

MEDICAL IMAGE SEGMENTATION AND
DIFFUSION WEIGHTED MAGNETIC RESONANCE IMAGE ANALYSIS

By
WEIHONG GUO

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2007

© 2007
by
Weihong Guo

To my family

ACKNOWLEDGMENTS

I would first and foremost like to express my deep gratitude to my advisor professor Yunmei Chen, for everything she has done for me during my doctoral study. This dissertation would not be possible without her. She provided invaluable advice on research and life. Dr. Chen introduced me to the field of medical imaging and always helped and encouraged me. I have been very lucky to work with her. Her enthusiasm about research stimulated my interests, her insight and experience have guided me through my research. I would also like to thank my other excellent committee members, Dr. David Groisser, Dr. Yijun Liu, Dr. Murali Rao, Dr. David Wilson for providing their advice and discussion. I also appreciate Dr. Yijun Liu, Dr. Guojun He from Brain Institute and Dr. Zhizhou Wang for happy cooperation. They provided a lot of invaluable research motivation and data.

I thank Feng Huang, one of my peers, for helping me with numerical implementation in my early stage research. Without his help, I could not have my scientific computation ability. I also thank Hemant Tagare from Yale University, Rachid Deriche from INRIA, Sheshadri Thiruvankadam(who is now at UCLA) for excellent comments and discussion on my work.

I never would have been able to have finished the dissertation were it not for the unwavering love and support of my family: my mum Zhongying, my husband Qingguo, my brother Weidong, my sister Shuangxia. Their support and encouragement were my source of strength. I want to especially thank my husband who teaches me the nature of true love. He is an extremely supportive man, I always appreciate his enthusiasm for life and research. I am so grateful about my 10-month-old daughter Tienna, who teaches me how wonderful life is by her sunny smiles.

I want to thank the editor in the Editorial office, Dr. David Groisser and Dr. David Wilson for their careful and thorough corrections to my thesis. I would also like to thank the staff and the faculty members who have been very patient and helpful during my study here.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS	4
LIST OF TABLES	7
LIST OF FIGURES	8
ABSTRACT	10
CHAPTER	
1 INTRODUCTION	12
1.1 Background	12
1.1.1 Medical Image Segmentation	12
1.1.2 Diffusion Weighted Image analysis	19
1.1.3 Neuron Fiber Tracts Reconstruction based on Smooth Tensor Field	23
1.2 Contributions	24
2 USING PRIOR SHAPE AND POINTS IN MEDICAL IMAGE SEGMENTATION	29
2.1 Description of the proposed model	29
2.2 Numerical scheme	33
2.3 Validation and Application to Echo Cardiovascular Ultrasound Images	34
2.4 Conclusion	37
3 USING NONPARAMETRIC DENSITY ESTIMATION TO SMOOTH AND SEGMENT IMAGES SIMULTANEOUSLY	41
3.1 Proposed Model	41
3.2 Numerical Implementation	45
3.3 Validation and Application to T1 Magnetic Resonance Image	49
3.4 An Existence Theorem for the Model	54
3.5 Conclusion	59
4 ESTIMATION,SMOOTHING AND CHARACTERIZATION OF APPARENT DIFFUSION COEFFICIENT	61
4.1 Introduction	61
4.2 Model Description	62
4.3 Characterization of anisotropy	68
4.4 Numerical Implementation Issues	69
4.5 Validation and Application to Diffusion Weighted Images(DWI)	73
4.5.1 Analysis of simulated data	73
4.5.2 Analysis of human MRI data	74
4.6 An Existence Theorem for the Model	77
4.7 Conclusion	82

5	ESTIMATION,SMOOTHING AND CHARACTERIZATION OF APPARENT DIFFUSION COEFFICIENT(A SECOND APPROACH)	87
5.1	Introduction	87
5.2	New Approximation Model for ADC Profiles	88
5.3	Use of CRE to Characterize Anisotropy	92
5.4	Conclusion	97
6	RECONSTRUCTION OF INTRA-VOXEL STRUCTURE FROM DIFFUSION WEIGHTED IMAGES	98
6.1	Determination of Fiber Directions	98
6.2	Validation and Application to HARD Weighted Images	99
6.3	Conclusion	102
7	RECONSTRUCT OF WHITE MATTER FIBER TRACES USING MULTI-TENSOR DEFLECTION IN DWI	104
7.1	Introduction	104
7.2	Recovery of Multi-Tensor Field in HARD MRI	104
7.3	White Matter Fiber Tractography	106
7.4	Experimental Results	107
7.5	Conclusion	111
8	FAST SEGMENTATION OF WHITE MATTER FIBER TRACTS BASED ON GEOMETRIC FLOWS	113
8.1	Introduction	113
8.2	Model	113
8.3	Experimental Results	116
8.3.1	Synthetic results	116
8.3.2	Human brain results	117
8.4	Conclusion	117
	REFERENCES	118
	BIOGRAPHICAL SKETCH	127

LIST OF TABLES

<u>Table</u>	<u>page</u>
3-1 Segmentation Accuracy	47
4-1 List of S_0 and $A_{l,m}$'s for two regions	73

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
2-1 Comparing segmentation results of a synthetic image	36
2-2 Compare segmentation results of one cardiac ultrasound image.	38
2-3 Compare segmentation results of another cardiac ultrasound image.	39
3-1 Compare results of model(3–5) based on two choices of $r(x)$	46
3-2 Segmentation results of a clean image and that of a noisy version	48
3-3 Segmentation result of a noisy T_1 human brain image	49
3-4 Compare three models in segmenting images with lower and higher level of noises	51
3-5 Comparison of graphs of SA obtained from three models	52
3-6 Compare segmentation and smoothing results between two models with fixed and adaptive radii	53
4-1 Comparing shapes of d	77
4-2 Comparison of A_{20}	83
4-3 Images of FA and R_2	84
4-4 Zoomed FA and A_{20}	85
4-5 Classification of voxels based on d	86
5-1 Comparison of ADC's obtained from two models	91
5-2 Images of R_2 by different models	93
5-3 An example for shapes of ADC for isotropic, one-fiber and two-fiber voxels . . .	94
5-4 Different measures	95
5-5 Characterization	96
6-1 Fiber direction color map	101
6-2 Shape of d with orientations	102
6-3 Fiber direction field	103
7-1 FA image of the first channel	108
7-2 Comparison between TEND and MTEND	109
7-3 Comparison between MTEND and TEND at internal capsule	110

8-1	Illustration of surface normal and evolving curve	114
8-2	Segmentation of a 2-D tensor field using two methods	116
8-3	Comparison of corpus callosum segmentation results from two models	117

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

MEDICAL IMAGE SEGMENTATION AND
DIFFUSION WEIGHTED MAGNETIC RESONANCE IMAGE ANALYSIS

By

Weihong Guo

May 2007

Chair: Yunmei Chen

Major: Mathematics

Medical image segmentation plays an important role in diagnosis, surgical planning, navigation, and various medical evaluations. Medical images are frequently corrupted by high levels of noise, signal dropout and poor contrast along boundaries. Sometimes, their intensity might have multi-modal distribution. In this dissertation, I will present one method to segment images that are corrupted by noise and some dropout. In the model presented, prior points together with prior shape information are incorporated into a joint segmentation and registration model in both a variational framework and in level set formulation. This technique is applied to segment cardiac ultrasound images. A second model, which is based on applying non-parametric density approximation to simultaneously segment and smooth noisy medical images without adding extra smoothing terms, is presented. My goal is to develop a powerful and robust algorithm to locate objects with interiors having a complex multi-modal intensity distribution and/or high noise level. The model was applied to the problem of segmenting T1 weighted magnetic resonance images.

Diffusion weighted images render non-invasive in vivo information about how water diffuses into a 3D intricate representation of tissues. My work provides histological and anatomical information about tissue structure, composition, architecture, and organization. I have proposed several models to reconstruct human brain white matter fiber tracts, to recover intra-voxel structure, to classify intra-voxel diffusion, to estimate,

smooth and characterize apparent diffusion coefficient profiles. A geometric flow is designed to segment the main core of white matter fiber tracts in diffusion tensor images.

CHAPTER 1 INTRODUCTION

1.1 Background

1.1.1 Medical Image Segmentation

Segmentation of anatomical structures from medical images has very important applications in diagnosis, surgical planning, navigation, and medical image analysis. A multitude of segmentation algorithms have been proposed to tackle these types of problems. They might be classified as edge-based, region-based, or a combination of these two. The edge-based methods (e.g. [1–6]) rely on the information of the edges, such as high magnitude of image gradient. The region-based methods (e.g. [7–10]) make use of homogeneity on the statistics of the regions being segmented. The algorithm developed in [11] integrates gradient and region information within a deformable boundary finding framework. The Geodesic Active Region models proposed in [12–14] integrate the edge and region-based segmentation methods into a variational approach.

The analysis of medical images is frequently complicated by noise, dropout, confusing anatomical structures, motion, poor contrast along boundaries, non-uniformity of regional intensities and multi-modal intensity distribution. It is hard to find one model to segment images with all the problems mentioned above. We will first provide some general literature review on image segmentation models, then focus on models dealing with images that are corrupted by noise and having multi-modal distribution. Finally, we review some existing models that segment images corrupted by noise, dropout, and poor contrast along boundaries.

Let $I_0 : \Omega \rightarrow R$ be a given bounded image-function defined on an open and bounded region Ω which is assumed to be a subset of R^2 for the purpose of illustration. But any dimensional case could be similarly considered. The segmentation problem is to find a decomposition Ω_i 's of Ω and an optimal piecewise smooth approximation I of I_0 such that I varies smoothly within each Ω_i , and rapidly or discontinuously across the boundaries of Ω_i . Let C be a close smooth curve formed by boundaries between Ω_i 's, and let the length

of C be $|C|$. To solve this problem, Mumford-Shah(MS) [8] proposed to minimize the following functional:

$$\sum_i \int_{\Omega_i} (I_i - I_0)^2 dX + \alpha \int_{\Omega \setminus C} |\nabla I|^2 dX + \beta |C| \quad (1-1)$$

The first and second parts are called data fidelity and smoothness terms, respectively. If the parameter α goes to ∞ , then the approximation I will be forced to be constant C_i in each Ω_i . Piecewise smooth MS is then reduced to piecewise constant MS:

$$\sum_i \int_{\Omega_i} (C_i - I_0)^2 dX + \beta |C| \quad (1-2)$$

While the functional is elegant itself, it is difficult in practice to find a solution as the functional is non-convex. L. Ambrosio et al.[15] approximated piecewise smooth MS via Γ -convergence. A. Tsai et al.[10] solved piecewise smooth MS using parametric curve evolution directly.

Chan and Vese [16, 17] implemented both piecewise constant and piecewise smooth MS using a geometric implicit framework-level set which was invented and continuously advanced by S. Osher and J. Sethian[18]. The level set method is used to represent C implicitly and to express each subregion. The main advantage of the level set representation is that topological changes, such as merging and pinching off of contours can be captured naturally through smooth changes to the level set function.

For illustration conciseness, we only consider the 2-phase case: I is composed of an object and background (i.e. $i = 1, 2$). For an image with more than 2 phases, the model could easily be extended by using more level set functions. The curve C is represented by the zero level set of a Lipschitz function ϕ by $C = \{x | \phi(x) = 0\}$. The object(A) and the background($\Omega \setminus A$) are represented by $\{x | \phi(x) > 0\}$ and $\{x | \phi(x) < 0\}$, respectively. A Heaviside function defined as $H(\phi) = 1$ when $\phi > 0$ and 0 elsewhere is used to distinguish A and $\Omega \setminus A$. Thus, A corresponds to the region where $H(\phi) > 0$, while Ω_2 corresponds to the region where $1 - H(\phi) > 0$. Let δ be the Delta function.

Chan-Vese model, the level set formulation of model (1-2)(2-phase case) is to minimize the functional:

$$\int_{\Omega} H(\phi)(C_1 - I_0)^2 dx + \int_{\Omega} (1 - H(\phi))(C_2 - I_0)^2 dx + \beta \int_{\Omega} |DH(\phi)| dx, \quad (1-3)$$

where the third term is the total variation of $H(\phi)$.

Due to its simplicity and robustness, the Chan-Vese model(1-3) has become quite popular and has been adopted in many applications. It has been expanded to segmenting vector-valued images, texture images etc. In [19], A. Yezzi et al. also use the level set framework to segment images.

The variational formulation in the Chan-Vese model is non-convex and a typical gradient-descent implementation is not guaranteed to converge to the global minimum and thus can get stuck in local minima. One trick [16, 17] is to set efficient initials. One example is a large number of small close contours uniformly distributed in the image. In this way, contours are initiated everywhere in the image such that chance of capturing the global minimum is enhanced. Chan and Esedoglu et al. [20, 21] provide another novel and fundamentally different approach. It basically convexifies the objective function by taking advantage of the implicit geometric properties of the variational models.

The traditional level set functions are Lipschitz continuous, and a reinitialization procedure is sometimes needed to prevent the level set function to from becoming too flat. Recently, Lie et al. [22] represent interfaces using piecewise constant level set functions, in which each level set function can only take two values at convergence, e.g., the function ϕ can only equal 1 or -1. Some of the properties of standard level set methods are preserved in the proposed model, while others are not. It is claimed that the new method provides as good results as methods using continuous level set functions. The Reinitialization procedure is removed in the new method. Tai et al. [23] also introduce a model that only requires minimizing the Chan-Vese functional(1-3) with respect to the level set functions, without estimating the constants C_i .

There are several ways to handle images with multiple phases. One widely used idea is to use $\log_2 n$ level sets to distinguish n phases in an image [17], but this can lead to difficulty in setting initials. Gao et al. [24] recursively use one level set in a hierarchical way: first segment into two regions, then segment each region into two new regions, and so on. See [25] for other related recent developments.

A generalization to piecewise constant MS is that instead of approximating I by a constant C_i in each region Ω_i , one can generally approximate the intensity in each region Ω_i by a probabilistic model $P(I(X)|\lambda_i)$ with a parameter vector λ_i . By Bayesian analysis, we have the posterior distribution $P(\lambda_i|I(X)) \propto P(I(X)|\lambda_i) \cdot P(\lambda_i)$. When choose the prior distribution $P(\lambda_i)$ to be uniform distribution, $P(\lambda_i|I(X)) \propto P(I(X)|\lambda_i)$.

The first term in (1-2) is then replaced by so called description cost and (1-2) becomes:

$$-\sum_i \int_{\Omega_i} \log P(\lambda_i|I_i(X)) dX + \beta|C| \quad (1-4)$$

Zhu et al.[7] and Rousson et al. [26] chose $P(I_i(X)|\lambda_i)$ to be a Gaussian distribution and permit for each region Ω_i not only a different mean C_i but also a different variance σ_i :

$$P(I_i(X)|C_i, \sigma_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(\frac{-(I_i(X) - C_i)^2}{2\sigma_i^2}\right) \quad (1-5)$$

We would like to mention that when all the Ω_i 's are the same, model(1-5) is equivalent with model (1-2). We call model(1-5) a global Gaussian probability density function(pdf) based method since it assumes all the pixels in one region Ω_i share the same mean and variance.

Model (1-4) has its restriction in real application since it is based on a specific assumption of the intensity distribution, while usually we do not know whether this assumption is reasonable. Especially, when there exists high level of noise and/or complex multi-modal intensity function, one single global Gaussian pdf or any other single parametric distribution is not enough. J. Kim et al.[27] presented an information-theoretic

approach to image segmentation. They cast the segmentation problem as the maximization of the mutual information between the region labels and the image pixel intensities, subject to a constraint on the total length of the region boundaries. They first assume that the probability densities associated with the image pixel intensities with each region are completely unknown priors, then formulate the problem based on nonparametric density estimates. Due to the nonparametric structure, the model did not require the image regions to have a particular type of probability distribution and did not require the extraction and use of a particular statistic. Huang et al. [28] used nonparametric kernel-based approximation of the intensity probability density function to capture texture, this texture information is then incorporated into a variational framework for image segmentation. Abd-Almagee et al. [29] introduced a general framework for driving an active deformable model using nonparametric estimation of the pdf. A Parzen window approach with a Gaussian kernel function was used as a parameter-free estimator for the pdf's for both the target to be segmented and the surrounding background. In Chapter 3, a new model based nonparametric density estimation is provided to segment and smooth images simultaneously.

For images that have noise, dropout, and poor contrast along boundaries, it is not enough to use image information itself to obtain the desired results. Recently, various approaches, including deterministic setting and probabilistic context, have been developed to use prior shape information in image segmentation (see [30–39]). A survey of methods, which incorporate prior knowledge into deformable models in medical image analysis is provided in [40]. The approaches presented in the papers [8, 9, 41–43] are closely related to our current work since they incorporate the statistical shape knowledge into either edge-based or region-based segmentation techniques. Leventon, Grimson and Faugeras [41] extended the technique of geometric active contours by incorporating shape information into the evolution process. A segmentation result was first obtained by the curve evolution driven by a force depending on image gradient and curvature. Later, the shape prior was

used to make a correction by maximizing a posterior estimate of shape and pose. Chen et al. [42] modified the techniques discussed by Leventon et al. [41] so that their variational method incorporates both high image gradients and shape information into the energy term of a geodesic active contour model. The shape term can also be used to recover a similarity transformation that maps the evolving interface to the prior shape. Cremers, Schnörr and Weickert [9] incorporated statistical shape knowledge into the Mumford-Shah segmentation scheme [8] by minimizing a functional that includes the Mumford-Shah energy and the shape energy corresponding to the Gaussian probability. Recently, in [43] Rousson and Paragios introduced an energy functional that constrains the level set representations to follow a shape global consistency while preserving the ability to capture local deformation.

The experimental results have shown that all these models have their own strengths, and provide promising results for particular applications. However, due to the accuracy and efficiency requirement for medical image analysis, and the complexity of medical images, as well as the variability of the shapes of anatomic structures, the question of how best to use shape prior to get a better segmentation will remain challenging. One of the most difficult problems is to determine local shape variations from the prior shape. One solution may be the use of nonrigid registration to assist segmentation [44, 45], but this approach requires reliable region or edge information in the image to estimate the velocity field. This technique is also computationally expensive.

The goal of the method presented in chapter 2 is to find an improved way to locate boundaries when they may vary widely from a prior shape and when they are to be detected in an image which may have significant signal loss. In particular, we are motivated by the problem of detecting the boundary of the left ventricle of the heart in echocardiographic image sequences. Cardiac ultrasound images are plagued with noise, signal dropout, and confusing intracavitary structures. Moreover, the shape of the boundary of the myocardium varies extensively from one patient image to another, and is

not equivalent to the “average” of a set of training shapes. More importantly, for patients with cardiac disease the shape distortion of the myocardium from the “normal shape” cannot be ignored.

To get a desirable segmentation in some cardiac ultrasound images more prior knowledge than the expected shape is needed. For instance, in [46] Dias and Leitaó used temporal information from time sequence of images to assist the determination of the inner and outer contours in the areas of low contrast in echocardiographic images. In this note we consider the cases in which users have the knowledge about the locations of a few points on the boundary of the object of interest. We will use this information as a constraint in addition to shape constraint to control the evolution of active contours. Our basic idea is to extend the segmentation algorithm developed in [42] by incorporating the information on the location of a few “key” points in addition to shape prior into geometric active contours.

The idea of matching nonequivalent shapes by the combination of a rigid transformation and a point-wise local deformation developed in the papers [43, 47, 48] will be applied to our formulation. In [48] Soatto and Yezzi view a general deformation as the composition of a finite dimensional group action (e.g. rigid or affine transformation) and a local deformation, and introduced a notion of “shape average” as the entity that separates a group action from a deformation. In [47] Paragios, Rousson and Ramesh proposed a variational framework for global as well as local shape registration. Their optimization criterion includes a global (rigid, affine) transformation together with local pixel-wise deformation. A similar idea was also used by Paragios et al. [43] to define shape prior models in terms of level set representations. These ideas can be used to simultaneously approximate, register, and track nonequivalent shapes as they move and deform through time. However, the question of how to determine the local deformation has not been considered in these previous works.

Paragios [49] introduced a framework for user-interaction within the propagation of curves using level set representations. The user-interaction term is introduced in the form of an evolving shape prior that transforms the user-edits to level set based propagation constraints. The work in [49] and in my study are based on the idea that the user interactive edits can be used as a constraint to correct local discrepancies. However, the formulations of the constraint in his work and my work are different(See Chapter 2 for details).

In chapter 2 we first briefly review geometric active contours and the model developed in [42]. We then propose the new model for incorporating both prior shape and points in active contours.

1.1.2 Diffusion Weighted Image analysis

Diffusion-weighted magnetic resonance imaging(MRI) (DW-MRI, shorten as DWI) adds to conventional MRI the capability of measuring the random motion of water molecules in tissue, referred as diffusion. The motion of water molecules can be free or restricted depending on the tissue structures. In tissues containing a large number of fibers (such as cardiac muscle and brain white matter) water diffusion is fastest along the direction that a fiber is pointing, but slowest in the direction perpendicular to it. This characteristic of the diffusion is termed as anisotropy. In tissues that contain few fibers water diffuses isotropically. DWI renders non-invasively such complex in vivo information about how water diffuses into a 3D intricate representation of tissues, and provides profound histological and anatomical information about tissue structure, composition, architecture, and organization. Changes in these tissue properties can often be correlated with processes that occur in development, degeneration, disease, and aging, so this technique has become more and more widely applied ([50–53]). One specific example is that Gupta et al [54] use DWI to search and quantify the extent of abnormality beyond the obvious lesions seen on the T_2 and fluid-attenuation inversion recovery (FLAIR) magnetic resonance images in patients with chronic traumatic brain injury with and

without epilepsy. Another example is that Roh et al. [55] uses DWI to detect acute multiple brain infarcts. They use the fact that DWI is superior to conventional MRI in identification of small new ischemic lesions and discrimination of recent infarcts from old ones.

The diffusion of water molecules in tissues over a time interval t can be described by a probability density function $p_t(\mathbf{r})$, which gives the probability that a water molecule will diffusion by \mathbf{r} . Since $p_t(\mathbf{r})$ is largest in the directions of least hindrance to diffusion and smaller in other directions, the information about $p_t(\mathbf{r})$ reveals fiber orientations and leads to meaningful inferences about the microstructure of tissues.

The density function $p_t(\mathbf{r})$ is related to DWI echo signal $s(\mathbf{q})$ via a Fourier transformation (FT) with respect to \mathbf{q} , which represents diffusion sensitizing gradient, by

$$s(\mathbf{q}) = s_0 \int p_t(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r}, \quad (1-6)$$

where s_0 is MRI signal in the absence of any gradient. Therefore, $p_t(\mathbf{r})$ can be estimated from the inverse FT of $s(\mathbf{q})/s_0$. Recently, Tuch et al. [56] introduced the method of higher angular resolution diffusion(HARD) MRI, and Wedeen et al. [57] succeed in acquiring 500 measurements of $s(\mathbf{q})$ in each scan to perform a fast FT inversion. However, this method requires a large number of measurements of $s(\mathbf{q})$ over a wide range of \mathbf{q} in order to perform a stable inverse FT.

A more common approach to estimate $p_t(\mathbf{r})$ from much sparser set of measurements $s(\mathbf{q})$ is assuming $p_t(\mathbf{r})$ to be a Gaussian. For Gaussian diffusion,

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

where D is called the diffusion tensor. Inserting this to equation (1-6) it yields

$$s(\mathbf{q}) = s_0 e^{-b\mathbf{u}^T D \mathbf{u}}, \quad (1-7)$$

where $\mathbf{u} = \mathbf{q}/|\mathbf{q}|$, diffusion weighting factor $b = \gamma^2 \delta^2 |q|^2 (\Delta - \delta/3)$. Here γ is the gyromagnetic ratio, and δ is the duration of two magnetic field gradient pulses with a separation time Δ in the use of Stejskal-Tanner pulsed gradient spin echo method [58]. In this case

$$d(\mathbf{u}) = \mathbf{u}^T D \mathbf{u}.$$

The principal eigenvector of D indicates diffusion direction of the diffusion. The fractional anisotropy (FA) defined as

$$FA = \sqrt{\frac{3}{2}} \sqrt{\frac{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}{(\lambda_1 + \lambda_2 + \lambda_3)^2}}, \quad (1-8)$$

where λ_i ($i = 1, 2, 3$) are the eigenvalues of D , has become the most widely used measure of diffusion anisotropy [53]. This is known as diffusion tensor imaging (DTI), and in particular useful for creating white matter fiber tracts [59–63].

However, it has been recognized that the single Gaussian model is inappropriate for assessing multiple fiber tract orientations, when complex tissue structure is found within a voxel [57, 60, 64–68]. A simple extension to non-Gaussian diffusion is to assume that the multiple compartments within a voxel are in slow exchange and the diffusion within each compartment is a Gaussian [65, 66, 69–71]. Under these assumption the diffusion can be modelled by a mixture of n Gaussians:

$$p_t(\mathbf{r}) = \sum_{i=1}^n f_i ((4\pi t)^3 \det(D_i))^{-1/2} e^{-\frac{\mathbf{r}^T D_i^{-1} \mathbf{r}}{4t}}, \quad (1-9)$$

where f_i is the volume fraction of the voxel with the diffusion tensor D_i , $f_i \geq 0$, $\sum_i f_i = 1$, and t is the diffusion time. Inserting (1-9) into equation (1-6) yields

$$s(\mathbf{q}) = s_0 \sum_{i=1}^n f_i e^{-b \mathbf{u}^T D_i \mathbf{u}}, \quad (1-10)$$

To estimate D_i and f_i , at least $7n - 1$ measurements $s(\mathbf{q})$ plus s_0 are required. In [69–71] the model of a mixture of two Gaussians were used to estimate the PDF. This estimation requires at least 13 diffusion weighted images from 13 different directions.

