UC Irvine

UC Irvine Electronic Theses and Dissertations

Title

Novel approach for characterizing properties of nerve fiber bundles in Central Nervous System

Permalink

https://escholarship.org/uc/item/7mp5k41t

Author

Vakilna, Yash Shashank

Publication Date

2018

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, availalbe at https://creativecommons.org/licenses/by/4.0/

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA,
IRVINE

Novel approach for characterizing properties of nerve fiber bundles in Central Nervous System

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Biomedical Engineering

by

Yash Shashank Vakilna

Thesis Committee:
Professor Frithjof Kruggel, Chair
Professor Lydia Min-Ying Su
Assistant Professor Beth Lopour

2018

© 2018 Yash Shashank Vakilna

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

Variable Name	Description
MRI	Magnetic Resonance Imaging
DWI	Diffusion Weighted Imaging
DTI	Diffusion Tensor Imaging
DT	Diffusion Tensor
MR	Magnetic Resonance
NMR	Nuclear Magnetic Resonance
SMT	Spherical Mean Technique
GM	Grey Matter
WM	White Matter
CSF	Cerebrospinal Fluid
FA	Fractional Anisotropy
ADC	Apparent Diffusion Coefficients
DSI	Diffusion Spectrum Imaging
QBI	Q-Ball Imaging
GMM	Gaussian Mixture Modeling
HCP	Human Connectome Project
FSL	FMRIB Software Library
FMRIB	Functional Magnetic Resonance Imaging of Brain
EPI	Echo Planar Imaging
BBR	Boundary Based Registration
FLIRT	FMRIBs Linear Image Registration Too
dODF	diffusion Orientation Probability Distribution Function
NIfTI	Neuroimaging Informatics Technology Initiative
EM	Expectation Maximization
SVD	Singular Value Decomposition
PVE	Partial Volume Effect

ACKNOWLEDGMENTS

I would like to express my immense gratitude to my committee chair Dr. Frithjof Kruggel for his unwavering support, encouragement, guidance throughout my tenure in UCI. His teaching sparked in me an appreciation for statistics. I would also like to thank Dr. Beth Lopour and Dr. Lydia Su for their valuable time and suggestions and for being a part of my thesis committee.

I also have deepest gratitude for Dr. Gregory Brewer for being a true mentor. I am eternally grateful to my parents, Shashank Vakilna and Trushna Vakilna, without whom nothing would have been possible. Last but not the least I would like to thank my dearest friends Anupam, Shivani and Shashank, I don't think I would have been able to complete my journey without them.

ABSTRACT OF THE THESIS

Novel approach for characterizing properties of nerve fiber bundles in Central Nervous System

By

Yash Shashank Vakilna

Master of Science in Biomedical Engineering

University of California, Irvine, 2018

Professor Frithjof Kruggel, Chair

Spherical Mean technique (SMT) is a novel method of quantifying the diffusion properties of the nerve fibers bundles in the central nervous system. It does this by calculating the spherical mean of the diffusion signal and fitting it to a parametric equation to obtain per voxel diffusion coefficients. We used Expectation - Maximization to obtain Gaussian Mixture Models (GMM) to find distinct clusters in per voxel coefficient space. We found that the diffusion properties of all the white matter fibers were clustered into a single Gaussian distribution in 867 brain volume samples. This implies that the diffusion properties of the white matter fibers are relatively homogeneous. Then, we checked this result by comparing the clusters obtained using GMM with tissue classification outputs obtained by clustering Fractional Anisotropy (obtained using Diffusion Tensor modeling), T1 weighted image intensity and B0 image intensity for 867 brain volume samples; we observed that the specific clusters of per voxel diffusion coefficients obtained using GMM represent specific tissue types (grey matter fibers, white matter fibers, cerebrospinal fluid). Since the parameters derived from SMT represent the physical diffusion properties that are independent of micro-

scopic fiber orientation and the distribution of diffusion coefficients of white matter can be modeled by a single Gaussian distribution, we can conclude that the diffusion properties of all white matter fiber are homogeneous.

Chapter 1

Introduction

Since its discovery in 1968 by Stejskal and Tanner [35], Diffusion Weighted Imaging (DWI) has been used extensively as a tool to reveal properties of white matter fibers. Advanced computational methods were developed to use this information to elucidate structural connectivity in the brain. DWI is a Magnetic Resonance Imaging (MRI) modality that sensitizes the acquired signal to the Brownian motion of the water molecules undergoing diffusion. The diffusion of water molecules is hindered by cellular structures such as myelin, cell membranes, intra-cellular organelles, which cause them to diffuse non-uniformly in certain directions. This phenomenon, called diffusion anisotropy, gives us valuable information about diagnostically relevant features such as cell size, shape and density.

In this study, we characterized the nerve fibers by calculating their per voxel diffusion coefficients. This problem is highly relevant in numerous clinical applications areas such as clinical neurology [37]. Changes in white diffusivity can be used to differentiate tumor grades [40] and quantify aging processes. Also, it helps in evaluating

neuronal damage following cardiac arrest or traumatic brain injury and also in lesion localization [38].

Over the years, numerous image and data processing techniques have been devised to extract relevant information from diffusion weighted signal. Diffusion Tensor Imaging (DTI) is the most widely used technique in a clinical setting. Quantitative measures, such as Fractional Anisotropy (FA) and Apparent Diffusion Coefficient (ADC), derived from tensor modeling are widely used for diagnostic purposes. However, the principal problem encountered by these measures is that they assume that all the fibers in a voxel are aligned in the same direction. Therefore, the diffusion tensor model works well in the region consisting of only one fiber bundle. However, it is unreliable in regions with several populations which are aligned along intersecting axes [7, 19, 11]. Several strategies were proposed to fix this problem. Newer signal acquisition methods, such as High Angular Resolution Diffusion Imaging (HARDI) [7], and increasingly complex biophysical models, such as Diffusion Spectrum Imaging, Q-Ball Imaging [37], spherical deconvolution representation [16], etc., were proposed. Unfortunately, all these novel data processing schemes require unusually high acquisition time, or complex gradient sequences which makes their adoption in clinical setting significantly harder [19]. Therefore, we are using parameters derived from Spherical Mean Technique (SMT). The advantages of using this technique are twofold. First, the per voxel diffusion coefficients estimated by SMT do not require information about intra-voxel parameters such fiber microanatomy, inter-fiber space inside the voxel, or the degree of myelination [19]. Second, this technique utilizes conventional pulse sequences and the only requirement is multiple acquisition with at least two b-values. As a result, it can be easily adopted in a clinical setting.

This study is an attempt to characterize nerve fibers using per voxel diffusion coeffi-

cients derived using SMT. These parameters are clustered into GM or WM and this classification was tested against tissue type classification obtained using DT modeling. This dissertation is organized as follows. The next chapter provides basic theoretical concepts necessary for understanding this technique. Additionally, it also describes popular alternative techniques for processing diffusion weighted data. The subsequent chapter elaborates on the methodology of data processing and focuses on details of implementation such as pre-processing, optimization parameters, and clustering algorithms, namely Gaussian Mixture Modeling and K-means clustering. The next chapter describes the results of the data processing algorithms, along with the tests performed to validate the output of this scheme. Finally, the thesis concludes by summarizing the work and provides future direction for further study in this area.

Chapter 2

Background

2.1 Magnetic Resonance Imaging

Nuclear Magnetic Resonance (NMR) was first observed by Isidor Rabi in 1940. It quickly became one of the most popular analytical tools to study the molecular structure in chemistry and biochemistry. In early 1970, researchers realized that magnetic field gradients can be used to localize NMR signals and to generate images displaying the magnetic properties of protons [29]. This discovery, coupled with the development of body-sized magnets and advances in computed tomography, sparked a significant interest among clinicians to use NMR to image the human body to detect pathological findings inside the body non-invasively.

Magnetic Resonance (MR) imaging utilizes the presence of atoms with unary protons in the form of water molecules in the human body. These protons have an intrinsic magnetic moment which, under normal ambient conditions, cancels each other's mag-

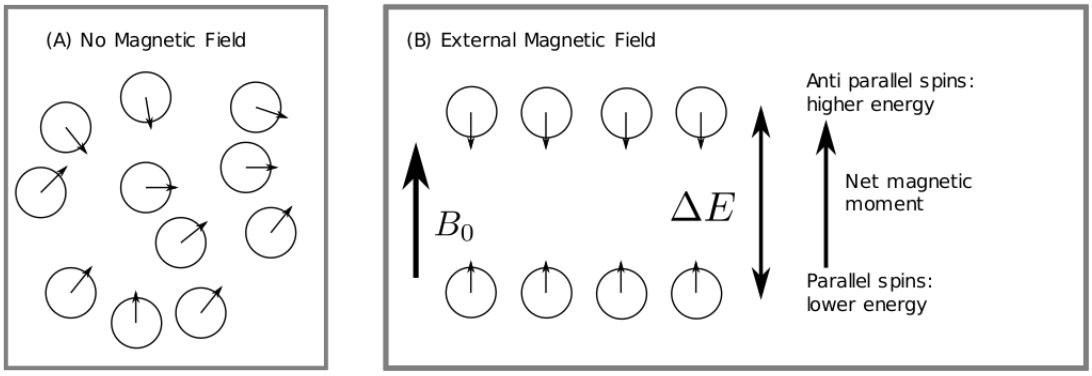

Figure 2.1: A. Magnetic moments are randomly oriented in absence of magnetic field. B. Magnetic moments align in the direction of external magnetic field [29]

netic moments resulting in net zero magnetic moment. During MR image acquisition a strong, uniform, external magnetic field aligns these magnetic moments in an either parallel or anti-parallel direction with respect to the magnetic field (Fig.2.1). This alignment is perturbed by inducing oscillation in the protons by application of radio frequency pulse resonant to the Larmor frequency, given by Eqn. (2.1).

$$\omega_{\text{Larmor}} = \gamma B_0 \tag{2.1}$$

where B_0 is the external magnetic field.

These oscillations are short lived and decay quickly. This decay gives information about the medium in which these protons are oscillating. There are two time scales associated with the decay of these oscillations, $T1$ and $T2$

- $T1$ (longitudinal relaxation times) is the time constant associated with rate at which excited protons return to equilibrium. It is a measure of the time taken for spinning protons to realign with the external magnetic field. This phenomenon

is called *spin-lattice relaxation* and can be expressed mathematically as:

$$M_z = M_0 \left(1 - e^{-t/T1}\right) \tag{2.2}$$

where M_z is the longitudinal magnetization at time t for a sample and $T1$ is the time needed for the recovery of $1 - \frac{1}{e}$ fraction of M_z. The modality of MRI which is sensitive to the $T1$ time constants is called T1-weighted imaging.

- $T2$ (longitudinal relaxation time) is the time constant that determines the rate of dephasing of the excited protons caused due to the Brownian motion of water molecules. This dephasing, called free induction decay, is a key physical phenomenon that gives us significant information about the medium. It can be expressed mathematically as:

$$M_{xy} = M_0 e^{-t/T2} \tag{2.3}$$

where M_{xy} is the transverse magnetic moment at time t for a sample with M_0 transverse magnetic moment at $t = 0$. $T2$ is the time when $M_{xy} = \frac{1}{e} M_0$. The modality of MRI which relies on different $T2$ time constants as contrast is called T2-weighted imaging.

