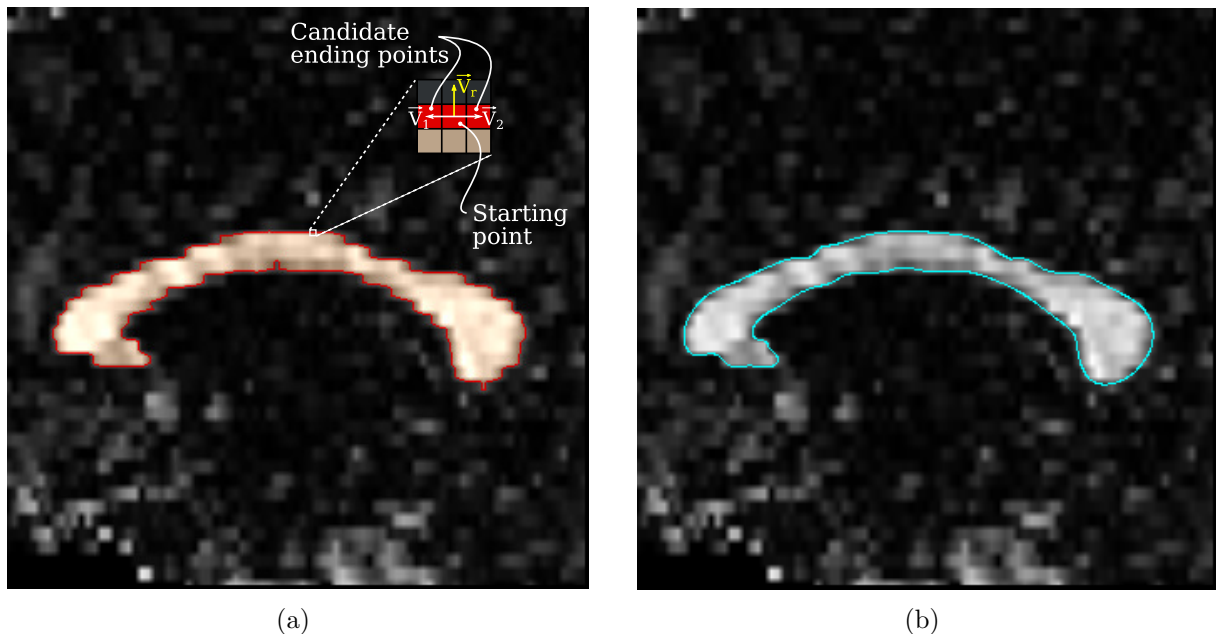(a)   (b)

Figure 4.3: Estimating the spline function that describe the corpus callosum shape: (a) edge pixels computed from the corpus callosum segmentation; and (b) the estimated spline. The detail in Fig. 4.3(a) show a randomly selected starting point and the candidate ending points. The choice of which candidate will be the ending point is done by an analysis of the vectors $\vec{V}_1$ and $\vec{V}_2$ in relation to the reference vector $\vec{V}_r$.

The edge pixels $\mathcal{E}$ must be sorted as a requirement for the spline estimation. In order to do the sorting we create a weighted graph $G_e = \{\mathcal{E}, E, W\}$ where the node set is the edge pixels $\mathcal{E}$, and the graph edges are defined according to an 8-neighbor adjacency. The weight associated with an edge $e_i e_j$ is given by the Euclidean distance between the pixels e_i and e_j. Then, an arbitrary node e_{start} is selected to be the beginning of the shape curve. The selection is random, but only nodes with degree two are eligible. Next, we select the node $e_{end} \sim e_{start}$ that will be the ending of the shape curve. There are only two options as the degree of e_{start} is two. The only constraint that defines which node to select is whether it produces a clockwise or anticlockwise oriented curve. In this work, we arbitrarily defined that the curve must be clockwise oriented. The procedure we defined to identify which neighbor of e_{start} generates a clockwise oriented curve was to analyze the region around the starting pixel with a 3×3 mask, as illustrated in Fig. 4.3(a). In this procedure we compute the vectors $\vec{V}_1 = e_1 - e_{start}$ and $\vec{V}_{ref} = e_b - e_{start}$, where e_1 is one of the two candidate pixels, chosen randomly, and e_b is a point computed by the average location of all background pixels in the mask. Then, we rotate the reference vector $\vec{V}_{ref}$ clockwise by the angle $\alpha = \arccos\left(\dfrac{\vec{V}_{ref} \cdot \vec{V}_1}{||\vec{V}_{ref}||\,||\vec{V}_1||}\right)$, obtaining the rotated vector $\vec{V}_r$. If $\vec{V}_r \not\parallel \vec{V}_1$, then $e_{end} = e_1$, otherwise e_{end} will be the other candidate pixel. Once defined the nodes that are the beginning and ending of the curve the edge $e_{start} e_{end}$ is removed from the graph. Finally, the sequence of points that are used to predict the spline is the correspondent nodes in the path between e_{start} and e_{end} with the lowest cost, computed using the Dijkstra algorithm [11].

Improving the robustness of the predicted shape

Although we succeed on building a continuous curve from segmentation in discreet space, the predicted curve is subject to the roughness of the segmentation and also to fluctuations present in the segmentation boundary due to partial volume effect. This usually leads to a sharp curve with abrupt turns, besides diminishing the reliability of the shape representation because an image of the same subject under a different spatial transformation could have sensitive differences in the shape representation. Therefore, in order to obtain a smoother and more reliable curve that better represents the corpus callosum shape, we combine several other curves, each one computed from a randomly transformed weighted-FA image using the standard approach described previously. The transformation used on each image was a single rotation by a normally distributed random angle $\hat{\alpha}$, with $\sigma = 5°$, which was enough to introduce perturbations is the corpus callosum segmentation, generating slightly different versions of the same shape representation as demonstrated in the examples in Fig. 4.4. Then, an arbitrary n number of points are

computed for each version of the corpus callosum shape, equally distributed along the domain of each spline, generating a tuple of points. Finally, an inverse transformation is applied to these points to match the original space, and then a final tuple of points is calculated as the average of the set of tuples. This final tuple of points is used to estimate the final smoothed version of the corpus callosum shape, as the example shown in Fig. 4.3(b). This last version of the corpus callosum shape is the one that is going to be employed in the next steps that define the corpus callosum median axis.

4.1.2   Points of interest in the corpus callosum boundary

This step is primal for tracing the corpus callosum median axis because it defines its endings. The illustration in Fig. 4.5 show the ideal location of the median axis endings for a hypothetical corpus callosum shape. Given that the corpus callosum present natural variability of shapes among distinct subjects, finding the correct location of the median axis endings is not a trivial task. The method we used to identify such interest points is based on an angular function derived from the predicted corpus callosum shape, where regional peaks mark the location of the interest points.

The angular function $\Phi^s(u)$ is derived from the spline function $r(u)$ that represents the corpus callosum shape. Let $\phi(u, s)$ be the angle between the vectors $\vec{V_a} = r(u - s) - r(u)$ and $\vec{V_b} = r(u + s) - r(u)$, computed as

$$\phi(u, s) = \arccos\left(\frac{\vec{V_a} \cdot \vec{V_b}}{||\vec{V_a}||\ ||\vec{V_b}||}\right). \tag{4.3}$$

Then the angular function is given as

$$\Phi^s(u) = \begin{cases} \phi(u, s) & \text{if } \vec{V_a}\ R(\phi(u, s))\ ||\ \vec{V_b} \\ 2\pi - \phi(u, s) & \text{otherwise} \end{cases}, \tag{4.4}$$

where $R(\phi)$ is a clockwise rotation matrix. Observe that the domain of the spline function is circular in the range $[0, 1]$, and the operation $u+s$ is computed accordingly. In addition, s defines the scale of the angular function as it is the distance between the reference point $r(u)$ and the neighbor points $r(u-s)$ and $r(u+s)$ that define the vectors used to compute the angle. Smaller s values lead to angular functions that capture more abrupt changes in the shape boundary, while bigger s values are more insensitive to small local variations and better describe the overall shape represented by the spline function. Another interesting property of the angular functions is their invariance to uniform scale [30].

The identification of the interest points is done simply by finding the regional
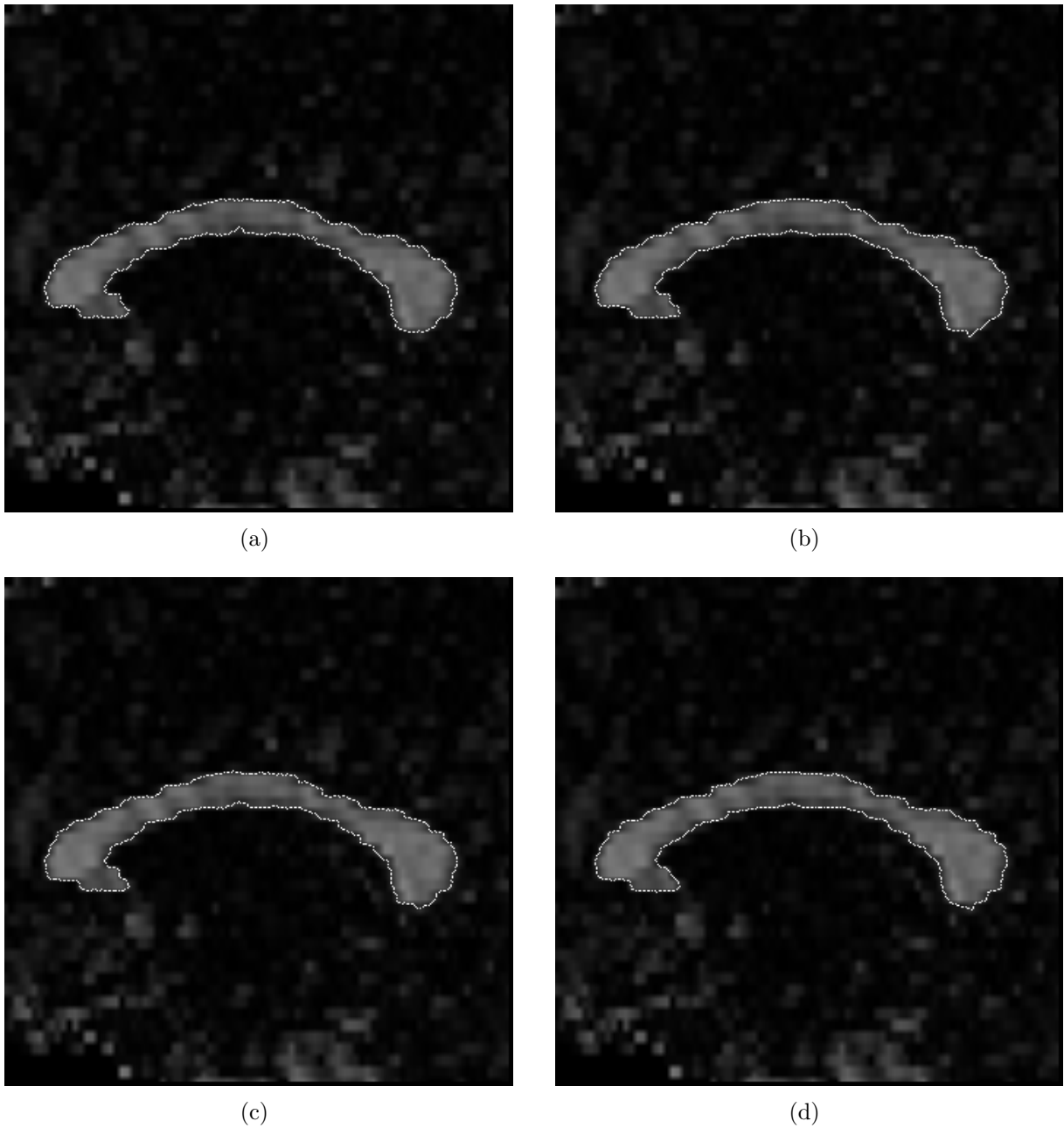
(a) (b)