One of the alternatives to estimate $p_t(\mathbf{r})$ and characterize diffusion anisotropy is using apparent diffusion coefficient(ADC) profiles $d(\theta, \phi)$, which are related to observed DWI signals through the Stejskal-Tanner equation:

$$s(\mathbf{q}) = s_0 e^{-bd(\theta, \phi)}, \quad (1-11)$$

where (θ, ϕ) ($0 \leq \theta < \pi$, $0 \leq \phi < 2\pi$) represents the direction of \mathbf{q} . For Gaussian diffusion $d(\mathbf{u}) = b\mathbf{u}^T D \mathbf{u}$, where \mathbf{u} is the normalized \mathbf{q} . The trace, eigenvalues and functions of eigenvalues of D can be used to characterize the anisotropy and directional properties of the diffusion. For non-Gaussian diffusion the spherical harmonic approximation of the ADC profiles estimated from HARD data has been used for characterization of diffusion anisotropy. This technique was first introduced by Frank [67], also studied by Alexander et al. [72]. In the work of [67, 72] $d(\mathbf{x}, \theta, \phi)$ was computed from HARD raw data via the linear form of (1-11):

$$d(\mathbf{q}) = -\frac{1}{b} \log \frac{s(\mathbf{q})}{s_0} \quad (1-12)$$

and represented by a truncated spherical harmonic series(SHS):

$$d(\mathbf{x}, \theta, \phi) = \sum_{l=0,2,\dots,l_{max}} \sum_{m=-l}^l A_{l,m}(\mathbf{x}) Y_{l,m}(\theta, \phi), \quad (1-13)$$

where $Y_{l,m}(\theta, \phi) : S^2 \rightarrow C$ are the spherical harmonics series and C denotes the set of complex numbers. The odd-order term coefficients in the SHS are set to be zero, since the HARD measurements are made by a series of 3-d rotation, and $d(\theta, \phi)$ is real and has antipodal symmetry. Then, the coefficients $A_{l,m}(\mathbf{x})$'s were used to characterize the diffusion anisotropy. In their algorithm, basically, the voxels with the significant 4th order ($l = 4$) components in SHS are characterized as anisotropic with two-fiber orientations (shorten as two-fibers), while voxels with the significant 2nd order ($l = 2$) but not the 4th order components are classified as anisotropic with single fiber orientation (shorten as one-fiber), which is equivalent to the DTI model. Voxels with the significant 0th order ($l=0$) but not the 2nd and 4th order components are classified as isotropic. The

truncated order is getting higher as the structure complexity increases. Their experimental results showed that non-Gaussian profiles arise consistently in various brain regions where complex tissue structure is known to exist.

Since the ADC profiles can be used to characterize the diffusion anisotropy, and to estimate $p_t(\mathbf{r})$ through the combination of (1–6) and (1–11), it is of great significance to develop models for better estimation of the ADC profiles from DW MR measurements. In general the raw HARD MRI data are noisy. Computing the coefficients directly from the raw data often provides poor estimates. As a result, it will lead to inaccurate or false characterization of the diffusion and consequently lead to incorrect fiber tracking. The goal of Chapter 4 and Chapter 5 are to present two novel variational frameworks for simultaneous smoothing and estimation of non-Gaussian ADC profiles from HARD MRI.

1.1.3 Neuron Fiber Tracts Reconstruction based on Smooth Tensor Field

The assessment of connectivity and the reconstruction of 3D curves representing fiber traces are useful for basic neuroanatomical research and for disease detection. Most normal brain functions require that specific cortical regions communicate with each other through fiber pathways. Diffusion imaging is based on magnetic resonance imaging technique which was introduced in the mid 1980s [73] and provides a very sensitive probe for detecting biological tissues architecture. The key concept that is of primary importance for diffusion imaging is that diffusion in biological tissues reflects their structure and their architecture at a microscopic scale. For instance, Brownian motion is highly influenced in the presence of tissues, such as cerebral white matter or the annulus fibrosus of inter-vertebral discs. Measuring, at each voxel, that very same motion along a number of sampling directions (at least six, up to several hundreds) provides an exquisite insight into the local orientation of fibers.

There are currently several different approaches for reconstruction of white matter traces, they can be roughly divided into four categories: (1) line propagation algorithms; (2) surface propagation algorithms; (3) global energy minimization to find the energetically

most favorable path between two predetermined pixels; (4) solving a diffusion equation. We will discuss line propagation algorithms in details. Line propagation algorithms use local information for each step of propagation. The main differences among techniques in this class stem from the kind of local information being considered and the way information from neighboring pixels is incorporated.

The main local information used in most classical algorithms for recovering brain connectivity mapping from DTI data is the principal eigenvector(PE) of the diffusion tensor[63, 74, 75]. PE successfully determines the fiber direction in cases where there is a single fiber direction in each voxel, and is therefore adequate for reconstructing large trace systems. However, with voxel sizes typical of diffusion acquisitions($10 - 30mm^3$), there is significant partial volume averaging of fiber direction in anatomical regions of both research and clinical interest, such as the association fibers near the cortex. Moreover, image noise will influence the direction of the major eigenvector. And as the degree of anisotropy decreases, the uncertainty in the major eigenvector increases, in which situation the tracking may be erroneous. When a diffusion tensor is planar shaped, PE even does not make sense. Westin et al [76] and Lazar et al[77] used the entire tensor to deflect the estimated fiber trajectory. This algorithm is called tensor deflection(TEND). The deflection term is better than PE in the sense that the previous one is less sensitive to image noise and is less erroneous in situation of degenerated anisotropy. But it still has the problem of partial volume averaging of fiber direction.

1.2 Contributions

- Using prior shape and points in medical image segmentation

In chapter 2 we will provide a model to segment images that are corrupted by noise, dropout and poor contrast along boundaries. We will employ prior shape and prior points information into our algorithm by viewing the evolution of an active contour as a deformation of the interface. This deformation consists of a rigid transformation and a local deformation. We will use the “average shape” to determine the rigid

transformation that better maps the interface to the prior shape, and use the image gradient and a few “key” points to determine the local deformation that provides more accurate segmentation. Based on these thoughts we propose a variational framework that is able to incorporate prior knowledge of the expected shape and a few points that boundary should pass through.

We modify the energy function of geodesic active contour so that it depends on the image gradient and prior shape, as well as a few prior points. We only need a few points since we have information on expected shape. The modified energy function provides a satisfactory segmentation despite the presence of both large shape distortions and image dropout. To combine both prior shape (a global constraint) and prior points (a local constraint) into a single variational framework, we use a level set formulation. Further, we report experimental results on synthetic and ultrasound images.

- A nonparametric adaptive method for simultaneous image smoothing and segmentation
In chapter 3, a model is developed to segment images that have complex multi-modal intensity distribution and/or have high level noise. In this situation, it is not wise to set up models based on a specific parametric assumption about image intensities. Instead, we use nonparametric density estimation to create a model that is able to simultaneously segment and smooth images without adding extra smoothing terms. At each voxel, intensities of voxels in its neighborhood are used to obtain a nonparametric estimate of its intensity distribution. Neighborhood sizes are chosen adaptively based on image gradients. We then cast the segmentation problem as the minimization of the negative log likelihood, subject to a constraint on the total length of the region boundaries. Image smoothing is automatically fulfilled due to the nonparametric density estimation. Rather than expensive Gaussian kernels, a quadratic kernel is applied in the estimate to save calculation. The optimization problem is solved by deriving the associated gradient flows and applying curve

evolution techniques. We use level set methods to implement the resulting evolution.

We demonstrate the superiority of proposed model over other models by showing segmentation results from various images with different levels and types of noise.

The proposed model is able to solve a variety of challenging image segmentation.

- Estimation, smoothing and characterization of apparent diffusion coefficient using two different methods

First, in Chapter 4 we present a new variational framework for simultaneous smoothing and estimation of apparent diffusion coefficient (ADC) profiles from HARD MRI. The model approximates the ADC profiles at each voxel by a 4th order spherical harmonic series (SHS). The coefficients in SHS are obtained by solving a constrained minimization problem. The smoothing with feature preserved is achieved by minimizing a variable exponent, linear growth functional, and the data constraint is determined by the original Stejskal-Tanner equation. The antipodal symmetry and positiveness of the ADC are accommodated in the model. We use these coefficients and variance of the ADC profiles from its mean to classify the diffusion in each voxel as isotropic, anisotropic with single fiber orientation, or two fiber orientations. The proposed model has been applied to both simulated data and HARD MRI human brain data. The experiments demonstrated the effectiveness of our method in estimation and smoothing of ADC profiles and in enhancement of diffusion anisotropy. Further characterization of non-Gaussian diffusion based on the proposed model showed a consistency between our results and known neuroanatomy. Secondly, in Chapter 5 we present another approximation for the ADC of non-Gaussian water diffusion with at most two fiber orientations within a voxel. The proposed model approximates ADC profiles by product of two spherical harmonic series (SHS) up to order 2 from High Angular Resolution Diffusion-weighted (HARD) MRI data. The coefficients of SHS are estimated and regularized simultaneously by solving a constrained minimization problem. An equivalent but non-constrained version of

the approach is also provided to reduce the complexity and increase the efficiency in computation. Compare to the first method, this one requires less measurements but provides comparable results. Moreover we use the Cumulative Residual Entropy (CRE) as a measurement to characterize diffusion anisotropy. By using CRE we can get reasonable results using two thresholds, while the existing methods either only can be used to characterize Gaussian diffusion or need more measurements and thresholds to classify anisotropic diffusion with two fiber orientations. The experiments on HARD MRI human brain data indicate the effectiveness of the method in the recovery of ADC profiles. The characterization of diffusion based on the proposed method shows a consistency between our results and known neuroanatomy.

- Reconstruction of Intra-voxel Structure from Diffusion Weighted Images

In chapter 6, we present a new variational method for recovering the intra-voxel structure under the assumption that $p_t(\mathbf{r})$ is a mixture of two Gaussians: $p_t(\mathbf{r}) = \sum_{i=1}^2 f_i((4\pi t)^3 \det(D_i))^{-1/2} e^{\frac{-\mathbf{r}^T D_i^{-1} \mathbf{r}}{4t}}$. Our approach differs from the existing methods in the following aspects. First, we recover each *field* $D_i(\mathbf{x})$ or $f_i(\mathbf{x})$ globally by simultaneous smoothing and data fitting, rather than estimating them from (1-10) with $n = 2$ in each isolated voxel, which leads to an ill-posed problem. Second, we recover the ADC profile $d(\mathbf{x}, \theta, \phi)$ in SH representation using method introduced in chapter 4 from the noisy HARD data before estimating $D_i(\mathbf{x})$ and $f_i(\mathbf{x})$. The recovered d and the voxel classification on diffusion anisotropy from d are incorporated into our energy function to enhance the accuracy of the estimates. Third, we apply the biGaussian model to all the voxels in the field, rather than the voxels where the Gaussian model only fits poorly. Since both the constraint of $f_1 \approx 1$ on the region of strong one-fiber diffusion, and the regularization for f_i and D_i are built in the model, the single fiber and multi-fiber diffusions can be separated

automatically by the model solution. This approach should be less sensitive to the error in voxel classification.

- **Reconstruct White Matter Fiber Traces Using Multi-Tensor Deflection In DWI**

In chapter 7, we will provide a new line propagation algorithm based on smooth multi-tensor field.

We assume there are up to two diffusion channels at each voxel. A variational framework for 3D simultaneous smoothing and recovering of multi-diffusion tensor field as well as a novel multi-tensor deflection(MTEND) algorithm for extracting white matter fiber traces based on multi-tensor field are provided. MTEND keeps all the advantages of TEND and has two additional good properties: first, problem of partial volume averaging is automatically solved as it is based on a multi-tensor field; second, it uses a dynamically adjusted step size to keep total curvature of traces low, to appropriately terminate tracking and to increase algorithm efficiency.

Fiber traces are colored using Laplacian eigenmaps. By applying the proposed model to synthetic data and human brain high angular resolution diffusion magnetic resonance images(MRI) data of several subjects, we show the effectiveness of the model in recovering intra-voxel multi-fiber diffusion and inter-voxel fiber traces.

Superiority of the proposed model over existing models are also demonstrated.

- **Diffusion Tensor Image Segmentation**

In chapter 8, a 3D geometric flow is designed to segment the main core of fiber tracts in diffusion tensor magnetic resonance images. The proposed model designed a new external force that depended on two magnitudes: similarity of diffusion tensors and the coincidence level of the evolving surface normal with the tensor field. The new model, based on this new external force, was able to segment white matter fiber bundles in DT-MRI more accurately and efficiently.

CHAPTER 2

USING PRIOR SHAPE AND POINTS IN MEDICAL IMAGE SEGMENTATION

Copy right [2003] lecture notes on computer science [78]. Reprinted with permission.

2.1 Description of the proposed model

In this section we present our variational approach for joint segmentation and registration using a prior shape and the locations of a few points. The main idea of our model is to propagate a curve/surface by a velocity that depends on the image gradients, prior shape and locations, so that the propagation stops when the active contour/surface forms a shape similar to the shape prior, arrives at high gradients, and passes through the prior points. The notion of shape is independent of translation, rotation, and scaling. In particular, two contours C_1 and C_2 have the same shape, if there exists a scale μ , a rotation matrix R (with respect to an angle θ), and a translation T , such that C_1 coincides with $\mu RC_2 + T$.

To begin the description of the proposed model, we first briefly review the geodesic active contour model in [4, 5], and the active contour with a shape prior in [42].

Let $C(p) = (x(p), y(p))$ ($p \in [0, 1]$) be a differentiable parameterized curve in an image I . The geometric active contour model minimizes the energy function:

$$E(C) = \int_0^1 g(|\nabla I|)(C(p)) |C'(p)| dp, \quad (2-1)$$

where

$$g(|\nabla I|) = \frac{1}{1 + \beta |\nabla G_\sigma * I|^2}, \quad (2-2)$$

with a parameter $\beta > 0$, and $G_\sigma(x) = \frac{1}{\sigma} e^{-\frac{|x|^2}{4\sigma^2}}$. The minimum of this energy functional occurs when the trace of the curve is over points of high gradient in the image. Because object boundaries are often defined by such points, the active contour becomes stationary at the boundary. In its level set formulation this model can handle topological change. However, since this algorithm requires a high image gradient to be present along the

boundary of the object to stop the curve evolution, it may “leak” through the “gaps” in the boundary, where it is not salient.

To overcome this problem a variational model was proposed in [42] that incorporates prior shape information in geodesic active contours. The feature of this model is the creation of a shape term in the energy functional (2-1). If $C^*(p)$ ($p \in [0, 1]$) is a curve representing the expected shape of the boundary of interest, then the energy functional to be minimized in [42] is

$$E(C, \mu, R, T) = \int_0^1 \{g(|\nabla I|)(C(p)) + \frac{\lambda}{2} d_{C^*}^2(\mu RC(p) + T)\} |C'(p)| dp, \quad (2-3)$$

where (μ, R, T) are similarity transformation parameters, and $d_{C^*}(x, y)$ is the distance of the point (x, y) to C^* .

The first term in the energy functional (2-3) is the same as the energy functional for geodesic active contours, which measures the amount of high gradient under the trace of the curve. The second term is the shape related energy, that measures the disparity in shape between the interface and the prior. The constant $\lambda > 0$ is a parameter, which balances the influence from the image gradient and shape. The curve C and the transformation parameters μ , R and T evolve to minimize $E(C, \mu, R, T)$. At the stationary point, the contour C lies over points of high gradient in the image and forms a shape close to C^* , and μ , R and T determine the “best” alignment of C to $\mu RC(p) + T$.

The experimental results in [42] showed their model is able to get a satisfactory segmentation in the presence of gaps, even when the gaps are a substantial fraction of the overall boundary, if the shape of interest is similar to the expected shape. However, if some parts of the boundary are not visible, and the shape of boundary of the object has relatively larger geometric distortion from the prior, as shown later, model (2-3) can not provide a desired segmentation, since the knowledge of the expected shape does not provide correct information about how the gaps should be bridged.

In this note we intend to incorporate extra prior knowledge of the location of a few points on the boundary into segmentation process. These points are given at the location where the edge information is not salient and the local shape has relatively large variation from the prior one. Our idea is the creation of a new energy term, that measures the distances of the prior points from the interface, into the energy functional of the active contours in [42]. Similar to the approach developed in [42], to get a smooth curve C that captures higher gradients we minimize the arc-length of C in the conformal metric $ds = g(|\nabla I|)(C(p))|C'(p)|dp$, where $g(|\nabla I|)$ is defined in (2-2). To capture the shape prior C^* , we find a curve C and the transformation (μ, R, T) , such that the curve $\mu RC + T$ and C^* are “best” aligned. To capture the prior points we minimize the distances of each prior point x_i ($i = 1, \dots, m$) from the curve C .

We present our model in a variational level set formulation. First, as well known, the level set method initiated in [18] allows for cusps, corners, and automatic topological changes, Secondly, it is more convenient to compute the distances of the prior points to the interface by using the level set form of the interface.

Let the contour C be the zero level set of a Lipschitz function u such that $C = \{x \in R^N : u(x) = 0\}$, with (inside of C) $= \{x \in R^N : u(x) > 0\}$, and (outside C) $= \{x \in R^N : u(x) < 0\}$. Let H be the Heaviside function $H(z) = 1$, if $z \geq 0$, otherwise $H(z) = 0$, and δ be the Dirac measure concentrated at 0 (i.e. $\delta(z) = H'(z)$ in the sense of distribution). Then, the length of the zero level set of u in the conformal metric $ds = g(|\nabla I|)|C'(p)|dp$ can be computed by $\int_{\Omega} g(|\nabla I|)|\nabla H(u)| = \int_{\Omega} \delta(u)g(|\nabla I|)|Du|$. The disparity in shape between the zero level set of u and C^* can be evaluated by $\int_{\Omega} \delta(u)d_{C^*}^2(\mu Rx + T)dx$, where the distance function d_{C^*} is the same as that in (2-3). Moreover, the constraint for contour C passing through the prior points x_1, \dots, x_m can be simply represented by $u(x_i) = 0$, ($i = 1, \dots, m$).

Now let $f_{\gamma}(x)$ be a smooth function defined on the image domain Ω , such that $0 \leq f_{\gamma}(x) \leq 1$, $f_{\gamma}(x) = 1$ for $x = x_i$ ($i = 1, \dots, m$). The function $f_{\gamma}(x)$ can be obtained

by the convolution of f and η_γ , where f is a function taking value one on the points x_1, \dots, x_m and zero elsewhere, and η_γ is a mollifier with parameter γ such that $f_\gamma(x)$ is only nonzero in neighborhood of x_i 's. Now we can formulate our new variational approach as

$$\begin{aligned} \min_{u, \mu, R, T} E(u, \mu, R, T) = & \min_{u, \mu, R, T} \left\{ \alpha \int_{\Omega} \delta(u) g(|\nabla I|) |\nabla u| \right. \\ & \left. + \frac{\beta}{2} \int_{\Omega} \delta(u) d_{C^*}^2(\mu R x + T) |\nabla u| + \frac{1}{2} \int_{\Omega} f_\gamma(x) u^2(x) \right\} dx, \end{aligned} \quad (2-4)$$

where $\alpha > 0$ and $\beta > 0$ are parameters. The first two terms in this energy functional are the same as those in (2-3), which tend to lead the interface arriving at a location where the magnitude of image gradient is high, and to form a shape similar to the prior. The last term tends to lead the interface to pass through the given points, since minimizing the third term in (2-4) with sufficiently small γ leads u to be close to zero at the given points. Note that $f_\gamma(x)$ is non-zero only on the γ neighborhood of the given points, so the third term doesn't affect much the shape of the contour outside the γ neighborhood of the given points.

This model performs a joint segmentation and registration. The segmentation is assisted by the registration between the interface and shape prior. This registration is non-rigid that consists of a global transformation (rigid) and a local deformation. The global transformation is determined by minimizing the second term in (2-4), while the local deformation, is controlled by minimizing the first and last terms in (2-4).

The evolution equations associated with the Euler-Lagrange equations for (2-4) are

$$\frac{\partial u}{\partial t} = \delta(u) \operatorname{div} \left(\left(\alpha g + \frac{\beta}{2} d^2 \right) \frac{\nabla u}{|\nabla u|} \right) - f_\gamma u, \quad (2-5)$$

$$\frac{\partial u}{\partial n} = 0, \quad x \in \partial\Omega, \quad t > 0; \quad u(x, 0) = u_0(x), \quad x \in \Omega, \quad (2-6)$$

$$\frac{\partial \mu}{\partial t} = -\beta \int_{\Omega} \delta(u) d \nabla d \cdot (Rx) |\nabla u| dx, \quad t > 0, \quad \mu(0) = \mu_0, \quad (2-7)$$

$$\frac{\partial \theta}{\partial t} = -a \int_{\Omega} \delta(u) \mu d \nabla d \cdot \left(\frac{dR}{d\theta} x \right) |\nabla u|, \quad t > 0, \quad \theta(0) = \theta_0, \quad (2-8)$$

$$\frac{\partial T}{\partial t} = -\beta \int_{\Omega} \delta(u) d \nabla d |\nabla u| dx, \quad t > 0, \quad T(0) = T_0, \quad (2-9)$$

where $d = d_{C^*}$, R is the rotation matrix in terms of the angle θ , and the function d is evaluated at $\mu Rx + T$.

2.2 Numerical scheme

We solve the proposed model (2-4) by finding the steady state solution(s) to the evolution problem (2-5)-(2-9). To solve the equations (2-5)-(2-9) numerically, as in [16], we replace δ in (2-4)-(2-9) by a slightly regularized versions of them, denoted by:

$$H_{\epsilon}(z) = \begin{cases} 1 & \text{if } z > \epsilon \\ 0 & \text{if } z < -\epsilon \\ \frac{1}{2} \left[1 + \frac{z}{\epsilon} + \frac{1}{\pi} \sin\left(\frac{\pi z}{\epsilon}\right) \right] & \text{if } |z| \leq \epsilon \end{cases}$$

$$\delta_{\epsilon}(z) = H'_{\epsilon}(z) = \begin{cases} 0 & \text{if } |z| > \epsilon \\ \frac{1}{2\epsilon} \left[1 + \cos\left(\frac{\pi z}{\epsilon}\right) \right] & \text{if } |z| \leq \epsilon \end{cases}$$

To discretize the equation of u , here we use a implicit finite difference scheme rather than the one used in [42]. Let h be the step size, and $(x_i, y_i) = (ih, jh)$ be the grid points, for $1 \leq i, j \leq M$. Let $u_{i,j}^n = u(t_n, x_i, y_j)$ be an approximation of $u(t, x, y)$. The time derivative u_t at (i, j, t_n) is approximated by the forward difference scheme: $u_t(i, j, t_n) = \frac{u_{i,j}^{n+1} - u_{i,j}^n}{\Delta t}$, where Δt is the time step.

We adopt the algorithm for the discretization of the divergence operator from [79], and the implicit iteration from [80]. Knowing u^n , we compute u^{n+1} by using the following discretization and linearization scheme of (2-5):

$$\begin{aligned} \frac{u_{i,j}^{n+1} - u_{i,j}^n}{\Delta t} &= \delta_\epsilon(u_{i,j}^n) \{ (\alpha g_{i,j}^n + \frac{\beta}{2} d_{i,j}^{n-2}) \operatorname{div}(\frac{\nabla u_{i,j}^{n+1}}{|\nabla u_{i,j}^{n+1}|}) \\ &+ \alpha \frac{(\nabla g \cdot \nabla u)_{i,j}^{n+1}}{|\nabla u_{i,j}^{n+1}|} + \beta d_{i,j}^n \frac{(\nabla d \cdot \nabla u)_{i,j}^{n+1}}{|\nabla u_{i,j}^{n+1}|} \} - f_\gamma(u_{i,j}^n) u_{i,j}^{n+1}, \end{aligned}$$

where

$$\operatorname{div}(\frac{\nabla u_{i,j}^{n+1}}{|\nabla u_{i,j}^{n+1}|}) = \Delta_-^x \left(\frac{\Delta_+^x u_{i,j}^{n+1}}{\sqrt{(\Delta_+^x u_{i,j}^{n+1})^2 + (\Delta_+^y u_{i,j}^{n+1})^2}} \right) + \Delta_-^y \left(\frac{\Delta_+^y u_{i,j}^{n+1}}{\sqrt{(\Delta_+^x u_{i,j}^{n+1})^2 + (\Delta_+^y u_{i,j}^{n+1})^2}} \right).$$

When we compute $u_{i,j}^{n+1}$, $u_{i+1,j}^{n+1}$ and $u_{i,j+1}^{n+1}$ were replaced by $u_{i+1,j}^n$ and $u_{i,j+1}^n$, respectively, since they are unknown. We use the forward or backward finite difference schemes adaptively to approximate $\nabla d \cdot \nabla u$ and $\nabla g \cdot \nabla u$. That is

$$\begin{aligned} (\nabla d \cdot \nabla u)_{i,j}^n &= (\max(\Delta^x d_{i,j}^n, 0) \Delta_+^x u_{i,j}^n + \min(\Delta^x d_{i,j}^n, 0) \Delta_-^x u_{i,j}^n \\ &+ \max(\Delta^y d_{i,j}^n, 0) \Delta_+^y u_{i,j}^n + \min(\Delta^y d_{i,j}^n, 0) \Delta_-^y u_{i,j}^n). \end{aligned}$$

The term $\nabla g \cdot \nabla u$ is approximated in the same way.

To keep the signed distance function near the front (zero level set of the evolving u) the technique of reinitialization developed and applied in [37, 80, 81] is also used in our computation. This procedure is made by using a new function $v(x)$, which is the steady state solution to the equation

$$\frac{\partial v}{\partial s} = \operatorname{sign}(u(\cdot, t))(1 - |\nabla v|), v(\cdot, 0) = u(\cdot, t),$$

as $u(\cdot, t)$ for the next iteration $t + \Delta t$.

The equation (2-4)-(2-9) are discretized as in [42] by using finite difference.

2.3 Validation and Application to Echo Cardiovascular Ultrasound Images

In this section we report our experimental results on both synthetic and ultrasound images.

The aim of our first experiment is to verify that the active contour with the prior shape and points can fill in the “gaps” in a boundary in a meaningful way.