2.2 Diffusion Weighted Imaging

Diffusion can be described as the tendency of molecules to travel along the concentration gradient. However, Einstein realized that the molecules undergo random Brow-

nian motion even in the absence of any concentration gradient. The phenomenon, called self-diffusion, involves a linear time dependence of the mean squared displacement of a particle:

$$\langle r^2 \rangle = 6Dt \tag{2.4}$$

where, $\langle r^2 \rangle$ is the mean squared displacement of the molecule in three dimensions, D is the diffusion coefficient and t is the diffusion time. This results in the probability distribution function $p(\mathbf{r}, t)$ of motion a freely diffusing molecules that follows Gaussian distribution, as shown in Eqn. (2.5).

$$p(\mathbf{r}, t) = \frac{1}{\sqrt{(4\pi t)^3 D}} \exp\left(\frac{-\mathbf{r}^{\mathrm{T}}\mathbf{r}}{4tD}\right) \tag{2.5}$$

where, D is the diffusion coefficient, $\mathbf{r}$ signifies the displacement of the molecule from the origin in time t.

In 1965, Stejskal and Tanner realized that an MRI artifact, which results in slight dephasing of spins caused by the Brownian motion of molecules, can be used to quantify the motion of water molecules inside a tissue sample. In their seminal paper, they introduced Diffusion weighted imaging (DWI) [35].

The diffusion of water molecules is dependent on their surroundings. For example, in the axons the water molecules can readily diffuse in the direction parallel to the axons, as compared to the perpendicular direction. As a result, diffusion weighted imaging can be used to track the nerve fiber bundles in the brain and is widely used to study the structural connectivity and fiber tractography [23]. Consequently, DWI is used to detect early identification of acute ischemia, several types of brain tumors, white

matter diseases, breast cancer, hepato-biliary pancreatic cancers, bowel disorders, and genito-urinary diseases [14].

For diffusion weighted MR signal acquisition, the homogeneity of the magnetic field is varied linearly by a pulsed field gradient. Since the Larmor precession frequency is proportional to the magnet strength, the protons begin to precess at different rates, resulting in dephasing and signal loss. This is followed by refocusing pulse of the same magnitude but with opposite direction. This refocusing will not be perfect for the protons that have moved during the time interval between the pulses, which leads to signal attenuation. This attenuation of the signal due to the diffusion process is governed by the following equation.

$$\frac{S}{S_0} = \exp\left(-\gamma^2 G^2 \delta^2 \left(\Delta - \frac{\delta}{3}\right)\right) = \exp(-bD) \tag{2.6}$$

where γ is the gyromagnetic ratio, G is the strength of the gradient pulse, δ is the duration of the pulse, Δ is the time between the two pulse, D is the diffusion coefficient, and b is the b-value of signal acquisition [6]. The above equation can be modified to calculate the Apparent Diffusion Coefficient (ADC) [21].

$$D = -\frac{1}{b} \ln \frac{S}{S_0} \tag{2.7}$$

One of the limitations of acquiring diffusion weighted MR signal is that diffusion is measured only in the direction parallel to DWI gradient, which necessitates multiple 3-D image acquisitions. Over the years, numerous mathematical procedures have been proposed to increase the accuracy of determining diffusion characteristics per voxel from the 3-D brain volumes.

2.2.1 Scalar Model and Apparent Diffusion Coefficient

The scalar model is the simplest extension of Eqn. (2.6). Three diffusion images (S_x, S_y, S_z) are acquired and so three ADC (ADC_x, ADC_y, ADC_z) are calculated along the three orthogonal directions by solving the equation three times followed by taking the geometric average to obtain pseudo ADC [6].

$$\langle ADC \rangle = \frac{ADC_x + ADC_y + ADC_z}{3} \tag{2.8}$$

This is model is computationally extremely efficient and is still a part of the portfolio of MRI used in clinical emergency [7, 22]. However, it is clear that diffusion is a 3D phenomenon and a mathematical framework is required to perform 3D diffusion imaging independent of (x, y, z) axes.

2.2.2 Diffusion Tensor Imaging (DTI)

The Gaussian distribution in Eqn. (2.5) describes the motion of water molecules in the absence of any hindrance. However, the water molecules in nerve fibre bundles encounter several barriers in the form of cellular structures structures such as organelles and cell membrane. This causes three dimensional direction dependent diffusion of the molecules which cannot be described using a scalar ADC. As a result the scalar model in Eqn. (2.6) is extended by replacing ADC with Diffusion Tensor (DT) to capture the anisotropy of diffusion of the water diffusion [3, 4]. DT is a 3×3,

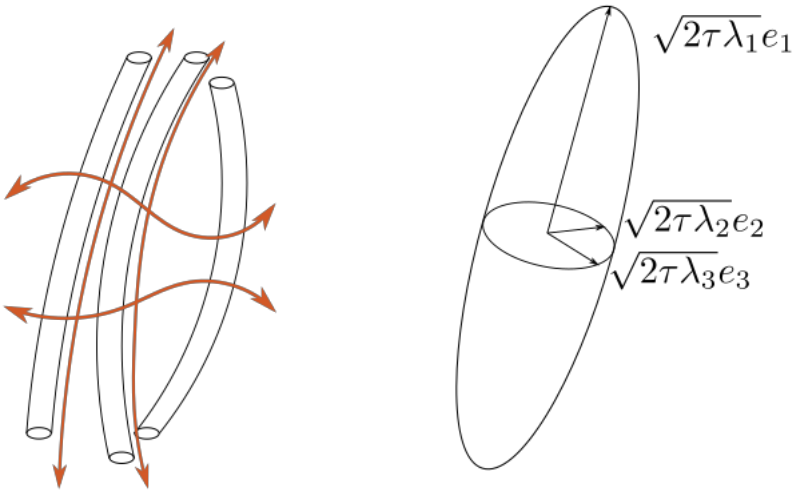

Figure 2.2: The diffusion tensor model models diffusion with ellipsoids [7]

.

symmetric positive semi-definite matrix.

$$D = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{pmatrix} \tag{2.9}$$

By multiplying $\mathbf{g}$, the unit gradient direction, with each side of the tensor, we obtain Eqn. (2.10).

$$\mathbf{g^T D g} = \begin{pmatrix} g_x, g_y, g_z \end{pmatrix} \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{pmatrix} \begin{pmatrix} g_x \\ g_y \\ g_z \end{pmatrix}$$

$$= D_{xx}{g_x}^2 + D_{yy}{g_y}^2 + D_{zz}{g_z}^2 + 2D_{xy}g_xg_y + 2D_{xz}g_xg_z + 2D_{yz}g_yg_z \quad (2.10)$$

The above equation represents an ellipse in 3-D as illustrated in Fig. 2.2, signifying that the diffusion process in each voxel is modeled as a single Gaussian distribution. Using this model to extend Eqn (2.6), we get:

$$S(\mathbf{g}, b) = S_0 \exp(-b\mathbf{g^T D g}) \quad (2.11)$$

The above equation contains six unknown coefficients and therefore requires six diffusion weighted image volumes to compute. We can obtain the three eigenvalues ($\lambda_1 \geq \lambda_2 \geq \lambda_3$) and three eigenvectors (e_1, e_2, e_3) by performing singular value decomposition on DT. These eigenvalues and eigenvectors fully describe the geometric and diffusion properties of the tensor [7]. e_1 gives the principal direction of DT, and along with the other two eigenvectors, span the orthogonal planes. Several rotationally invariant scalar values that quantify diffusion characteristics, can be extracted from this model; following equations describe some such examples:

$$\text{Fractional Anisotropy (FA)} : \frac{3}{2}.\sqrt{\frac{(\lambda_1 - \lambda_2)^2 + (\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_2)^2}{(\lambda_1)^2 + (\lambda_2)^2 + (\lambda_3)^2}} \quad (2.12)$$

$$\text{Mean Diffusivity (MD)}: \frac{(\lambda_1) + (\lambda_2) + (\lambda_3)}{3} \quad (2.13)$$

$$\text{Radial Diffusivity (RD)} = \frac{\lambda_2 + \lambda_3}{2} \tag{2.14}$$

Since diffusion characteristics arise from constrained diffusion of water molecules due to intact cellular structures such as cell membrane and myelin sheath [5], change in these diffusion characteristics is indicative of a structural abnormality. Therefore, these quantities are widely used in the clinical settings to diagnose structural abnormalities such as ischemic brain injury, adult stroke, and inflammation [10]. Advantages of DTI is its short acquisition time and its availability in all existing clinical systems. However, for voxels which have multiple fiber populations, the Gaussian diffusion model assumption is violated. This means, in the worst case, DTI does not accurately measure the diffusion process for 66% of the voxels [7] .

2.2.3 Diffusion Spectrum Imaging and Q-ball imaging

In conventional MRI ($T1$ or $T2$ weighted), the MR signal is phase and frequency encoded by applying gradients in specific directions and intensities and at specific times. As a result, the raw image acquired is in the spatial frequency domain, which can be explained mathematically as [27]:

$$S = \iiint f(\mathbf{r}).\exp\left(-j2\pi(\mathbf{k^T x})\right)d\mathbf{x} \tag{2.15}$$

where, $f(\mathbf{x})$ is the spatio-temporally varying MR signal, and $d\mathbf{x}$ is the differential volume element in the patient and $j = \sqrt{-1}$. A Fourier transform is applied to the

raw image to obtain a processed human readable image. A similar idea is used in diffusion weighted imaging wherein diffusion data is acquired in q-space

$$S(\mathbf{q},t) = S_0 \int_{R^3} p(\mathbf{r},t).\exp(-j2\pi\mathbf{q}^{\mathrm{T}}\mathbf{r})d\mathbf{r} \tag{2.16}$$

where, $p(\mathbf{r},t)$ is called *diffusion propagator*, and $\mathbf{q}$ is the wave vector.

For every voxel, **q**-space represents this 3D coordinate system, which upon Fourier transform, yields a probability distribution function, which is called diffusion propagator. This mathematical model of measuring diffusion is called **q**-space imaging.