(c) (d)

Figure 4.4: Intermediary corpus callosum shape estimations used to generate the final smoothed shape. Note that despite all shape estimations are good representation for the corpus callosum shape, there are small variations in all of the estimations. It the final version of the shape estimation those variations are combined, resulting in a smoother curve that is much more suitable for the process of defining the corpus callosum median axis.

maximums in the angular descriptor. However, since the spline that defines the corpus callosum shape starts at a random boundary point the domain of one shape is not equivalent to another. Therefore, a circular registration of the angular descriptor is required in order to align the domain of distinct shape curves. The registration consists of a 1-D circular translation transform that is applied to the spline domain. The registration
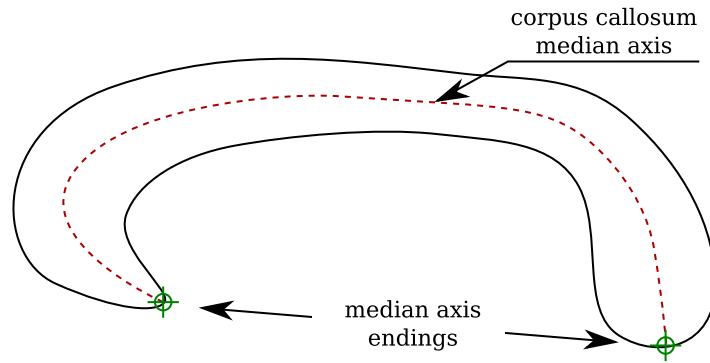
Figure 4.5: Corpus callosum median axis schema for a hypothetical corpus callosum segmentation performed at an also hypothetical mid–callosal plane. This schema shows where the median axis is expected to be. The positions of the endings are very relevant for finding the correct corpus callosum median axis.

reference is a template created by averaging a sample of angular descriptors with scale $s = 0.1$, which provide enough insensitivity to small shape variations. The criterion for registering a given corpus callosum angular descriptor, also with scale $s = 0.1$, is the maximum positive correlation to the template.

The registration to a template allows us to estimate the location of interest points in the corpus callosum boundary because it has a common general shape, despite lesser natural variations, and the locations at the template and the registered descriptor are correspondent. Thus, the estimated location of the interest points can be manually marked in the template, as it is static. As shown in Fig. 4.6, the static AP^s and PP^s, respectively related to the anterior and posterior interest points, are close to the final interest points AP and PP. The final AP and PP are then defined by the regional maximums around AP^s and PP^s in the angular functions with scale $s = 0.01$ and $s = 0.1$, respectively.

The distinction in the angular function scale for defining each interest point is due to particular features we aim at each case. In the PP case we want it to be a general terminal location more aligned with the corpus callosum skeleton, which is introduced in the next Section 4.1.3, and independent of more local shape characteristics. As for the AP, we wish to find exactly the corpus callosum anterior tip, which usually has a more abrupt turn than most other locations in the corpus callosum boundary. Unfortunately, the criterion used for the AP lack robustness and in some cases it fails to find the correct location. However, this is a minor problem because this issue is very easy to correct with supervision.
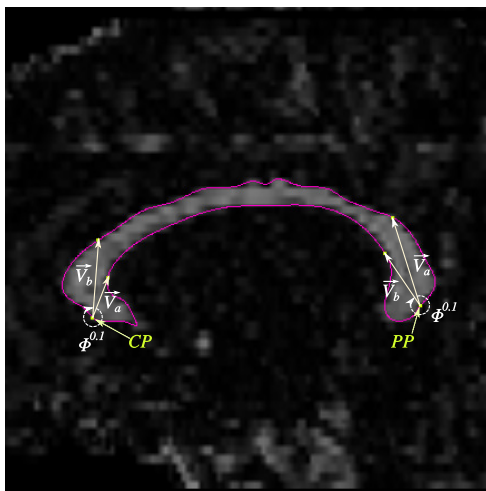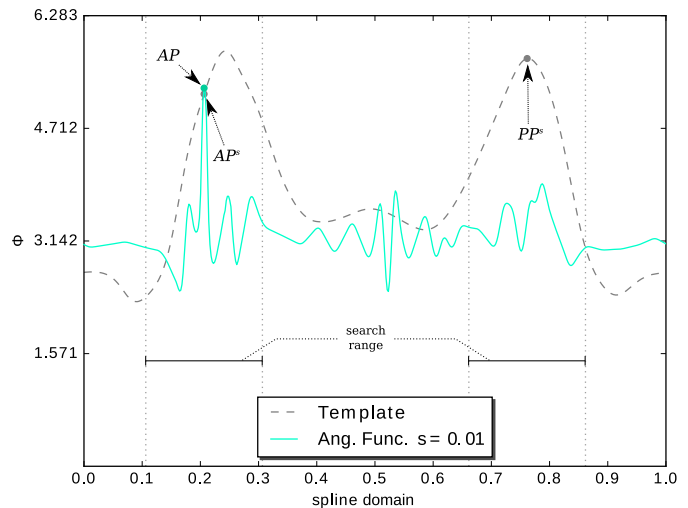
(a)
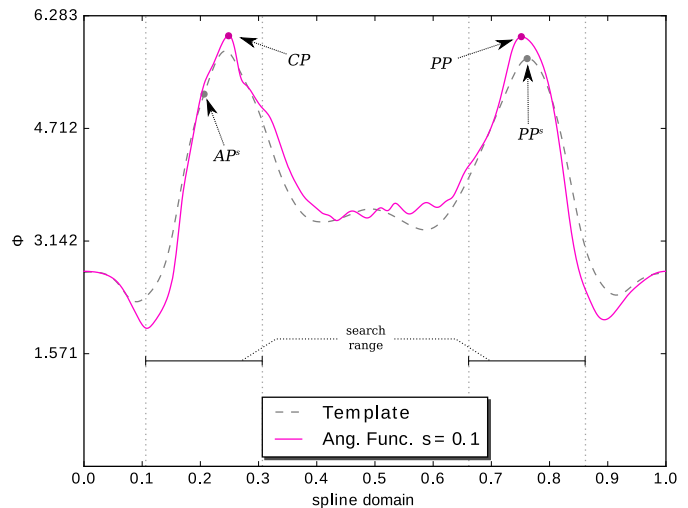
(b)
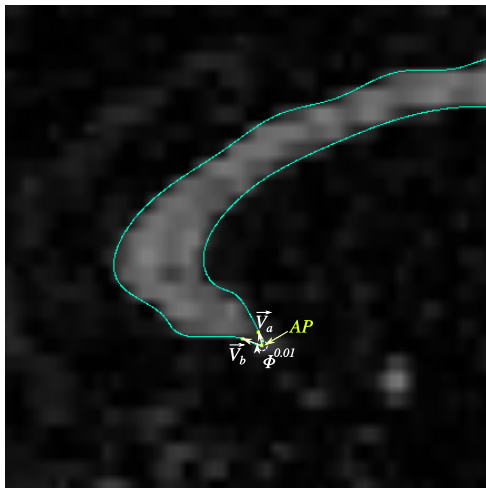
(c)

(d)

Figure 4.6: Estimation of interest points by using angular functions: (a) and (c) show the locations of the estimated interest points AP, CP, and PP, along with the vectors and angles considered for the generation of the angular functions; (b) and (d) show the angular functions respectively with $s = 0.1$ and $s = 0.01$, together with a template angular function ($s = 0.01$), and the static prototype points AP' and PP'.

4.1.3 Corpus callosum median axis

The median axis can be thought as the corpus callosum backbone, and it defines the location from where the signature values are extracted. It lies in the mid–callosal plane and traverses the corpus callosum from the anterior tip to the posterior extreme as illustrated in Fig. 4.5. It is similar to the skeleton of the 2–D corpus callosum shape, but the known skeletonization techniques would fail to produce the intended median axis because of the strict requirements for the skeleton endings. Therefore, we have developed

a specific method for this task by combining a skeletonization technique with a procedure that extends the skeleton to the marked interest points.

To compute the skeleton of the 2–D corpus callosum shape we have used a Euclidean distance transform [23] $D : \mathcal{P} \to \mathbb{R}$ relative to the shape boundary, being $\mathcal{P}$ strictly the set of points inside the corpus callosum shape, including the boundary ones. Next, we defined a weighted graph $G^d = (\mathcal{P}, E, W)$ where the edges in E are determined by an 8-neighborhood between the elements of $\mathcal{P}$. The weights were computed as $W(p_1, p_2) = \max(D) - \max(D(x_1), D(x_2))$, where p_1 and $p_2 \in \mathcal{P}$ and $\max(D)$ is the global maximum of the distance transform D. Then, the skeleton is defined as the lowest cost path between the nodes that correspond to the interest points AP and PP found in the previous step. The path is computed using the Dijkstra algorithm [11] adapted to store the shortest path information. Unfortunately, both skeleton extremes are not quite what is expected to match the desired median axis, as shown in Fig. 4.7(a). To correct this issue we remove the undesired parts of the skeleton, which are the extremes. The locations where the skeleton is cropped are defined through an analysis of the distance profile, which is a function $D^p : \mathcal{S} \to \mathbb{R}$, where $\mathcal{S} \subset \mathcal{X}$ is a tuple with the skeleton points ordered from anterior to posterior, and $D^p(s) = D(s)$. Thus, the skeleton final endings are defined as the first and last local maximums in the smoothed distance profile $\bar{D}^p$, which is a mean filtered version of the distance profile D^p. Figure 4.7(b) show the raw and smoothed distance profiles for the skeleton in Fig. 4.7(a), with the marked cropping locations.