Figure 2-1a shows a typical binary image I with three points and an ellipse superimposed. The ellipse and points are used as the prior shape and points in this experiment, respectively. The object to be segmented is partially occluded, and the shape of its boundary is not equivalent to the prior shape. We want to determine whether or not the active contour with the prior shape and points can use the partial boundary to aid the process of filling in the rest.

The active contour was initialized by the solid curve displayed in figure 2-1c. Evolving the active contour according to (2-5)-(2-9) with the parameters $\alpha = 250$, $\beta = 15$, $\sigma = 0.5$ (in $g(x)$), $\mu_0 = 1$, $\theta_0 = 0$, $T_0 = (0, 0)$, we get the stationary contour C (the dotted one) in figure 1c, and the transformation parameters $\mu = 0.91$, $\theta = -0.14$, and $T = (-0.5, 0.3)(pixels)$. We can see that even though complete gradient information is not available the contour C captures the high gradient in the image I , passes through three prior points, and forms a shape similar but not the same as the prior shape. To show the advantage of using prior points we compared the segmentation results obtained by using model (2-4) and (2-3). Figure 2-1 B) shows the segmentation result by using model (2-3). In figure 2-1 B) the solid contour is the initial contour, and the dotted one is the segmentation result. Since the prior points are not incorporated in the model (2-3), the segmented contour only captures the prior shape and high gradients. It can't accurately capture local shape variations.

The aim of the second experiment is to segment the endocardium (the inner boundary of the myocardium surrounding the left ventricle) in an apical two-chamber image of the heart (see Figure 2-2 A) for a typical image). The endocardium is not completely visible in the image, and its shape is not the same as the “average shape” (the shape prior). Our task is to determine the endocardium using “average shape” and five points given by an expert.

The prior shape is created by the same way as that in [42]. It is obtained by averaging the aligned contours in a training set. The alignment of two contours C_1 and C_2

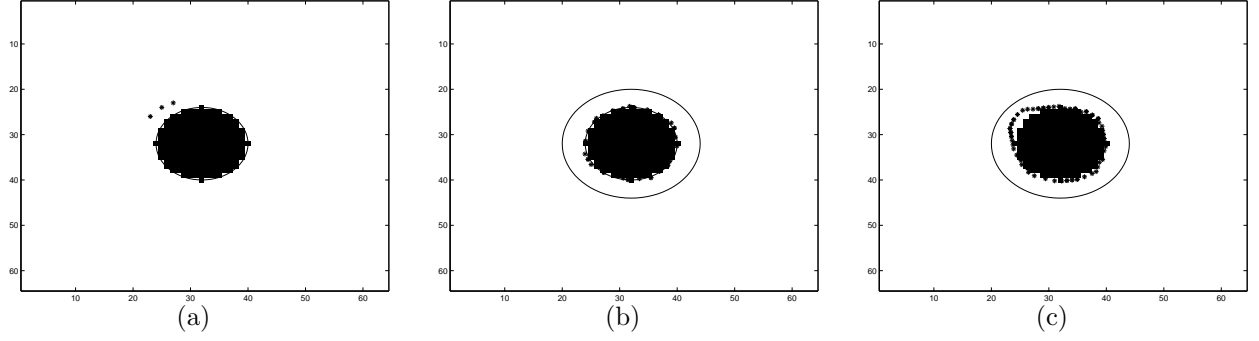


Figure 2-1. Comparing segmentation results of a synthetic image. A) An image with the prior shape and three points superimposed. B) The segmentation result using model (2-3) (dotted) and the initial contour (solid). C) The segmentation result using model (2-4) (dotted) and the initial contour (solid).

is made by finding the “best” scaling constant μ , rotation matrix R and translation vector T such that the overlapping area of the interiors of C_1 and $\mu RC_2 + T$ is maximized. If the shapes of the curves in the training set have a large variation, a clustering technique is required to group these curves into several groups. The shape priors for each group are obtained using this technique.

For this particular problem to create the prior shape, an expert echocardiographer traced endocardial boundaries on 112 image sequences for 66 patients. After the boundaries were clustered, the average was computed. Figure 2-2b shows the “average contour” for one of the clusters (the dotted contour), the endocardium outlined by an expert (the solid contour), and five points on the expert’s contour. The prior points are usually given at the location where the image gradients are low, and the local shape distortions are larger. Figure 2-2 C) presents the image $|\nabla G_\sigma * I|$. From this image we can see the dropout of image information at several parts of the endocardium.

To segment the endocardium in the image shown in Figure 2-2 D) (it is the same as in Figure 2-2 A), the the active contour was initialized as the contour shown in Figure 2-2 A). This contour was evolved according to the equations (2-5)-(2-9) and it finally stopped at the location of the dotted contour in Figure 2-2 D). We also obtained the transformation parameters $\mu = 1.0024$, $\theta = -0.1710$, and $T = (-24.6163, 32.3165)$

(pixels). The solid contour in Figure 2-2 D) is the expert’s endocardium. Observe that the segmentation is close to the expert’s contour. To see the shape variation between the solution of (2-4) and “average shape” we aligned the solution of (2-4) to the “average shape” using the solutions (μ, R, T) of (2-4). Figure 2-2 E) shows the disparity in shape between these two contours. The dotted contour is the transformed solution of (2-4), and the solid one is the “average shape”. From Figures 2-2d and 2-2 E) we can see that our active contour formed a shape different from the prior one in order to capture the high gradients and given points. Figure 2-2 F) provides the segmentation result obtained by using model (2-3). In this figure the dotted contour is the solution of (2-3), and the solid contour is the expert’s endocardium. Comparing Figure 2-2 D) with Figure 2-2 F), note that the solution of (2-4) is closer to the expert’s contour than the solution of (2-3). Figure 2-2 G) presents the shape comparison between the solution of (2-3) and prior shape. In Figure 2-2 G) the solid contour is the “average shape”, and the dotted one is the transformed solution of (2-3), (the transformation parameters are the solution of (2-3)). From Figures 2-2 F) and 2-2g we see that the solution of (2-3) can only capture the high image gradients and the “average shape”, but it can’t provide as desirable a segmentation result as the expert’s endocardium.

The last experiment is a repetition of the second experiment on a second apical 2-chamber cardiac ultrasound image. We list the figures below for the results of this experiment in the same order as above. The segmentation C is given in figure 2-3d represented by the dotted contour, the transformation parameters are $\mu = 0.9883$, $\theta = -0.1981$, and $T = (-28.7246, 49.1484)$.

2.4 Conclusion

In this chapter we proposed the addition of prior points to an active contour with shape in a variational framework and in level set formulation. The key idea was to introduce an energy term which measured the image gradients, the closeness of the shape between the active contour and a prior shape, as well as the distance of the prior point

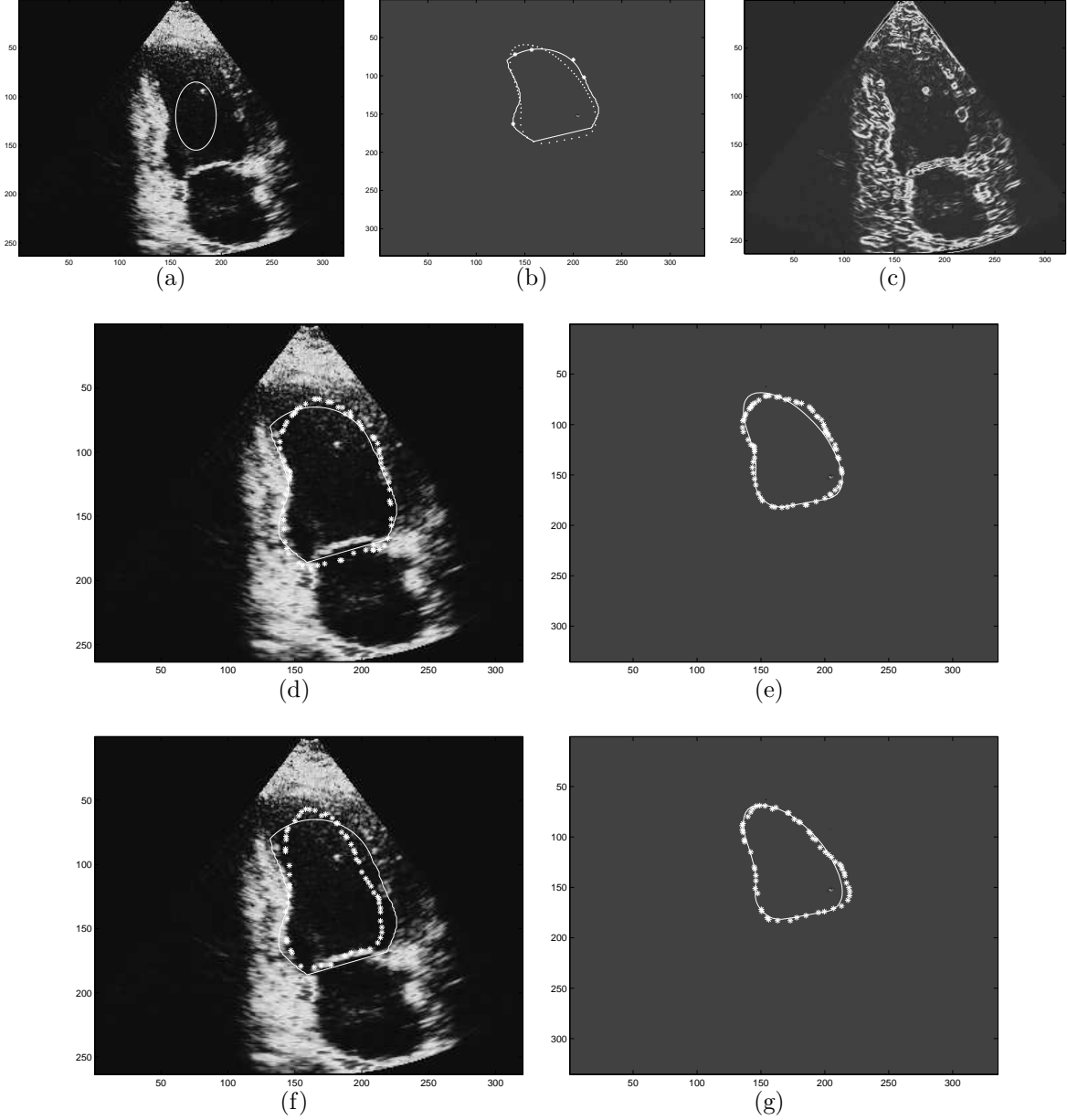


Figure 2-2. Compare segmentation results of one cardiac ultrasound image. A) A typical 2-chamber ultrasound image with an initial contour; B). Expert's endocardium (solid contour), "average shape" (dotted contour), and five points on the expert's contour. C). The image $|\nabla G_\sigma * I|$. D). The endocardium segmented by using model (2-4) (dotted) and the expert's contour (solid). E) The transformed solution of (2-4) (dotted), and the "average shape". F). The endocardium segmented by using model (2-3) (dotted) and the expert's contour (solid). G) The transformed solution of (2-3) (dotted), and the "average shape".

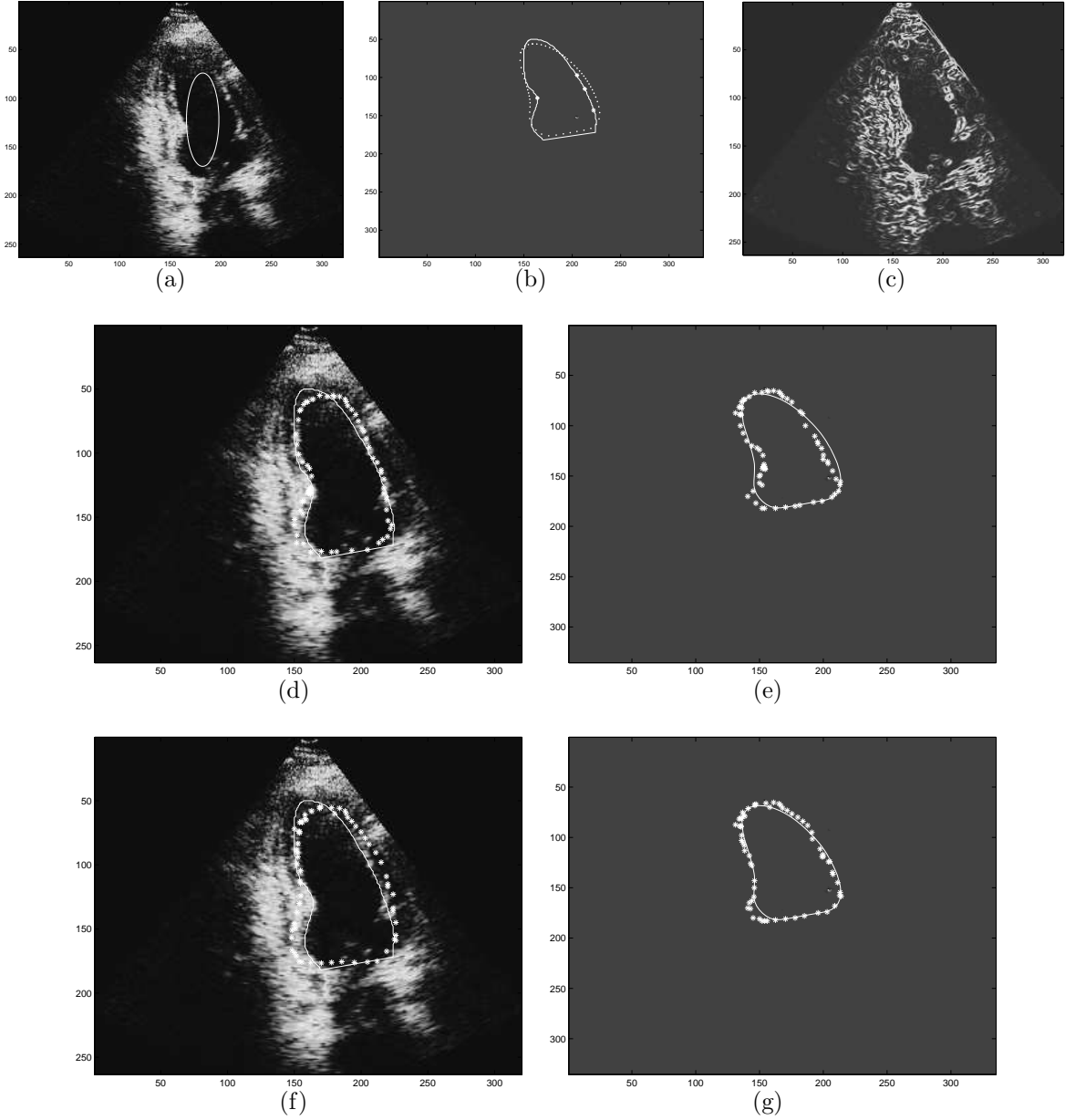


Figure 2-3. Compare segmentation results of another cardiac ultrasound image. A) a typical 2-chamber ultrasound image with an initial contour. B) Expert's endocardium (solid contour), "average shape" (dotted contour), and five points on the expert's contour. C) The image $|\nabla G_\sigma * I|$. D) The endocardium segmented by using model (2-4) (dotted) and the expert's contour (solid). E) The transformed solution of (2-4) (dotted), and the "average shape". F) The endocardium segmented by using model (2-3) (dotted) and the expert's contour (solid). G) The transformed solution of (2-3) (dotted), and the "average shape".

from the active contour. We used an implicit numerical scheme to solve the minimization problem. In the experiments with application to cardiac ultrasound images, the active contour could segment images in which the complete boundary was missing and the shape of the boundary has relatively large distortion from the prior. Besides the segmentation, the algorithm also provides estimates of translation, rotation, scale that map the active contour to the prior shape. These estimates are useful in aligning images.

CHAPTER 3

USING NONPARAMETRIC DENSITY ESTIMATION TO SMOOTH AND SEGMENT IMAGES SIMULTANEOUSLY

3.1 Proposed Model

A new model which could simultaneously segment and smooth images without using smoothness term as the second part in (1-1) is provided here.

For a recap of nonparametric density estimation, let $X \sim f(x)$ be a random variable, $\{X_1, X_2, \dots, X_n\}$ be a set of realization of X . A kernel method estimation of $f(x)$ would be $\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - X_i)$, where $K_h = \frac{K(\cdot/h)}{h}$ and $K(\cdot)$ is a kernel function symmetric about y axis and maximized at $x = 0$, and h is called bandwidth. Please refer [82, 83] for more detailed knowledge about nonparametric methods.

For notation simplicity, we denote by I_i the restriction of function I to Ω_i . To ensure applicability of the nonparametric density estimation idea, we make two assumptions. Firstly, at each voxel x , we treat $I_i(x)$ as a random variable. This is reasonable in the sense that even for the same subject, different scans would provide different data I_0 due to motion of the subject, noise involved in the data collection process etc. Randomness of I_0 leads to that of $I_i(x)$'s. Secondly, we assume that $I_i(x)$ has the same distribution as that of the $I_0(y)$'s, for all y that are in a small neighborhood of x and are inside region Ω_i . Note this is more general than in model (1-4), where $I_i(x)$ is assumed to have the same distribution as that of the $I_i(y)$'s, for all the y 's inside region Ω_i . Under the above two assumptions, we approximate the distribution of intensity $I_i(x)$ using intensities of $I_0(y)$, for y stays in the intersection of Ω_i and a neighborhood of x , denoted by $B(x, r(x))$, which is a ball centered at x with radius $r(x)$. The estimated distribution is called $\hat{f}_i(x)$:

$$\hat{f}_i(x) = \frac{1}{n} \sum_{j=1}^n K_h(I_i(x) - I_0(y_j)), y_j \in B(x, r(x)) \cap \Omega_i \quad (3-1)$$

In real implementation, sum and integral are treated equivalently, so for notation simplicity, we denote

$$\hat{f}_i(x) \propto \int_{B(x,r(x)) \cap \Omega_i} K_h(I_i(x) - I_0(y)) dy \quad (3-2)$$

Under the independence assumption, probability density function of I_i in the domain Ω_i would be $\prod_{x \in \Omega} \hat{f}_i(x)$. Joint probability density function of I in the whole domain Ω would be $\prod_i \prod_{x \in \Omega} \hat{f}_i(x)$. We still call logarithm of this function as log likelihood even though no parameters are involved. The proposed model is to find C , and $I_i(x)$'s for each point $x \in \Omega_i$, to minimize the following functional:

$$-\sum_i \int_{\Omega_i} \log \left(\int_{B(x,r(x)) \cap \Omega_i} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(I_i(x)-I_0(y))^2}{2\sigma^2}} dy \right) dx + \beta|C| \quad (3-3)$$

The first term is negative log likelihood, in which K_h is chosen as a Gaussian kernel function and bandwidth h is set to be 1. For each point x in each subregion Ω_i , the first part maximizes the log likelihood for all y 's that lie in the intersection of Ω_i and a neighborhood $B(x, r(x))$ of x . This is equivalent to some extent to trying to force $I_i(x) - I_0(y)$ to be 0, i.e., intensity $I_i(x)$ to be close to $I_0(y)$'s for all y 's lying in the intersection of Ω_i and neighborhood $B(x, r(x))$ of x . Roughly speaking, $I_i(x)$ will be close to mean value of I_0 in a small neighborhood that is inside Ω_i , which will smooth images $I_i(x)$ without crossing boundaries. Detailed proof of a simplified version will be provided in section 3.4. Thus the first part basically smooths each I_i inside Ω_i without crossing the boundary. Therefore, the first term not only prevents I_i 's from being too far away from I_0 , but also smooths I_i 's. This implies it works similarly to the first two parts in piecewise smooth MS(1-1). We would like to mention that convolution also smooths image through taking average, but convolution will blur images at boundaries and also cause shifting of boundaries. The reason is that $I_i(x)$ is set to be average of $I_0(y)$'s for all y 's lying in a neighborhood of x , this neighborhood is not restricted inside Ω_i . Results of model (3-3) will not cause these problems.

The second term works the same as the third term in (1-1) to obtain a smooth boundary. The proposed model is able to simultaneously segment and smooth images without using extra smoothness terms, and this also saves annoying work on choosing appropriate parameter to balance data fidelity term and smoothness term.

Gaussian kernel could be replaced by any bell-shaped kernels that are maximized at 0. By bell-shaped, we mean the graph of the kernel function is increasing on the left hand side of 0 and then decreasing on the right hand side of 0. Computation complexity would be decreased a lot if the Gaussian kernel involving expensive exponential computation is replaced by kernels involving simple calculation. Quadratic kernel is a good example. On example for quadratic kernel is $P(u) = \frac{3}{4}(1 - u^2)\chi_{[-1,1]}(u)$, where $\chi_{[-1,1]}(u)$ is the characteristic function of $[-1, 1]$. An alternative model would be to minimize the following functional with respect to $0 \leq I_1, I_2 \leq 1$ and C :

$$-\sum_i \int_{\Omega_i} \log \left(\int_{B(x, r(x)) \cap \Omega_i} \frac{3}{4}(1 - (I_i(x) - I_0(y))^2) dy \right) dx + \beta |C| \quad (3-4)$$

To force $I_i(x) - I_0(y)$ to be in interval $[-1, 1]$, which is the support of the quadratic kernel mentioned above, a linear transformation, which does not change segmentation results, is required to re-scale range of the initial image I_0 to $[a, b]$ for any $1 > b > a > 0$ ($a = .1, b = .9$ would be a choice). This condition is also required to guarantee existence of solutions. We can show that it is not enough to re-scale range of I_0 to $[0, 1]$, it has to be a proper subset of $[0, 1]$. Ranges of I_1 and I_2 are restricted to $[0, 1]$. In section 3.4, we will show that (3-4) is actually a non-constrained problem, i.e., the constraints will be satisfied automatically for solutions.

Obviously, the first term is minimized when for each point x in each subregion Ω_i , $I_i(x)$ is equal to $I_0(y)$'s for all y 's lying in neighborhood $B(x, r(x))$ of x and Ω_i . Hence it would work similarly as model (3-3) but require much less computation. Actually, (3-4) is equivalent to (3-3) when all the $\sigma(x)$'s are set to be a constant. Note this does not

mean (3-4) is returned to the piecewise constant MS (1-2) even though both of them have no variance involved. (3-4) forces $I_i(x)$ to be close to mean of $I_0(y)$'s for all y 's in the intersection of $B(x, r(x))$ and Ω_i . This is much more general and applicable than piecewise constant MS which enforces $I_i(x)$ to be close to a constant which turns out to be mean value of $I_0(y)$'s for all y 's in the whole Ω_i . This generality enables (3-4) to handle images which involve multi-modal distributions.

The level set formulation of model (3-4) is as follows. By using Heaviside function, all the integrals in the model are over Ω , so implementation is much easier:

$$\begin{aligned}
& - \int_{\Omega} H(\phi(x)) \log \left(\int_{\Omega} \chi_{B(x, r(x))}(y) H(\phi(y)) \frac{3}{4} (1 - (I_1(x) - I_0(y))^2) dy \right) dx \\
& - \int_{\Omega} (1 - H(\phi(x))) \log \left(\int_{\Omega} \chi_{B(x, r(x))}(y) (1 - H(\phi(y))) \frac{3}{4} (1 - (I_2(x) - I_0(y))^2) dy \right) dx \\
& + \beta \int_{\Omega} \delta(\phi(x)) |\nabla \phi(x)| dx
\end{aligned} \tag{3-5}$$

where $\chi_{B(x, r(x))}(y)$ is the characteristic function of $B(x, r(x))$. Ball size $r(x)$ is adaptively dependent on image gradient at x : when image gradient is high, i.e., at locations near boundaries, radii are smaller, while radii are bigger at more homogeneous regions. As a result, smoothing speed of I_1, I_2 are higher at homogenous regions and lower near boundaries to keep fine structure. Therefore, the smoothed images will be sharp. In comparison, in [84], ball sizes are fixed for all x , so it is harder to choose one uniform ball size that works for all the locations; smoothing speed will be the same at all locations. If a too small ball size is chosen, noise in the homogenous region can not be removed sufficiently, while if a too big ball size is selected, fine structures near boundaries will disappear. Similarly, model (1-1) also simultaneously smooths and segments images, but it assigns a fixed coefficient α for all x , so smoothing speed will be similar to what was observed in [84]. A comparison between models using variable radii and fixed radii will be shown in section 3.2.

3.2 Numerical Implementation

We will discuss some numerical implementation details of (3-5). First, radius $r(x)$ is chosen in the following way:

$$r(x) = \frac{M}{1 + N|\nabla\tilde{I}_0(x)|} \quad (3-6)$$

where \tilde{I}_0 is a smoothed version of the given initial image I_0 obtained through convolving it with a smooth kernel. Another straightforward choice is to replace I_0 by I , smoother version of I_0 that is updating during iterations:

$$r(x) = \frac{M}{1 + N|\nabla\tilde{I}(x)|} \quad (3-7)$$

where \tilde{I} is a smoothed version of the updating image I obtained through convolving it with a smooth kernel, we still convolve I even though it is smoother than I_0 . The reason is during the iteration, especially in the early stage, it is not smooth enough. In Figure3-1, we compare segmentation and smoothing results of model (3-5) with radii defined as in (3-6)(3-7). Figure3-1A) is a clean plane image, a speckle noise with parameter .05 is added to it, the resulting noisy image is shown in B). The second and the third row demonstrates segmentation and smoothing results based on (3-6)(3-7) respectively. No significant difference in both segmentation and smoothing results is visualized. Therefore, we will use (3-6) for all the following experiments. Also, it is easier to prove existence of model (3-5) for $r(x)$ defined as in (3-6), section 3.4 provides details.