Diffusion Spectrum Imaging (DSI) is an implementation of **q**-space imaging in 3-D using the Cartesian sampling of a large number of **q**-space points as seen in Fig. 2.3. It does not rely on any particular model or hypothesis and therefore, is only limited by the resolution of data acquisition in **q**-space. However, this type of imaging is clinically not feasible as the time required for acquiring an image is too long. **Q-ball imaging** (QBI) tries to combine the best of both model free and model driven strategies. A voxel is assumed to have the compartment inside voxel roughly resembling a pin cushion, consisting of a set of straight and thin pipes with impermeable walls [11]. The diffusion orientation probability distribution (dODF) is estimated using only a single shell sampling of **q**-space as demonstrated in Fig. 2.3. dODF captures the angular content of the diffusion propagator.

$$\psi(\theta,\phi) = \int_0^\infty p(r,\theta,\Phi)r^2 dr \tag{2.17}$$

where ψ is the dODF, p is the diffusion propagator. r,θ,ϕ are the spherical coordinates, where (θ,ϕ) obey the physics convention ($\theta \in [0,\pi]$ is the polar angle and

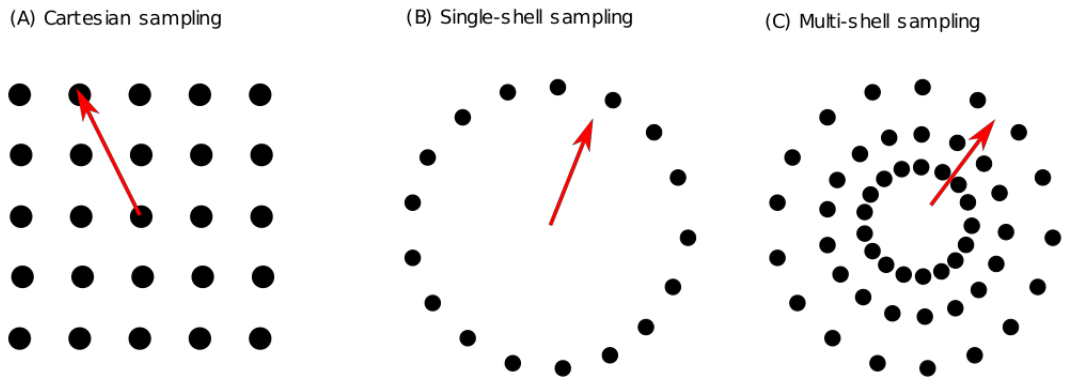

Figure 2.3: Sampling scheme in **q**-space. (a) Cartesian sampling dedicated to diffusion DSI. (b) Single-shell spherical sampling dedicated to Q-ball imaging. (c) Multiple shell spherical sampling scheme dedicated to SMT [7]

.

$\phi \in [0, 2\pi)$ is the azimuthal angle).

dODF is estimated by using a mathematical technique called Funk-Radon transform. Intuitively, the Funk-Radon transform at a given spherical point can be understood as computing the great circle integral of the diffusion signal on the sphere defined by the plane through the origin perpendicular to the point of evaluation as demonstrated in Fig. 2.4.

As mentioned in the previous section, DTI is unable to resolve intra-voxel orientational heterogeneity, such as, nerve fiber crossing, bending, or twisting. This is because it assumes a single diffusion compartment per voxel, and therefore it can have one maximum only. Since QBI does not require the diffusion process in a voxel to be Gaussian or multi-Gaussian distributed, it can resolve the intra-voxel orientation heterogeneity.

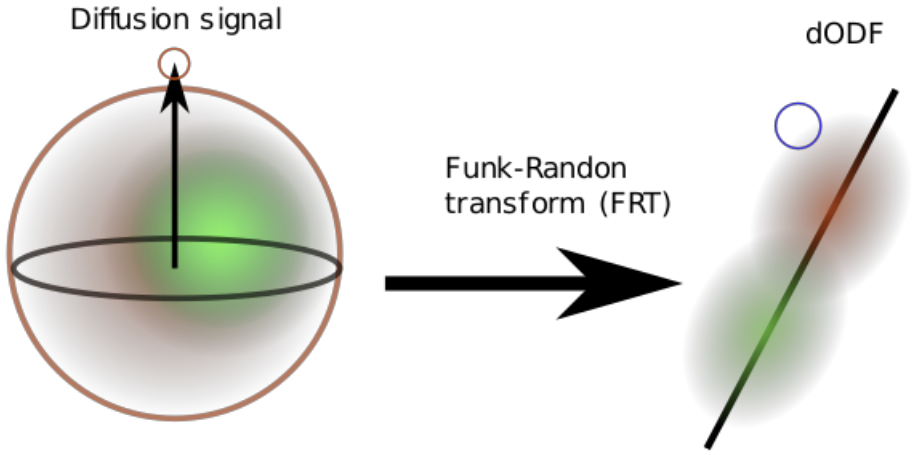

Figure 2.4: Funk-Radon transform. Great circle integrals are computed from the measured signal [7]

2.2.4 Spherical Deconvolution Technique

Just as Fourier series form a complete orthonormal basis in Cartesian space, spherical harmonics form a complete orthonormal basis set for all the functions over a sphere. Therefore, the spherical harmonic functions can be used to represent any bounded function over a sphere [36]. Similarly, rotational harmonics form a complete orthonormal basis over the space of pure rotations.

Using these basis sets, spherical deconvolution is performed as an ensemble of rotations acting on the function defined over a sphere. Spherical convolution works under the assumption that the measured diffusion (S) signal is the result of convolution between the distribution of fiber orientations (R) and the single-fiber response function (F) [7]. As shown in Fig. 2.5, nth order spherical harmonic decomposition is

formulated as:

$$S^n = R^n F^n \tag{2.18}$$

The single-fiber response function is estimated empirically by using voxels with an FA value above 0.7 [7] using DTI model [36], e.g., as found in the corpus callosum. Once R is known, F can be calculated by inverting each R^n to recover F^n.

$$F^n = (R^{\mathrm{T}} R)^{-1} S^n \tag{2.19}$$

Since spherical deconvolution requires a deconvolution kernel, it cannot be considered as a model-free technique. Since the technique does not assume any model of diffusion, it does not suffer from problems which are encountered by DTI. However, the linear solution to the spherical deconvolution suffers from significant instability at the higher harmonic order which leads to negative fiber orientation values and spurious peaks [7]. Hence, spherical deconvolution is constrained by using nonlinear methods to deal with this instability.

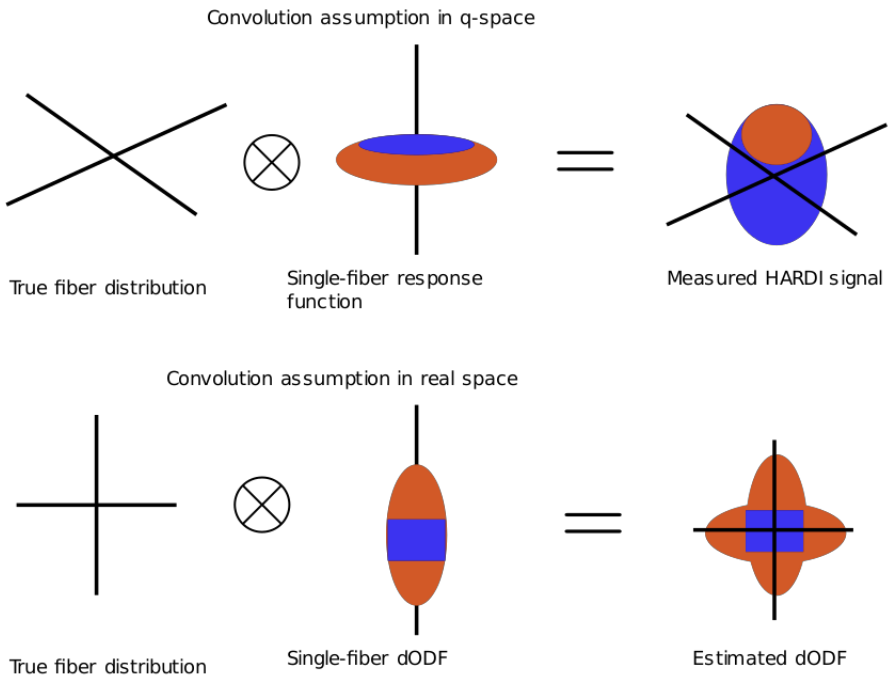

Figure 2.5: Spherical deconvolution technique. Model formulation in q-space and real space [7]

Chapter 3

Methodology

3.1 Overview of Data Processing

In this study, we used a pipeline of data processing steps on the acquired MRI data. These processes can be roughly divided into following categories:

- Preprocessing: Cleaning the data for the purpose of analysis so that the artifacts acquired during data generation do not affect the results. Following steps were implemented during preprocessing:
    1. B0 intensity normalization.
    2. Distortion correction due to eddy currents and subject motion.
    3. Registration of B0 to T1 weighted images.
    4. Removal of Rician noise and brain volume extraction.
- Spherical mean estimation: Extracting meaningful parameters using parametric

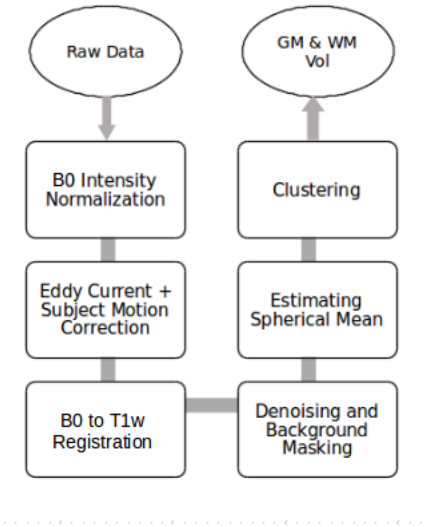

Figure 3.1: Graphical representation of the data processing protocol

equations.

- Clustering: Generating meaningful insights from the data by processing them using unsupervised clustering methods.

This section starts with detailing the acquisition DWI data used for this study; this is followed by explanation of the methods used for its preprocessing. The details of T2 weighted imaging acquisition and preprocessing are explained in the Appendix. This is followed by description of Spherical Mean Technique (SMT) and calculation of per voxel diffusion coefficient. The chapters ends with an explanation of clustering methods used here.

3.2 Data Overview

Data were provided by the Human Connectome Project (HCP), WU-Minn Consortium (available online at https://humanconnectome.org). The datasets were acquired

using Siemens 3 Tesla Skyra scanner equipped with a 32 channel phased array head coil and a customized SC72 insert with gradient strength up to 100 mT/m.

The diffusion datasets had three b-shells of 1000, 2000, and 3000 s/mm^2 with 90 gradient directions each, which were acquired using Stejskal–Tanner sequence. Echo time (TE) of 89 ms and repetition time (TR) of 5.5 s were matched across all the shells, which were interleaved during acquisition. A pair of apodized sinc RF pulses was used for excitation and refocusing, with flip angles empirically set to 78 and 160 degrees, respectively, with a multiband factor of three. The diffusion data was acquired with phase encoding in both right-to-left and left-to-right directions. Additionally, six images with zero diffusion weighting (B0) were acquired, which were evenly distributed across the experiment. The brain volume was divided into 111 slices as the data was acquired; each in the form of a 144 $\times$ 168 image matrix with a field of view (FOV) 210 $\times$ 180 mm^2 [13].

3.3 Preprocessing

The raw image acquired by the MRI machine is known to have low anatomical accuracy and internal consistency. The causes for the distortions are varied and may include magnetic field inhomogeneity caused due to magnetic susceptibility (explained in details later), eddy currents induced due to rapidly changing magnetic field [2], subject motion and many more. Therefore, preprocessing of the data is necessary before analysis.