The final median axis is estimated using the skeleton points together with the interest points AP and PP, and with a few control points added between the AP and the skeleton starting point SSP, which is the anterior terminal point of the skeleton. The purpose of these extra control points is to ensure the median axis to have a proper curve in the anterior extreme that match our expectations about its form and location. These control points are defined using a recursive procedure that involves three points: the AP, the SSP and a curve controller point CP defined through the same process used for the interest points, except that in this case it uses the anterior fixed location as reference and the angular function where the maximum is defined has scale $s = 0.1$. The location of these three points is illustrated in Fig. 4.7(a) for a sample corpus callosum shape. Then, each iteration of the recursive process computes both a new curve controller point $CP' = \dfrac{CP + SSP}{2}$ and a new anterior point $AP' = \dfrac{AP + CP'}{2}$. The process repeats for a predefined number of steps n, which in this work was arbitrarily defined as $n = 5$. It is important to note that only the AP' are added as control points for defining the median axis. After the determination of these extra control points the median axis is finally computed as a continuous curve represented by a spline function that starts in the
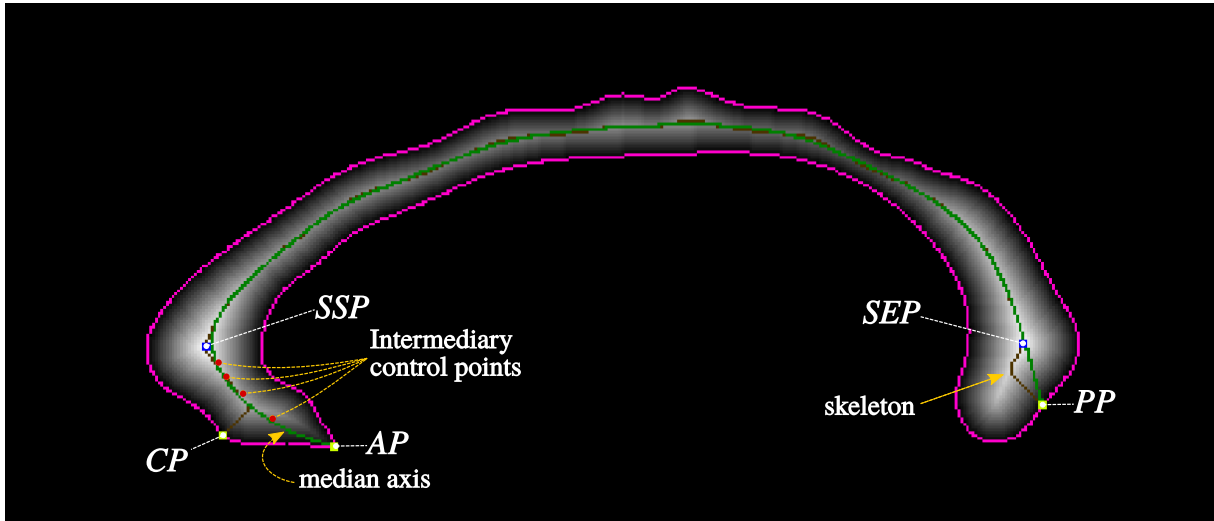
(a)
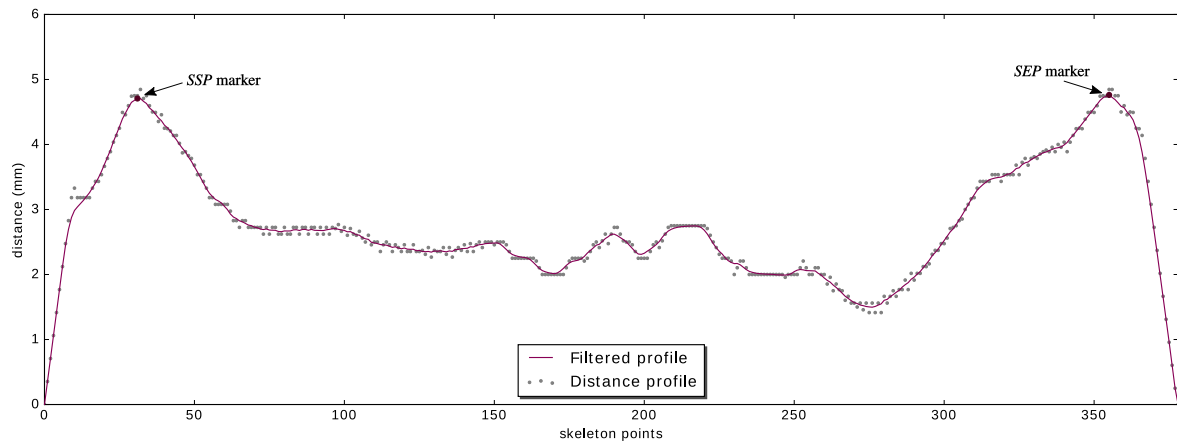
(b)

Figure 4.7: Using the Euclidean distance transform to determine the median axis: (a) show the transform for a sample corpus callosum shape, with elements related to the procedure; (b) plot of the distances along the corpus callosum extension, which is used to identify the useful portion of the skeleton.

AP and ends in the PP.

4.1.4 Signature generation

A signature can be computed from any scalar data associated with the image voxels, which in this work we demonstrate using DTI derived scalar features like the FA and MD. However, the scalar values used to generate a signature can also be from the raw DWI values or even from other image modalities like T1-weighted, in which case a transformation of the median axis according to a multi–modality registration would be required.

The median axis provides the reference for reading the scalar values. Once the median axis representation is continuous, the signature can have any number of points. Of course, there are limits due to the nature of discrete image space, and a very large number of sample points would not be practical. Thus, the number of points must be defined according to the spatial resolution of the image. For this project, we have defined the number of points that form the signature to be $sn = 120$.

The signature purpose is to describe the microstructure features along the corpus callosum extension. If we read the scalar values at each reference point, i. e., the points along the median axis, we would be failing to describe the overall corpus callosum features because the reading would represent only the local scalar values that coincide with the median axis. In addition, the generated signature would also be more susceptible to the effects of noise from the image. Therefore, the reading for each reference point considers values in a region around the point, being the reading itself statistical measurements computed from the values in the correspondent region. In this work, the region is defined by a 3–D isotropic Gaussian function, which gives more weight for values closer to the reference point than for the ones more distant. Thus, to each reference point, there is a sample of weighted values from which the statistical measurements can be computed.

To ensure the signature is computed strictly from values that belong to the corpus callosum another of weights are employing together with the Gaussian ones. These weights are defined by a smoothed 3–D corpus callosum segmentation, where the binary mask was filtered using a Gaussian filter to smooth the boundaries. The smoothed mask gives lower weights for voxels at the boundary of the segmentation and provides some compensation for wrongly classified voxels. Thus, let $b : \mathbb{R}^3 \to \mathbb{Z}_{[0,1]}$ be a binary image that represents the corpus callosum 3–D segmentation, and $\mathcal{G}^\sigma$ be an isotropic 3–D Gaussian function, then

$$\mathrm{B} = b \cdot \mathcal{G}^\sigma, \tag{4.5}$$

is the weighted mask used in conjunction with the Gaussian weights. In this work, we have used $\sigma = 1mm$ for the weighted mask.

Sample distribution

To keep the method generic we are not going to assume any particular distribution for the samples related to each reference point. Instead, we represent the sample distribution by means of an estimated empirical distribution function (EDF), which is obtained by polynomial regression over the raw data. The EDF's domains are the probabilities, in the range [0,1], at which a sample element lesser or equal a value can occur, as illustrated in Fig. 4.8. It is important to notice that the EDF take into consideration

the weights of each sample element since it is built as an accumulated histogram of the weighted sample values, where the weights relate to the number of occurrences for the associated values.
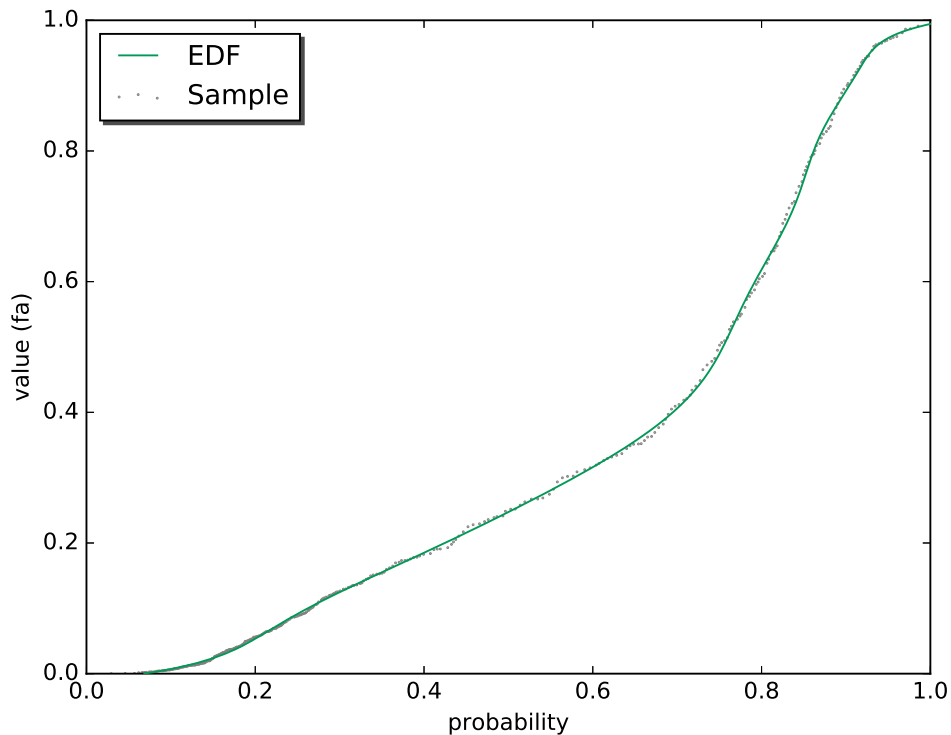

Figure 4.8: EDF estimated from an arbitrary sample of FA values.

Particularly for this work, the EDF was defined as a function of the probabilities instead of the values. The EDF as defined in this work can directly provide percentile values, which may be useful for some analysis and for visualization. Besides the direct percentile estimation, we can also obtain an equivalent non-weighted sample, which is computationally relevant because it is the only acceptable input in a third–party function used to compare two empirical samples as explained in Section 4.2.

4.2   Signature analysis

In this Section, we explain how to compare signatures from distinct images and the concepts involved in this procedure. We also perform some experiments in an 80 images dataset in order to observe the behavior of the signature and how is the relationship between signatures.

4.2.1 How many partitions?

The usual practice in the analysis of the corpus callosum microstructure is to divide its area into a number of regions and compare the distribution in correspondent regions from distinct images. Many division schemata were already proposed and each has its own particularities, but their common purpose is to match corpus callosum sections to specific cortical areas [44, 18, 17]. The signature we propose in this work is somehow analogous to a partitioning schema in the sense it is composed of several samples collected along the corpus callosum, but in the signature method, there is no a priori correlation to any cortical regions as there is in the traditional partition methods. Instead, the reference points are evenly spaced along the corpus callosum axis and are many more than the usual number of partitions. Further, the influence radius is the same for all reference points, although the size of the sample can be different due to the segmentation constraint, wherein the partitioning methods the partitions have distinct sizes that are determined according to a study [44, 17]. Therefore, it is clear that although both approaches aim at providing means for a more detailed analysis of the corpus callosum characteristics the goals of each approach are quite distinct. Ultimately, the number of reference points used for generating the signature is only significant for defining its resolution, and have no inherent relations to cortical areas of the brain as in the partitioning methods. The purpose of the signature is to provide a more general description of the corpus callosum characteristics, which can then be used for further analysis, including the ones related to cortical brain areas as in the traditional partitioning methods. It is possible to notice that the shape of the signature follows a very distinctive pattern, as shown in Fig. 4.9. The existence of this pattern conforms with the fundamental argument that justifies the traditional partitioning analysis, which states that the fiber bundles interconnecting specific cortical areas of the brain traverse specific portions of the corpus callosum and may present distinct properties, as axon diameter and density [2].