M is an integer chosen based on noise level of the input image I_0 : M is chosen to be higher if the noisy level is high, and chosen smaller otherwise. N is chosen to be small all the time to keep and enhance boundaries while smoothing. s is set to be the 85th percentile of $|\nabla\tilde{I}_0(x)|$. This means we assume 15% of the points belong to the boundaries. 15% is a very conservative number; usually, in one image, number of points on the boundary will not be that many. Experimental results are not very sensitive to this

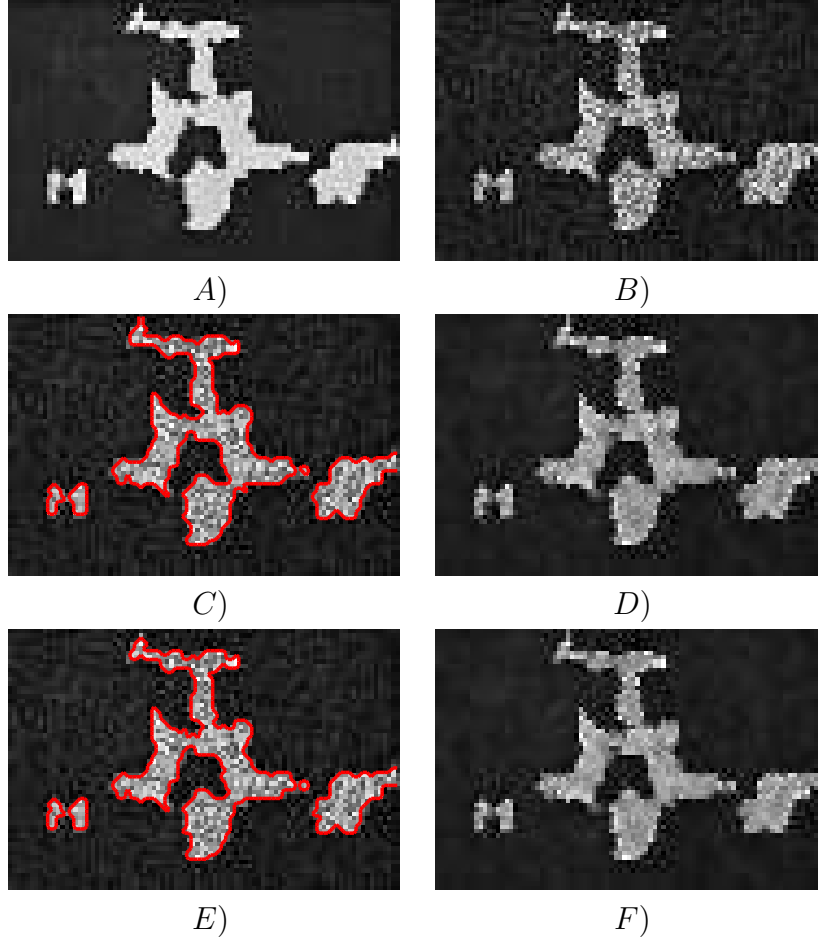


Figure 3-1. Compare results of model(3-5) based on two choices of $r(x)$ A) A clean plane image(60×90). B) a noisy plane image obtained by adding speckle noise with parameter .05 to A). C)-D): segmentation and smoothing results of B) using model(3-5) with radius as defined in (3-6) respectively. E)-F) segmentation and smoothing results of B) using model(3-5) with radius as defined in (3-7) respectively. For all the results, time step size = .2, $\beta = .01$, $M = 1$, $N=0$, converge in 20 iterations.

parameter; the general rule to adjust it is: set it to be higher if you prefer to keep and enhance fine structures more than to smooth images, while make it smaller if there are not that many fine structures in a quite noisy image.

The second issue is to use regularized version of Heaviside function and Delta function to make the functional differentiable with respect to ϕ . We use those defined in [16].

Thirdly, due to the difficulty of finding derivative of the functional with respect to ϕ directly, the functional is totally discretized. The discretized version of (3–5) is then implemented using an additive operator splitting(AOS) algorithm with boundary conditions $\frac{\partial \phi}{\partial n} = 0$ and $\frac{\partial I_i}{\partial n} = 0$ for $i = 1, 2$. AOS [85–87] is a semi-implicit scheme which is stable for large step sizes. It is at least ten times more efficient than the widely used explicit schemes. It guarantees equal treatment of all coordinate axes, can be implemented easily in arbitrary dimensions, and its computational complexity and memory requirement are linear in the number of pixels.

Finally, initial I_1, I_2 and ϕ could be chosen extremely flexible; even simple constant initials would work for most of the images we have tried. This removes the work of creating the distance function of an initial curve for initial ϕ in regular level set based method. The algorithm usually converges in fewer than 25 iterations, making reinitialization of level set function ϕ unnecessary. Parameter β affects experimental results and needs to be tuned for each image. But it does not vary much among different images.

Table 3-1: Segmentation Accuracy

Gaussian Noise	0.01	0.05	0.1	0.25
SA	0.996	0.979	0.957	0.868
Salt Pepper Noise	0.05	0.15	0.25	0.50
SA	0.996	0.988	0.970	0.885

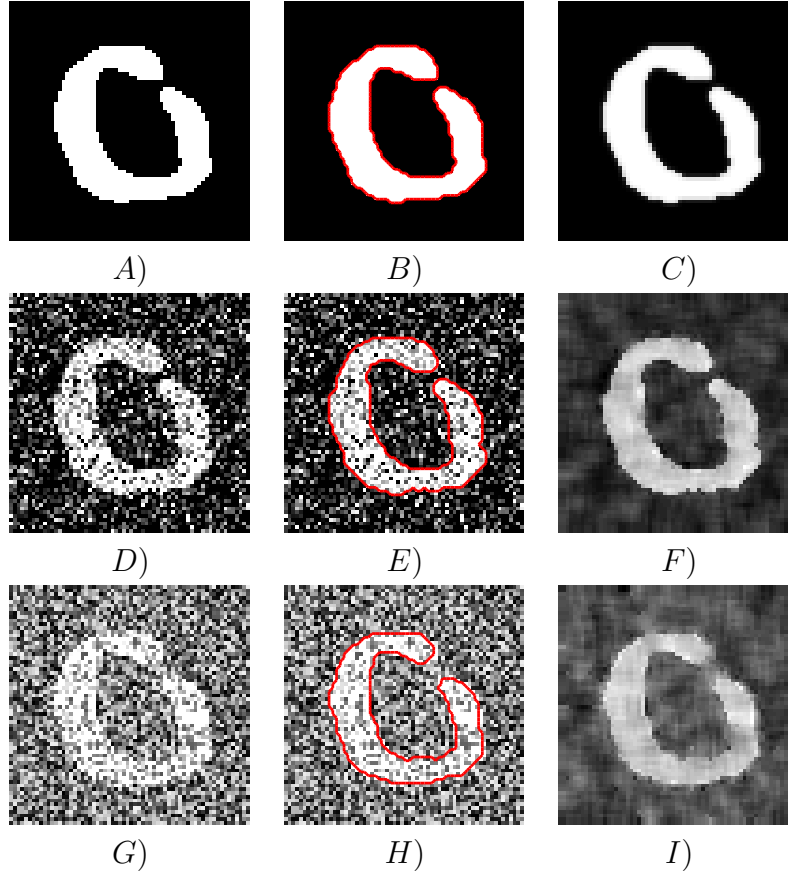


Figure 3-2. Segmentation results of a clean image and that of a noisy version. A)-C) Original cartoon C-shaped image, segmentation result and the smooth output resp. D)-F) An image obtained by adding Gaussian noise with variance 0.05 to A), segmentation result and smoothed version. G)-I) An image obtained by adding 0.03 speckle noise to A), segmentation ,smoothed version. For all the results, time step size = .1, $\beta = .01$, $M = 2$, $N=1$, converge in 20 iterations.

3.3 Validation and Application to T1 Magnetic Resonance Image

We provide three sets of experimental results. The goal of the first set is to test the ability of model (3-5) to simultaneously segment and smooth noisy images, and the second set is to compare segmentation ability of (3-5) with that of piecewise constant MS and global Gaussian pdf based method. Finally, segmentation and smoothing capability of model (3-5) that uses adaptive radii is compared with model in [84] that uses fixed radius. No comparison between (3-5) and piecewise smooth MS is provided as we think model in [84] is much easier to implement, and involves less parameters but works somehow equivalently as piecewise smooth MS does.

The first set of experiments is based on a clean synthetic C-shaped cartoon image with only two phases(Figure 3-2A)). We add different levels of random Gaussian noise and salt & pepper noise respectively to Figure3-2A) to create different noisy images. Model (3-5) is then applied to each of the noisy image and the clean cartoon image with parameter $\beta = .2$. A segmentation accuracy(SA) measure is defined to quantitatively observe how accurate the model is in segmenting a noisy image compared to the original cartoon image. We treat segmentation result of the clean cartoon image as ground truth. Then define SA to be the ratio of number of pixels sharing the same segmentation with the ground truth over total pixel number. SA is then between 0 and 1, the closer it is to 1, the better the model is in segmenting this noisy image.

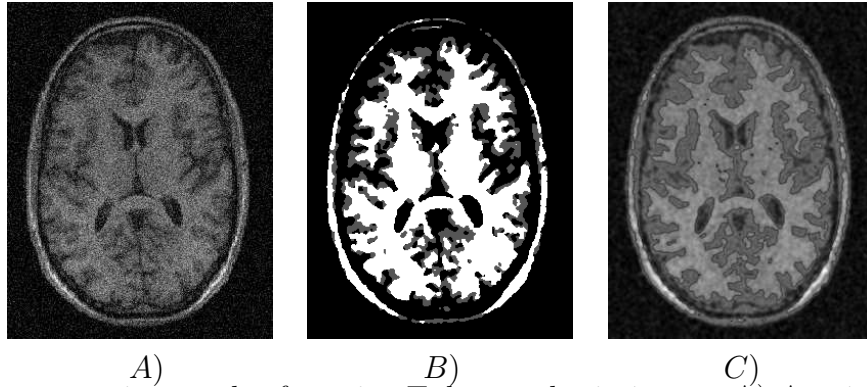


Figure 3-3. Segmentation result of a noisy T_1 human brain image. A) A noisy T_1 image. B) Segmentation result based on (3-5): black: background and CSF; gray:gray matter; white:white matter. C) Recovered smooth T_1 image.

Table 3-1 lists SA 's corresponding to images obtained by adding random Gaussian noise, with mean 0 but different variances, and salt & pepper noise with different parameters respectively. Each of the SA shown is the average value of two SA 's corresponding to two different images obtained through adding same type and strength of random noise at two different times. It is observed from Table 1 that when variance of Gaussian noise is lower than .1 and parameter of salt & pepper noise is less than 0.25 respectively, the segmentation accuracy could be higher than 95%. And an obvious trend is that as noise level increases, SA decreases as expected.

Figure 3-2 shows segmentation and smoothing results based on the clean C-shaped cartoon image (Figure 3-2A)) and two noisy versions of it: (Figure 3-2D)) which is obtained by adding a random Gaussian noise with zero mean and variance 0.05, (Figure 3-2G)) which is obtained by adding speckle noise with parameter 0.03. Their boundaries depicted in red curves which are found by applying model (3-5) are superimposed on them and shown in Figure 3-2B)E)H) respectively. It is observed that they are almost the same, actually, they coincide up to 99.6%. Figure 3-2C)F)I) demonstrate the smoothed version of Figure 3-2A)D)G) respectively. The recovered images are much smoother than the original noisy images. Figure 3-3 demonstrates segmentation and smoothing results on a real medical image. Figure 3-3A) is T1 weighted MRI brain image that involves four phases: Background, cerebral spinal fluid(CSF), Gray Matter and White Matter. Noise with unknown type and length is involved, distribution of intensity is of unknown type also, model (3-5) is applied to this image, segmentation and smoothing results are shown in Figure 3-3B)C) respectively. The segmentation result is quite reasonable, and it is obviously that Figure 3-3C) is significantly sharper than Figure 3-3A). We use hierarchical level set method recursively as Gao et al. mentioned in [24].

The second set of experiments is based on a plane image shown in Figure 3-4A). We add random Gaussian noise with mean 0 and variance 0.01,0.05 to it to create two noisy images which are shown in Figure 3-4B) and Figure 3-4C) respectively. We apply piecewise

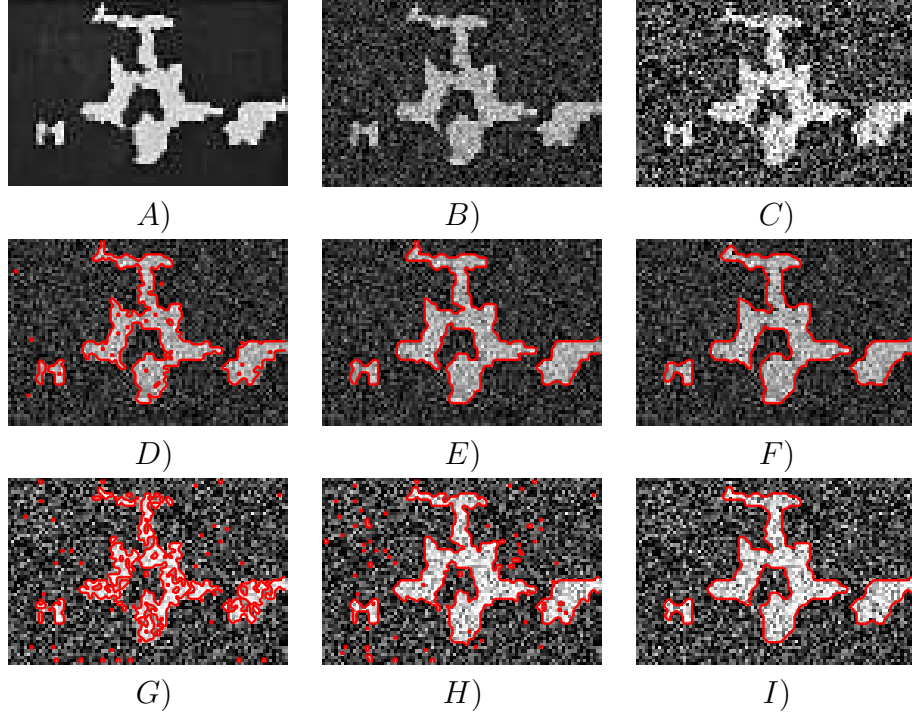


Figure 3-4. Compare three models in segmenting images with lower and higher level of noises. A) A plane image. B) A noisy plane image obtained by adding a random Gaussian noise with mean 0, variance 0.01 to A). C) A noisy plane image obtained by adding a random Gaussian noise with mean 0, variance 0.05 to A). D)-F) Segmentation results (red curve) of B) superimposed on B) based on piecewise constant MS, global Gaussian pdf and (3-5) respectively. G)-I) Segmentation results (red curve) of C) superimposed on B) based on piecewise constant MS, global Gaussian pdf and (3-5) respectively. For all the results, time step size = .1, $\beta = .01$, $M = 1$, $N=0$, converge in 20 iterations.

constant MS, global Gaussian pdf based method and proposed model (3-5) to Figure3-4B) respectively and obtained results(red curves) superimposed on Figure3-4B) as shown in Figure3-4 D)-F) respectively. Visually, not much difference among them is detected. Thus all the three models are able to segment images with low level noise. But after applying these three models to the noisier image Figure3-4C), from their results superimposed on Figure3-4C) as shown in Figure3-4G)-I) respectively, huge differences are observed. Moreover, it is obvious that result based on (3-5) is the best as it simultaneously segments and smooths the image so that boundaries of the noisy speckles are removed and the boundary of the plane is kept well. Result of piecewise constant MS is the worst: in order to remove boundaries of speckles, it pays the price of obtaining bad plane boundary. Result of global Gaussian pdf based model keeps plane boundary well but gains more boundaries of speckles as it allows variation of intensity in Ω_i 's.

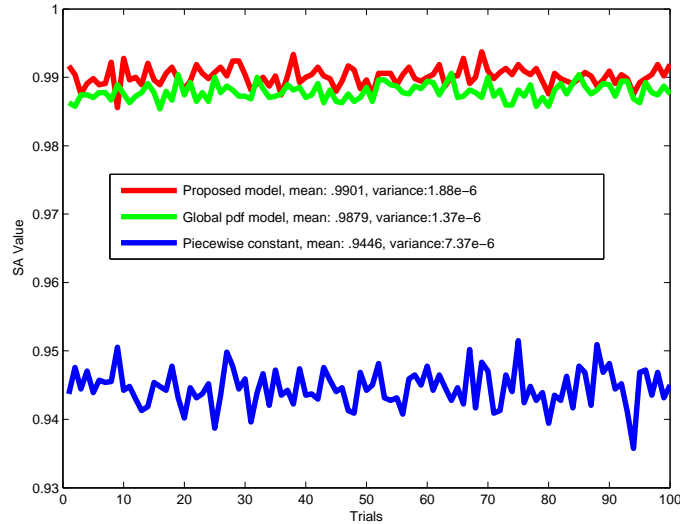


Figure 3-5. Graphs of SA obtained from three models: red: proposed model, green: global Gaussian pdf based model, blue: piecewise constant MS model.

A quantitative comparison of the three models is provided through SA. We add random Gaussian noise with mean 0 and variance 0.05 100 times to the clean plane image Figure3-4B) to get 100 noisy images. The three models are then applied to these 100 noisy images to obtain 100 SA's for each model, Figure3-5 shows the graphs of SA for

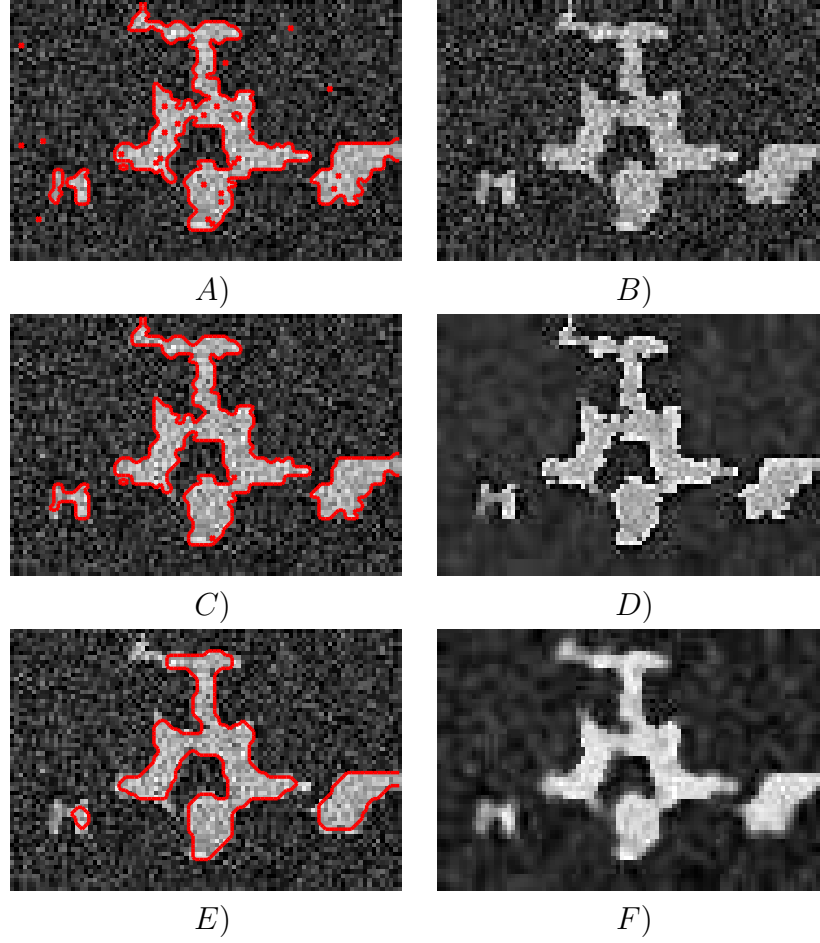


Figure 3-6. Compare segmentation and smoothing results between two models with fixed and adaptive radii A)-B) Segmentation and Smoothing results of Figure3-4B) using fixed radius 0. C)-D) Segmentation and Smoothing results of figure3-4B) using adaptive radii defined in (3-6). E)-F) Segmentation and Smoothing results of figure3-4B) using fixed radius 1. For all the results, time step size = .1, $\beta = .01$, converge in 20 iterations.

the three models. Mean value and variance of the 100 SA's obtained from the proposed model (shown in red contour), global Gaussian model (shown in green contour) and piecewise constant model (shown in blue contour) are $(.9901, 1.88 \cdot 10^{-6})$, $(.9446, 7.37 \cdot 10^{-6})$, $(.9879, 1.37 \cdot 10^{-6})$. The proposed model has the highest mean and smaller variance, while the piecewise constant has the lowest mean and smallest variance. It is obvious that the proposed model and the Global Gaussian pdf based model are significantly better than piecewise constant MS, while the proposed model is slightly better than the Global Gaussian pdf based model. But an advantage of the proposed model over the Global Gaussian pdf based model is that the former one can recover smooth image but the later cannot.

Thirdly, we compare model (3-5) with the model in [84] that uses fixed radius. In Figure 3-6 segmentation and smoothing results of model (3-5) and the model in [84] with two different radii are shown. The first row corresponds to that of model in [84] with fixed radius 0, the second row depicts that of model (3-5) with radius equals to 1 at relatively homogenous regions, 0 near boundaries. The third row shows that of model in [84] with fixed radius 1. It is obviously that the second row gives good results in both segmentation and smoothing. In Figure 3-6B), regions away from boundaries are smoothed well, near boundaries, image has discontinuity. In comparison, when use the model in [84] with a fixed small radius (see the first row), boundaries of fine structures are kept well but boundaries of the noisy background are also caught, recovered image is not smooth enough either. Oppositely, when a bigger fixed radius is chosen (the third row), images are smoothed enough, however, but we pay the cost of losing fine structures. This illustrates the advantage of model (3-5).

3.4 An Existence Theorem for the Model

To prove existence, we will use some facts about BV space([88],[89],[90]):

Let Ω be an open, bounded and connected set in \mathbb{R}^N . We say $u \in L^1(\Omega)$ is a function of bounded variation, denoted by $u \in BV(\Omega)$, if

$$|Du|(\Omega) = \sup_f \left\{ \int_{\Omega} u \operatorname{div} f dx \mid f \in C_c^1(\Omega; \mathbb{R}^N), |f| \leq 1 \right\} < \infty.$$

$|Du|(\Omega)$ is also denoted by $\int_{\Omega} |\nabla \chi_{\Omega}|$. Characteristic function of a set $A \subset \mathbb{R}^N$ is in $BV(\Omega)$ if and only if its perimeter $\operatorname{per}_{\Omega}(A) := |D\chi_A|(\Omega)$ is finite, where χ_A is the characteristic function of A .

The space $BV(\Omega)$ is a Banach space endowed with the norm:

$$\|u\|_{BV(\Omega)} = \|u\|_{L^1(\Omega)} + |Du|(\Omega).$$

Compactness theorem in BV

Assume Ω is an open, bounded and connected set with Lipschitz boundary. If $\{u_n\}_{n \geq 1}$ is a bounded sequence in $BV(\Omega)$, then there exists a subsequence $\{u_{n_j}\}$ of $\{u_n\}$ in $BV(\Omega)$ and a function $u \in BV(\Omega)$, such that $u_j \rightarrow u$ strongly in $L^1(\Omega)$ as $j \rightarrow \infty$.

Lower Semi-Continuity of BV-norm in L^1

If $u, \{u_j\} \subset BV(\Omega)$ and $u_j \rightarrow u$ in $L^1(\Omega)$, then

$$|Du|(\Omega) \leq \liminf_{j \rightarrow \infty} |Du_{n_j}|(\Omega).$$

To prove the existence of solution(s) to model (3–5) concisely, we assume image domain Ω is decomposed into two regions A and $\Omega \setminus A$. Proof would be similar if there are more subregions, in which case, there will be just more similar additive terms in the energy functional E . Let $A = \{x \in \Omega \mid H(x) > 0\}$, then $\chi_A = H(\phi)$, using ideas as in [16], model (3–5) could be reformulated as:

$$\begin{aligned}
& \min_{\chi_A \in BV(\Omega), 0 < I_1, I_2 < 1} E(\chi_A, I_1, I_2) = \beta \int_{\Omega} |D\chi_A| \\
& - \int_{\Omega} \chi_A(x) \log \left(\int_{B(x, r(x))} \chi_A(y) \frac{3}{4} (1 - (I_1(x) - I_0(y))^2) dy \right) dx \\
& - \int_{\Omega} (1 - \chi_A(x)) \log \left(\int_{B(x, r(x))} (1 - \chi_A(y)) \frac{3}{4} (1 - (I_2(x) - I_0(y))^2) dy \right) dx \quad (3-8)
\end{aligned}$$

where the first term is the length term represented as total variation of characteristic function χ_A . In BV space notation, it is $|D\chi_A|(\Omega)$. Euler-Lagrange equations of $E(\chi_A, I_1, I_2)$ with respect to I_1 and I_2 are

$$I_1(x) \int_{B(x, r(x))} \chi_A(y) dy - \int_{B(x, r(x))} \chi_A(y) I_0(y) dy = 0, x \in A \quad (3-9)$$

$$I_2(x) \int_{B(x, r(x))} (1 - \chi_A(y)) dy - \int_{B(x, r(x))} (1 - \chi_A(y)) I_0(y) dy = 0, x \in \Omega \setminus A \quad (3-10)$$

We may assume $\int_{B(x, r(x))} \chi_A(y) dy > 0$ for all $x \in A$, then

$$I_1(x) = \frac{\int_{B(x, r(x))} \chi_A(y) I_0(y) dy}{\int_{B(x, r(x))} \chi_A(y) dy}, x \in A \quad (3-11)$$

If $\int_{B(x, r(x))} \chi_A(y) dy = 0$, which is called the degenerate case, $\int_{B(x, r(x))} \chi_A(y) I_0(y) dy$ would also equal to 0, (3-9) is satisfied automatically for any choice of $I_1(x)$ between 0 and 1.