In order to remove these distortions, we employed a diffusion preprocessing pipeline, published by the Human Connectome Project (HCP) [8]. Next, we used custom

C++ algorithms to remove distortions due to noise, as explained in [19], to mask the background and extract the brain from the image.

3.3.1 HCP Preprocessing Pipelines

These tools are implemented in the form of bash scripts that interact with the FMRIB Software Library (FSL).

3.3.1.1 Intensity normalization and susceptibility induced distortions removal

Diffusion images were acquired using Echo Planar Imaging (EPI), and therefore, are sensitive to non-zero off-resonance magnetic fields. Such fields are the result of susceptibility distribution of the subject's head and of eddy currents generated due to rapid switching of the diffusion gradients. Here, susceptibility is a measure of the extent to which a substance gets magnetized when placed in magnetic field.

The "topup" tool calculates the induced susceptibility field. It uses two images with opposite phase encoding direction and estimates the field that when applied to the two volumes (of opposite phases encoding direction) will maximize the similarity of the un-warped volume. In the pipeline, "topup" normalizes the intensity of mean B0 image across six diffusion volumes.

3.3.1.2 Eddy current induced field inhomogeneity and head motion correction

This tool simultaneously calculates the effects of diffusion eddy currents and movement of the head. It does this by modeling the diffusion signal as a Gaussian process. All the field distortions, including the distortion calculated by "topup", are fed to a Gaussian process predictor and the corrections are applied in a single resampling step [2].

3.3.1.3 Registration to T1 weighted image

During the final step, the pipeline calculates the transform between the diffusion and the structural space. The B0 image is registered to the T1 weighted image using translation and rotation. This is done using FMRIB's Linear Image Registration Tool (FLIRT) using Boundary Based Registration (BBR) cost function .

3.3.2 Change of Format

Further analysis of the data was performed using BRIAN 3.0.0 Signal and Image processing libraries implemented in C++. Therefore, we had to convert the images from the Neuroimaging Informatics Technology Initiative (NIfTI) format to the BRIAN format [20].

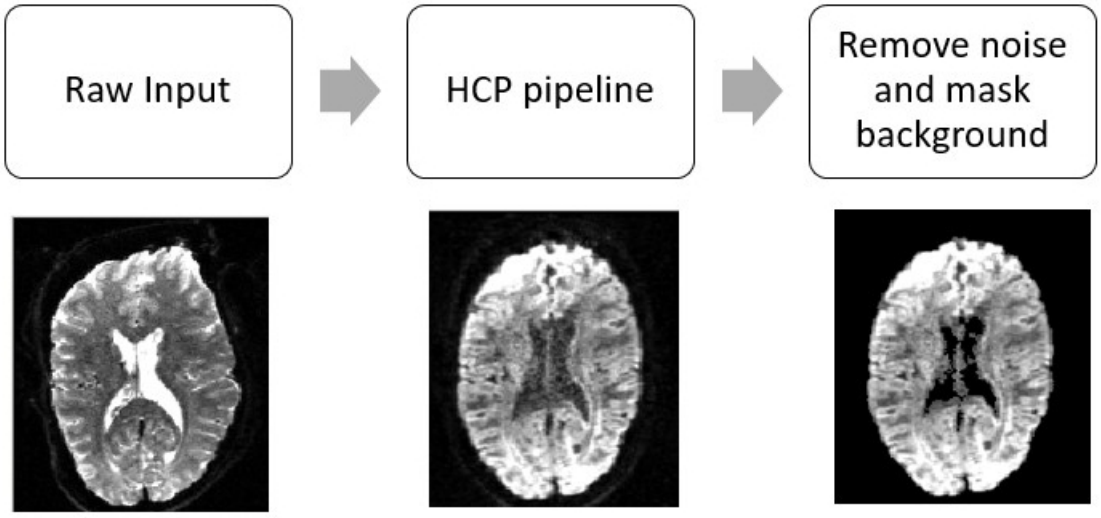

Figure 3.2: Representative example of preprocessing

3.3.3 De-noising

The noise of the MRI signal conforms to a Rician distribution $R(E, \varsigma)$ [9]. So, in order to minimize the potential effect of noise induced bias, the following correction was made:

$$\hat{E} = \arg\min_{E \geq 0} \left(S - \sqrt{\frac{\pi\varsigma^2}{2}} L_{1/2}\left(-\frac{E^2}{2\varsigma^2}\right)\right)^2 \tag{3.1}$$

where, $\hat{E}$ is the adjusted signal. The second term on the right-hand side stands for the mean of the Rician distribution, and L_n denotes the nth Laguerre polynomial. ς is estimated by computing the median of the noise level estimated voxel by voxel from the B0 images using a maximum-likelihood approach [33].

3.3.4 Remove Background

Following steps were taken to remove the background of the image [34]:

1. The image was binarized using an empirical threshold at the 65th percentile of intensities of B0 image.
2. First, morphological opening operation was applied, which was followed by applying morphological closing operation.
3. Selecting the biggest component of the binarized image using C18 face and edge connectivity as a mask.
4. Removing all the voxels of DW image outside the mask.

3.4 Spherical Mean Technique

The macroscopic diffusion weighted MR signal is primarily impacted by two biophysical effects – the microfiber orientation and microscopic anisotropy inside a voxel. One of the major aims of current diffusion MRI research is to disentangle the impact of these two effects [18]. The major insight of SMT is that the spherical mean of diffusion MR signal only depends on the b-value of acquisition and is independent of the fiber orientation distribution. Moreover, it has an additional advantage that it does not involve complex gradient waveform or magic-angle spinning, which are currently unavailable in clinical setting, it simply requires conventional pulse gradient sequence featuring two or more b-shells [19]. Consider two axons with different

diffusion properties; the result signal can be calculated using Eqn. (2.11):

$$S(b, \mathbf{g}) = S_0(a_1 \exp(-b\mathbf{g}^\mathbf{T}\mathbf{D_1}\mathbf{g} + a_2 \exp(-b\mathbf{g}^\mathbf{T}\mathbf{D_2}\mathbf{g})) \tag{3.2}$$

where D_n is the diffusion tensor and a_n is the volume fraction of the nth axon. Now, for axially symmetric volumes the diffusion tensor can be written as:

$$D_n = \begin{pmatrix} \lambda_\perp & 0 & 0 \\ 0 & \lambda_\perp & 0 \\ 0 & 0 & \lambda_\parallel \end{pmatrix} \tag{3.3}$$

where $\lambda_\perp$and $\lambda_\parallel$ are the eigenvalue for the eigenvector parallel and perpendicular to the nth axon, respectively, with the assumption that $0 \leq \lambda_\perp \leq \lambda_\parallel \leq \lambda_{\text{free}}$, where the free water diffusion λ_{free} is about 3.05 μm^2/ms at 37°C [28]. These are often referred to as transverse and longitudinal diffusion coefficients, respectively. Using Eqn. (3.3), diffusion measured at a particular direction (in the diagonal coordinate system with coordinates (α, ψ)) relative to an arbitrary reference direction in the plane perpendicular to the axon, we get (3.4).

$$\begin{aligned} D(\alpha, \psi) &= \begin{pmatrix} \sin\alpha\ \cos\psi & \sin\alpha\ \sin\psi & \cos\alpha \end{pmatrix} \begin{pmatrix} \lambda_\perp & 0 & 0 \\ 0 & \lambda_\perp & 0 \\ 0 & 0 & \lambda_\parallel \end{pmatrix} \begin{pmatrix} \sin\alpha\ \cos\psi \\ \sin\alpha\ \sin\psi \\ \cos\alpha \end{pmatrix} \\ &= \sin^2\alpha\lambda_\perp + \cos^2\alpha\lambda_\parallel \\ &= \lambda_\perp + \cos^2\alpha(\lambda_\parallel - \lambda_\perp) \end{aligned} \tag{3.4}$$

where α is the polar angle and ψ is the azimuthal angle.

Assuming both the axons in Eqn. (3.2) have the same longitudinal and transverse diffusion coefficients ($\lambda_\perp$,$\lambda_\parallel$) but different fiber axes, the signal measured at an arbitrary point becomes

$$S(b,\mathbf{g}) = S_0(\exp(-b\lambda_\perp)(a_1 \exp(\lambda_\parallel - \lambda_\perp)\cos^2\alpha_1 + a_2 \exp(\lambda_\parallel - \lambda_\perp)\cos^2\alpha_2)) \quad (3.5)$$

where α_n is the angle between the diffusion gradient direction and the nth axon.

The above result can be generalized easily for an arbitrary angular distribution of fibers $P(\theta', \phi')$ as shown in Eqn. (3.6).

$$S(b,\theta,\phi) = S_0 \exp(-b\lambda_\perp) \int_0^{2\pi} \int_0^{\pi} P(\theta',\phi') \exp(-b(\lambda_\parallel - \lambda_\perp)\cos^2\alpha_{RR'}) \sin\theta' \, d\theta' \, d\phi' \quad (3.6)$$

where θ' and ϕ' are the polar and azimuthal angle, respectively, of the axon in axis direction **R'**. Similarly, θ and ϕ are the polar and azimuthal angles of the diffusion gradient direction **R**, and $\alpha_{RR'}$ is the angle between **R** and **R'** as shown in Eqn. (3.6). This can be interpreted as $S(\theta,\phi)$ being proportional to the convolution of $P(\theta,\phi)$ with the kernel

$$\exp(-b\theta_\perp)\exp(-b(\lambda_\parallel - \lambda_\perp)\cos^2\alpha)$$

which is the impulse response of a single fiber [1]. Transforming Eqn. (3.6) into

spherical harmonics space as shown in Eqn. (3.7) and Eqn. (3.8), we get:

$$S(\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} s_{lm} Y_{lm}(\theta,\phi) \tag{3.7}$$

$$P(\theta',\phi') = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} p_{lm} Y_{lm}(\theta',\phi') \tag{3.8}$$

where,

$$s_{lm} = \int_0^{2\pi} \int_0^{\pi} Y^*_{lm}(\theta,\phi) S(\theta,\phi) \sin\theta d\theta d\phi \tag{3.9}$$

$$p_{lm} = \int_0^{2\pi} \int_0^{\pi} Y^*_{lm}(\theta',\phi') P(\theta',\phi') \sin\theta' d\theta' d\phi' \tag{3.10}$$

As a result, we get simple relation,

$$s_{lm} = S_0 \; c_l \; p_{lm} \tag{3.11}$$

Now in order to calculate p_{lm}, let us substitute $l = m = 0$ in Eqn. (3.10) [1]:

$$\begin{aligned} p_{00} &= \frac{1}{\sqrt{4\pi}} \int_0^{2\pi} \int_0^{\pi} P(\theta',\phi') \sin\theta' \; d\theta' \; d\phi' \\ &= \frac{1}{\sqrt{4\pi}} \end{aligned} \tag{3.12}$$

Similarly,

$$s_{00} = \frac{1}{\sqrt{4\pi}} \int_0^{2\pi} \int_0^{\pi} S(\theta, \phi) \sin\theta \; d\theta \; d\phi$$
$$= \sqrt{4\pi} \; \overline{S} \tag{3.13}$$

$$c_0 = 2\pi^{3/2} \frac{\mathrm{erf}(\sqrt{b(\lambda_\parallel - \lambda_\perp)})}{\sqrt{b(\lambda_\parallel - \lambda_\perp)}} \exp\left(-b\lambda_\perp\right) \tag{3.14}$$

After substituting Eqs. (3.12), (3.13), and (3.14) into Eqn. (3.11) for $l = 0$, we get

$$\overline{e}_b(\lambda_\parallel, \lambda_\perp) = \frac{\overline{S}(b, \lambda_\parallel, \lambda_\perp)}{S_0} = \exp\left(-b\lambda_\perp\right) \frac{\mathrm{erf}(\sqrt{b(\lambda_\parallel - \lambda_\perp)})}{2\sqrt{b(\lambda_\parallel - \lambda_\perp)}} \tag{3.15}$$

Since Eqn.(3.15) has two unknowns, we require at least two DW images with different b values. The image acquisition is similar to Fig. 2.3.