In addition to the differences between the methods already discussed in this Section it is worth mentioning another fundamental one related to sample selection. While the traditional partitioning methods perform a hard and mutually exclusive classification of the corpus callosum data elements for each of its predefined partitions, the samples that generate the signature are weighted and overlap with other samples from neighboring reference points. These weighting factors were designed to potentially reduce miss–classification issues and also the influence of partial volume effects from data elements in the corpus callosum structure boundary. And lastly, the samples for the signature method are selected from a 3–D subspace of the image volume, whereas in the partitioning methods the sample selection is restricted are restricted to data elements in a 2–D slice.
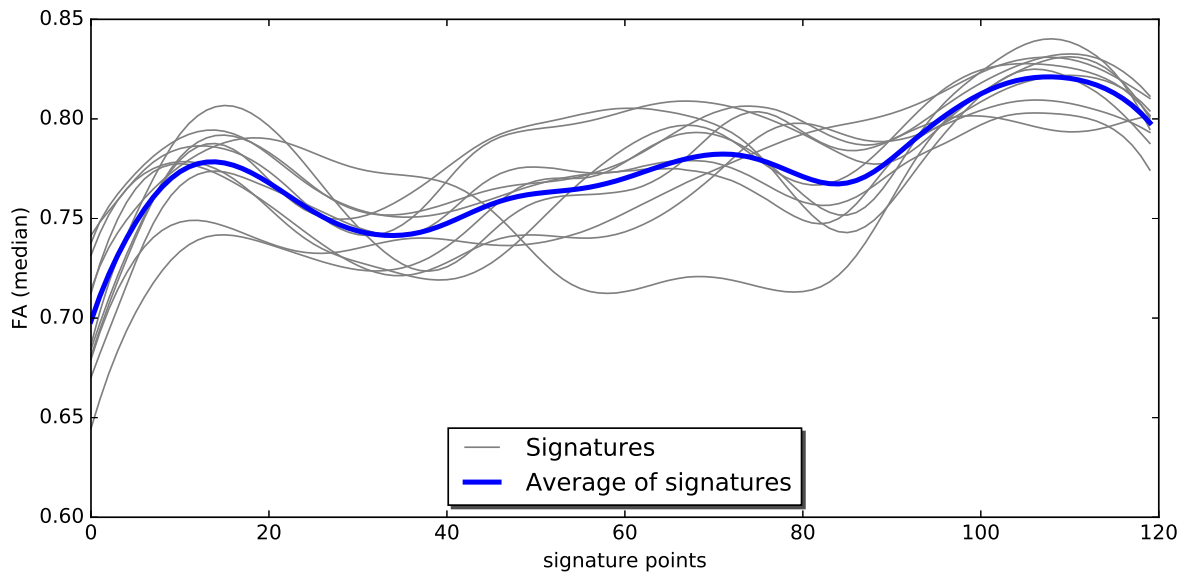
Figure 4.9: The visual profile (median) of ten selected FA signatures, with the correspondent average profile. The pattern present lower values roughly at the transitions from the genu to the body, and from the body to the splenium of the corpus callosum.

4.2.2   Comparison between signatures

Signatures provide means for doing partial comparisons between corpora callosa, i. e., reference points from one signature are directly related to reference points from a second signature as they should be representing equivalent physical areas of the corpora callosa associated to them. However, even with perfectly defined median axes, the occurrence of miss match between what should be considered equivalent areas is still possible because we are dealing with biological tissue, which despite presenting consistent patterns lacks on precision. Therefore, a matching between the signatures must be done before proceeding with the comparison to enhance the alignment of the most likely equivalent areas of the corpora callosa being analyzed. That is done by a 1–D registration procedure of the signatures as functions of the distribution median values. The registration optimizes translation and scale to the maximum correlation between the functions. After this process, the points from one of the signatures are recalculated according to the registration parameters. There is a range expected for the registration parameters, which is $[0.8, 1.2]$ for the scale and $[-0.2n, 0.2n]$ for the translation, where n is the number of reference points. If the translation or scale exceed the expected range the registration procedure is considered invalid and the original signatures are used instead. The registration procedure can be seen as a fine–tuning adjustment to enhance the likelihood of the reference points from distinct signatures to match correspondent regions of the corpora callosa based on its own microstructure characteristics, exploring the concept that the correspondent regions in normal conditions should have similar characteristics.

Now that the reference points from both signatures are in optimal position, the comparison process is reduced to comparing each correspondent pair of reference points. In particular, we are interested in testing whether the sample from a reference point in one signature can be considered to be from the same distribution as the sample from the correspondent reference point in the other signature. Therefore, what we want is a statistical test that will reject or fail to reject the hypothesis that both samples are drawn from the same distribution, considering some confidence level. This is a well–known field of study in statistics and there are many methods available for performing such a test. The choice of which methods would be better will depend on the nature of the distribution. Since we are proposing a generic method we can't know this a priori, so our choices are limited to methods that can be applied to any distribution. Further, we need a method able to compare two EDF instead of one EDF against a theoretical distribution. Therefore, the method we suggest and also the one we use in our experiments is the k–sample Anderson–Darling test [3], which was designed to test empirical distribution functions and has all the characteristics we require for keeping our method generic.

While it is very interesting to have localized tests along the signature to see where the distributions differ considerably, an overall similarity measurement between two signatures is also very useful and required. The measure we propose is simply the rating between the amount of signature reference point pairs for which the test failed to reject the null hypothesis over the total number of reference points from the signature. This is a straight forward approach that also offers the convenience of having the value constrained between 0 and 1, which can also be interpreted as between 0% to 100%.

4.2.3   Comparing signatures in practice

In this Section, we perform experiments to test the signature comparison within a population of healthy subjects. The goal is to assess fundamental characteristics of the signature comparison process in order to unveil the potential usage of signatures in studies involving the corpus callosum. For simplicity, we are restricting the experiments to signatures of the fractional anisotropy. The dataset is the same used in the experiments for testing the MCP method, described in Chapter 3. It is worth to remember that dataset has 80 images from healthy subjects with ages ranging from 8 to 60 years old, and with a larger proportion (72.5%) of female subjects. Major age concentrations are around 25 and 45 yo respectively, and there are only two subjects with ages under 10 years old. There are also three pairs of images that were acquired from the same subjects with a time delta of 6 months.

Our approach in this experiment was to perform an all–to–all overall similarity

comparison of the FA signatures from the images in the dataset. The results from all those comparisons were modeled as a weighted complete 80 nodes graph, where the edge weights are the overall similarity measurement between the signatures associated with the edge nodes. The graph model is suitable for our analysis because we are interested in investigating if there are clusters in our dataset. Our first analysis aim at having an insight about the distribution of similarity measurements for all the population. Therefore, we have placed all the weights from the graph in a histogram, which is shown in Fig. 4.10(a). We can observe that a significantly larger portion of the comparison pairs have similarity under 50%, being the highest count in the range from 0% to 10% similarity while the lowest count is in the range from $90\% to 100\%$ similarity. Although our dataset is considerably heterogeneous in regards to the subject's ages, all subjects are healthy and we were expecting a higher count of high similarity pairs. This result indicates that the signature is sensitive to natural variations in the subject's individual characteristics.

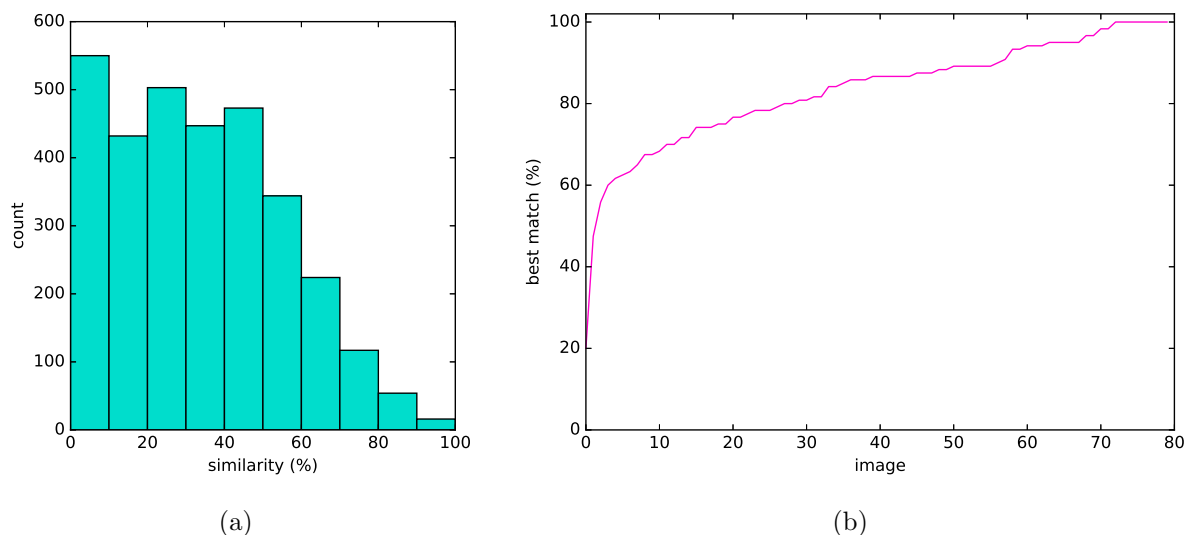$\qquad$ (a) $\qquad\qquad\qquad\qquad\qquad\qquad$ (b)

Figure 4.10: Results of the comparison of FA signatures for the whole dataset, using the defined similarity function: (a) histogram of the similarity values for each possible comparison pair of signatures in the dataset; (b) highest similarity value each signature had to another signature in the dataset.

The plot from Fig. 4.10(b) show the highest similarity measurement for each image in the dataset. Despite the low count on high similarity pairs of signatures, it shows that more than half signatures have at least one peer that has more than 80% similarity. Further, almost 10% of the signatures has at least one connection that is 100% similar and only three signatures have all connections bellow 60% similarity. Therefore, most signatures have at least one reasonably good representative in the population, regardless of the high count of very low similarity connections.
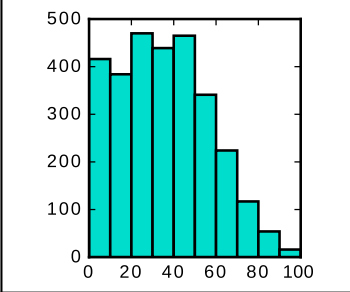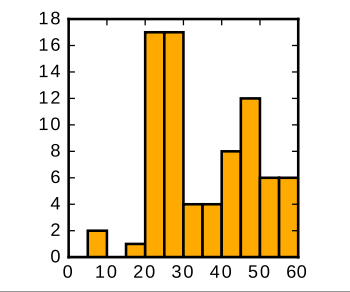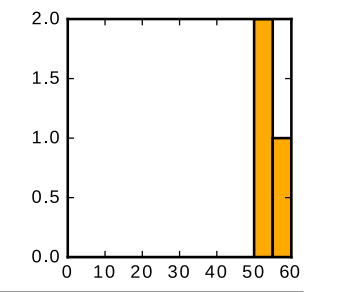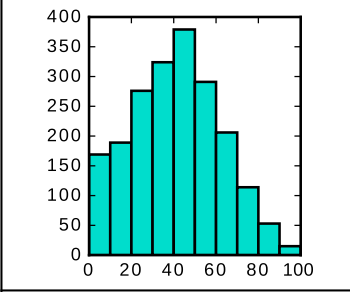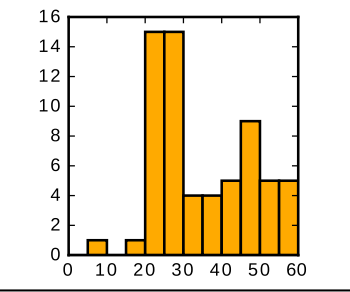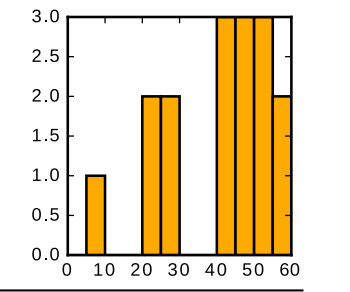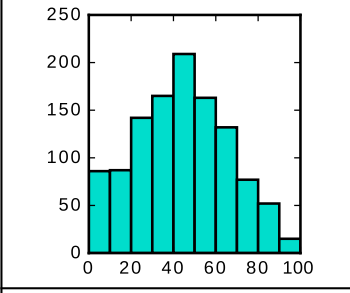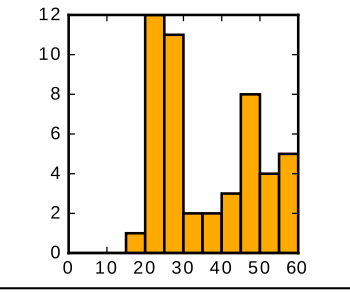
In the next analysis, our goal is to investigate the existence of high similarity clusters in our dataset. The approach is to eliminate from the graph all edges for which

the weight is lower than a given threshold t, and subsequently identify the connected components defined by the remaining edges. We have performed the clustering process using $t = \{60\%, 70\%, 80\%, 90\%\}$, and the main results are being presented in Table 4.1, where the first column shows a histogram for all weights in the subgraph formed by the largest connected component (cluster), the second column shows a histogram of the subject's ages associated with the nodes in the largest connected component, and the last column is a histogram of the ages from every other node that is not in the largest connected component. This table also shows the number of nodes in the largest connected component for each threshold, and also the proportion of female subjects in the largest connected component and its complement.