Similar analysis for $I_2(x)$ gives

$$I_2(x) = \frac{\int_{B(x, r(x))} (1 - \chi_A(y)) I_0(y) dy}{\int_{B(x, r(x))} (1 - \chi_A(y)) dy}, x \in \Omega \setminus A \quad (3-12)$$

The solution of I_1, I_2 must be related with χ_A by (3-11)(3-12). Moreover, it is observed from equations (3-11)(3-12) that if $I_0(x) \in (.1, .9)$, for all $x \in \Omega$, then,

$$\begin{aligned}
.1 \int_{B(x, r(x))} \chi_A(y) dy & \leq \int_{B(x, r(x))} \chi_A(y) I_0(y) dy \leq .9 \int_{B(x, r(x))} \chi_A(y) dy, \text{ from which we get} \\
.1 & \leq \frac{\int_{B(x, r(x))} \chi_A(y) I_0(y) dy}{\int_{B(x, r(x))} \chi_A(y) dy} \leq \frac{\int_{B(x, r(x))} \chi_A(y) dy}{\int_{B(x, r(x))} \chi_A(y) dy} \leq .9, \text{ i.e., } 0.1 \leq I_1(x) \leq .9, \text{ therefore } 0 < I_1(x) < 1.
\end{aligned}$$

We can also prove $0 < I_2(x) < 1$. Hence $I_1(x), I_2(x)$ will be in $(0, 1)$ automatically for all $x \in \Omega$. Thus, we do not need to worry about the constraints of I_1, I_2 .

Because the unknown functions I_1 and I_2 have explicit representations as functions of the unknown χ_A , we can consider the energy E in χ_A only: $E(\chi_A, I_1(\chi_A), I_2(\chi_A))$. Therefore, we obtain the minimization problem:

$$\min_{\chi_A} E(\chi_A), \chi_A \in \{0, 1\}dx - a.e. \quad (3-13)$$

among characteristic functions of sets A with finite perimeter in Ω , i.e., characteristic functions of bounded variation([89]).

We now state our existence result:

Theorem

If Ω is open, bounded and connected with Lipschitz boundary, $0.1 < I_0(x) < .9$ for all $x \in \Omega$, then $\inf_{\chi_A} E(\chi_A), \chi_A \in BV(\Omega), \chi_A \in \{0, 1\}dx - a.e$ has a minimizer.

Proof: Observe that

$$\begin{aligned} & \left| - \int_{\Omega} \chi_A(x) \log \left(\int_{B(x, r(x))} \chi_A(y) \frac{3}{4} (1 - (I_1(x) - I_0(y))^2) dy \right) dx \right| \\ & \leq \int_{\Omega} |\chi_A(x)| \cdot \left| \log \left(\int_{B(x, r(x))} |\chi_A(y) (1 - (I_1(x) - I_0(y))^2)| dy \right) \right| dx \\ & \leq \int_{\Omega} |\log |B(x, r(x))|| dx \leq \int_{\Omega} |\log |\Omega|| dx = |\Omega| \cdot |\log |\Omega|| \end{aligned}$$

Similarly,

$$\begin{aligned} & \left| - \int_{\Omega} (1 - \chi_A(x)) \log \left(\int_{B(x, r(x))} (1 - \chi_A(y)) \frac{3}{4} (1 - (I_1(x) - I_0(y))^2) dy \right) dx \right| \\ & \leq \int_{\Omega} |1 - \chi_A(x)| \cdot \left| \log \left(\int_{B(x, r(x))} |(1 - \chi_A(y)) (1 - (I_1(x) - I_0(y))^2)| dy \right) \right| dx \\ & \leq \int_{\Omega} |\log |B(x, r(x))|| dx \leq \int_{\Omega} |\log |\Omega|| dx = |\Omega| \cdot |\log |\Omega|| \end{aligned}$$

Thus the energy functional $E(\chi_A)$ is bouned below, so there exists a minimizing sequence $\{\chi_{A_n}\}$ such that

$$\lim E(\chi_{A_n}) = \liminf_{\chi_A \in BV(\Omega)} E(\chi_A).$$

By boundness of $\{E(\chi_{A_n})\}$ we have $\{|D\chi_{A_n}|\}$ bounded above because the second and the third terms in $\{E(\chi_{A_n})\}$ are bounded below. Moreover, $\int_{\Omega} \chi_{A_n} dx \leq \int_{\Omega} 1 dx = |\Omega| < \infty$, so $\{\chi_{A_n}\}$ is bounded in $L^1(\Omega)$. Therefore $\{\chi_{A_n}\}$ is a bounded sequence in $BV(\Omega)$, so by compactness of BV space and lower semi-continuity of the BV-norm in L^1 topology, there exists a subsequence $\{\chi_{A_{n_j}}\}$ and a $u \in BV(\Omega)$, such that $\chi_{A_{n_j}} \rightarrow u$ strongly in $L^1(\Omega)$ as $j \rightarrow \infty$. From this we deduce that there exists a subsequence, still denoted by A_{n_j} , such that $\chi_{A_{n_j}} \rightarrow u$ a.e. in Ω , and since $\chi_{A_{n_j}}$ takes values of either 0 or 1, $u = \chi_D$ for some D with finite perimeter in Ω . By compactness theorem in BV space and lower semi-continuity in L^1 , we have

$$|Du|(\Omega) \leq \liminf_{j \rightarrow \infty} |D\chi_{A_{n_j}}|(\Omega). \quad (3-14)$$

Note

$$I_{1_n}(x) = \frac{\int_{B(x,r(x))} \chi_{A_{n_j}}(y) I_0(y) dy}{\int_{B(x,r(x))} \chi_{A_{n_j}}(y) dy}, x \in A_{n_j}$$

$$I_{2_n}(x) = \frac{\int_{B(x,r(x))} (1 - \chi_{A_{n_j}}(y)) I_0(y) dy}{\int_{B(x,r(x))} (1 - \chi_{A_{n_j}}(y)) dy}, x \in \Omega \setminus A_{n_j}$$

Since $\chi_{A_{n_j}} \rightarrow \chi_A$ pointwisely a.e, and $.1 < I_0 < .9$, we have $0.1 \leq \chi_{A_{n_j}}(y) I_0(y) \leq .9$, so by the Dominated Convergence Theorem, $I_{1_n}(x)$ converges to $I_1(x) := \frac{\int_{B(x,r(x))} \chi_D(y) I_0(y) dy}{\int_{B(x,r(x))} \chi_A(y) dy}$. The range of $I_1(x)$ is then obviously in $(0, 1)$ because $.1 \leq I_0 \leq .9$. It is straightforward to show that $I_2(x)$, defined as $\frac{\int_{B(x,r(x))} (1 - \chi_D(y)) I_0(y) dy}{\int_{B(x,r(x))} (1 - \chi_D(y)) dy}$ (limit of $I_{2_n}(x)$) has range in $(0, 1)$ as well.

Regarding the second term in (3-8),

$$\chi_{A_{n_j}}(y) \frac{3}{4} (1 - (I_{1_n}(x) - I_0(y))^2) \rightarrow \chi_D(y) \frac{3}{4} (1 - (I_1(x) - I_0(y))^2)$$

strongly as $j \rightarrow \infty$ and all the terms are less than or equal to 1, so by the Dominated Convergence Theorem,

$$\int_{B(x,r(x))} \chi_{A_{n_j}}(y) \frac{3}{4} (1 - (I_{1_n}(x) - I_0(y))^2) dy \rightarrow \int_{B(x,r(x))} \chi_D(y) \frac{3}{4} (1 - (I_1(x) - I_0(y))^2) dy$$

pointwise. This implies $\chi_{A_{n_j}}(x) \log \left(\int_{B(x,r(x))} \chi_{A_{n_j}}(y) \frac{3}{4} (1 - (I_1(x) - I_0(y))^2) dy \right)$ converges to $\chi_D(x) \log \left(\int_{B(x,r(x))} \chi_D(y) \frac{3}{4} (1 - (I_1(x) - I_0(y))^2) dy \right)$ pointwise. Because all the terms are less than or equal to $|\log(|\Omega|)|$, then apply Dominant convergence theorem again, we obtain

$$\begin{aligned} & \lim_{j \rightarrow \infty} - \int_{\Omega} \chi_{A_{n_j}}(x) \log \left(\int_{B(x,r(x))} \chi_{A_{n_j}}(y) \frac{3}{4} (1 - (I_1(x) - I_0(y))^2) dy \right) dx \\ &= - \int_{\Omega} \chi_D(x) \log \left(\int_{B(x,r(x))} \chi_D(y) \frac{3}{4} (1 - (I_1(x) - I_0(y))^2) dy \right) dx \end{aligned}$$

Similarly we have continuity of the third term in the energy functional. Continuity of these two terms together with (3–14) implies

$$E(\chi_D) \leq \liminf_{j \rightarrow \infty} E(\chi_{A_{n_j}})$$

Hence χ_D is a minimizer.

Remark: $I_1(x), I_2(x)$ turn out to be mean values of I_0 in a neighborhood of x that is restricted to lie inside subregions A and $\Omega \setminus A$ respectively. This is comparable but different from piecewise constant Mumford-Shah model in which intensity in each subregion is approximated by a constant that is in fact the mean value of the original image intensity in the whole subregion. In comparison, model (3–8) is more general and is able to smooth and segment images that have high level noise involved.

3.5 Conclusion

A new model that utilized both intensity and edge information was introduced to simultaneously segment and smooth images without using extra smoothness terms. Due to lack of knowledge regarding image intensity distribution, we applied nonparametric method to estimate intensity distribution, using intensity information from neighboring points. Sizes of neighborhood were chosen adaptively based on image gradient. Segmentation and smoothing results are obtained through minimizing negative log likelihood and also the length of the objects boundaries. The minimization problem was implemented using semi-implicit iterative scheme. During the iteration, segmentation results were updated based on better smoothing images, smoothing results were then

updated by using the better edge information derived from the better segmentation results. Segmentation and smoothing procedures were improving each other in one single model. The adaptive neighborhood sizes allowed for variate smoothing levels for different locations, such that edges were enhanced.

Experimental results on several sets of synthetic and real medical images with different types and levels of noise demonstrated the potential of the proposed model in simultaneously segmenting and smoothing images. Comparison with another three models showed the advantage of the proposed model.

CHAPTER 4

ESTIMATION, SMOOTHING AND CHARACTERIZATION OF APPARENT DIFFUSION COEFFICIENT

4.1 Introduction

There is growing interest in diffusion tensor denoising and reconstruction from DTI data. There are several popular approaches: (1). Smoothing the raw data $s(\mathbf{q})$ then estimating the diffusion tensor from the smoothed raw data ([91–93]); (2). Smoothing the principal diffusion direction after the diffusion tensor has been estimated from the raw noisy measurements ([94–98]); (3). Smoothing tensor-valued data while preserving the positive definite property of D [92, 96, 99–101].

However, very little research reported in literature to date on HARD data analysis considers denoising problem in the reconstruction of the ADC profiles when the HARD raw data is noisy. To improve the accuracy of the estimation, in this paper we present a novel model that has the ability of simultaneously smoothing and estimating the ADC profile $d(\mathbf{x}, \theta, \phi)$ from the noisy HARD measurements $s(\mathbf{x}, \theta, \phi)$ while preserving the relevant features, and the positiveness and antipodal symmetry constraints of $d(\mathbf{x}, \theta, \phi)$. The basic idea of our approach is to approximate the ADC profiles at each voxel by a 4th order SHS($l_{max} = 4$ in (1–13)):

$$d(\mathbf{x}, \theta, \phi) = \sum_{l=0,2,4} \sum_{m=-l}^l A_{l,m}(\mathbf{x}) Y_{l,m}(\theta, \phi) \quad (4-1)$$

whose coefficients are determined by solving a constrained minimization problem. This minimization problem minimizes a non-standard growth functional to perform a feature preserved regularization, while it minimizes the data fidelity term. Notice, there are 15 unknown complex valued functions $A_{l,m}$ involved. Since $d(\theta, \phi)$ is real and $Y_{l,m}$ satisfies $Y_{l,-m} = (-1)^m \overline{Y_{l,m}}$, each complex valued $A_{l,m}$ is constrained by $A_{l,-m} = (-1)^m \overline{A_{l,m}}$, where $\overline{A_{l,m}}$ denotes the complex conjugate of $A_{l,m}$. This constraint transforms the 15 unknown complex valued functions in (4-1) to 15 real valued functions: $A_{l,0}(\mathbf{x})$, ($l = 0, 2, 4$), $Re A_{l,m}(\mathbf{x})$, $Im A_{l,m}(\mathbf{x})$, ($l = 2, 4; m = 1, \dots, l$).

This chapter elaborates a work on how to obtain all those 15 coefficients at each voxel. This method differs from the existing approaches developed in [67] and [72] mainly in the aspect of the determination of the $A_{l,m}(\mathbf{x})$'s in (1–13). In ([67]) the $A_{l,m}(\mathbf{x})$'s (l is even) are determined by

$$A_{l,m}(\mathbf{x}) = \int_0^{2\pi} \int_0^\pi -\frac{1}{b} \log \frac{s(\mathbf{q})}{s_0} Y_{l,m}(\theta, \phi) \sin\theta d\theta d\phi, \quad (4-2)$$

and in [72] they are estimated as the least-squares solutions of

$$-\frac{1}{b} \log \frac{s(\mathbf{q})}{s_0} = \sum_{l=0}^{l_{max}} \sum_{m=-l}^l A_{l,m} Y_{l,m}(\theta, \phi). \quad (4-3)$$

In this chapter, the estimation of $A_{l,m}(\mathbf{x})$'s is not performed individually at each isolated voxel, but a process of joint estimation and regularization across the entire volume. The joint estimation and regularization not only guarantees the wellposedness of the proposed model, but also enhances the accuracy of the estimation since the HARD data are noisy. Moreover, in this paper we provide more detailed method to characterize the diffusion anisotropy, which uses not only the information of $A_{l,m}(\mathbf{x})$'s as in ([67, 72]), but also the variation of $d(\theta, \phi)$ about its mean. Our experimental results showed the effectiveness of the model in the estimation and enhancement of anisotropy of the ADC profile. The characterization of the diffusion anisotropy based on the reconstructed ADC profiles using the proposed model is consistent with the known fiber anatomy.

4.2 Model Description

In this section we will present a variational framework for simultaneous smoothing and estimation of the ADC profile $d(x, \theta, \phi)$ from the noisy HARD measurements $s(\mathbf{q})$. This model minimizes an energy functional consisting of two parts: one is the cost for regularizing d , and the other is the cost for fitting d to the HARD measurements through the original Stejskal-Tanner equation (1–11). To explain the basic idea of our method, we focus our attention on the cases where there are at most two fibers passing through a

single voxel. The same idea can be applied to the cases where there are more fibers within a voxel.

The challenge in regularizing d comes from two aspects. First, d is defined on $\Omega \times S^2$ rather than $\Omega \times R^2$, hence, the derivatives for (θ, ϕ) should be along the sphere. Secondly, the regularized d has to preserve the antipodal symmetry property with respect to (θ, ϕ) . Considering these facts we adopt the idea developed in [67, 72] that approximates d by its SHS consisting of only even order components up to order 4, i.e.

$$d(x, \theta, \phi) = \sum_{l=0,2,4} \sum_{m=-l}^l A_{l,m}(\mathbf{x}) Y_{l,m}(\theta, \phi). \quad (4-4)$$

The expression in (4-4) ensures the smoothness and antipodal symmetry property of $d(\mathbf{x}, \theta, \phi)$ in terms of (θ, ϕ) , this is easy to see from the definition of $Y_{l,m}(\theta, \phi)$. For the cases where possibly k fibers cross in a single voxel, the sum in (4-4) should be replaced by $\sum_{l=0,2,\dots,2k}$.

Now the problem of regularization and estimation of $d(x, \theta, \phi)$ reduces to that for the 15 complex valued functions $A_{l,m}(x)$ ($l = 0, 2, 4$ and $m = -l, \dots, l$) in (4-4). Since $d(\theta, \phi)$ at each voxel is a real valued function, and $Y_{l,m}$ satisfies $Y_{l,-m} = (-1)^m Y_{l,m}^*$, $A_{l,m}$ should be constrained by

$$A_{l,-m} = (-1)^m A_{l,m}^*.$$

This constraint reduces the number of the unknown coefficients $A_{l,m}$ in (4-4) to 15 real valued functions. They are

$$A_{l,0}(x), \quad (l = 0, 2, 4), \quad \text{Re}A_{l,m}(x), \quad \text{Im}A_{l,m}(x), \quad (l = 2, 4 \text{ and } m = 1, \dots, l). \quad (4-5)$$

By using (4-5), we can rewrite (4-4) as

$$d(x, \theta, \phi) = \sum_{l=0,2,4} A_{l,0}(\mathbf{x}) Y_{l,0}(\theta, \phi) + 2 \sum_{l=2,4} \sum_{m=1}^l (\text{Re}A_{l,m}(x) \text{Re}Y_{l,m}(\theta, \phi) - \text{Im}A_{l,m}(x) \text{Im}Y_{l,m}(\theta, \phi)), \quad (4-6)$$

where ReF and ImF represent the real and imaginary part of a function F respectively. Now the problem of regularizing and estimating d reduces to smoothing and estimation of 15 functions in (4–5) simultaneously.

There are many choices of regularizing operators to smooth the 15 functions in (4–5). Total Variation (TV) based regularization, first proposed by Rudin, Osher and Fatemi[79], proved to be an invaluable tool for feature preserving smoothing. However, it sometimes causes a staircase effect making restored image blocky, and even containing 'false edges' [102, 103]. An improvement, that combines the TV based smoothing with isotropic smoothing, was given by Chambolle and Lions [103]. Their model minimizes the TV norm when the magnitude of the image gradient is larger, and the L^2 norm of the image gradient if it is smaller. However, this model is sensitive to the choice of the threshold which separates the TV based and isotropic smoothing. To further improve Chambolle and Lions' model and make the model having an ability to self adjust diffusion property, recently, certain nonstandard diffusion models based on minimizing $L^{p(x)}$ norm of image gradient have been developed [102, 104]. To recover an image u from an observed image I in [102] the diffusion was governed by minimizing

$$\min_u \int_{\Omega} |\nabla u|^{p(|\nabla u|)}$$

where $p(s)$ is monotonically decreasing function and $\lim_{s \rightarrow 0} p(s) = 2$, $\lim_{s \rightarrow \infty} p(s) = 1$. In [104] the diffusion was performed through minimizing

$$\min_u \int_{\Omega} \phi(x, Du)$$

where $\phi(x, r) := \frac{1}{p(x)}|r|^{p(x)}$ if $|r| \leq a$, and $\phi(x, r) := |r| - \frac{ap(x) - a^{p(x)}}{p(x)}$, if $|r| > a$ for threshold a . In this model $p(x)$ is chosen as

$$p(x) = p(|\nabla I|) = 1 + \frac{1}{1 + k|\nabla G_{\sigma} * I|^2},$$

where $k, \sigma > 0$ are parameters, G_σ is the Gaussian kernel. Both models in [102] and [104] are able to self adjust diffusion range from isotropic to TV-based depending on image gradient. At the locations with higher image gradients ($p = 1$), the diffusion is TV based and strictly tangential to the edges ([79, 103, 105]). In homogeneous regions the image gradients are very small ($p = 2$), the diffusion is essentially isotropic. At all other locations, the image gradient forces $1 < p < 2$, and the diffusion is between isotropic and total variation based and varies depending on the local properties of the image. This self adjusting ability enables these models to effectively preserve features when images are smoothed.

Applying the idea of minimizing functionals with variable exponent to the problem of regularizing the coefficients $A_{l,m}$'s in (4-4), we propose to minimize

$$E_1(A_{l,m}) := \int_{\Omega} \sum_{l=0,2,4} \sum_{m=-l}^l \phi_{l,m}(x, DA_{l,m}), \quad (4-7)$$

where

$$\phi_{l,m}(x, r) := \begin{cases} \frac{1}{p_{l,m}(x)} \left| \frac{r}{M_{l,m}} \right|^{p_{l,m}(x)}, & |r| \leq M_{l,m} \\ \left| \frac{r}{M_{l,m}} \right| - \left(1 - \frac{1}{p_{l,m}(x)}\right), & |r| > M_{l,m} \end{cases} \quad (4-8)$$

We would like to point out that for a function $A_{l,m} \in BV$, $DA_{l,m}$ is a measure, the definition of (4-7) is not obvious. This will be discussed in existence section 4.6 below.

In (4-8)

$$p_{l,m} = 1 + \frac{1}{1 + k|\nabla G_\sigma * a_{l,m}(x)|^2}, \quad (4-9)$$

and $a_{l,m}$ is the least-squares solution of

$$-\frac{1}{b} \log \frac{s(x, \theta, \phi)}{s_0(x)} = \sum_{l=0,2,4} \sum_{m=-l}^l a_{l,m}(x) Y_{l,m}(\theta, \phi). \quad (4-10)$$

In real application, $M_{l,m}$ is picked as 90th percentile of r . Experimental results are not sensitive to this parameter.

Again we would like to point out that E_1 only needs to include 15 terms corresponding to the 15 real valued functions in (4-5). Here we write it in terms of

$A_{l,m}$ in order to shorten the expression of the formula. If using $A_{l,m}$ instead of $a_{l,m}$ in (4-9) we may get better numerical results, since $p_{l,m}$ would depend on updated $A_{l,m}$ rather than the fixed $a_{l,m}$ in the iterations to minimize (4-7). However, it gives difficulty in the study of the existence of solutions.

Since $d(x, \theta, \phi)$ is related to the HARD measurements $s(x, \theta, \phi)$ and $s_0(x)$ through the Stejskal-Tanner equation (1-11), the estimation of the $A_{l,m}$'s is based on the original Stejskal-Tanner equation (1-11) rather than its (log) linearized form (1-12), that is,

$$E_2(A_{l,m}) := \frac{1}{2} \int_{\Omega} \int_0^{2\pi} \int_0^{\pi} |s(\mathbf{x}, \theta, \phi) - s_0(x)e^{-bd(x,\theta,\phi)}|^2 \sin\theta d\theta d\phi dx, \quad (4-11)$$

where d is determined in (4-4). As observed in [106] when the signal to noise ratio is low the linearized model gives different results.

Finally, to simultaneously regularize and estimate the ADC $d(x, \theta, \phi)$, our model minimizes the energy function

$$E(A_{l,m}) := \lambda E_1(A_{l,m}) + E_2(A_{l,m}), \quad (4-12)$$

with respect to $A_{l,m}$ ($l = 0, 2, 4$ and $m = -l, \dots, l$) in the space of $BV(\Omega)$, (in fact, only 15 functions in (4-5) are needed), and subject to the constraint:

$$d(x, \theta, \phi) \geq 0. \quad (4-13)$$

In (4-12), (4-13), $s(x, \theta, \phi)$ and $s_0(\mathbf{x})$ are the noisy HARD measurements (real valued), $d(x, \theta, \phi)$ is the SHS given in (4-4), $\Omega \subset R^3$ is the image domain, $\lambda > 0$ is a parameter which could be different for different A_{lm} . E_1 and E_2 are given in (4-7) and (4-11), respectively.

Before we derive the Euler-Lagrange equations for our model (4-12), (4-13), we would like to point out that if the measurements satisfy the condition $s(x, \theta, \phi) \leq s_0(x)$, the solution of (4-12) meets the constraint (4-13) automatically. Therefore we can treat our model as an unconstrained minimization. This is given in the following lemma.

Lemma: Under the assumption that

$$s(x, \theta, \phi) \leq s_0(x), \quad \text{for all } \mathbf{x} \in \Omega, \quad 0 \leq \theta < \pi, \quad 0 \leq \phi < 2\pi \quad (4-14)$$

the minimizer of (4-12) always satisfies the constraint (4-13).

Proof: Let $A_{l,m}(x)$ ($l = 0, 2, 4$ and $m = -l, \dots, l$) be the minimizer of (4-12) in $BV(\Omega)$, and $d(x, \theta, \phi)$ be the function defined in (4-4) associated with these optimal $A_{l,m}(x)$'s. Given $x \in \Omega$, if $d(\mathbf{x}, \theta, \phi) < 0$ for some $0 \leq \theta < \pi, 0 \leq \phi < 2\pi$, then define $\hat{d}(x, \theta, \phi) = 0$, otherwise, define $\hat{d}(x, \theta, \phi) := d(x, \theta, \phi)$.

Correspondingly,

$$\hat{A}_{l,m}(x) := \int_0^{2\pi} \int_0^\pi \hat{d}(x, \theta, \phi) Y_{l,m}(\theta, \phi) \sin\theta d\theta d\phi.$$

Then, using the orthonormality of the spherical harmonics and the definition of d , we have

$$\hat{A}_{l,m}(x) = \begin{cases} A_{l,m}(x), & \text{if } d(x, \theta, \phi) \geq 0, \forall 0 \leq \theta < \pi, 0 \leq \phi < 2\pi \\ 0, & \text{if } d(x, \theta, \phi) < 0 \end{cases} \quad (4-15)$$

This implies that

$$\phi_{l,m}(x, D\hat{A}_{l,m}) \leq \phi_{l,m}(x, DA_{l,m}),$$

hence

$$E_1(\hat{A}_{l,m}) \leq E_1(A_{l,m}).$$

Moreover, it is easy to obtain

$$E_2(\hat{d}) \leq E_2(d),$$

if (4-14) holds. From the last two inequality above, we obtain that $E(\hat{d}) \leq E(d)$. This contradicts to the fact that d minimizes energy functional (4-12).

Now we give the evolution equations associated with the Euler-Lagrange (EL) equations for (4-12): for $l = 0, 2, 4$ and $m = -l, \dots, l$,

$$\frac{\partial A_{l,m}}{\partial t} = \lambda \operatorname{div}(\partial_r \phi_{l,m})(x, DA_{l,m}) - b \int_0^{2\pi} \int_0^\pi s_0 e^{-bd} (s - s_0 e^{-bd}) Y_{l,m} \sin\theta d\theta d\phi, \quad (4-16)$$

with the initial and boundary conditions:

$$A_{l,m} = a_{l,m}, \quad \text{on } \Omega \times \{t = 0\},$$

$$(\partial_r \phi_{l,m})(x, DA_{l,m}) \cdot n = 0 \quad \text{on } \partial\Omega \times \mathbb{R}^+.$$

In the above EL equation n is the unit outward normal to the boundary of Ω , and

$\partial_r \phi_{l,m}(x, r)$ is a continuously differentiable function in r , and

$$\partial_r \phi_{l,m}(x, r) := \begin{cases} \frac{1}{M_{l,m}^p(x)} |r|^{p(x)-2} r, & |r| \leq M_{l,m} \\ \frac{1}{M_{l,m}} |r|^{-1} r, & |r| > M_{l,m} \end{cases} \quad (4-17)$$

$\partial_r \phi_{l,m}$ can also be written as

$$\partial_r \phi_{l,m}(x, r) := \frac{1}{M_{l,m}^q(x)} |r|^{q(x)-2} r, \quad (4-18)$$

where $q(x) = p(x)$ if $|r| \leq M_{l,m}$, and $q(x) = 1$ if $|r| > M_{l,m}$.