Now $\lambda_\parallel$ and $\lambda_\perp$ are determined by performing constrained least-squares optimization [19] .

$$\min_{0 \le \lambda_\perp \le \lambda_\parallel \le \lambda_{\mathrm{free}}} \sum_{i=1}^{n} (\hat{e}_{b_i} - \bar{e}_{b_i}(\lambda_\parallel, \lambda_\perp))^2 \tag{3.16}$$

Fractional Anistropy can be used as a metric to calculate anisotropy and can be calculated using longitudinal and transverse diffusion coefficients as shown in Eqn. (3.17).

$$\mathrm{FA} = \sqrt{\frac{3}{2} \frac{(\lambda_\perp - \hat{\lambda})^2 + 2(\lambda_\perp - \hat{\lambda})^2}{{\lambda_\parallel}^2 + 2{\lambda_\perp}^2}} \tag{3.17}$$

3.5 Unsupervised Clustering

Unsupervised machine learning are sets of algorithms that deduce a function that describes the structure of the unlabeled data. Clustering algorithms is a subset of unsupervised machine learning algorithms, which infer the structure of data by grouping sets of objects based on a specific notion of similarity. This technique is often used in exploratory data analysis in various fields such as pattern recognition, information retrieval, bioinformatics, data compression, and computer graphics.

Since we did not have access to any labeled data sets relevant to our problem, we used unsupervised clustering methods – K-means clustering (K-means) and Gaussian mixture Modeling (GMM) using Expectation-Maximization (E-M) approach. The motivation behind choosing GMM is that quantities that are expected to be the sum of number independent processes often have distributions that are nearly Gaussian [25] and the human brain, being a biological system, is expected to have various independent processes that might influence diffusion weighted MR signal. K-means clustering was chosen because of its simplicity and speed.

3.5.1 K-Means Clustering

K-means clustering is used to partition a set of observed data into K disjoint clusters. Consider a set of N observations $\{\mathbf{x}_1, \cdots \mathbf{x}_N\}$, each an instance of a D dimensional Euclidean variable $\mathbf{x}$. These observations are to be divided into K disjoint clusters. Each cluster can be intuitively thought of as a subset of observed points which are closer to each other than to other observed points. The closeness is quantified by small distances between points within a cluster. Each cluster is represented by a

representative point from the cluster, $\boldsymbol{\mu}_k$. K-means clustering finds a grouping of points into K clusters such that the sum of the distance of each point $\mathbf{x}_n$ from the representative point for its assigned cluster is minimized.

Let r_{nk} be binary variable which is 1 if the point $\mathbf{x}_n$ is assigned to cluster k and is 0 otherwise. The objective function to be minimized, J, can then be expressed as follows.

$$J = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} ||\mathbf{x}_n - \boldsymbol{\mu}_k||^2 \tag{3.18}$$

Here $||\mathbf{x}_n - \boldsymbol{\mu}_k||^2$ represents the Euclidean distance between $\mathbf{x}_n$ and $\boldsymbol{\mu}_k$.

This function, called the *distortion function*, represents the sum of squares of distances of each observed point from the representative point of its assigned cluster, $\boldsymbol{\mu}_k$. The goal is to find values of r_{nk} and $\boldsymbol{\mu}_k$ which minimize the value of J.

K-means clustering proceeds through several iterations of two steps, after an initialization step. In the initialization step, initial values for $\boldsymbol{\mu}_k$ are chosen for $k \in \{1, \cdots K\}$.

1. The distortion function J is minimized with respect to the binary variables r_{nk} for $k \in \{1, \cdots K\}$ and $n \in \{1, \cdots N\}$, keeping the values of $\boldsymbol{\mu}_k$ fixed.

2. The distortion function is minimized with respect to the representative points $\boldsymbol{\mu}_k$ for $k \in \{1, \cdots K\}$ keeping r_{nk} fixed.

These steps are repeated until convergence, e.g., when cluster assignments do not change anymore.

3.5.2 Gaussian Mixture Modeling using Expectation Maximization

Gaussian mixture modeling is used to fit a linear combination (or 'mixture') of different Gaussian distribution to observed data. The probability distribution for a mixture of K Gaussian distributions can be written as a linear superposition of Gaussian distributions of the following form:

$$p(\mathbf{x}) = \sum_{k=1}^{K} w_k \left(\frac{1}{(2\pi)^{D/2}|\boldsymbol{\sigma}_k|^{1/2}} \exp\left(-\tfrac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_k)^{\mathrm{T}}\boldsymbol{\sigma}_k^{-1}(\mathbf{x}-\boldsymbol{\mu}_k)\right) \right) \tag{3.19}$$

Here w_k represents the weight of the kth Gaussian distribution. These weights are typically unknown. Therefore it is useful to introduce a latent, unobserved variable, $\mathbf{z}$, drawn from a Dirichlet process. $\mathbf{z}_n$ is a K dimensional binary random variable with $z_{nk} = 1$ if $\mathbf{x}_n$ is drawn from the kth Gaussian distribution and $z_{nk} = 0$ otherwise. The values of $\mathbf{z}_n$ satisfy $z_{nk} \in \{0, 1\}$ and $\sum_k z_{kn} = 1$. The joint distribution $p(\mathbf{x}, \mathbf{z})$ is expressed using the marginal distribution $p(\mathbf{z})$ and a conditional distribution $p(\mathbf{x}|\mathbf{z})$. The marginal distribution $p(\mathbf{z})$ is defined to be equal to the mixing weights.

$$p(z_k = 1) = w_k \tag{3.20}$$

These weights satisfy $0 \leq w_k \leq 1$ and $\sum_k w_k = 1$. Thus, the marginal distribution $p(\mathbf{z})$ can be written as follows.

$$p(\mathbf{z}) = \prod_{k=1}^{K} (w_k)^{z_k} \tag{3.21}$$

Moreover, the conditional distribution $p(\mathbf{x}|\mathbf{z})$ can be written as follows.

$$p(\mathbf{x}|\mathbf{z}) = \prod_{k=1}^{K} \left(\frac{1}{(2\pi)^{D/2}|\boldsymbol{\sigma}_k|^{1/2}} \exp\left(-\tfrac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_k)^{\mathrm{T}}\boldsymbol{\sigma}_k^{-1}(\mathbf{x}-\boldsymbol{\mu}_k)\right) \right)^{z_k} \tag{3.22}$$

The joint distribution $p(\mathbf{x}, \mathbf{z})$ is the product of the marginal $p(\mathbf{z})$ and the conditional $p(\mathbf{x}|\mathbf{z})$. The marginal distribution for $\mathbf{x}$, $p(\mathbf{x})$ can be obtained by summing over $\mathbf{z}$ as follows.

$$p(\mathbf{x}) = \sum_{\mathbf{z}} p(\mathbf{z})p(\mathbf{x}|\mathbf{z}) = \sum_{k=1}^{K} w_k \left(\frac{1}{(2\pi)^{D/2}|\boldsymbol{\sigma}_k|^{1/2}} \exp\left(-\tfrac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_k)^{\mathrm{T}}\boldsymbol{\sigma}_k^{-1}(\mathbf{x}-\boldsymbol{\mu}_k)\right) \right) \tag{3.23}$$

Consider now N observations of D dimensional Euclidean variable $\mathbf{x}$. Let $\mathbf{X}$ be a $N \times D$ dimensional vector representing all the observations. Estimating the parameters $w_k, \boldsymbol{\mu}_k, \boldsymbol{\sigma}_k$ which maximize the likelihood of observing these points is done by maximizing the likelihood function, which represents the conditional probability of observing the data $\mathbf{X}$ given the parameters of the model $\mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\sigma}$

$$p(\mathbf{X}|\mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\sigma}) = \sum_{n=1}^{N}\sum_{k=1}^{K} w_k \left(\frac{1}{(2\pi)^{D/2}|\boldsymbol{\sigma}_k|^{1/2}} \exp\left(-\tfrac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_k)^{\mathrm{T}}\boldsymbol{\sigma}_k^{-1}(\mathbf{x}-\boldsymbol{\mu}_k)\right) \right) \tag{3.24}$$

In practice, it is useful to maximize the log of this function, that is the log of the likelihood $\ln\left(p(\mathbf{X}|\mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\sigma})\right)$ with respect to parameters $\mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\sigma}$. This is done through the Expectation-Maximization algorithm, which involves two steps, similar to the K-means algorithm. Initial values for $\boldsymbol{\mu}_k$ and $\boldsymbol{\sigma}_k$ are chosen and the following two steps are repeated.

1. In the Expectation step, the values of w_k are determined to maximize the log likelihood function for fixed values of $\boldsymbol{\mu}_k$ and $\boldsymbol{\sigma}_k$.

2. In the Maximization step, the values of $\boldsymbol{\mu}_k$ and $\boldsymbol{\sigma}_k$ are determined to maximize the log of the likelihood function for fixed values of w_k.

These steps are repeated until convergence, e.g., when the log likelihood function reaches a steady value.

Chapter 4

Experiments

Now, we describe the implementation details of calculating per voxel diffusion coefficient using SMT and clustering them to obtain a classification into homogeneous compartments. Afterwards, these tissue classifications were compared against tissue classifications found by clustering Fractional Anisotropy (FA) calculated using DT modeling the diffusion process, the T2 weighted image intensity and B0 image.

Next, we describe the output of SMT by demonstrating a representative example and calculating a 2D histogram of the per voxel diffusion coefficients. Then, we cluster the diffusion coefficients and compare the performance of K-means clustering and GMM against the output obtained by clustering FA, B0 image intensity and T2 weighted image intensity. We accomplished this by creating a probability map that describes the probability of each voxel being in gray matter or white matter.

4.1 Calculating per Voxel Diffusion Coefficients using SMT

After preprocessing, per voxel longitudinal and transverse diffusion coefficients were estimated by using the following recipe:

1. The input array of the diffusion weighted brain volumes were segregated according to their b-values.

2. For every b shell, the average of signal intensity corresponding to all the direction was computed to find the expected spherical mean ($\overline{e}_b$) .