The first observation we made about the clustering results is that our dataset has mainly only one large cluster, and in all thresholds, there is no secondary cluster larger than 5 nodes. In fact, most nodes not in the largest cluster remain isolated after the threshold. That was the reason we tailore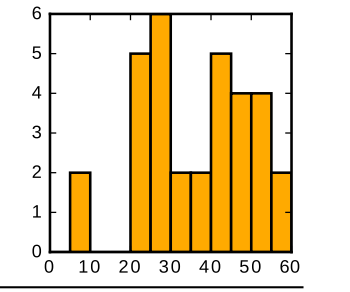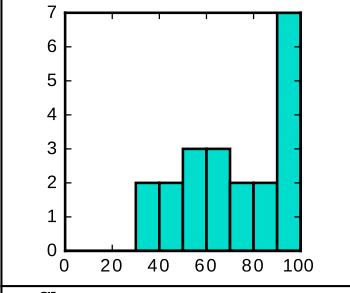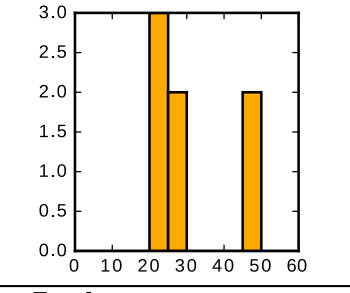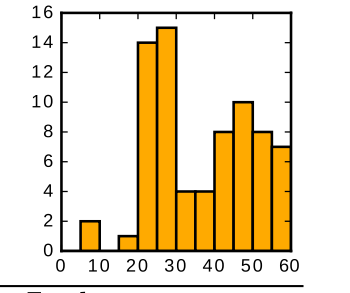d the cluster analysis as a bipartition of the graph by placing in one side the largest cluster and in the other side all the remaining nodes. As shown in Table 4.1 the largest cluster has more than half graph nodes for all thresholds, except the 90% one. As the largest cluster become more selective and its size decreases we can notice significant changes in the distribution of similarity values. The removal of only three nodes from the graph by using the $t = 60\%$ threshold was enough to reduce the count of extremely low similarity connections, being a major concentration now in the range from 20% to 50% of similarity. For the thresholds $t = \{70\%, 80\%\}$ the changes were considerably more evident, with a significant reduction in the count of connections with similarity bellow 40%, turning the distributions into normal–like with mean around 40% and 50% similarity. Finally, the $t = 90\%$ threshold promote a radical change in the distribution of similarity values, being the highest count now between 90% and 100% similarity.

Although the smaller cluster from the 90% similarity threshold has a high similarity interconnectivity, there are still some low similarity weights associated with the cluster's edges. The measurement itself could be responsible for such an outcome, or at least have an influence over it. For that reason, we have conducted another experiment to assess the capacity for the proposed similarity measurement to describe the distance between signatures. The hypothesis is that two signatures S_a and S_b that are 100% similar to each other represent the exact same point in space, and a third signature S_c should have the exact same distance from each of the signatures S_a and S_b. Thus, we have calculated the differences in similarity from every signature in the dataset to all pairs of signatures that scored 100% in similarity. The results from this experiment are shown in Fig. 4.11, and it is clear that our proposed measure does not attend the established hypothesis. Indeed, even though two signatures can be considered 100% similar, they

Table 4.1: Cluster analysis based on the similarity between signatures. Each line correspond to a threshold value used to eliminate edges from the graph.

t	Inter-similarity from largest cluster	Ages from nodes in the largest cluster	Ages from complementary nodes
60%			
	Cluster size: **77**	Female proportion: **72.7%**	Female proportion: **66.7%**
70%			
	Cluster size: **64**	Female proportion: **75%**	Female proportion: **62.5%**
80%			
	Cluster size: **48**	Female proportion: **70.8%**	Female proportion: **75%**
90%			
	Cluster size: **7**	Female proportion: **57.1%**	Female proportion: **74%**

don't necessarily represent the same point in space as the test on each of the signature's dimensions has a tolerance. Further, the output of the test between pairs of signature reference points is binary and based on a threshold, which can explain how in some cases the differences were as high as 50%. Nonetheless, the majority of the cases were bellow

10% of difference in similarity, being only a few cases over 25%. Therefore, we consider our similarity measure to be valid for the purposes of this study, and it can also be valid for many others as well. However, we advise caution on its usage as this proposed similarity measure does not fit well as a distance measure.
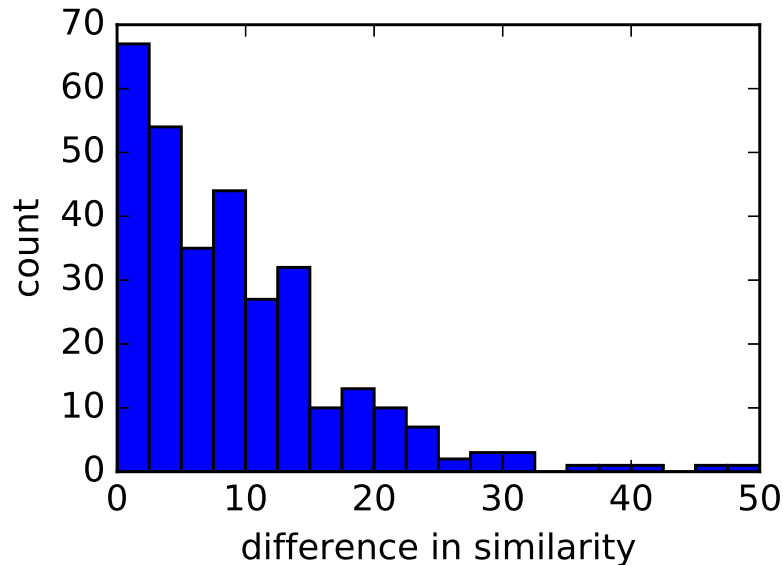

Figure 4.11: Let A, B, and C be three distinct signatures, where the similarity between A and B is 100%. Then, the difference in similarity is defined by the absolute difference of the similarity value between A and C, and the similarity value between B and C. The plot shown here is the result of comparing all pairs of signatures that scored 100% similarity against all other signatures in the dataset.

Let's resume the analysis of the results shown in Table 4.1. The second and third columns provide us with the distributions of ages from subjects in the major cluster and its complement, respectively. In addition, the proportion of female subjects in each group are also presented. For the first thresholds $t = \{60\%, 70\%\}$ there is a clear tendency for proportionally more elder subjects in the group of nodes that were disconnected from the main cluster. As for $t = 80\%$, this trend is very subtle, but it is still possible to notice a higher proportion of elder subjects in the complementary group relative to the population distribution. Finally, the largest cluster for $t = 90\%$ has 5 subjects with ages between 20 and 30 years old, and two subjects with ages between 45 and 50 years old. The two children in the dataset were disconnected respectively on the thresholds $t = 70\%$ and $t = 80\%$. It is also possible to notice a trend in the proportion of female subjects, where for the first two thresholds $t = \{60\%, 70\%\}$ there is a higher proportion of females in the larger cluster, and for the last thresholds $t = \{80\%, 90\%\}$ this relation is inverted. There are several studies that relate particular brain characteristics to the age and sex of the subject [44, 29, 35, 25], including the corpus callosum microstructure. Although there are conflicting results and conclusions in such studies, many confirm that age and

sex have an influence over the corpus callosum microstructure. The results we found in the clustering experiment are aligned to the results of such studies and, if this hypothesis is true, it means that the signature method has the potential to identify distinct groups, even though it is very sensitive to individual characteristics. However, these experimental findings are not suitable to confirm or deny the hypothesis that sex and age have an influence over the corpus callosum microstructure because our dataset was not prepared for testing this hypothesis, nor was our goal to do so.

Individual cases

In this Section, some individual cases of interest are shown for more insight into the signature method. Particularly interesting are the signatures from the three image pairs that were acquired from the same subjects, shown in Fig. 4.12. In these cases, the signature pairs from the same subjects are obviously expected to be very similar. The results show that despite the null hypothesis being rejected in a portion of the signatures for the cases presented in Fig. 4.12(a) and 4.12(c), the overall shape of the visual signature match and overlap at several points. Further, each one of the three subjects in these cases has a particular visual signature shape. The mismatch between the signatures are concentrated in one single location instead of spreading across the entire signature space, and the difference in the median values are larger where the mismatch occurred.

Figure 4.13 show the visual representation of the three signatures with a better score in the dataset, along with the correspondent estimated corpora callosa shapes and median axis drawn upon FA maps. The score is an average of the ten most similar connections relative to each signature. Notably, the visual representations of the three selected signatures are very similar. The values of the anterior and posterior regions of the corpus callosum are almost at the same level, and in all three cases, the visual signature values are around 0.7. The ages of the subjects associated with these signatures are between 20 and 30 years old, which is in agreement with the results from the clustering experiment. Despite the similarity between the selected signatures, it is evident that the respective associated corpora callosa are very distinct in shape.