4.3 Characterization of anisotropy

In [67] the $|A_{l,m}(x)|$ ($l = 0, 2, 4$ and $m = -l, \dots, l$) in the truncated SHS (4-4) are used to characterize the diffusion anisotropy at each voxel x . Our experimental results, however, indicate this information alone is insufficient to separate isotropic diffusion, one-fiber diffusion, and multi-fiber diffusion within a voxel. We propose to combine the information from $|A_{l,m}|$ with the variances of $d(\phi, \theta)$ about its mean value to characterize the diffusion anisotropy. We outline our algorithm as follows:

(1). If

$$R_0 := |A_{0,0}| / \sum_{l=0,2,4} \sum_{m=-l}^l |A_{l,m}|, \quad (4-19)$$

is large, or the variance of $d(\theta, \phi)$ about its mean is small, the diffusion at such voxels is classified as isotropic.

(2). For the remaining voxels, if

$$R_2 := \sum_{m=-2}^{m=2} |A_{2,m}| / \sum_{l=2,4} \sum_{m=-l}^l |A_{l,m}| \quad (4-20)$$

is large, the diffusion at such voxels is characterized as one-fiber diffusion. Figure 4-3 D) presents an intensity-coded image of R_2 in a brain slice through the external capsule, an important structure of the human white matter. In Figure 4-3D) those voxels of a high intensity (bright regions on the image) are characterized as one-fiber diffusion.

(3). For each uncharacterized voxel after the above two steps, search the directions (θ, ϕ) , where $d(\theta, \phi)$ attains its local maxima. Note, $d(\theta, \phi)$ is antipodal symmetric, i.e., $d(\theta, \phi) = d(\pi - \theta, \phi + \pi)$, we mod out this symmetry when count the number of local maxima. Then we compute the weights for the local maxima (say we have 3 local maxima):

$$W_i := \frac{d(\theta_i, \phi_i) - d_{min}}{\sum_{i=1}^3 d(\theta_i, \phi_i) - 3d_{min}},$$

where (θ_i, ϕ_i) ($i = 1, 2, 3$) are the directions in which d attains its local maxima. If one of the weights is significant, it is considered as one fiber diffusion. If two weights are similar but much larger than the third one, it is viewed as two-fiber diffusion, if all three weights are similar, d can be considered either three-fiber diffusion or isotropic diffusion. In our experiment we restrict ourselves to the cases where we only distinguish isotropic, one-fiber or two-fiber diffusions. Under this restriction if three weights are similar, we include this voxel in the class of isotropic diffusion. Figure 5-5A) shows our classification of isotropic diffusion (dark region), one-fiber diffusion (gray region), and two-fiber diffusion (bright region) in the same slice as in Figure 4-3.

4.4 Numerical Implementation Issues

To efficiently solve the Euler-Lagrange equations (4-16), we use Additive Operator Splitting(AOS) algorithm for the diffusion operator (see [86, 107]). By using this algorithm, the computational and storage cost is linear in the number of voxels, and the computational efficiency can be increased by a factor of 10 under realistic accuracy requirements([86]). The algorithm is ready to be modified to a parallel version.

To avoid the complicated notation, we use X to represent any $A_{l,m}$ in the Euler-Lagrange equations, and write the algorithm for only one of the equations (4-16) in the system, since each equation has the same structure as others.

We use semi-implicit finite difference scheme:

$$\begin{aligned} \frac{X_{i,j}^{(n+1)} - X_{i,j}^{(n)}}{\tau} &= f(X_{i,j}^{(n)}) + \lambda \operatorname{div} \left(\frac{\nabla X_{i,j}^{(n+1)}}{M^{q_{ij}} |\nabla X_{i,j}^{(n)}|^{2-q_{ij}}} \right) \\ &= f(X_{i,j}^{(n)}) + \frac{-\lambda \ln M}{M^{q_{ij}}} \frac{\nabla q_{ij} \cdot \nabla X_{i,j}^{(n+1)}}{|\nabla X_{i,j}^{(n)}|^{2-q_{ij}}} + \frac{\lambda}{M^{q_{ij}}} \operatorname{div} \left(\frac{\nabla X_{i,j}^{(n+1)}}{|\nabla X_{i,j}^{(n)}|^{2-q_{ij}}} \right) \end{aligned} \quad (4-21)$$

Here X can be replaced by one of $A_{l,m}$'s with $l = 0, 2, 4$, $m = -l \cdots l$, and f is a function of results from last iteration, namely, f is a function of all $A_{l,m}^{(n)}$'s. $q(x) = p(x)$ if $|\nabla X| \leq M$, and $q(x) = 1$ if $|\nabla X| > M$ for some fixed constant M , which was chosen based on initial value of X , so M might be different for different $A_{l,m}$'s.

For simplicity of formulas, we define:

$$\begin{aligned} \Delta_-^x X_{i,j} &= X_{i,j} - X_{i-1,j}, \quad \Delta_+^x X_{i,j} = X_{i+1,j} - X_{i,j}, \quad \Delta^x X_{i,j} = X_{i+1,j} - X_{i-1,j} \\ \Delta_+^y X_{i,j} &= X_{i,j+1} - X_{i,j}, \quad \Delta_-^y X_{i,j} = X_{i,j} - X_{i,j-1}, \quad \Delta^y X_{i,j} = X_{i,j+1} - X_{i,j-1} \end{aligned}$$

Adopting a discretization of the divergence operator from [16], one can write (4-21)

as:

$$\begin{aligned} \frac{X_{i,j}^{(n+1)} - X_{i,j}^{(n)}}{\tau} &= f(X_{i,j}^{(n)}) - \frac{\lambda \ln M}{M^{q_{ij}}} \frac{[\Delta^x q_{ij}, \Delta^y q_{ij}]}{2h} \cdot \frac{[\Delta^x X_{i,j}^{(n+1)}, \Delta^y X_{i,j}^{(n+1)}] / (2h)}{\left(\frac{(\Delta^x X_{i,j}^{(n)})^2}{(2h)^2} + \frac{(\Delta^y X_{i,j}^{(n)})^2}{(2h)^2} \right)^{\frac{2-q_{ij}}{2}}} + \frac{\lambda}{M^{q_{ij}} h^2} \cdot \\ &\quad \left[\Delta_-^x \left(\frac{\Delta_+^x X_{i,j}^{(n+1)}}{\left(\frac{(\Delta_+^x X_{i,j}^{(n)})^2}{h^2} + \frac{(\Delta^y X_{i,j}^{(n)})^2}{(2h)^2} \right)^{\frac{2-q_{ij}}{2}}} \right) + \Delta_-^y \left(\frac{\Delta_+^y X_{i,j}^{(n+1)}}{\left(\frac{(\Delta_+^y X_{i,j}^{(n)})^2}{h^2} + \frac{(\Delta^x X_{i,j}^{(n)})^2}{(2h)^2} \right)^{\frac{2-q_{ij}}{2}}} \right) \right] \\ &= f(X_{i,j}^{(n)}) + (C_{i,j} - G_{i,j}) X_{i-1,j}^{(n+1)} - (C_{i,j} + D_{i,j}) X_{i,j}^{(n+1)} + (D_{i,j} + G_{i,j}) X_{i+1,j}^{(n+1)} + \\ &\quad + (E_{i,j} - H_{i,j}) X_{i,j-1}^{(n+1)} - (E_{i,j} + F_{i,j}) X_{i,j}^{(n+1)} + (F_{i,j} + H_{i,j}) X_{i,j+1}^{(n+1)} \end{aligned} \quad (4-22)$$

Where C, D, E and F are from divergence operation, while G and H are generated by dot product, in detail:

$$\begin{aligned}
C_{i,j} &= \frac{\lambda}{M^{q_{ij}} h^2} \left[\frac{(X_{i,j}^{(n)} - X_{i-1,j}^{(n)})^2}{h^2} + \frac{(X_{i-1,j+1}^{(n)} - X_{i-1,j-1}^{(n)})^2}{(2h)^2} \right]^{\frac{q_{i-1,j}-2}{2}} \\
D_{i,j} &= \frac{\lambda}{M^{q_{ij}} h^2} \left[\frac{(X_{i+1,j}^{(n)} - X_{i,j}^{(n)})^2}{h^2} + \frac{(X_{i,j+1}^{(n)} - X_{i,j-1}^{(n)})^2}{(2h)^2} \right]^{\frac{q_{ij}-2}{2}} \\
E_{i,j} &= \frac{\lambda}{M^{q_{ij}} h^2} \left[\frac{(X_{i+1,j-1}^{(n)} - X_{i-1,j-1}^{(n)})^2}{(2h)^2} + \frac{(X_{i,j}^{(n)} - X_{i,j-1}^{(n)})^2}{(h)^2} \right]^{\frac{q_{i,j-1}-2}{2}} \\
F_{i,j} &= \frac{\lambda}{M^{q_{ij}} h^2} \left[\frac{(X_{i+1,j}^{(n)} - X_{i-1,j}^{(n)})^2}{(2h)^2} + \frac{(X_{i,j+1}^{(n)} - X_{i,j}^{(n)})^2}{(h)^2} \right]^{\frac{q_{ij}-2}{2}} \\
G_{i,j} &= -\frac{\lambda \ln M(q_{i+1,j} - q_{i-1,j})}{M^{q_{ij}} (2h)^2} \left[\frac{(X_{i+1,j}^{(n)} - X_{i-1,j}^{(n)})^2}{(2h)^2} + \frac{(X_{i,j+1}^{(n)} - X_{i,j-1}^{(n)})^2}{(2h)^2} \right]^{\frac{q_{ij}-2}{2}} \\
H_{i,j} &= -\frac{\lambda \ln M(q_{i,j+1} - q_{i,j-1})}{(2h)^2} \left[\frac{(X_{i+1,j}^{(n)} - X_{i-1,j}^{(n)})^2}{(2h)^2} + \frac{(X_{i,j+1}^{(n)} - X_{i,j-1}^{(n)})^2}{(2h)^2} \right]^{\frac{q_{ij}-2}{2}}
\end{aligned}$$

Solving (4-22) would involve matrix inverse operation, which would become more and more complicated and dramatically expensive as dimension increases if we solve it directly. Instead, here we use Additive Operator Splitting (AOS) algorithm, which allows us to reformat system (4-22) into the 2 following systems:

$$\begin{aligned}
&\frac{\bar{X}_{i,j}^{(n+1)} - X_{i,j}^{(n)}}{\tau} = f(X_{i,j}^{(n)}) + \\
&2 \left[(C_{i,j} - G_{i,j}) \bar{X}_{i-1,j}^{(n+1)} - (C_{i,j} + D_{i,j}) \bar{X}_{i,j}^{(n+1)} + (D_{i,j} + G_{i,j}) \bar{X}_{i+1,j}^{(n+1)} \right] \quad (4-23)
\end{aligned}$$

$$\begin{aligned}
&\frac{\bar{\bar{X}}_{i,j}^{(n+1)} - X_{i,j}^{(n)}}{\tau} = f(X_{i,j}^{(n)}) + \\
&2 \left[(E_{i,j} - H_{i,j}) \bar{\bar{X}}_{i,j-1}^{(n+1)} - (E_{i,j} + F_{i,j}) \bar{\bar{X}}_{i,j}^{(n+1)} + (F_{i,j} + H_{i,j}) \bar{\bar{X}}_{i,j+1}^{(n+1)} \right] \quad (4-24)
\end{aligned}$$

and

$$X_{i,j}^{(n+1)} = \frac{\bar{X}_{i,j}^{(n+1)} + \bar{\bar{X}}_{i,j}^{(n+1)}}{2}$$

To accommodate the boundary condition $\frac{\partial X}{\partial n} = 0$ for the $M \times N$ matrix X , one needs to have:

$$X_{1,j}^{(n+1)} = X_{2,j}^{(n+1)}, X_{M-1,j}^{(n+1)} = X_{M,j}^{(n+1)}$$

$$X_{i,1}^{(n+1)} = X_{i,2}^{(n+1)}, X_{i,N-1}^{(n+1)} = X_{i,N}^{(n+1)}$$

Then (4-23) and (4-24) correspond to linear systems in matrix-vector notation:

$$A_1 \bar{X}^{(n+1)} = \tau f(\bar{X}^{(n)}) + \bar{X}^{(n)}$$

$$A_2 \bar{\bar{X}}^{(n+1)} = \tau f(\bar{\bar{X}}^{(n)}) + \bar{\bar{X}}^{(n)}$$

where \bar{X} and $\bar{\bar{X}}$ are $(M-2)(N-2) \times 1$ vectors formed by columns and transpose of rows of the original matrix X respectively, both A_1 and A_2 are $(M-2)(N-2) \times (M-2)(N-2)$ matrices, specifically, A_1 is a tri-diagonal matrix that repeats a $(M-2) \times (M-2)$ tri-diagonal matrix $(N-2)^2$ times diagonally, and A_2 is a tri-diagonal matrix that repeats a $(N-2) \times (N-2)$ tri-diagonal matrix $(M-2)^2$ times. They are defined as:

$$A_1 = I - 2\tau.$$

$$\begin{bmatrix} -D_{2,2} - G_{2,2} & D_{2,2} + G_{2,2} & 0 & \cdot & \cdot & \cdot \\ C_{3,2} - G_{3,2} & -C_{3,2} - D_{3,2} & D_{3,2} + G_{3,2} & 0 & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & D_{M-2,N-1} + G_{M-2,N-1} & \cdot \\ \cdot & \cdot & 0 & C_{M-1,N-1} - G_{M-1,N-1} & -C_{M-1,N-1} + G_{M-1,N-1} & \cdot \end{bmatrix}$$

$$A_2 = I - 2\tau.$$

$$\begin{bmatrix} -F_{2,2} - H_{2,2} & F_{2,2} + H_{2,2} & 0 & \cdot & \cdot & \cdot \\ E_{2,3} - H_{2,3} & -E_{2,3} - F_{2,3} & F_{2,3} + H_{2,3} & 0 & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & F_{M-1,N-2} + H_{M-1,N-2} & \cdot \\ \cdot & \cdot & 0 & E_{M-1,N-1} - H_{M-1,N-1} & -E_{M-1,N-1} + H_{M-1,N-1} & \cdot \end{bmatrix}$$

$$\bar{X} = [X_{2,2} \ X_{3,2} \cdots X_{M-1,2} \ X_{2,3} \cdots X_{M-2,N-1} \ X_{M-1,N-1}]^T$$

$$\bar{\bar{X}} = [X_{2,2} \ X_{2,3} \cdots X_{2,N-1} \ X_{3,2} \cdots X_{M-1,N-2} \ X_{M-1,N-1}]^T$$

Since both A_1 and A_2 are tri-diagonal matrices, one can get their inverses efficiently by using Thomas Algorithm([108]).

4.5 Validation and Application to Diffusion Weighted Images(DWI)

In this section we present our experimental results on the application of the proposed model (4-12)-(4-13) to simulated data and a set of HARD MRI data from the human brain.

4.5.1 Analysis of simulated data

The aim of our experiment on the simulated data is to test whether our model can efficiently reconstruct a regularized ADC profile from the noisy HARD measurements. We simulated an ADC profile on a 2D lattice of size 8×4 . The volume consists of two homogeneous regions, values of S_0 and all the $A_{l,m}$'s were shown in table 4-1.

Table 4-1: List of S_0 and $A_{l,m}$'s for two regions

Region	1	2
S_0	414	547
$A_{0,0}$	5.21×10^{-3}	1.43×10^{-2}
$A_{2,0}$	-1.17×10^{-3}	-2.68×10^{-3}
$ReA_{2,1}$	-4.37×10^{-5}	0
$ReA_{2,2}$	1.43×10^{-3}	0
$ImA_{2,1}$	3.64×10^{-5}	0
$ImA_{2,2}$	3.28×10^{-5}	0
$A_{4,0}$	-3.15×10^{-5}	8.4×10^{-6}
$ReA_{4,1}$	-1.56×10^{-4}	0
$ReA_{4,2}$	1.02×10^{-4}	0
$ReA_{4,3}$	6.30×10^{-5}	0
$ReA_{4,4}$	-8.54×10^{-5}	-1.73×10^{-3}
$ImA_{4,1}$	-8.01×10^{-5}	0
$ImA_{4,2}$	$0.9961.55 \times 10^{-4}$	0
$ImA_{4,3}$	1.41×10^{-5}	0
$ImA_{4,4}$	3.63×10^{-5}	0

In Figure 4-1 we displayed the true, noisy, and recovered ADC profiles $d(x, \theta, \phi)$ for the synthetic data with size 8×4 . The ADC profile $d(x, \theta, \phi)$ was computed by (4-6) based on these simulated data, and the corresponding $s_{true}(x, \theta, \phi)$ was constructed via (1-11) with $b = 1000s/mm^2$. Then the noisy HARD MRI signal $s(x, \theta, \phi)$ was generated by adding a zero mean Gaussian noise with standard deviation $\sigma = 0.15$. Figure 4-1B) shows the

ADC profile d computed by (4-6), where the coefficients of the SHS are the least-squares solutions of (4-3) with noisy s .

We then applied our model (4-12)-(4-13) to the noisy $s(x, \theta, \phi)$ to test the effectiveness of the model, with $\lambda_{0,0} = 4$, $\lambda_{2,m} = 40(m = -2 \dots 2)$, $\lambda_{4,m} = 60(m = -4 \dots 4)$. By solving the system of equations (4-16) in 2.5 seconds on computer with PIV 2.8GHZ CPU and 2G RAM using Matlab script code, we obtained 15 reconstructed functions as in (4-5). Using these $A_{l,m}$ (the solutions of (4-16)) we computed $d(x, \theta, \phi)$ via (4-6). The reconstructed $d(x, \theta, \phi)$ is shown in Figure 4-1C). Comparing these three figures, it is clear that the noisy measurements s have changed Figure 4-1A), the original shapes of d , into Figure 4-1B). After applying our model (4-12)-(4-13) to reconstruct the ADC profiles, the shapes of d in Figure 4-1A) were recovered, as shown in Figure 4-1C). These simulated results demonstrate that our model is effective in simultaneously regularizing and recovering ADC profiles.

4.5.2 Analysis of human MRI data

The second test is to reconstruct and characterize ADC profiles $d(x, \theta, \phi)$ from human HARD MRI data.

The raw DWI data, usually contains a certain level of noise, were obtained on a GE 3.0 Tesla scanner using a single shot spin-echo EPI sequence. The scanning parameters for the DWI acquisition are: repetition time (TR)=1000ms, echo time (TE) =85ms, the field of view (FOV)=220 mm x 220 mm. 24 axial sections covering the entire brain with the slice thickness=3.8 mm and the intersection gap=1.2 mm. The diffusion-sensitizing gradient encoding is applied in fifty-five directions (selected for the HARD MRI acquisition) with $b = 1000s/mm^2$. Thus, a total of fifty-six diffusion-weighted images, with a matrix size of 256 x 256, were obtained for each slice section. We applied model (4-12) to these data to compute the ADC profiles in the entire brain volume. By solving a system of equations (4-16) we obtained all the coefficients $A_{l,m}$'s in (4-5), and determined $d(x, \theta, \phi)$ using (4-6).

Then, we used these $A_{l,m}(x)$ to calculate R_0 and R_2 defined in (4-19) and (4-20) respectively, as well as the variance $\sigma(x)$ of $d(x, \theta, \phi)$ about its mean: $\sigma(x) = \int_0^\pi \int_0^{2\pi} (d(x, \theta, \phi) - \sum_{i=1}^{55} d(\mathbf{x}, \theta_i, \phi_i)/55)^2 d\theta d\phi$. Based on results from the HARD MRI data of this particular patient, we characterized the diffusion anisotropy according to the following procedure. If $R_0(x) > 0.856$, or $\sigma(\mathbf{x}) < 19.65$ the diffusion at x is classified as isotropic. For the remaining voxels if $R_2(x) > 0.75$, the diffusion at such voxels is considered as one-fiber diffusion. For uncharacterized voxels from these two steps we further classified them by the principles stated in the section 4.3. The selection of the thresholds mentioned above for R_0 , R_2 and σ involves experts' input and large sample experiments. Experimental results definitely depend on these thresholds, but not sensitively.

Figure 4-2 presents $A_{2,0}(x)$, one of the coefficients in (4-6), for the particular slice in the volume. The images $A_{2,0}(x)$ in Figure 4-2A) and 4-2B) are estimated by using (4-2) and solving (4-12), respectively.

Figure 4-3 Compares FA and three $R_2(x)$'s with $A_{l,m}(\mathbf{x})$'s obtained from three different models for the same slice as shown in Figure 4-2. Figure 4-3A) displays the FA image obtained by using advanced system software from GE. The $A_{l,m}(\mathbf{x})$'s used to obtain $R_2(\mathbf{x})$ in Figure 4-3B) are directly computed from (4-2). Those used to obtain $R_2(\mathbf{x})$ in Figures. 4-3C) and 4-3D) are the least-squares solutions of (4-3) and the solutions of (4-12), respectively. In Figures. 4-3C) and 4-3D) the voxels with high levels of intensities (red, yellow, yellow-light blue) are characterized as one-fiber diffusion.

Although the FA image in Figure 4-3A) is obtained based on a conventional DTI model (1-7), it is still comparable with the R_2 map, since single tensor diffusion characterized by SHS representation from the HARD images agrees with that characterized by the DTI model. However, in DTI a voxel with a low intensity of FA indicates isotropic diffusion, while using our algorithm, multi-fibers diffusion may occur at the location with the low value of R_2 .

The ability to characterize anisotropic diffusion is enhanced by this algorithm, as shown in Figures. 4-3A)-4-3D). Figure 4-3B) indicates again that the estimates of $A_{l,m}$ directly from the \log signals usually are not good. Even the least-squares solution of (4-3) is not always effective. This can be seen by comparing the anatomic region inside the red square of Figures. 4-3C) and 4-3D), which are zoomed in Figures. 4-4A) and 4-4B), respectively. There is a dark broken line showing on the map of the external capsule (arrow to the right on Figure 4-4A), this same region was recovered by the proposed model and characterized by the third step in our algorithm as two-fiber anisotropic diffusion (arrow to the right in Figure 4-4B). (The model solutions reduced the value of R_0 , increased the values of R_1 slightly, and made the 3rd step in our characterization to be applied). Our results also show the connection in a cortical associative tract (arrow to the left in Figures. 4-4B), however, this connection was not mapped out on Figure 4-3C) or the zoomed image in Figure 4-4A). In fact this connection was not mapped out on Figures 4-3A)-B) either. All these mapped connections are consistent with known neuroanatomy. Combined together, our results indicate that our proposed model for joint recovery and smoothing of the ADC profiles has an advantage over existing models for enhancing the ability to characterize diffusion anisotropy.

Figure 5-5A) shows a partition of isotropic, one-fiber, and two-fiber diffusion for the same slice used in Figure 4-4. The two-fiber, one-fiber, and isotropic diffusion regions were further characterized by the white, gray, and black regions, respectively. The region inside the white square in Figure 5-5A), which is the same one squared in Figures. 4-3C) and 4-3D), is zoomed in Figure 4-4C). It is seen that the two arrayed voxels in Figure 4-4B) are classified as two-fiber diffusion. The characterization of the anisotropy on the voxels and their neighborhoods is consistent with the known fiber anatomy.

Figure 5-5 B) represents the shapes of $d(x, \theta, \phi)$ at three particular voxels (upper, middle and lower rows). The d in all three voxels is computed using (4-6). However, the $A_{l,m}(\mathbf{x})$ used in computing d on the left column are the least-squares solutions of (4-3),

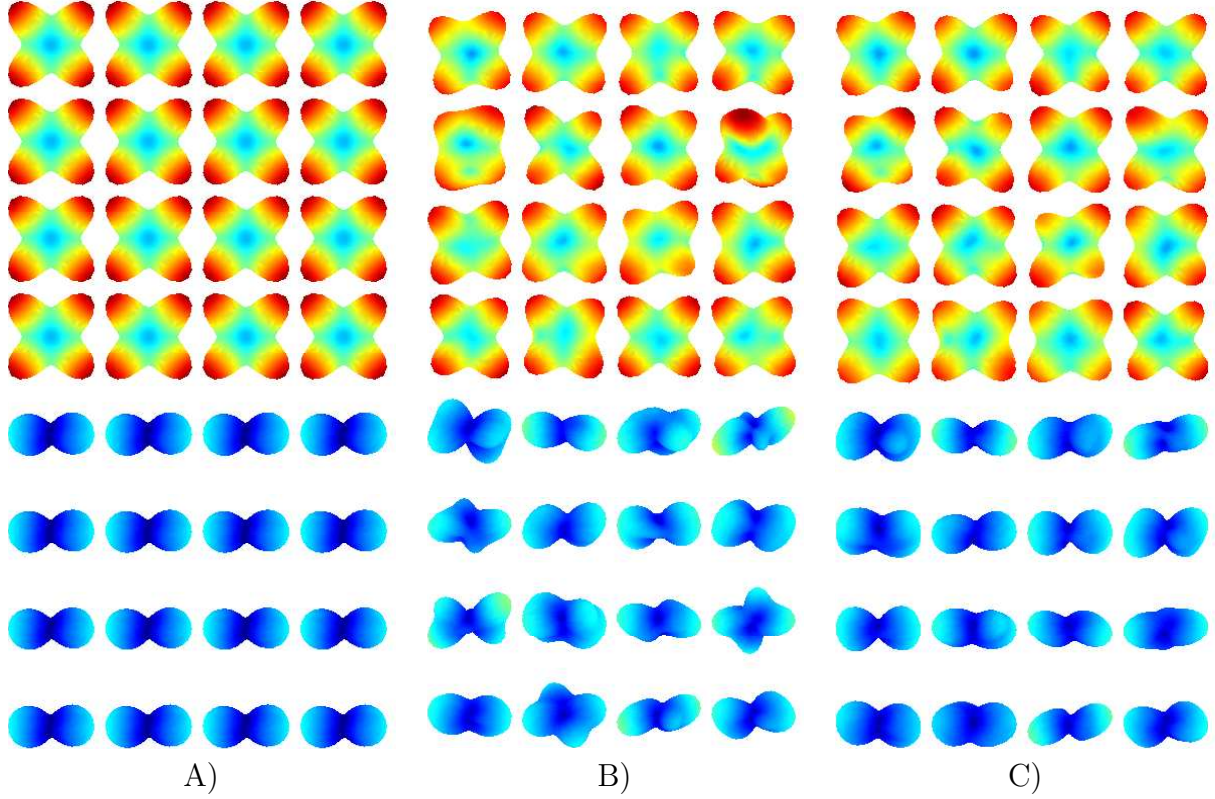


Figure 4-1. Comparing shapes of d . A) True d . B) The d generated by (4-6), with $A_{l,m}$'s the least square solution of (4-10) with the noisy measurement s . C) Recovered d by applying model (4-12).

while in the right column they are the solutions of the proposed model (4-12). The first and second rows show two voxels that can be characterized as isotropic diffusion before denoising, but as two-fiber diffusion after applying model (4-12). These two voxels are the same voxels as in Figure 4-4 directed by arrows. The lower row of Figure 5-5 B) shows the one-fiber diffusion was enhanced after applying our model.