3. A thousand random samples of longitudinal and transverse diffusion coefficients were generated by randomly sampling uniformly in the ranges $0 \leq \lambda_\perp \leq \lambda_\parallel \leq \lambda_{\text{free}}$. These per voxel diffusion coefficients were used to calculate spherical mean according to Eqn. (3.15).

4. Constrained least-square minimization was performed as described in Eqn. (3.16). This minimization was implemented using Monte Carlo method by computing spherical mean estimates from the 1,000 samples of per voxel diffusion coefficients that were randomly sampled, as explained in step three. The error between expected and estimated spherical mean ($\sum_b (\hat{e_b} - \bar{e_b})^2$) was computed. Consequently, the longitudinal and transverse diffusion coefficients corresponding to the spherical mean estimate with minimum error were chosen.

The output of SMT was obtained in the form of two (longitudinal and transverse diffusion coefficient) images. The recipe described above was implemented using

custom C++ programs while making use of an existing library of functions (BRIAN 3.0.0 -Signal and Image processing tools [20]). Initially, the program took about 90 minutes per patient on an Intel Core i7 processor 4500U (two cores running at 1.8 GHz). In order to decrease this computation time, a look up table was pre-computed for the function $\mathrm{erf}(\sqrt{b(x)})/\sqrt{b(x)}$ for a precision of 10^{-6}, which reduced the processing time to about 15 min.

4.2 Clustering

Gaussian Mixture Modeling (GMM) and K-means clustering [26] were implemented after computing the log transform of the transverse diffusion coefficients. As explained in the previous section, the Expectation-Maximization algorithm was used to estimate the model parameters. K-means clustering is a straightforward technique, though it is sensitive to its starting points. Therefore, clustering was performed 20 times using 20 randomly chosen starting points. In both the cases, we chose the number of clusters to be four, based on evaluating the 2-D histogram of the per axon diffusion coefficients.

Exploratory data analysis was performed using the R statistical system. Clustering methods were implemented using libraries - "fpc" (K-means) [12] and "mclust" (GMM) [32]. The package "ggplot2" was used for visualization [39]. However, C++ programs were used to speed up the clustering process. The implementation of these algorithm made use of existing functions in BRIAN 3.0.0 library [20].

4.3 Comparison with DTI

In order to test the output of clustering of per voxel diffusion coefficients obtained using SMT, the compartments containing gray matter (GM), white matter (WM), and cerebro-spinal fluid (CSF) were identified. This was achieved by modeling the microscopic diffusion process as Diffusion Tensor (DT). Implementation details are as follows:

1. Fractional Anisotropy (FA) was obtained for all the voxels by performing calculations based on Eqn. (2.12) using eigenvalues obtained by performing Singular Value Decomposition (SVD) on DT .

2. T2 weighted image was registered to DW image with no diffusion gradient (B0 image) because of difference in resolution.

3. Brain volume was clustered using GMM (number of cluster = 4) on a vector containing T2-weighted image intensity, B0 image intensity and FA for all voxels.

4. The clusters were sorted in increasing order of their mean FA values. The first cluster was labeled as CSF, the second as GM and third and fourth as WM.

The tissue classifications obtained using this method are henceforth referred to as output obtained using DT modeling.

Chapter 5

Results

Fig. 5.1 shows a representative example of the output of SMT. The top and bottom row represent longitudinal and transverse per voxel diffusion coefficients, respectively, in the axial, coronal, and sagittal planes. After collecting the distribution of per voxel diffusion coefficients in a joint histogram (Fig. 5.2 (a)), we can identify the presence of certain data clusters. In order to check the consistency of this observation, we analyzed 867 brain volumes. Fig.5.2 (b) represents the average distribution of these per voxel coefficients across these volumes and allows us to identify the data cluster of the per voxel coefficients at locations similar to Fig. 5.3 (a).

Next, we explored the clustering of per voxel diffusion coefficients to classify into compartments gray matter, white matter and cerebro-spinal fluid.

Fig. 5.3 illustrates the output of the clustering of per voxel diffusion coefficients. The left and right panels show the output of Gaussian Mixture Model and K-means clustering, respectively, in a representative example. Each color represents a specific

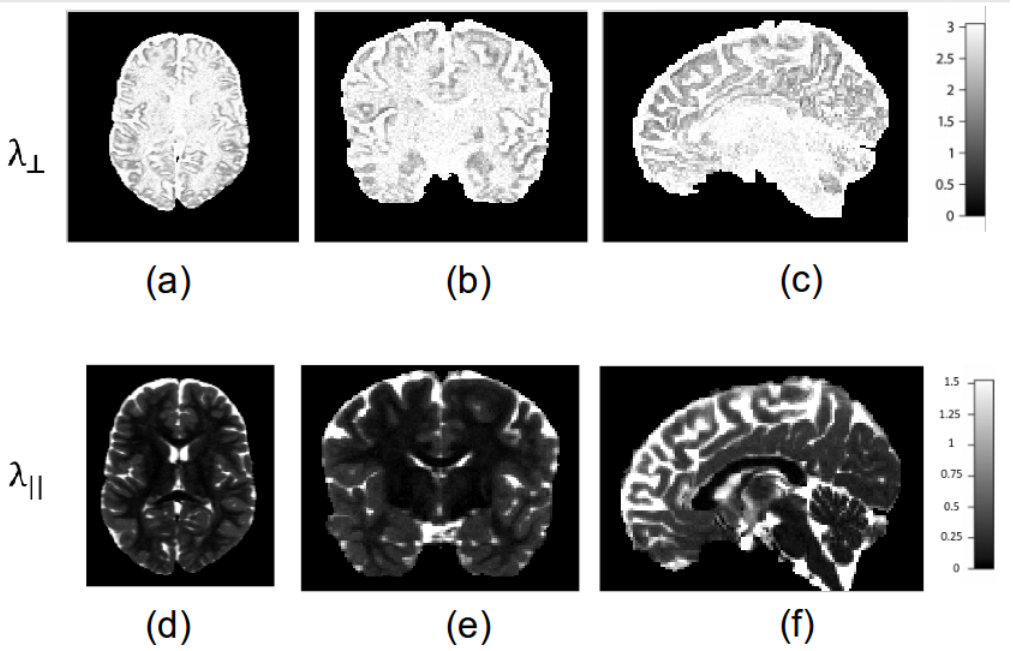

Figure 5.1: The upper row maps the longitudinal microscopic diffusion coefficients $\lambda_{\parallel}$ and the bottom row map the microscopic transverse diffusion coefficients $\lambda_{\perp}$of the axons, shown in the axial, coronal, and sagittal plane (from left to right).

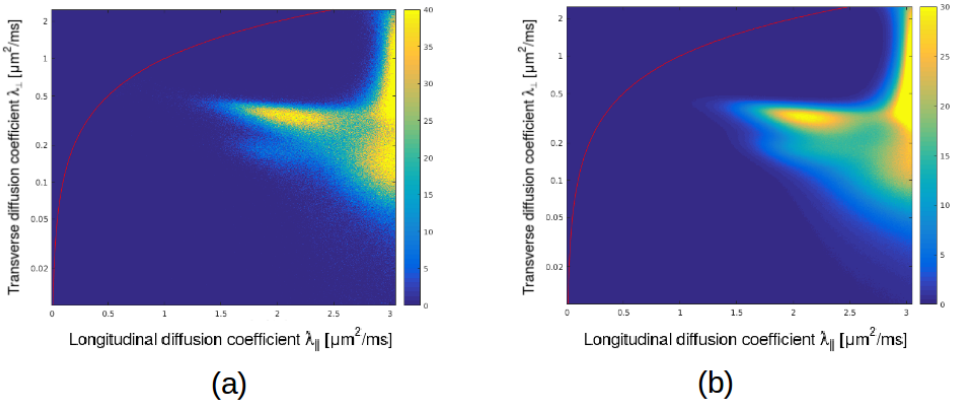

Figure 5.2: (a) Histogram of per-voxel diffusion coefficients of the brain volume shown in Fig. 5.1. The red line represents isotropic microscopic diffusion with $\lambda_{\perp} = \lambda_{\parallel}$. (b) Histogram of per voxel diffusion coefficients averaged across 867 brain volumes. The red line represents isotropic microscopic diffusion with $\lambda_{\perp} = \lambda_{\parallel}$.

cluster across all the images. The lower panel shows the parameter distribution of longitudinal and transverse coefficients in relation to cluster boundaries. From the figure we can identify the first cluster as CSF, the second as GM and the fourth as WM and the third as voxels containing a mixture of GM and CSF (PVE).

Compartment classifications were compared with results obtained from DT modeling. Fig. 5.4 shows that the tissue classifications obtained using DT modeling resembles more closely to the tissue classifications identified by clustering per voxel coefficients obtained using GMM than they do to the tissue classifications obtained using K-means.

Fig. 5.6 and Fig. 5.5 represent the likelihood of each voxel in brain volume space being in GM and WM, respectively. The probability maps were generated by averaging 867 tissue classifications output, similar to Fig. 5.4. The result obtained by comparing the probability maps supports the result obtained by comparing the representative example, i.e., the tissue classification outputs generated by using DT modeling resemble more closely to the tissue classification outputs generated by performing GMM on the per voxel diffusion coefficients than they do to the output generated by performing K-means clustering on per voxel diffusion coefficients. Therefore, we can conclude that white matter tissue classification obtained using GMM is accurate. Also a significant observation is that all the diffusion coefficients obtained from all the white matter bundles can be fit into a single Gaussian distribution. Therefore, we can also conclude that diffusion properties of white matter bundles are homogeneous.

K-means clustering performs a hard assignment of each point to a cluster, whereas Gaussian Mixture Modeling performs a soft assignment of each point to a cluster. In GMM, each point has a weight of membership to each cluster during every iteration.