Finally, the three worst cases, defined as the ones with the lowest similarity for their best connection to the dataset, are shown in Fig. 4.14. As in the best cases, the corpora callosa estimated shapes and median axis are also shown in Fig. 4.14. One of the most evident aspects of these cases is the lower FA values for the anterior region of the corpus callosum in relation to its posterior region, which is very distinct from the best cases shown in Fig. 4.13. In addition, the FA median values are in general lower, and directly correlated with the selection criterion. As occurred in the best cases example, the ages

(a) 85%   (b) 100%   (c) 87%

(d) 85%   (e) 100%   (f) 87%

(g) 85%   (h) 100%   (i) 87%
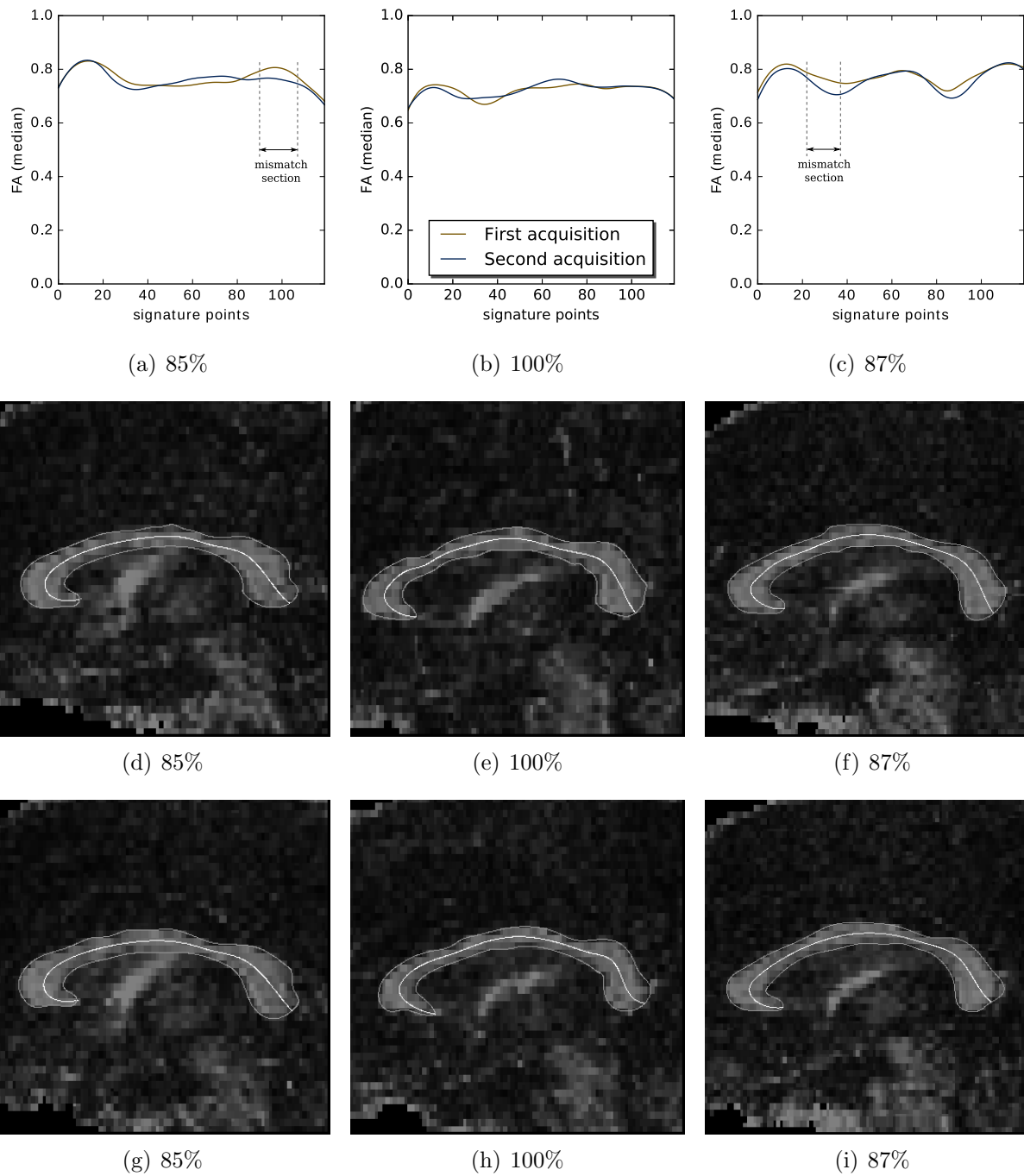
Figure 4.12: FA visual signatures and median axes computed for three subjects whose had two images in the dataset acquired at distinct times. Each column relate to a same subject.

from the three worst cases are also aligned with results from the clustering experiment, which are in the range from 50 to 60 years old. In fact, the three selected worst cases are the signatures separated from the main cluster using the $t = 60\%$ threshold. Again, the corpora callosa shapes are very distinct, although in this case, we were not expecting it to be similar to each other.

(a) 83.7%   (b) 81.9%   (c) 81.4%

(d)   (e)   (f)

Figure 4.13: FA visual signatures and median axes computed for the three subjects with highest similarity scored to other subjects from the dataset.

4.3 Chapter conclusion

In this Chapter, we have presented a new method for extracting a corpus callosum signature. The method was designed to depend only on a DTI to completely estimate a corpus callosum median axis, which is the foundation for the signature extraction. Although the method is native for extracting signatures from DTI scalar maps, which can be any, it is also applicable to other image modalities through inter–modality registration, even though this application was not demonstrated in this work. The procedure for defining the median axis was designed to be robust to lesser variations in the corpus callosum shape in order to provide a reliable foundation for the signature extraction. The signature itself is a collection of weighted samples defined by 3–D Gaussian regions centered at reference points across the median axis. Therefore, the resulting signature carries a rich description of the corpus callosum features, which can be used as a framework for localized analysis that is independent of the inherent corpus callosum shape.

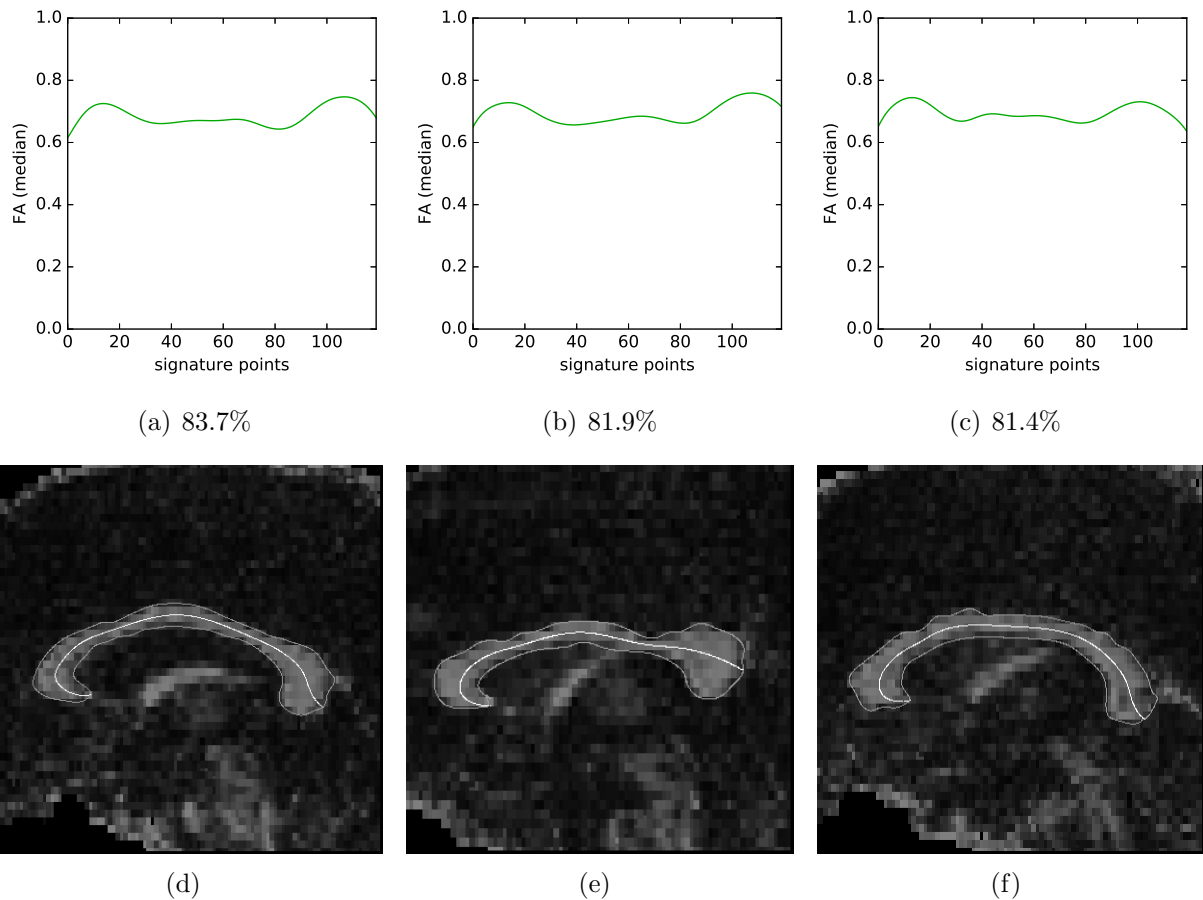The results obtained from running the signature extraction for the 80 images
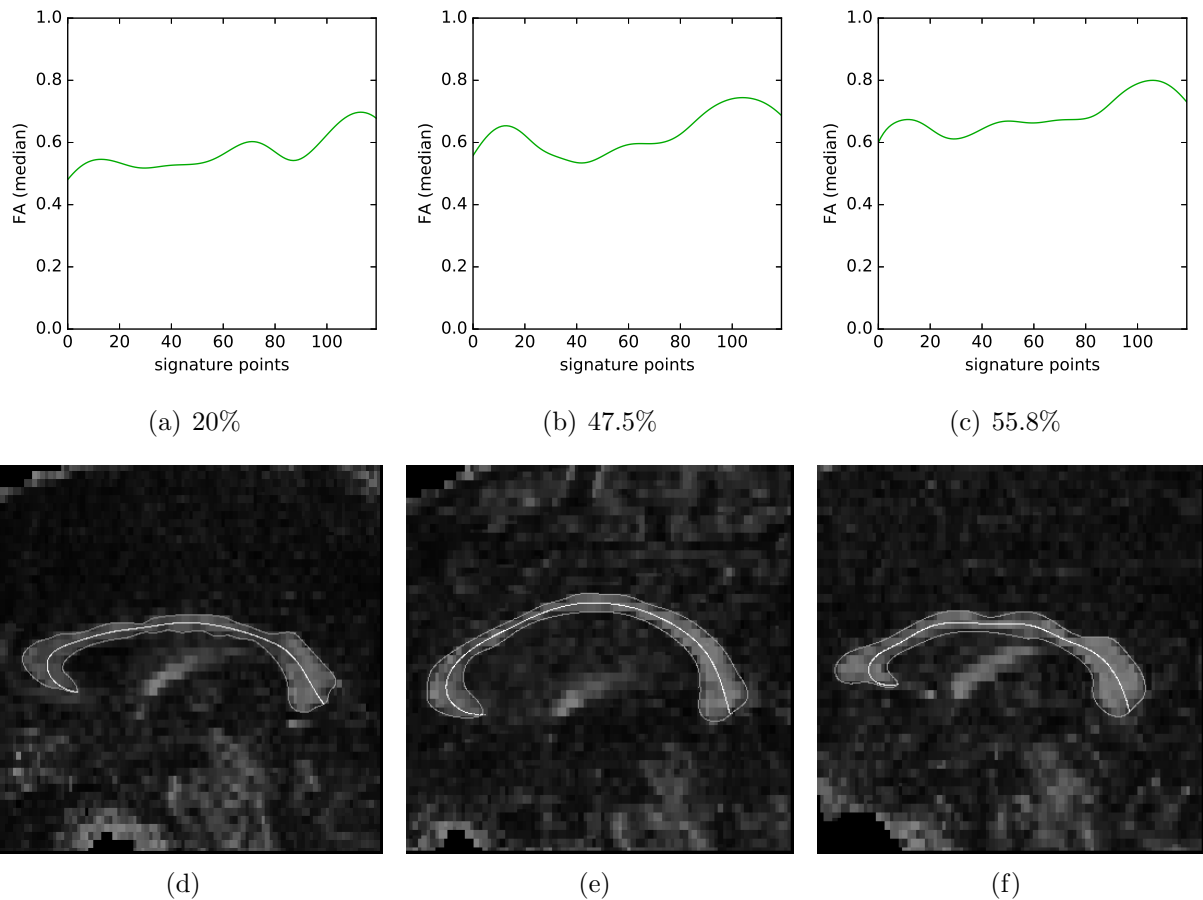
Figure 4.14: FA visual signatures and median axes computed for the three subjects with lowest similarity scored to other subjects from the dataset.

in our dataset revealed that the method we proposed is robust, and in all cases, the segmentation of the corpora callosa were of high quality. The only issue we encountered was for correctly identifying the anterior end of the median axis in two cases, which were near the expected location but not quite there. Although the method could be improved to fix that, this issue can be easily resolved by introducing an interactive step at the end of the median axis estimation procedure. In a production software, this step would be required for checking the effectiveness of the estimation anyways. Therefore, we are confident the method introduced in this work will provide a reliable reference for the corpus callosum in normal conditions[2], where correspondent corpus callosum regions can be successfully matched, despite how complex it is to analyze the data.