Solving $A_{l,m}$'s of size $15 \times 109 \times 86 \times 8$ from 4-D data of size $55 \times 109 \times 86 \times 8$ takes 46.2 seconds for each iteration on computer with PIV 2.8GHZ CPU and 2G RAM in Matlab script code.

4.6 An Existence Theorem for the Model

In this section we will discuss the existence of a solution to our minimization problem (4-12) using the idea developed in [104].

Recall that for a function $u \in BV(\Omega)$,

$$Du = \nabla u \cdot \mathcal{L}^n + D^s u$$

is a Radon measure, where ∇u is the density of the absolutely continuous part of Du with respect to the n -dimensional Lebesgue measure \mathcal{L}^n , and $D^s u$ is the singular part. To minimize (4-12) over the functions in $BV(\Omega)$, we first need to give a precise definition for E_1 .

Definition: For $A_{l,m} \in BV(\Omega)$, define

$$\int_{\Omega} \phi_{l,m}(x, DA_{l,m}) := \int_{\Omega} \phi_{l,m}(x, \nabla A_{l,m}) dx + \int_{\Omega} |D^s A_{l,m}|$$

where $\phi_{l,m}$ is defined as in (4-8), and $\int_{\Omega} |D^s A_{l,m}|$ is the total variation norm of $A_{l,m}$.

Then, our energy functional (4-12) is defined as

$$\begin{aligned} E(A_{l,m}) &= \lambda \int_{\Omega} \sum_{l=0,2,4} \sum_{m=-l}^l \phi_{l,m}(x, \nabla A_{l,m}) + \lambda \int_{\Omega} |D^s A_{l,m}| \\ &+ \frac{1}{2} \int_{\Omega} \int_0^{2\pi} \int_0^{\pi} |s(x, \theta, \phi) - s_0(x) e^{-bd(x, \theta, \phi)}|^2 \sin \theta d\theta d\phi dx. \end{aligned} \quad (4-25)$$

In the discussion of existence, without loss of generality, we set the parameter $\lambda = 1$ in (4-12) and threshold $M_{l,m} = 1$ in (4-8) to reduce the complexity in the formulation.

Next we will show lower semi-continuity of the energy functional (4-12) in L^1 , i.e. if for each l, m ($l = 0, 2, 4$ and $m = -l, \dots, l$), as $k \rightarrow \infty$,

$$A_{l,m}^k \rightarrow A_{l,m}^0 \quad \text{in } L^1(\Omega),$$

then

$$E(A_{l,m}^0) \leq \liminf_{k \rightarrow \infty} E(A_{l,m}^k), \quad (4-26)$$

To prove this we need the following lemma:

Lemma: Let

$$\phi(x, r) := \begin{cases} \frac{1}{p(x)} |r|^{p(x)}, & |r| \leq 1 \\ |r| - (1 - \frac{1}{p(x)}), & |r| > 1 \end{cases} \quad (4-27)$$

For $u \in BV(\Omega)$ denote

$$\Phi(u) := \int_{\Omega} \phi(x, Du),$$

and

$$\tilde{\Phi}(u) := \sup_{\substack{\psi \in C_0^1(\Omega, R^n) \\ |\psi| \leq 1}} \int_{\Omega} -u \operatorname{div} \psi - \frac{p(x) - 1}{p(x)} |\psi|^{\frac{p(x)}{p(x)-1}} dx.$$

Then,

$$\Phi(u) = \tilde{\Phi}(u) \quad (4-28)$$

Furthermore, Φ is lower semi-continuous on $L^1(\Omega)$, i.e. if $u_j, u \in BV(\Omega)$ satisfy $u_j \rightarrow u$ weakly in $L^1(\Omega)$ as $j \rightarrow \infty$ then

$$\Phi(u) \leq \liminf_{j \rightarrow \infty} \Phi(u_j).$$

Proof First note that for each $\psi \in C_0^1(\Omega, R^n)$, the map

$$u \mapsto \int_{\Omega} -u \operatorname{div} \psi - \frac{p(x) - 1}{p(x)} |\psi|^{\frac{p(x)}{p(x)-1}} dx$$

is continuous and affine on $L^1(\Omega)$. Therefore, $\tilde{\Phi}$ is convex and lower semi-continuous on $L^1(\Omega)$ and the domain of $\tilde{\Phi}$, $\{u \mid \tilde{\Phi}(u) < \infty\}$, is precisely $BV(\Omega)$.

Next we show (4-28). For $u \in BV(\Omega)$, we have that for each $\psi \in C_0^1(\Omega, R^n)$,

$$- \int_{\Omega} u \operatorname{div} \psi dx = \int_{\Omega} \nabla u \cdot \psi dx + \int_{\Omega} D^s u \cdot \psi$$

and so

$$\tilde{\Phi}(u) = \sup_{\substack{\psi \in C_0^1(\Omega, R^n) \\ |\psi| \leq 1}} \int_{\Omega} \nabla u \cdot \psi - \frac{p(x) - 1}{p(x)} |\psi|^{\frac{p(x)}{p(x)-1}} dx + \int_{\Omega} D^s u \cdot \psi$$

Since the measures dx and $D^s u$ are mutually singular, by a standard argument we can have

$$\tilde{\Phi}(u) = \sup_{\substack{\psi \in C_0^1(\Omega, \mathbb{R}^n) \\ |\psi| \leq 1}} \int_{\Omega} (\nabla u \cdot \psi - \frac{p(x) - 1}{p(x)} |\psi|^{\frac{p(x)}{p(x)-1}}) dx + \int_{\Omega} |D^s u|.$$

To prove (4-28) it only remains to show that

$$\int_{\Omega} \phi(x, \nabla u) dx = \sup_{\substack{\psi \in C_0^1(\Omega, \mathbb{R}^n) \\ |\psi| \leq 1}} \int_{\Omega} (\nabla u \cdot \psi - \frac{p(x) - 1}{p(x)} |\psi|^{\frac{p(x)}{p(x)-1}}) dx. \quad (4-29)$$

Since any $\rho \in L^\infty(\Omega, \mathbb{R}^n)$ can be approximated in measure by $\psi \in C_0^1(\Omega, \mathbb{R}^n)$, we have that

$$\sup_{\substack{\psi \in C_0^1(\Omega, \mathbb{R}^n) \\ |\psi| \leq 1}} \int_{\Omega} \nabla u \cdot \psi - \frac{p(x) - 1}{p(x)} |\psi|^{\frac{p(x)}{p(x)-1}} dx = \sup_{\substack{\rho \in L^\infty(\Omega, \mathbb{R}^n) \\ |\rho| \leq 1}} \int_{\Omega} \nabla u \cdot \rho - \frac{p(x) - 1}{p(x)} |\rho|^{\frac{p(x)}{p(x)-1}} dx. \quad (4-30)$$

Choosing $\rho(x) = 1_{\{|\nabla u| \leq 1\}} |\nabla u|^{p(x)-1} \frac{\nabla u}{|\nabla u|} + 1_{\{|\nabla u| > 1\}} \frac{\nabla u}{|\nabla u|}$, where χ_E is the indicator function on E , we see that the right hand side of (4-30) is

$$\geq \int_{\Omega} \frac{1}{p(x)} |\nabla u|^{p(x)} 1_{\{|\nabla u| \leq 1\}} + \left[|\nabla u| - \frac{p(x) - 1}{p(x)} \right] 1_{\{|\nabla u| > 1\}} dx = \int_{\Omega} \phi(x, \nabla u) dx. \quad (4-31)$$

To show the opposite inequality, we argue as follows. For any $\rho \in L^\infty(\Omega, \mathbb{R}^n)$, since $p(x) > 1$ we have that for almost all x ,

$$\nabla u(x) \cdot \rho(x) \leq \frac{1}{p(x)} |\nabla u|^{p(x)} + \frac{p(x) - 1}{p(x)} |\rho(x)|^{\frac{p(x)}{p(x)-1}}$$

In particular, if $|\nabla u| \leq 1$,

$$\nabla u(x) \cdot \rho(x) - \frac{p(x) - 1}{p(x)} |\rho(x)|^{\frac{p(x)}{p(x)-1}} \leq \frac{1}{p(x)} |\nabla u|^{p(x)} \quad (4-32)$$

If $|\nabla u| > 1$, noticing $p(x) > 1$ and $|\rho| \leq 1$ for almost all x we have that

$$\nabla u \cdot \rho = |\nabla u| \frac{\nabla u}{|\nabla u|} \cdot \rho \leq |\nabla u| \left[\frac{1}{p(x)} + \frac{p(x) - 1}{p(x)} |\rho|^{\frac{p(x)}{p(x)-1}} \right]$$

and so

$$\nabla u \cdot \rho - \frac{p(x) - 1}{p(x)} |\rho|^{\frac{p(x)}{p(x)-1}} \leq \frac{1}{p(x)} |\nabla u| + (|\nabla u| - 1) \frac{p(x) - 1}{p(x)} |\rho|^{\frac{p(x)}{p(x)-1}} \leq |\nabla u| - \frac{p(x) - 1}{p(x)} \quad (4-33)$$

Combining, (4-30), (4-31), (4-32), and (4-33), we have (4-29), and hence for all $u \in BV(\Omega)$, $\tilde{\Phi}(u) = \Phi(u)$.

Note that $\phi(x, r) = \phi_{l,m}(x, r)$ if $p(x) = p_{l,m}(x)$. A direct consequence of this lemma is that we have that E_1 in (4-7) is weakly lower semi-continuous in L^1 topology on $BV(\Omega)$ norm.

Furthermore, we can show that E_2 in (4-7) is lower semi-continuous on $L^1(\Omega)$.

Indeed, when

$$A_{l,m}^k \rightarrow A_{l,m}^0, \quad \text{in } L^1(\Omega), \quad \text{as } k \rightarrow \infty,$$

for all $l = 0, 2, 4$ and $m = -l, \dots, l$, $A_{l,m}^k \rightarrow A_{l,m}^0$ a.e. on Ω . Then, if $s(x, \theta, \phi) \in L^2(\Omega \times S^2)$ and $s_0(x) \in L^2(\Omega)$, by the Dominated Convergence Theorem we have

$$E_2(A_{l,m}^k) \rightarrow E_2(A_{l,m}^0).$$

Therefore, $E = E_1 + E_2$ is lower semi-continuous in L^1 topology on $BV(\Omega)$ norm, and (4-26) holds.

Now we can prove our existence results.

Theorem: Let Ω be a bounded open set of R^n . Assume that $s(x, \theta, \phi) \in L^2(\Omega \times S^2)$ and $s_0(x) \in L^2(\Omega)$. Then, there exists a solution consisting of functions $A_{l,m}^0$ ($l = 0, 2, 4$ and $m = -l, \dots, l$) to the minimization problem (4-12) over the space of $BV(\Omega)$.

Proof: Let $A_{l,m}^k$ ($l = 0, 2, 4$ and $m = -l, \dots, l$) be the minimizing sequences of (4-12) in $BV(\Omega)$. Then for each (l, m) the sequence $A_{l,m}^k$ is bounded in $BV(\Omega)$. From the compactness of $BV(\Omega)$ there exist subsequences of $A_{l,m}^k$ (still denoted by $A_{l,m}^k$) and functions $A_{l,m}^0 \in BV(\Omega)$ satisfying

$$A_{l,m}^k \rightarrow A_{l,m}^0 \quad \text{strongly in } L^1(\Omega).$$

By the lower semi-continuity of the energy functional on $L^1(\Omega)$ (see (4–26)), we have

$$E(A_{l,m}^0) \leq \liminf_{k \rightarrow \infty} E(A_{l,m}^k) \leq \inf_{A_{l,m} \in BV(\Omega)} E(A_{l,m}).$$

Hence, all these $A_{l,m}^0$ ($l = 0, 2, 4$ and $m = -l, \dots, l$) together form a solution to the minimization problem (4–12).

4.7 Conclusion

A novel variational framework was introduced for simultaneous smoothing and estimation of ADC profiles in the form of truncated SHS based on HARD MRI. Features of this model included minimizing a nonstandard growth function with nonlinear data fitting. Moreover, the constraints on the positivity and antipodal symmetry properties of d was also accommodated in the model. We also demonstrated our algorithm for using the variance of d from its mean and the coefficients of its truncated SHS approximation to characterize diffusion anisotropy.

Our experiments on both synthetic data and human HARD MRI data showed the effectiveness of the proposed model in the estimation of ADC profiles and the enhancement of the characterization of diffusion anisotropy. The characterization of non-Gaussian diffusion from the proposed method was consistent with known neuroanatomy.

The choice of the current parameters, however, may affect the results. Our choice was made based on the principle that classification for one-fiber diffusion from the model solution should agree with a priori knowledge of the fiber connections.

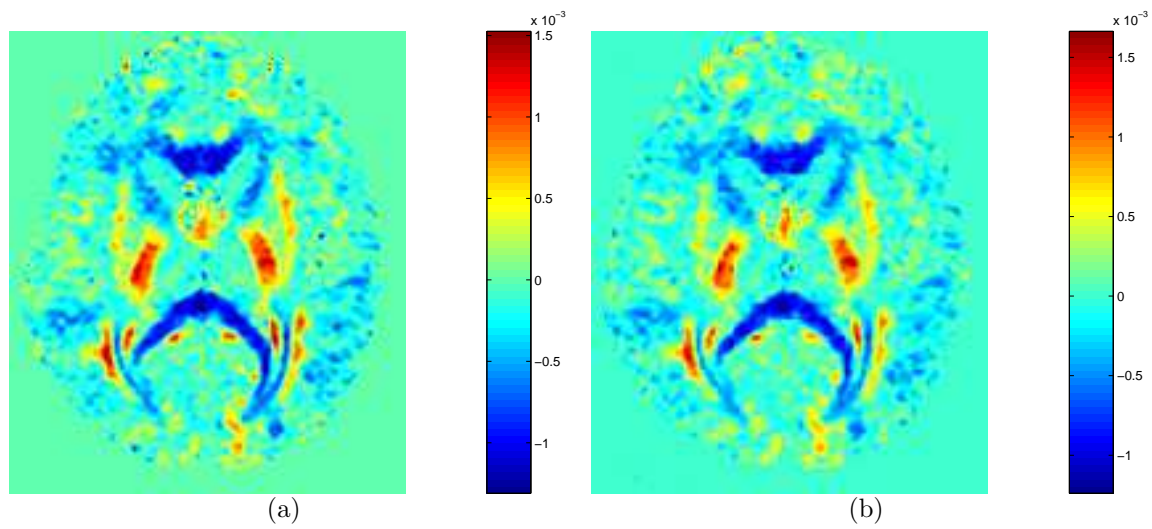


Figure 4-2. Comparison of A_{20} . A) A_{20} computed from (4-2). B) A_{20} obtained from model(4-12).

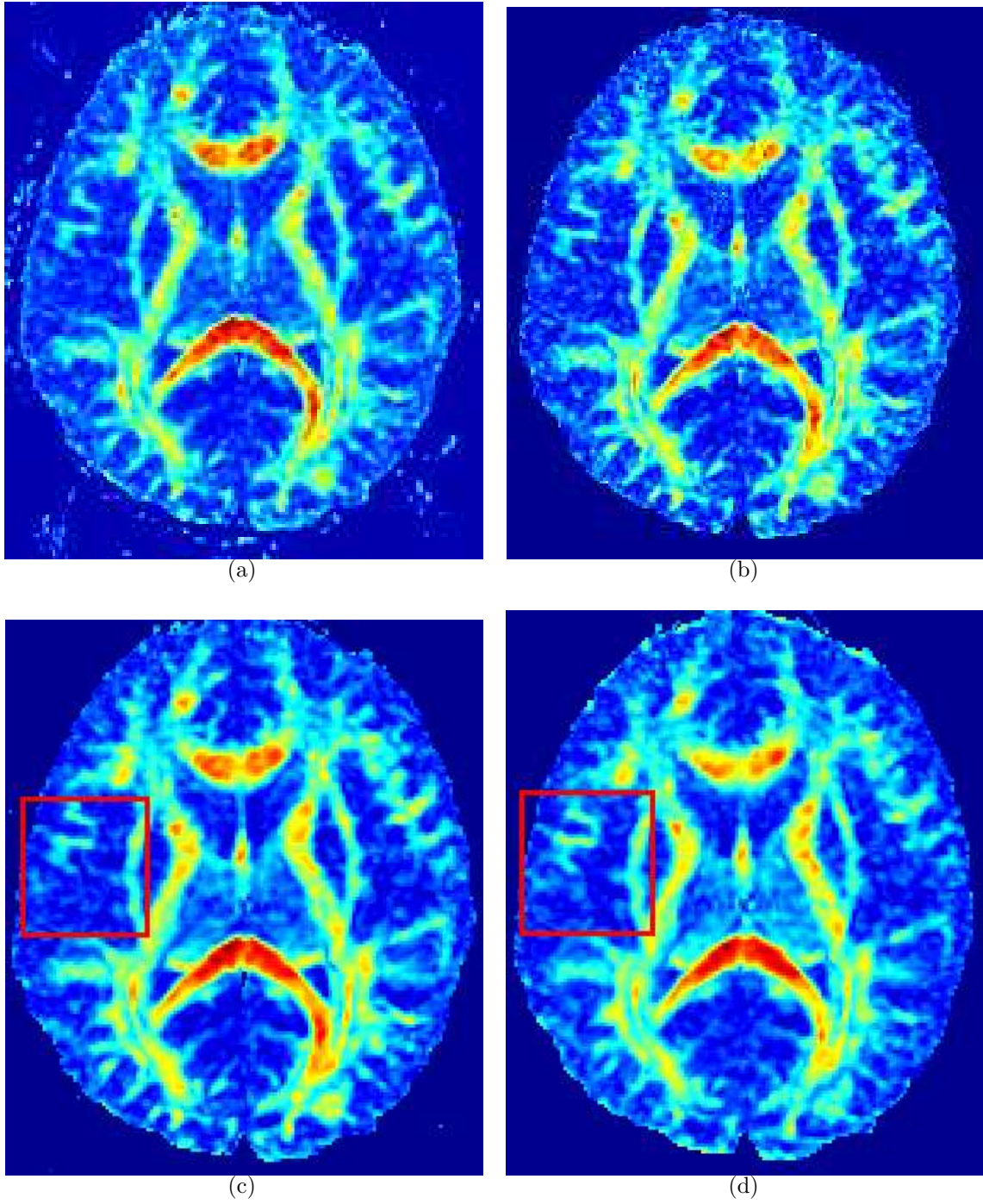


Figure 4-3. Images of FA and R_2 . A) FA from GE software. B)-D) R_2 with the $A_{l,m}$'s as the solutions of (4-2), least-squares solutions of (4-3), and model solutions, respectively.

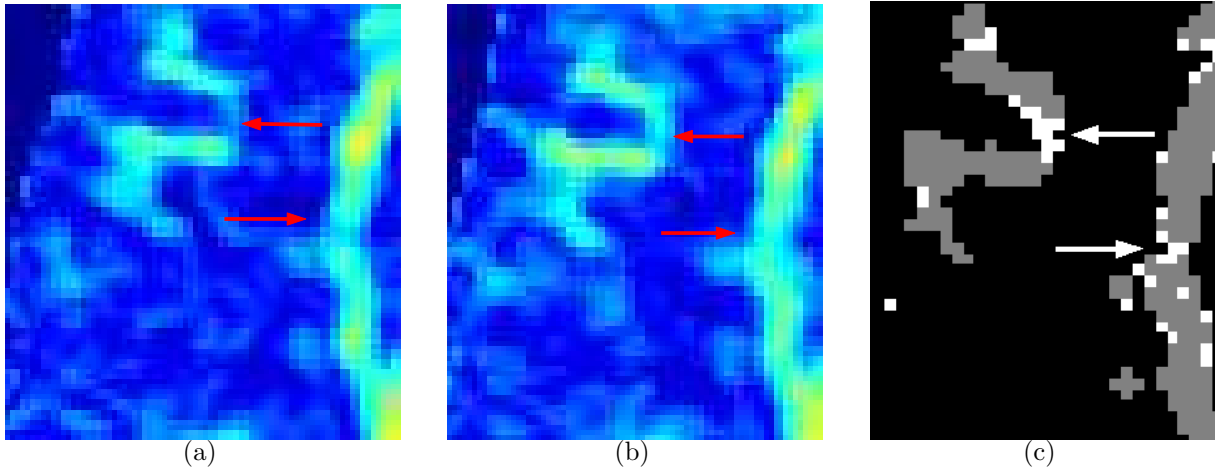


Figure 4-4. Zoomed FA and A_{20} . A)-B) Enlarged portions inside the red squares in Figures. 4-3 C) and 4-3D), respectively. C) Enlarged portions inside the white squares in Figure 5-5A).

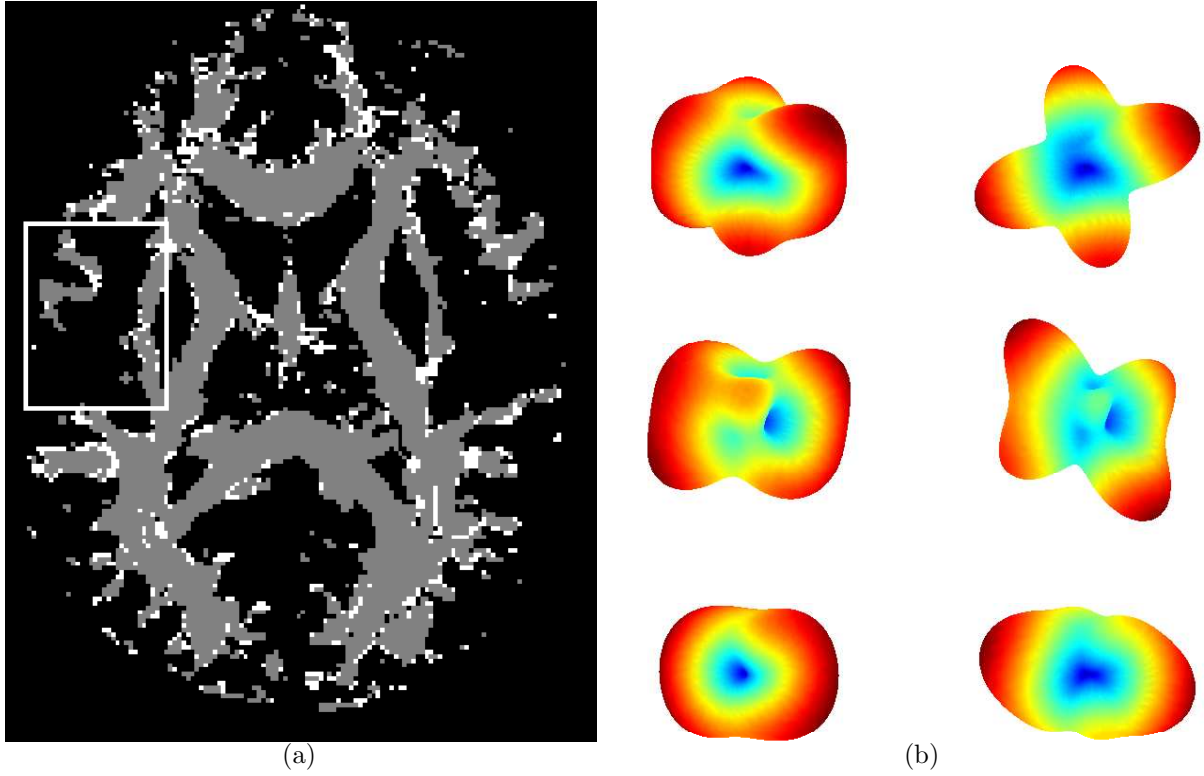


Figure 4-5. Classification of voxels based on d . A) Classification: white, gray, and black voxels are identified as two-fiber, one-fiber, and isotropic diffusion respectively. B) Shapes of $d(x, \theta, \phi)$ at three particular points (upper, middle and lower rows). The d is computed via (4-6). $A_{l,m}(\mathbf{x})$ used in (4-6) in the left columns are the least-squares solutions of (4-3), while in the right column are the solutions from our model.

CHAPTER 5

ESTIMATION, SMOOTHING AND CHARACTERIZATION OF APPARENT DIFFUSION COEFFICIENT (A SECOND APPROACH)

Copyright [2005] Lecture Notes on Computer Science [109]. Portions reprinted, with permission.