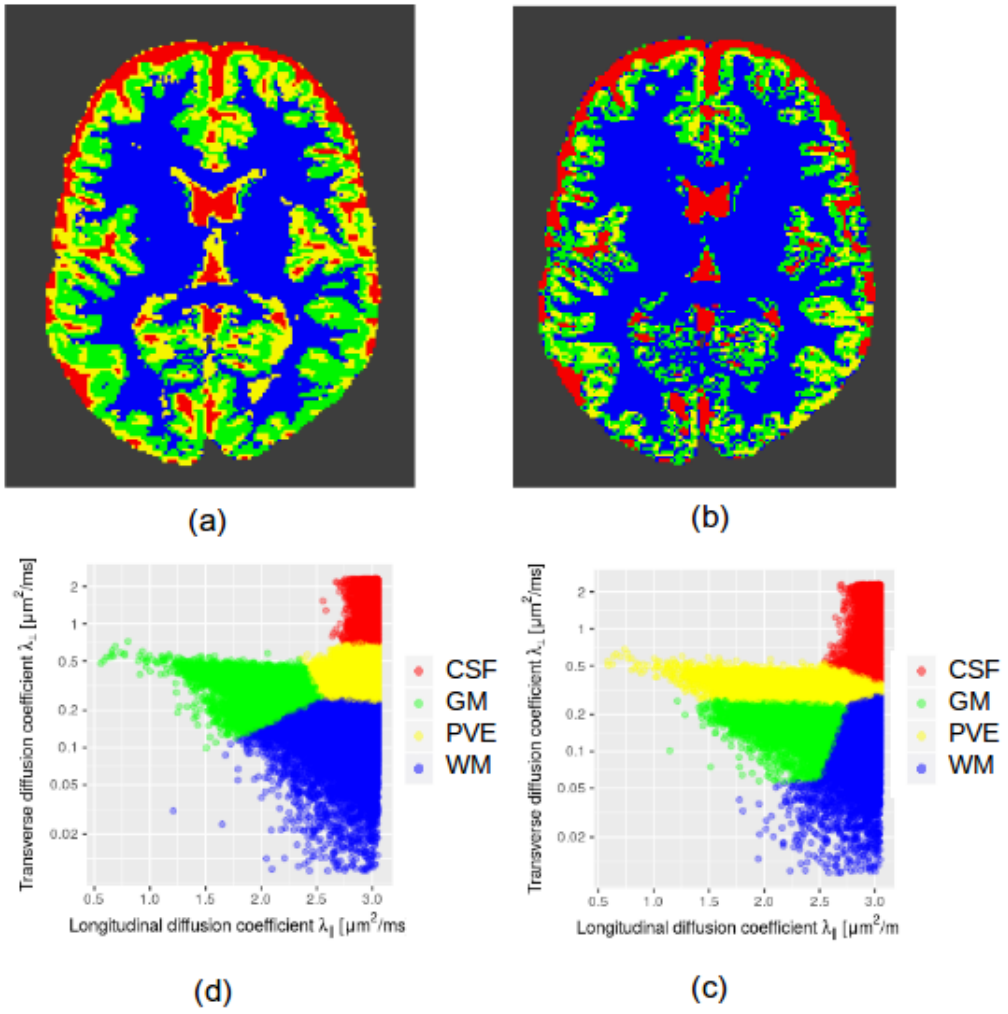

Figure 5.3: Output of Gaussian Mixture Model (left) and K-means clustering (right). Voxels in blue, green, red, yellow represent white matter (WM), gray matter (GM), cerebro-spinal fluid (CSF), and voxels with partial volume effect (PVE), respectively.

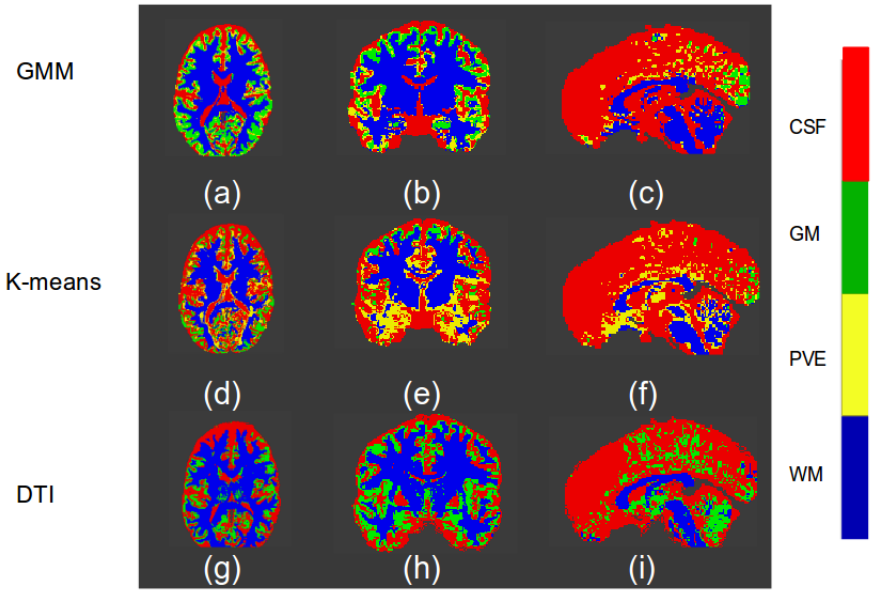

Figure 5.4: Comparison between the outputs of clustering per-axon diffusion coefficients obtained using SMT and tissue classifications using DTI. Axial, coronal, and sagittal sections (from left to right) of tissue classification obtained from SMT using GMM ((a)-(c)), K-means((d)-(f)) and DT modeling ((g)-(i)).

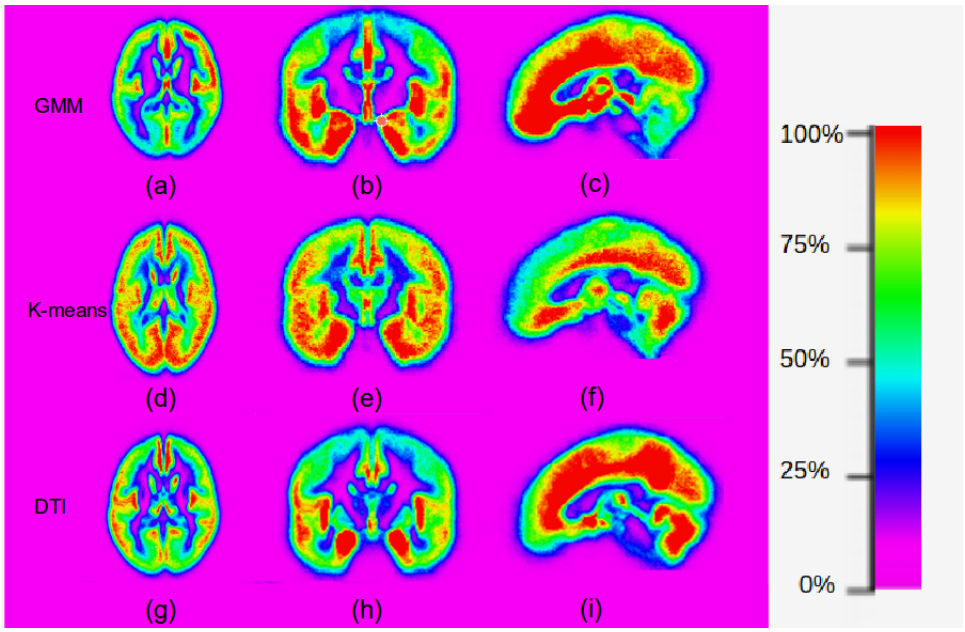

Figure 5.5: Probability maps comparing the likelihood of a voxel being in GM obtained by averaging the clusters of per-axon diffusion coefficients obtained using SMT and tissue classifications using DTI. Axial, coronal, and sagittal sections (from left to right) of tissue classification obtained from SMT using GMM ((a)-(c)), K-means((d)-(f)) and DT modeling ((g)-(i)).

In K-means, each point is on only cluster in every iteration. Furthermore, GMM has more parameters to fit the observed data than K-Means. Therefore GMM provides a more sophisticated model of the observed data. Since the biological data arises from various independent complex processes, GMM is better equipped to handle such data.

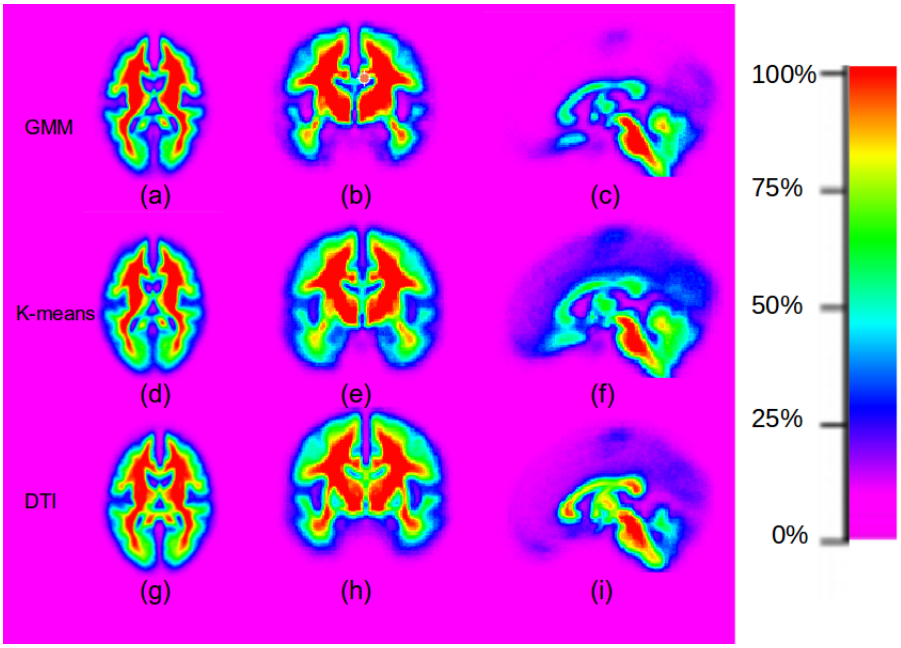

Figure 5.6: Probability maps comparing the likelihood of a voxel being in WM obtained by averaging the clusters of per-axon diffusion coefficients obtained using SMT and tissue classifications using DTI. Axial, coronal, and sagittal sections (from left to right) of tissue classification obtained from SMT using GMM ((a)-(c)), K-means((d)-(f)) and DT modeling ((g)-(i)).

Chapter 6

Summary

In this study, we characterized the diffusion properties of the nerve fibers in the central nervous system by calculating per voxel measures, namely, the longitudinal and transverse diffusion coefficients. These coefficients were estimated using the Spherical Mean Technique (SMT) from diffusion weighted images. We found that these coefficients are relatively homogeneous across white matter tracts and within gray matter. This was established by clustering per voxel diffusion coefficients (using Expectation - Maximization to fit Gaussian Mixture models) to obtain compartment classifications (WM, GM or CSF). Classification results were compared against those obtained by clustering FA values calculated from Diffusion Tensor modeling, T1 and B0 image intensity. This comparison was implemented by comparing probability maps obtained by averaging the tissue classification outputs of 867 subject data sets.

Looking at the 2D histogram containing GM and WM clusters reveals that WM clusters (mean values $(\mu(\lambda_{\parallel}), \mu(\lambda_{\perp})) = (2.70, 0.108)$ $\mu m^2/ms$) have higher anisotropy than GM cluster (mean values $(\mu(\lambda_{\parallel}), \mu(\lambda_{\perp})) = (2.110, 0.280)$ $\mu m^2/ms$). This is

consistent with the observations found in numerous other publications (e.g., [24, 5]). Higher anisotropy of WM could be a result of WM mostly containing axonal tract that interconnects cortical and neuronal cell bodies [30]. These structures could be considered akin to long cylinders, and therefore cause anisotropic diffusion of water. Conversely, GM primarily contain cell bodies and synapses, these structures do not contribute to anisotropy. We observed larger variance of per voxel diffusion coefficients of WM, as compares to that of G. This difference in diffusion coefficients, and therefore anisotropy, is known to primarily arise due to large range of axon calibers in WM [5].