The visual representation of the signature is a simplification that discards most of the data in each reference point but is powerful for providing quick insight into the conditions of the subject's tissue. The visual representation can be made available in a production software as a tool for specialists to visually and compare signatures, in addi-

---

[2]See Chapter 5 for a discussion about the method's limitations.

tion to more complete data analysis. One improvement that can be easily implemented in the visual representation of the signature would be to include the 25% and 75% quartiles, which together with the median can provide a more detailed description of the distributions in each signature reference point. In our experiments, we observed a recurrent pattern for the FA visual signatures, which is an "M"–like shape where a valley form in between the anterior, body and posterior corpus callosum regions. This pattern correlates with the corpus callosum histology reported in works like [44, 17], where the anterior and posterior regions are associated with more dense fiber bundles that would explain the relatively higher FA values. However, we noticed this pattern is hard to identify in some cases, and in others, the visual signature would just not follow the pattern at all. This reflects how the biological tissue can be sometimes distinct from what is expected, and reveal that the signature indeed can capture such distinctions.

Despite the individual variations, some degree of similarity to a group is also expected, and this can be observed from the individual cases presented in Section 4.2.3. The three better cases have very similar signature visual representations, with the FA median values at equivalent levels in general. Meanwhile, the three worse cases also show a similar pattern between them, but with FA median values at distinct levels. Further, the levels of the median values correlate with their ranks. Lastly, the three signature pairs from the same subjects show distinct patterns between then but have FA value at more similar levels than the worse cases. We suppose the distinction in the pattern for the same subjects cases is due to its selection being random in relation to their ranks to the group, i. e., the selection criteria were only based on which subjects had more than one scan present in the dataset. The cluster experiment result also corroborates that the signature can be used to identify groups of individuals. The results from that experiment showed a clear trend for grouping signatures from younger subjects, while the signatures associated with the elder ones were the most dissimilar to the group.

While the signature framework can provide a detailed description of the internal corpus callosum microstructure, analyzing such data has proven a challenging task. The reason is that each signature point is associated with an entire sample, and comparing samples is not a trivial process, despite the advancements in statistics we have. Our first challenge was to find a way to test an empirical sample against another empirical one. Most statistical tests we considered using were designed to test an empirical observation against a theoretical distribution. However, while the test we employed was suitable for our requirements, the binary output of the test is not adequate for defining a proper distance metric to compare signatures, which can be problematic for many studies that need to compare signatures. One example of how this issue can affect analysis can be seen in our same–subject experiment, where two out of the three image pairs from same subjects had similarity measure for about 85%, despite the median signature functions being very

similar in general. If using a proper distance metric we would expect a less significant difference in those cases. Nevertheless, we still consider the defined metric suitable for the purposes of this exploratory study. An extended discussion about how the signature analysis can be improved, along with suggestions for future related research is available in Chapter 5.

5. Conclusion

In this work, we have proposed a framework for performing feature extraction along the corpus callosum extension. Our method provides all the steps needed for establishing a reliable geometrical reference for the extraction of a corpus callosum signature, which holds a rich description of the structural features that are independent of its shape. According to our experimental results, the signatures are expected to fit into a pattern, but also present individual variations. The signatures are comparable as their reference points are expected to describe correspondent corpus callosum regions, although we haven't defined in this work a proper distance function. In addition, it is possible to derive profiles from the signatures for fast visual inspection, which will certainly be suitable for clinical applications. Likewise, it will also be possible to derive other descriptors from the signature, or a set of signatures, which are going to be valuable for several applications. Overall, our proposed method provides the basis for a new corpus callosum analysis approach that is inherently tridimensional and morphological independent.

One major contribution from this work was to bring into attention issues related to the usual MSP analysis approach that has been widely employed to assess the corpus callosum microstructure. Although the issues are evident and concerning, there seems to be a lack of discussion about the subject by the scientific community. In our understanding, the best way to address the issue was to propose an alternative approach for performing the microstructural analysis. Thus, by design, our proposed framework naturally solves the sampling issue of the traditional MSP method, generating data that covers a larger portion of the corpus callosum, thus generating a more representative description of its characteristics. Furthermore, because our framework detaches the data from the corpus callosum shape while still holding the relative location information, it makes partitioning analyses trivial to perform as there is no need to deal directly with the structure shape, and defining partitioning schemata is as simple as establishing breaking points in the generated signature. Given that different portions of the corpus callosum connect distinct cortical areas, this type of analysis is highly appropriate and is also naturally supported by our proposed framework. In our design we provide several unique solutions for establishing a robust geometrical reference for extracting the signature:

1. a fully automated method to estimate the corpus callosum symmetry plane based on the organization of its own internal fibers, which is effective and also very efficient;

2. a robust segmentation procedure with an estimation of the corpus callosum continuous shape, represented by a spline;

3. defining the critical extremes of the median axis by means of multiscalar angular

functions generated from the continuous structure shape, which is invariant to the scale and orientation of the shape, in addition to being robust to specific shape variations.

Finally, our design also provides how to extract the features, storing the estimated empirical distribution function for each reference point across the median axis.

5.1   Method limitations and potential improvements

In this Section we share our experience on developing the proposed framework, highlighting the known limitations and points that have room for improvements. The first topic we are going to address is about the conditions where the method is expected to work. Since the method is based on the own corpus callosum characteristics to establish the median axis, we do not expect the method to work properly in situations where the whole or parts of the corpus callosum is missing, such as in cases of corpus callosum agenesis or surgical removal. Furthermore, the signature extraction will also be compromised in cases where the brain present deformities around the corpus callosum, such as in [27]. In these conditions, a plane would not be adequate for representing a symmetrical separation of the corpus callosum, and therefore the estimated median axis would be incorrectly placed. Using a surface instead of a plane can be a solution to this problem if the study requires the assessment of brains in such conditions. Alternatively, it may be possible to estimate a corpus callosum median axis without relying upon a symmetry plane or surface at all. Although we haven't implemented any model that can proof this concept, from our knowledge about the fiber organization and anatomy of the corpus callosum we recognize it may be feasible to combine morphological techniques with a fiber symmetry function similar to the one we used in this work to estimate the median axis from a rough initial 3–D segmentation of the structure.

Because the median axis holds the base geometrical reference for the signature readings, a primal concern in our design was to ensure its estimation would be robust and invariant to the shape of the corpus callosum, which presents major variations amongst individuals. The proposed solution that relies on the shape angular function analysis was the best approach we found because it is invariant to the orientation of the shape and is able to describe the overall pattern of the shape, which in the case of the corpus callosum is well defined. In fact, this approach made it viable to define the anterior termination of the median axis exactly at the anterior corpus callosum tip, which seems to be the most logical location. Although the detection of this marker will fail in a few cases, this is a minor issue because at that stage a user inspection would be required nonetheless, and the

interaction for the few needed cases would be very easy for the user to perform. However, the real importance of using this location instead of a more robust one as in the posterior termination case is unknown, and may very well be insignificant. We recommend a study to assess this issue, and if the conclusion is that the extreme anterior tip of the corpus callosum is not relevant for any analysis, then a more robust location should be used instead.

In this work we have defined an isomorphic 3–D Gaussian with fixed dispersion to define the sampling area and weights for each reference point along the median axis. This approach gives more relevance to the scalar values closer to the reference points while still considering more distant values. However, there are a number of different approaches for defining the sampling area that is also possible solutions, and signatures extracted using different approaches may have distinct characteristics. For instance, an anisotropic Gaussian with a smaller radius parallel to the median axis would reduce the superposition of values from other reference points. Other solutions could consider the thickness of the corpus callosum to variate the Gaussian radius accordingly or use unweighted sampling. Therefore, we strongly recommend a study about the implications of using different sampling approaches. In addition, this study should also assess how the displacement of the reference points would affect the sample from each sampling approach in order to establish tolerance parameters for the median axis estimation.

5.2 Future related research

Although we have defined a similarity function to compare signatures, it is not a proper distance function and its purpose was to support the experiments we did in this work. Establishing a distance between signatures is a complex problem and will require a specific study. Particularly, most statistical methods to test two empirical distributions would fail to produce distance functions, since they rely on thresholding of a given score. Ideally, a proper distance function would be able to tell for each reference point pair how far apart they are, and which one is lower than the other. A half–solution would be to use a simplified descriptor for the samples, like for instance the median or the mean, and then use this to set a distance function.

The independence the signature has from the shape of the corpus callosum offer a unique opportunity to simplify the analysis of corpus callosum parts. In this sense, a study can be conducted to translate the commonly used partitioning schemata into the signature domain, where the geometrical partitions would be simply breaking points in the signature.

Currently, the correspondence between reference points from signatures is done by geometrical location with a fine tuning using a correlation of the FA median. However, if it is possible to determine through tractography which cortical areas are associated with each location of the corpus callosum, it will be possible to fine–tune the correspondence of reference points based on this information, and consequently, the comparison between signatures would have a stronger meaning.

Publications

- "Mid–callosal plane determination using preferred directions from diffusion tensor images" [9]: in this work, we report the early study we conducted about the symmetry of the corpus callosum internal fibers.

- "Web-based platform for collaborative medical imaging research" [41]: in this work, we present the Adessowiki as a web–based platform for collaborative medical imaging research.

- "Divergence Map from Diffusion Tensor Imaging: Concepts and Application to Corpus Callosum" [38]: in this work, we explore the application of the divergent operator to the vector field defined by the first eigenvectors from DTI.

Submissions

- "Corpus Callosum Fibers Symmetry Plane in Diffusion Tensor Images", originally submitted to IEEE Transactions on Medical Imaging, current being revised for resubmission. This work describes the mid–callosal plane estimation method and the results reported in Chapter 3 from this document.

Bibliography

[1] ABOITIZ, F., AND MONTIEL, J. One hundred million years of interhemispheric communication: the history of the corpus callosum. *Brazilian Journal of Medical and Biological Research 36*, 4 (2003), 409–420.