Since per voxel diffusion coefficients do not have large difference across WM bundles in a healthy subject (standard deviations ($\sigma(\lambda_{\parallel})$, $\sigma(\lambda_{\perp})$) = (0.263, 0.0473) μm^2/ms) and across healthy subjects (standard deviations ($\sigma(\lambda_{\parallel})$, $\sigma(\lambda_{\perp})$) = (0.225, 0.0499) μm^2/ms), these coefficients have a potential to be a robust diagnostic markers for structural damages to WM fibers. Also, as SMT does not assume any specific model of microscopic diffusion and is independent of microfiber architecture inside the voxel, compartments calculated by clustering of per voxel diffusion coefficients yield more accurate results than those obtained by DTI modeling, which is used more commonly in clinical settings. Moreover, SMT does not require any complicated non-conventional pulse sequence during acquisition, it can easily be adopted in clinical setting [19]. However, more research is required to test the sensitivity and diagnostic capability of per voxel diffusion coefficients. A future research direction could be testing the algorithm on data sets with diseased brain volume samples, and characterize their diffusion property. Moreover, we observed a significant number of voxels that could not be classified as either GM or WM due to the partial volume effect. Further research may focus on eliminating signal due to CSF by using fluid attenuated inversion recovery (FLAIR) sequence.

Bibliography

[1] A. W. Anderson. Measurement of fiber orientation distributions using high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 54(5):1194–1206, nov 2005.

[2] J. L. Andersson, S. Skare, and J. Ashburner. How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. *NeuroImage*, 20(2):870–888, oct 2003.

[3] P. J. Basser, J. Mattiello, and D. Lebihan. Estimation of the Effective Self-Diffusion Tensor from the NMR Spin Echo. *Journal of Magnetic Resonance, Series B*, 1994.

[4] P. J. Basser, J. Mattiello, and D. LeBihan. MR diffusion tensor spectroscopy and imaging. *Biophysical journal*, 66(1):259–267, 1994.

[5] C. Beaulieu. The basis of anisotropic water diffusion in the nervous system - a technical review. *NMR in Biomedicine*, 15(7-8):435–455, nov 2002.

[6] J. T. Bushberg, J. A. Seibert, E. M. J. Leidholdt, and J. M. Boone. *The Essential Physics of Medical Imaging*. Lippincott Williams & Wilkins, 2011.

[7] M. Descoteaux. High Angular Resolution Diffusion Imaging (HARDI). In *Wiley Encyclopedia of Electrical and Electronics Engineering*, pages 1–25. American Cancer Society, 2015.

[8] M. F. Glasser, S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, S. Jbabdi, M. Webster, J. R. Polimeni, D. C. Van Essen, M. Jenkinson, and W.-M. H. C. P. WU-Minn HCP Consortium. The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage*, 80:105–124, oct 2013.

[9] H. Gudbjartsson and S. Patz. The Rician distribution of noisy MRI data. *Magnetic resonance in medicine*, 34(6):910–914, dec 1995.

[10] R. Gupta, R. Trivedi, and R. Rathore. Review: Clinical application of diffusion tensor imaging. *Indian Journal of Radiology and Imaging*, 18(1):45, 2008.

[11] P. Hagmann, L. Jonasson, P. Maeder, J.-P. Thiran, V. J. Wedeen, and R. Meuli. Understanding Diffusion MR Imaging Techniques: From Scalar Diffusion-weighted Imaging to Diffusion Tensor Imaging and Beyond. *RadioGraphics*, 26(suppl_1):S205–S223, 2006.

[12] C. Hennig. fpc: Flexible procedures for clustering. *R package version*, 2(2):0–3, 2010.

[13] M. R. Hodge, W. Horton, T. Brown, R. Herrick, T. Olsen, M. E. Hileman, M. McKay, K. A. Archie, E. Cler, M. P. Harms, G. C. Burgess, M. F. Glasser, J. S. Elam, S. W. Curtiss, D. M. Barch, R. Oostenveld, L. J. Larson-Prior, K. Ugurbil, D. C. Van Essen, and D. S. Marcus. ConnectomeDB–Sharing human brain connectivity data. *NeuroImage*, 124(Pt B):1102–1107, jan 2016.

[14] T. Huisman. Diffusion-weighted and diffusion tensor imaging of the brain, made easy. *Cancer imaging : the official publication of the International Cancer Imaging Society*, 10 Spec no(1A):S163–71, oct 2010.

[15] M. Jenkinson, P. Bannister, M. Brady, and S. Smith. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2):825–841, 2002.

[16] S. N. Jespersen, C. D. Kroenke, L. Ostergaard, J. J. H. Ackerman, and D. A. Yablonskiy. Modeling dendrite density from magnetic resonance diffusion measurements. *Neuroimage*, 2007.

[17] J. Jovicich, S. Czanner, D. Greve, E. Haley, A. Van Der Kouwe, R. Gollub, D. Kennedy, F. Schmitt, G. Brown, J. MacFall, B. Fischl, and A. Dale. Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data. *NeuroImage*, 2006.

[18] E. Kaden, N. D. Kelm, R. P. Carson, M. D. Does, and D. C. Alexander. Multi-compartment microscopic diffusion imaging. *NeuroImage*, 139:346–359, 2016.

[19] E. Kaden, F. Kruggel, and D. C. Alexander. Quantitative mapping of the per-axon diffusion coefficients in brain white matter. *Magnetic Resonance in Medicine*, 75(4):1752–1763, 2016.

[20] F. Kruggel and G. Lohmann. BRIAN (Brain image analysis): A toolkit for the analysis of multimodal brain datasets. *CAR '96: Computer Assisted Radiology*, pages 323–328, 1996.

[21] D. Le Bihan and E. Breton. Imagerie de diffusion in-vivo par résonance magnétique nucléaire. *Comptes-Rendus de l'Académie des Sciences*, 93(5):27–34, dec 1985.

[22] A. Lerner, M. A. Mogensen, P. E. Kim, M. S. Shiroishi, D. H. Hwang, and M. Law. Clinical Applications of Diffusion Tensor Imaging. *World Neurosurgery*, 82(1):96–109, 2014.

[23] J. Li, Y. Shi, and A. W. Toga. Mapping Brain Anatomical Connectivity Using Diffusion Magnetic Resonance Imaging: Structural connectivity of the human brain. *IEEE signal processing magazine*, 33(3):36–51, may 2016.

[24] D. Liewald, R. Miller, N. Logothetis, H.-J. Wagner, and A. Schüz. Distribution of axon diameters in cortical white matter: an electron-microscopic study on three human brains and a macaque. *Biological cybernetics*, 108(5):541–57, oct 2014.

[25] A. Lyon. Why are Normal Distributions Normal? *The British Journal for the Philosophy of Science*, 65(3):621–649, sep 2014.

[26] J. MacQueen. On convergence of k-means and partitions with minimum average variance. In *Annals of Mathematical Statistics*, volume 36, page 1084, 1965.

[27] R. Mezrich. A perspective on K-space. *Radiology*, 195(2):297–315, may 1995.

[28] R. Mills. Self-diffusion in normal and heavy water in the range 1-45.deg. *The Journal of Physical Chemistry*, 77(5):685–688, mar 1973.

[29] I. Rabi, J. Zacharias, S. Millman, and P. Kusch. A New Method of Measuring Nuclear Magnetic Moment. *Physical Review*, 53:318, feb 1938.

[30] B. R. Ransom, M. P. Goldberg, K. Arai, and S. Baltan. White Matter Pathophysiology. *Stroke*, pages 113–128, jan 2016.

[31] J. K. Rilling, M. F. Glasser, S. Jbabdi, J. L. Andersson, and T. M. Preuss. Continuity, Divergence, and the Evolution of Brain Language Pathways. *Frontiers in Evolutionary Neuroscience*, 2012.

[32] L. Scrucca, M. Fop, T. B. Murphy, and A. E. Raftery. {mclust} 5: clustering, classification and density estimation using {G}aussian finite mixture models. *The {R} Journal*, 8(1):205–233, 2017.

[33] J. Sijbers and A. J. den Dekker. Maximum likelihood estimation of signal amplitude and noise variance from MR data. *Magnetic Resonance in Medicine*, 51(3):586–594, mar 2004.

[34] S. M. Smith. Fast robust automated brain extraction. *Human Brain Mapping*, 2002.

[35] E. O. Stejskal and J. E. Tanner. Spin Diffusion Measurements: Spin Echoes in the Presence of a TimeDependent Field Gradient. *The Journal of Chemical Physics*, 42(1):288–292, jan 1965.

[36] J.-D. Tournier, F. Calamante, D. G. Gadian, and A. Connelly. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage*, 23(3):1176–1185, nov 2004.

[37] D. S. Tuch. Q-ball imaging. *Magnetic Resonance in Medicine*, 52(6):1358–1372, dec 2004.

[38] A. W. Van Der Eerden, O. Khalilzadeh, V. Perlbarg, J. Dinkel, P. Sanchez, P. E. Vos, C.-E. Luyt, R. D. Stevens, N. Menjot de Champfleur, C. Delmaire, et al. White matter changes in comatose survivors of anoxic ischemic encephalopathy and traumatic brain injury: comparative diffusion-tensor imaging study. *Radiology*, 270(2):506–516, 2014.

[39] H. Wickham, W. Chang, et al. ggplot2: An implementation of the grammar of graphics. *R package version 0.7, URL: http://CRAN. R-project. org/package= ggplot2*, 2008.

[40] W. Yuan, S. K. Holland, B. V. Jones, K. Crone, and F. T. Mangano. Characterization of abnormal diffusion properties of supratentorial brain tumors: a preliminary diffusion tensor imaging study. *Journal of Neurosurgery: Pediatrics*, 1(4):263–269, 2008. PMID: 18377300.

Appendix A

Appendix Title

A.1 T2-Weighted Image Acquisition and Preprocessing

T2 weighted imaging datasets were acquired using the variable flip angle turbo spin-echo sequence (Siemens SPACE) with 0.7mm isotropic resolution. The acquisitions were performed using TR=3200 ms, TE=565 ms, BW=744 Hz per pixel, no fat suppression pulse, phase encoding undersampling factor GRAPPA=2, total turbo factor=314, FOV=224 mm, matrix=320, 256 sagittal slices in a single slab [8].

The HCP minimal preprocessing pipeline named PreFreesurfer was applied to T2 weighted images and the data processing protocol involved following steps.

1. MR gradient nonlinearity-induced distortion was corrected using using customized version of the gradient_nonlin_unwarp package available in Freesurfer

[17]. It is based on calculating warpfield that represents the spatial distortion of the image using a proprietary Siemens gradient coefficient file.

2. Any repeated runs of images were averaged after they were aligned with a 6 degrees of freedom (DOF) rigid body transformation using FMRIB Software Library (FSL)s FMRIB's Linear Image Registration Tool (FLIRT) [15].

3. The Montreal Neurological Institute(MNI) space template was used to align T1 weighted and T2 weighted images using a rigid 6 DOF transform, derived from a 12 DOF affine registration.

4. The brain is extracted using an initial linear and non-linear registration of the image to the MNI template [34].

5. The T2 weighted image is crossmodally registered to the T1 weighted image using FLIRTs boundary-based registration (BBR) cost function.

6. The intensity inhomogeneity correction approach was used to correct for B1 bias and some B1+ bias [31].