[2] ABOITIZ, F., SCHEIBEL, A. B., FISHER, R. S., AND ZAIDEL, E. Fiber composition of the human corpus callosum. *Brain research 598*, 1 (1992), 143–153. 01015.

[3] ANDERSON, T. W., AND DARLING, D. A. A test of goodness of fit. *Journal of the American Statistical Association 49*, 268 (1954), 765–769.

[4] APPENZELLER, S., RONDINA, J. M., LI, L. M., COSTALLAT, L. T. L., AND CENDES, F. Cerebral and corpus callosum atrophy in systemic lupus erythematosus. *Arthritis & Rheumatism 52*, 9 (Sept. 2005), 2783–2789.

[5] BARTELS, R. H., BEATTY, J. C., AND BARSKY, B. A. *An Introduction to Splines for Use in Computer Graphics and Geometric Modeling.* Morgan Kaufmann, 1987.

[6] BERGO, F. P., FALCÃO, A. X., YASUDA, C. L., AND RUPPERT, G. C. Fast, accurate and precise mid-sagittal plane location in 3d mr images of the brain. In *Biomedical Engineering Systems and Technologies.* Springer, 2009, pp. 278–290.

[7] CAMPBELL, J. S. W., SIDDIQI, K., RYMAR, V. V., SADIKOT, A. F., AND PIKE, G. B. Flow-based fiber tracking with diffusion tensor and q-ball data: Validation and comparison to principal diffusion direction techniques. *NeuroImage 27*, 4 (2005), 725 – 736.

[8] CAO, Q., SUN, L., GONG, G., LV, Y., CAO, X., SHUAI, L., ZHU, C., ZANG, Y., AND WANG, Y. The macrostructural and microstructural abnormalities of corpus callosum in children with attention deficit/hyperactivity disorder: A combined morphometric and diffusion tensor mri study. *Brain Research 1310* (Jan. 2010), 172–180.

[9] COSTA, A. L., RITTNER, L., LOTUFO, R. A., AND APPENZELLER, S. Mid-callosal plane determination using preferred directions from diffusion tensor images. In *Proc. of SPIE Vol. 9417* (Mar. 2015), B. Gimi and R. C. Molthen, Eds., p. 94172A.

[10] DI PAOLA, M., SPALLETTA, G., AND CALTAGIRONE, C. In vivo structural neuroanatomy of corpus callosum in alzheimer's disease and mild cognitive impairment using different mri techniques: a review. *Journal of Alzheimer's disease 20*, 1 (2010), 67–95.

[11] DIJKSTRA, E. W. A note on two problemsin connexionwith graphs. *Numerische Mathematik 1* (1959), 269–271.

[12] DORON, K., AND GAZZANIGA, M. Neuroimaging techniques offer new perspectives on callosal transfer and interhemispheric communication. *Cortex 44*, 8 (Sept. 2008), 1023–1029.

[13] DOWNHILL, J. E., BUCHSBAUM, M. S., WEI, T., SPIEGEL-COHEN, J., HAZLETT, E. A., HAZNEDAR, M. M., SILVERMAN, J., AND SIEVER, L. J. Shape and size of the corpus callosum in schizophrenia and schizotypal personality disorder. *Schizophrenia research 42*, 3 (2000), 193–208.

[14] FREITAS, P., RITTNER, L., APPENZELLER, S., AND LOTUFO, R. Watershed-based segmentation of the midsagittal section of the corpus callosum in diffusion mri. In *Graphics, Patterns and Images (Sibgrapi), 2011 24th SIBGRAPI Conference on* (2011), pp. 274–280.

[15] GARG, N., REDDEL, S. W., MILLER, D. H., CHATAWAY, J., RIMINTON, D. S., BARNETT, Y., MASTERS, L., BARNETT, M. H., AND HARDY, T. A. The corpus callosum in the diagnosis of multiple sclerosis and other CNS demyelinating and inflammatory diseases. *Journal of Neurology, Neurosurgery & Psychiatry* (Apr. 2015).

[16] GOZZI, M., NIELSON, D. M., LENROOT, R. K., OSTUNI, J. L., LUCKENBAUGH, D. A., THURM, A. E., GIEDD, J. N., AND SWEDO, S. E. A magnetization transfer imaging study of corpus callosum myelination in young children with autism. *Biological Psychiatry 72*, 3 (Aug. 2012), 215–220.

[17] HOFER, S., AND FRAHM, J. Topography of the human corpus callosum revisited—comprehensive fiber tractography using diffusion tensor magnetic resonance imaging. *NeuroImage 32*, 3 (Sept. 2006), 989–994. 00681.

[18] HUANG, H., JIANG, H., WAKANA, S., POETSCHER, L., ZHANG, J., MILLER, M. I., VAN ZIJL, P. C., AND MORI, S. Dti–based parcellation of white matter: Application to the corpus callosum. *Proc. Intl. Soc. Mag. Reson. Med., ISMRM, Berkeley* (2004).

[19] HUTCHINSON, A. D., MATHIAS, J. L., JACOBSON, B. L., RUZIC, L., BOND, A. N., AND BANICH, M. T. Relationship between intelligence and the size and composition of the corpus callosum. *Experimental Brain Research 192*, 3 (Jan. 2009), 455–464.

[20] JACOBSON, S., AND MARCUS, E. M. *Neuroanatomy for the neuroscientist.* Springer, New York, 2011.

[21] JENKINSON, M., BECKMANN, C. F., BEHRENS, T. E., WOOLRICH, M. W., AND SMITH, S. M. Fsl. *NeuroImage 62*, 2 (Aug. 2012), 782–790.

[22] JONES, E., OLIPHANT, T., PETERSON, P., ET AL. Scipy: Open source scientific tools for python, 2001–. [Online; accessed 2018/09].

[23] KIMMEL, R., KIRYATI, N., AND BRUCKSTEIN, A. M. Sub–pixel distance maps and weighted distance transforms. *Journal of Mathematical Imaging and Vision 6*, 2 (June 1996), 223–233.

[24] LACERDA, A. L., BRAMBILLA, P., SASSI, R. B., NICOLETTI, M. A., MALLINGER, A. G., FRANK, E., KUPFER, D. J., KESHAVAN, M. S., AND SOARES, J. C. Anatomical mri study of corpus callosum in unipolar depression. *Journal of Psychiatric Research 39*, 4 (July 2005), 347–354.

[25] LEBEL, C., CAVERHILL-GODKEWITSCH, S., AND BEAULIEU, C. Age-related regional variations of the corpus callosum identified by diffusion tensor tractography. *NeuroImage 52*, 1 (Aug. 2010), 20–31.

[26] LIANG, Z. *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective.* Wiley, John & Sons, Incorporated, 1999.

[27] LIU, Y., COLLINS, R. T., AND ROTHFUS, W. E. Robust midsagittal plane extraction from normal and pathological 3–d neuroradiology images. *Medical Imaging, IEEE Transactions on 20*, 3 (2001), 175–192.

[28] LUDERS, E., NARR, K. L., BILDER, R. M., THOMPSON, P. M., SZESZKO, P. R., HAMILTON, L., AND TOGA, A. W. Positive correlations between corpus callosum thickness and intelligence. *NeuroImage 37*, 4 (Oct. 2007), 1457–1464.

[29] MITCHELL, T. N., FREE, S. L., MERSCHHEMKE, M., LEMIEUX, L., SISODIYA, S. M., AND SHORVON, S. D. Reliable callosal measurement: population normative data confirm sex–related differences. *American Journal of Neuroradiology 24*, 3 (2003), 410–418.

[30] MOKHTARIAN, F., AND MACKWORTH, A. A theory of multiscale, curvature-based shape representation for planar curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence 14*, 8 (Aug. 1992), 789–805.

[31] MOLLINK, J., KLEINNIJENHUIS, M., SOTIROPOULOS, S. N., ANSORGE, O., JBABDI, S., AND MILLER, K. L. Diffusion restriction along fibres: How coherent is the corpus callosum. In *Proceedings of the 23rd Annual Meeting of the ISMRM, Toronto, ON, Canada* (2015), p. 2783.

[32] MORI, S. *Introduction to Diffusion Tensor Imaging.* Elsevier Science, 2007.

[33] NAZEM-ZADEH, M. R., SAKSENA, S., BABAJANI-FERMI, A., JIANG, Q., SOLTANIAN-ZADEH, H., ROSENBLUM, M., MIKKELSEN, T., AND JAIN, R. Segmentation of corpus callosum using diffusion tensor imaging: validation in patients with glioblastoma. *BMC Medical Imaging 12*, 1 (2012), 10.

[34] NIOGI, S. N., MUKHERJEE, P., AND MCCANDLISS, B. D. Diffusion tensor imaging segmentation of white matter structures using a reproducible objective quantification scheme (roqs). *NeuroImage 35*, 1 (Mar. 2007), 166–174.

[35] OTA, M., OBATA, T., AKINE, Y., ITO, H., IKEHIRA, H., ASADA, T., AND SUHARA, T. Age-related degeneration of corpus callosum measured with diffusion tensor imaging. *NeuroImage 31*, 4 (July 2006), 1445–1452.

[36] PARK, J.-S., YOON, U., KWAK, K.-C., SEO, S. W., KIM, S. I., NA, D. L., AND LEE, J.-M. The relationships between extent and microstructural properties of the midsagittal corpus callosum in human brain. *NeuroImage 56*, 1 (May 2011), 174–184.

[37] PAUL, L. K. Developmental malformation of the corpus callosum: a review of typical callosal development and examples of developmental disorders with callosal involvement. *Journal of neurodevelopmental disorders 3*, 1 (2011), 3–27.

[38] PINHEIRO, G., SOARES, G., COSTA, A. L., RITTNER, L., AND LOTUFO, R. A. Divergence map from diffusion tensor imaging: Concepts and application to corpus callosum. *IEEE - EMBC* (2016).

[39] PRIMA, S., AND WIEST-DAESSLÉ, N. Computation of the mid-sagittal plane in diffusion tensor mr brain images. In *Proc. of SPIE 6512, Medical Imaging 2007: Image Processing* (Mar. 2007), J. P. W. Pluim and J. M. Reinhardt, Eds., p. 65121I.

[40] PYKETT, I. L. Nmr imaging in medicine. *Scientific American 246*, 5 (1982), 78–91.

[41] RITTNER, L., BENTO, M. P., COSTA, A. L., SOUZA, R. M., MACHADO, R. C., AND LOTUFO, R. A. Web-based platform for collaborative medical imaging research. In *Proc. of SPIE Vol. 9417* (Mar. 2015), T. S. Cook and J. Zhang, Eds., p. 94180B.

[42] RITTNER, L., FREITAS, P. F., APPENZELLER, S., AND LOTUFO, R. D. A. Automatic dti-based parcellation of the corpus callosum through the watershed transform. *Revista Brasileira de Engenharia Biomédica 30*, 2 (June 2014), 132–143. 00000.

[43] STEGMANN, M. B., SKOGLUND, K., AND RYBERG, C. Mid-sagittal plane and mid-sagittal surface optimization in brain mri using a local symmetry measure. In *Proc.*

SPIE 5747, Medical Imaging 2005: Image Processing (Apr. 2005), J. M. Fitzpatrick and J. M. Reinhardt, Eds., p. 568.

[44] WITELSON, S. F. Hand and sex differences in the isthmus and genu of the human corpus callosum. a postmortem morphological study. *Brain 112* (June 1989), 799–835.