5.1 Introduction

To use SHS model (4-1) to approximate d in (1-11), and hence to detect two-fiber diffusion, at least 15 diffusion weighted measurements $s(\mathbf{q})$ over 15 carefully selected directions are required. However, to use the mixture model (1-9) with $n=2$ to detect two-fiber diffusion only 13 unknown functions: f , 6 entries of each of D_1, D_2 need to be solved. This motivated us to study what is the minimum number of the diffusion weighted measurements required for detecting diffusion with no more than two fiber orientations within a voxel, and what is the corresponding model to approximate the ADC profiles in this case. In this chapter we propose to approximate the ADC profiles from HARD MRI by the product of two second-order SHS's instead of a fourth-order SHS. We also show that the product of two second-order spherical SHS's describes only the diffusion with at most two fiber orientations, while the fourth-order SHS may describe the diffusion with three fiber orientations.

Moreover, in this chapter we will introduce an information measurement developed in [110], and termed as CRE (see definition (5-7)) to characterize the diffusion anisotropy. CRE differs from Shannon entropy in the aspect that Shannon entropy depends only on the probability of the event, while CRE depends also on the magnitude of the change of the random variable. We observed that isotropic diffusion has either no local minimum or many local minima with very small variation in the denoised $s(\mathbf{q})/s_0$, i.e., e^{-bd} profiles, in comparing with one fiber or two-fiber diffusions, which implies the corresponding CRE to be small. We also found that one fiber diffusion has only one local minimum with larger variation in the $s(\mathbf{q})/s_0$ profiles, which leads to larger CRE. Therefore, we propose

to properly threshold the CRE for the regularized $s(\mathbf{q})/s_0$ profiles to characterize the diffusion anisotropy.

5.2 New Approximation Model for ADC Profiles

In [66, 72, 111] to detect the diffusion with at most two fiber orientation the ADC profiles were represented by a truncated SHS up to order 4 in the form of (4-1). In [66] the coefficients $A_{l,m}$'s (l is even) were determined by inverse spherical harmonic transform of $-\frac{1}{b}\log\frac{s(\mathbf{q})}{s_0}$ and in [72] they were estimated as the least-squares solutions of

$$-\frac{1}{b}\log\frac{s(\mathbf{q})}{s_0} = \sum_{l=0}^{l_{max}} \sum_{m=-l}^l A_{l,m} Y_{l,m}(\theta, \phi). \quad (5-1)$$

Regularization on the raw data or $A_{l,m}$ wasn't considered in these two work. In [111] $A_{l,m}$'s were considered as a function of \mathbf{x} , and estimated and smoothed simultaneously by solving the following constrained minimization problem:

$$\begin{aligned} \min_{A_{l,m}(\mathbf{x}), \tilde{s}_0(\mathbf{x})} \int_{\Omega} \{ \sum_{l=0,2,4} \sum_{m=-l}^l |\nabla A_{l,m}(\mathbf{x})|^{p_{l,m}(\mathbf{x})} + |\nabla \tilde{s}_0(\mathbf{x})|^{p(\mathbf{x})} \} d\mathbf{x} \\ + \frac{\lambda}{2} \int_{\Omega} \{ \int_0^{2\pi} \int_0^{\pi} |s(\mathbf{x}, \mathbf{q}) - \tilde{s}_0(\mathbf{x}) e^{-bd(\mathbf{x}, \theta, \phi)}|^2 \sin\theta d\theta d\phi + |\tilde{s}_0 - s_0|^2 \} d\mathbf{x}, \end{aligned} \quad (5-2)$$

with the constraint $d > 0$. In this model $p_{l,m}(\mathbf{x}) = 1 + \frac{1}{1+k|\nabla G_{\sigma} * A_{l,m}|^2}$, $q(\mathbf{x}) = 1 + \frac{1}{1+k|\nabla G_{\sigma} * s_0|^2}$, and d takes the form (4-1). By the choice of $p_{l,m}$ and q , the regularization is total variation based near edges, isotropic in homogeneous regions, and between isotropic and total variation based depending on the local properties of the image at other locations. In this work since the ADC profile was approximated by (4-1), at least 15 measurements of $s(q)$ were required to estimate the 15 coefficients $A_{l,m}$.

However, the mixture model (1-9) with $n = 2$, which is also able to detect two-fiber diffusion, involves only 13 unknown functions. This motivates us to find a model that is able to detect non-Gaussian diffusion with the minimum number of unknowns. In this paper we only discuss the diffusion with no more than two fiber orientations within a voxel. The significance of this study is clear: a smaller number of unknowns lead to a

smaller number of required HARD measurements. This will significantly reduce the scan time and thus is important in clinical applications.

Our basic idea is to approximate the ADC profiles by the product of two second order SHS's instead of a SHS up to order four. This can be formulated as

$$d(\mathbf{x}, \theta, \phi) = \left(\sum_{l=0,2} \sum_{m=-l}^l b_{l,m}(\mathbf{x}) Y_{l,m}(\theta, \phi) \right) \cdot \left(\sum_{l=0,2} \sum_{m=-l}^l c_{l,m}(\mathbf{x}) Y_{l,m}(\theta, \phi) \right). \quad (5-3)$$

In this model there are only 12 unknowns: $b_{l,m}$, $c_{l,m}$ ($l = 0, 2$ and $-l \leq m \leq l$).

To estimate the ADC profile from the raw HARD MRI data, which usually contains a certain level of noise, we propose a simultaneous smoothing and estimation model similar to (5-2) for solving $b_{l,m}$, $c_{l,m}$, that is the following constrained minimization problem:

$$\begin{aligned} \min_{b_{l,m}(\mathbf{x}), c_{l,m}(\mathbf{x}), \tilde{s}_0(\mathbf{x})} & \int_{\Omega} \left\{ \sum_{l=0,2} \sum_{m=-l}^l \alpha (|\nabla b_{l,m}(\mathbf{x})| + |\nabla c_{l,m}(\mathbf{x})| + \beta |\nabla \tilde{s}_0(\mathbf{x})| \right\} d\mathbf{x} \\ & + \frac{1}{2} \int_{\Omega} \left\{ \int_0^{2\pi} \int_0^{\pi} |s(\mathbf{x}, \mathbf{q}) - \tilde{s}_0(\mathbf{x}) e^{-bd(\mathbf{x}, \theta, \phi)}|^2 \sin \theta d\theta d\phi + |\tilde{s}_0 - s_0|^2 \right\} d\mathbf{x}, \end{aligned} \quad (5-4)$$

with constraint $d \geq 0$, where d is in the form of (5-3). α, β are constants. The first 3 terms are the regularization terms for $b_{l,m}$, $c_{l,m}$ and s_0 respectively. The last two terms are the data fidelity terms based on the original Stejskal-Tanner equation(1-11).

Next, feasibility of this model will be explained. Let SA denote the space of even SHS of order 4, i.e., $SA = \{d : d(\theta, \phi) = A = \sum_{l=0,2,4} \sum_{m=-l}^l A_{l,m} Y_{l,m}\}$, let SBC be the space of products of two even SHS of order 2, i.e., $SBC = \{d : d(\theta, \phi) = B \cdot C = \sum_{l=0,2} \sum_{m=-l}^l b_{l,m} \cdot Y_{l,m}(\theta, \phi) \cdot \sum_{l=0,2} \sum_{m=-l}^l c_{l,m} Y_{l,m}(\theta, \phi)\}$. Since the product of two SHS of order 2 and 2 can be expressed as a linear combination of spherical harmonics of order less or equal to $2 + 2 = 4$, $SBC \subset SA$. But a simple dimension count, $\dim(SA)=15$, while dimension of SBC is less than or equal to 12. Then SBC is a proper subset of SA . So functions in SBC are less general than functions in SA . However, numerous experiments show that when a voxel is not more complicated than two-fiber diffusion, its ADC is well-approximated by a function in SBC . This is not true if 3-fiber or more complicated

diffusion is involved (see Figure 5-1). Therefore, if we focus only on characterizing at most two-fiber diffusion, which is the most common case, model (5-3) is reasonable and sufficient to represent ADC.

Model (5-4) is a minimization problem with constraint $d(\theta, \phi) \geq 0$ for all $0 \leq \theta < \pi, 0 \leq \phi < 2\pi$, which is usually difficult to implement. To improve the efficiency of computation we used the idea that any second order SHS $\sum_{l=0,2} \sum_{m=-l}^l b_{l,m} Y_{l,m}(\theta, \phi)$ is equivalent to a tensor model $u^T D u$ for some semi-positive definite 3×3 matrix D , where $u(\theta, \phi) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. This means that the coefficients $b_{l,m}$ ($l = 0, 2, m = -l, \dots, l$) in SHS and the entries $D(i, j)$, ($i, j = 1, \dots, 3$) in D can be computed from each other explicitly. Here are two examples: $b_{00} = \frac{2}{3}\sqrt{\pi}(D(1, 1) + D(2, 2) + D(3, 3))$. $D(1, 1) = -\frac{\sqrt{5}b_{20}-2b_{00}-\sqrt{30}Re(b_{22})}{4\sqrt{\pi}}$, where $Re(b_{22})$ is the real part of b_{22} . Hence, we could let $B = uD_1u^T, C = uD_2u^T$, and $d = (uD_1u^T)(uD_2u^T)$. Then for $i = 1, 2$ we can write $D_i = L_i L_i^T$, with L_i a lower triangular matrix to guarantee semi-positiveness of D_i . The ADC is finally approximated by

$$d(\mathbf{x}, \theta, \phi) = [u(\theta, \phi)L_1(\mathbf{x})L_1(\mathbf{x})^T u(\theta, \phi)^T][(u(\theta, \phi)L_2(\mathbf{x})L_2(\mathbf{x})^T u(\theta, \phi)^T)]. \quad (5-5)$$

Model (5-4) is then replaced by

$$\begin{aligned} \min_{L_1^{jk}(\mathbf{x}), L_2^{jk}(\mathbf{x}), \tilde{s}_0(\mathbf{x})} & \int_{\Omega} (\alpha \sum_{i=1}^2 \sum_{j=1}^3 \sum_{k=1}^j |\nabla L_i^{jk}| + \beta |\nabla \tilde{s}_0|) d\mathbf{x} \\ & \frac{1}{2} \int_{\Omega} \left\{ \int_0^{2\pi} \int_0^{\pi} |s - \tilde{s}_0 e^{-bd}|^2 \sin\theta d\theta d\phi + |\tilde{s}_0 - s_0|^2 \right\} d\mathbf{x}, \end{aligned} \quad (5-6)$$

where $d = (uL_1L_1^T u^T)(uL_2L_2^T u^T)$. All the $b_{l,m}, c_{l,m}, l = 0, 2, m = -l \dots l$ are smooth functions of $L_i^{jk}, i = 1, 2; j = 1, 2, 3; k \leq j$. Smoothness of L_i^{jk} guarantees that of $b_{l,m}$'s, $c_{l,m}$'s. The first term in model (5-6) thus works equivalently to the way the first two terms in model (5-4) do, while all the other terms are the same as those remaining in (5-4). Hence, (5-6) is equivalent to (5-4), but it is a non-constrained minimization problem and

is thus easy to implement. After we get L_1 and L_2 , $b_{l,m}$ and $c_{l,m}$ in (5-4) can be obtained by the one to one relation between them.

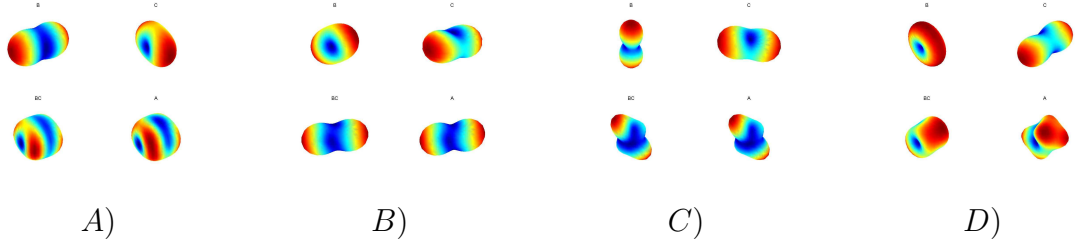


Figure 5-1. Comparison of the ADC's approximated by (4-1) and (5-3) in four cases: A) isotropic diffusion; B) one-fiber diffusion; C) two-fiber diffusion; D) three-fiber diffusion. In A)-D) from left to right, top to bottom, we show shapes of B , C , $B \cdot C$, and A , respectively.

We applied model(5-6) to a set of human brain HARD MRI data to reconstruct and characterize ADC profiles. The data set consisted of 55 diffusion weighted images $S_k : \Omega \rightarrow R, k = 1, \dots, 55$, and one image S_0 in the absence of a diffusion-sensitizing field gradient($b=0$ in (1-11)). 24 evenly spaced axial planes with 256×256 voxels in each slice are obtained using a 3T MRI scanner with single shot spin-echo EPI sequence. Slice thickness is $3.8mm$, gap between two consecutive slices is $1.2mm$, repetition time (TR) = $1000ms$, echo time (TE) = $85ms$ and $b = 1000s/mm^2$. The field of view (FOV) = $220mm \times 220mm$. We first applied model(5-6) to the data to get L_i , and then used L_i to compute $b_{l,m}$ and $c_{l,m}$, $l = 0, 2, m = -l \dots l$, and the ADC $d = B \cdot C$. For purpose of comparison, we also used the model (5-2) to estimate $A_{l,m}$ and get A . The comparison for the shapes of ADC in the form of $B \cdot C$ and A is demonstrated in Figure5-1A)-D) at four specific voxels. The diffusion at these 4 voxels are isotropic A), one-fiber B), two-fiber C), and three-fiber D), respectively. In each sub figure, the up left, up right, down left, down right ones are the shapes of B , C , $B \cdot C$ and A , respectively. It is evident that if the diffusion is isotropic, one-fiber or two-fiber, $B \cdot C$ and A are the same. However, if the diffusion is three-fiber, A can't be well approximated by $B \cdot C$.

To show the effectiveness of the proposed model in recovering ADC, in Figure 5-2A)-D) we compared images of R_2 (defined in section 5.3) with coefficient $A_{l,m}$ estimated by 4 different methods. The voxels with higher value of R_2 were considered as one-fiber diffusion. The $A_{l,m}$'s in A), B) and C) were estimated using least-squares method in [72], model (5-2), and model(5-6) with the diffusion-sensitizing gradient applied to 55 directions, respectively. The $A_{l,m}$'s in D) are estimated by the same way as that in C), but from the HARD data with 12 carefully chosen directions. The model (5-6) applied on 55 measurements worked as good as the model (5-2) in getting higher value of R_2 . Both of them worked better than the least-squares method that does not consider regularization. Although the result from 12 measurements was not as good as that from 55 measurements, they are still comparable. We will show in Figure 5-5A) and B) that the anisotropy characterization results based on the ADC presented in C) and D) are also close. These experimental results indicated that by using the proposed model the voxels with two-fiber diffusion can be detected reasonably well from 12 HARD measurements in carefully selected directions.

5.3 Use of CRE to Characterize Anisotropy

As mentioned, FA is only able to detect Gaussian diffusion. For non-Gaussian diffusion, Frank and Alexander et al. used the order of significant components in SHS to characterize anisotropy. They considered voxels with significant 4th order components as two-fiber diffusion. In [111] Chen et al. realized that such a voxel could have isotropic or one-fiber diffusion. They defined $R_0 := \frac{|A_{0,0}|}{\sum_{l=0,2,4} \sum_{m=-l}^l |A_{l,m}|}$, $R_2 := \frac{\sum_{m=-2}^{m=2} |A_{2,m}|}{\sum_{l=0,2,4} \sum_{m=-l}^l |A_{l,m}|}$. Higher values of R_0 and R_2 are correspond to isotropic and one-fiber diffusion, respectively. For other points, the number of local maxima of ADC, together with the weights of the variances at the local maxima, were used to classify voxels as isotropic, one-fiber or two-fiber diffusion. This procedure is more precise, but there are many criteria involved and thus more thresholds needed to be set subjectively. In this section, we will introduce a simple scheme using only one measurement, CRE, and two thresholds.

CRE is a measure of uncertainty/information in a random variable. If X is a random variable in R , the CRE of X is defined by

$$CRE(X) = - \int_{R_+} P(X > \lambda) \log P(X > \lambda) d\lambda, \quad (5-7)$$

where $R_+ = \{X \in R | X \geq 0\}$.

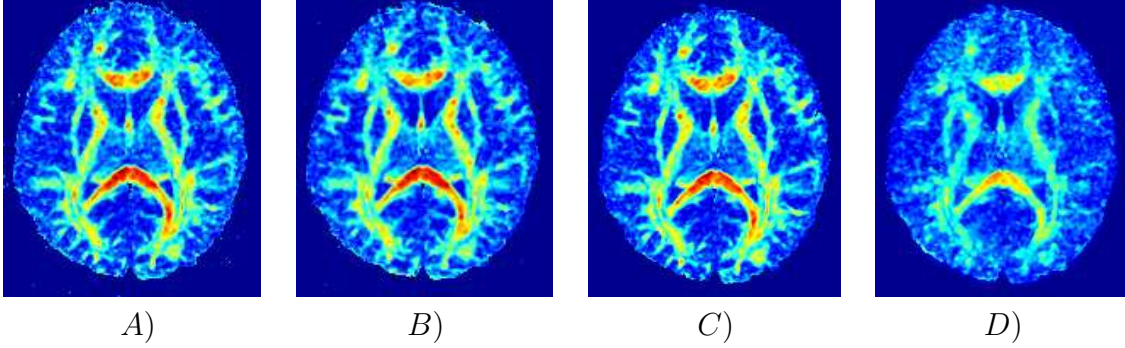


Figure 5-2. A)-D) are images of R_2 with $A_{l,m}$'s calculated using least-squares method, model (5-2), model (5-6) applied on 55 measurements, and model (5-6) applied on 12 measurements, respectively.

We use CRE of e^{-bd} rather than d to characterize diffusion anisotropy when d is recovered from HARD measurements through (5-6). The magnitude of ADC is usually on the order of 10^{-3} , while the magnitude of e^{-bd} is in the order of 10^{-1} , which is larger than that of ADC itself. Moreover, e^{-bd} is a smooth approximation of the data s/s_0 .

The weak convergence property of CRE proved in [111] makes empirical CRE computation based on the samples converges in the limit to the true CRE. This is not the case for the Shannon entropy. We define the empirical CRE of e^{-bd} as

$$CRE(e^{-bd(\theta,\phi)}) = - \sum_{i=2}^M P(e^{-bd(\theta,\phi)} > \lambda_i) \log P(e^{-bd(\theta,\phi)} > \lambda_i) \Delta \lambda_i \quad (5-8)$$

where $\{\lambda_1 < \lambda_2 < \dots < \lambda_M\}$ is range of e^{-bd} at voxel \mathbf{x} . $\Delta \lambda_i = \lambda_i - \lambda_{i-1}$ is the absolute difference between two adjacent values of e^{-bd} ; note this term is not shown in Shannon entropy. In most of the cases, the variation of e^{-bd} is the largest for one-fiber diffusion voxels, smaller for two-fiber diffusion and smallest for isotropic voxels. This also explains

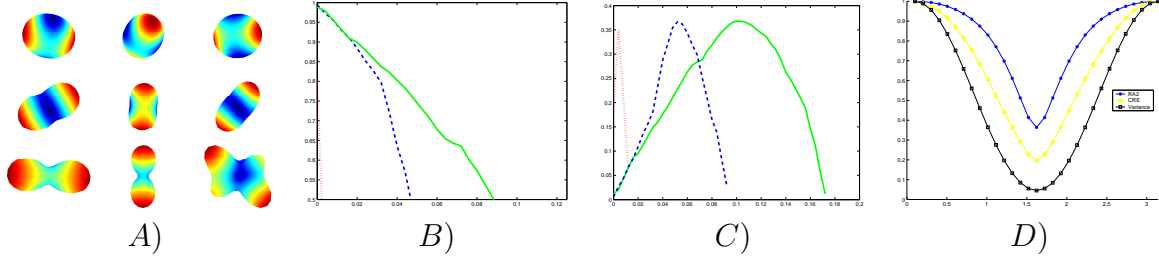


Figure 5-3. An example for shapes of ADC for isotropic, one-fiber and two-fiber voxels. A) Shapes of ADC at an isotropic (first row), one-fiber (second row) and two-fiber (last row). B)-C) Graphs of $F(\lambda)$, $-F(\lambda)\log F(\lambda)$ at three particular voxels: isotropic (red), one-fiber (green), two-fiber (blue). D) R_2 (blue), CRE (yellow), variance (black) as functions of rotation angle ψ used in constructing synthetic data.

why CRE is the largest for one-fiber, medium for two-fiber and smallest for isotropic diffusion voxels. In our experiment, we choose $M=1000$ uniformly distributed directions (θ, ϕ) in (5–8).

Define the decreasing distribution function $F(\lambda) := P(e^{-bd} > \lambda)$. Figure 5-3B) shows the graphs of $F(\lambda)$ at three pre-classified voxels: isotropic (red), one-fiber (green), two-fiber (blue). It is observed that the support and magnitude of $F(\lambda)$ are largest at the voxel with one-fiber diffusion, and smallest at that with isotropic diffusion. Figure 5-3C) demonstrates the graphs of $-F(\lambda)\log F(\lambda)$ at the same three voxels. It is evident that the area under the green curve (one-fiber) is much larger than that under the blue curve (two-fiber), while the area under the red curve is the smallest. Since CRE is exactly the area under curve $-F(\lambda)\log F(\lambda)$, we can see that the measure $CRE(e^{-bd})$ is the largest at the voxels with one-fiber diffusion, medium with two-fiber diffusion, and smallest with isotropic diffusion. Thus measure $CRE(e^{-bd})$ could be used to distinguish isotropic, one-fiber and two-fiber diffusion with two thresholds T_1 and T_2 , with $T_1 < T_2$. Set up 3 intervals: $(0, T_1)$, (T_1, T_2) , (T_2, ∞) . Voxels with CRE that fall into the first, second, and third intervals are classified as isotropic, two-fiber and one-fiber diffusion respectively.

Figure 5-3D) on synthetic data and Figure 5-4 on human brain HARD MRI data further show the strengths of CRE over the three popularly used measures, R_2 , FA and

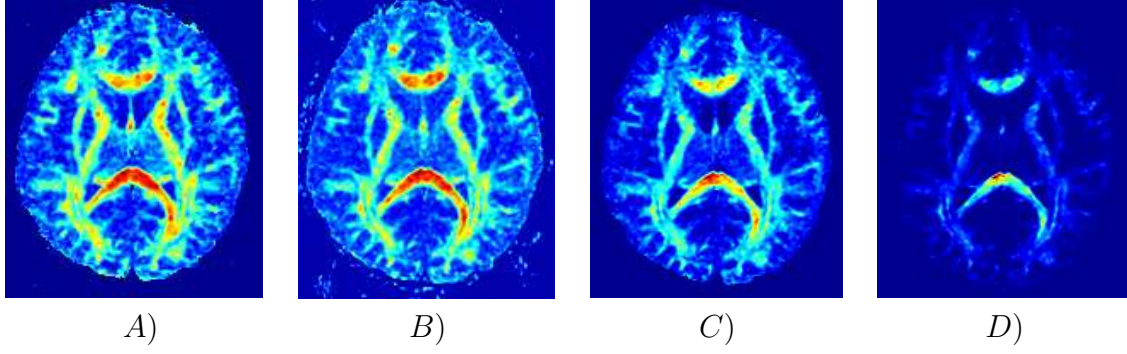


Figure 5-4. Images of four measures: A) R_2 ; B) FA; C) CRE of e^{-bd} ; D) Variance of e^{-bd} .

variance, in characterizing diffusion anisotropy. The human data are the same as that used in Figure 5-2. The synthetic data are constructed as follows: Set D_1 and D_2 to be two 3×3 diagonal matrices with diagonal elements 4×10^{-2} , 10^{-2} , 2×10^{-2} and 8×10^{-2} , 10^{-2} , 3×10^{-2} , respectively. Then fix D_1 but rotate principal eigenvector of D_2 about xy axis by angle ψ to get $D_2(\psi)$. Let $B(\theta, \phi) = u^T D_1 u$, $C_\psi(\theta, \phi) := u^T D_2(\psi) u$. We computed R_2 , FA and CRE, variance of $e^{-bB \cdot C}$ for various values of ψ and showed them in Figure 5-3D) in blue, yellow and black respectively. When ψ varies from 0 to $\pi/2$, $B \cdot C$ changes from a typical one fiber diffusion to a two fiber diffusion, and from $\pi/2$ to π $B \cdot C$ changes back to the same shape as $\psi = 0$. The graph of CRE shows the value of CRE decreases when $B \cdot C$ varies from one-fiber diffusion to two-fiber diffusion, and increases when $B \cdot C$ gradually changes from two-fiber diffusion backs to one-fiber diffusion.

R_2 cannot detect multi-fiber diffusion since it measures the significance of only the second order components in SHS. Nonsignificant difference between R_2 and FA is observed from the images in Figure 5-4A) and B). But CRE differs much from R_2 and FA. In Figure 5-3D) the graph CRE is much steeper than the others. In Figure 5-4, visually, contrast of CRE is much better than that of FA and R_2 . Furthermore, the smallness of magnitude of R_2 or FA is unable to distinguish between isotropic and two-fiber diffusion, while that of CRE does better job. Note, CRE is comparable to FA or R_2 in detecting isotropic and one-fiber diffusion.

Next we discuss from the theoretical point of view why CRE beats variance in characterizing diffusion anisotropy. Let X be a random variable, $Var(X)$ its variance. According to the proof in [112], $E(|X - E(X)|) \leq 2CRE(X)$. In our case X is e^{-bd} , whose magnitude is a multiple of $10^{-2} < 1$, so we have $Var(X) = E(|X - E(X)|^2) \leq E(|X - E(X)|) \leq 2CRE(X)$. Our experimental results show that magnitude of CRE is almost 10 times of that of $Var(X)$. Higher magnitude of CRE makes it less sensitive to rounding errors. Moreover, in Figure 5-4D), which representing the the variance of e^{-bd} , the Genu/Splenium of corpus callosum is so bright that regions besides it are not clearly visualized, so CRE is much better than variance visually.

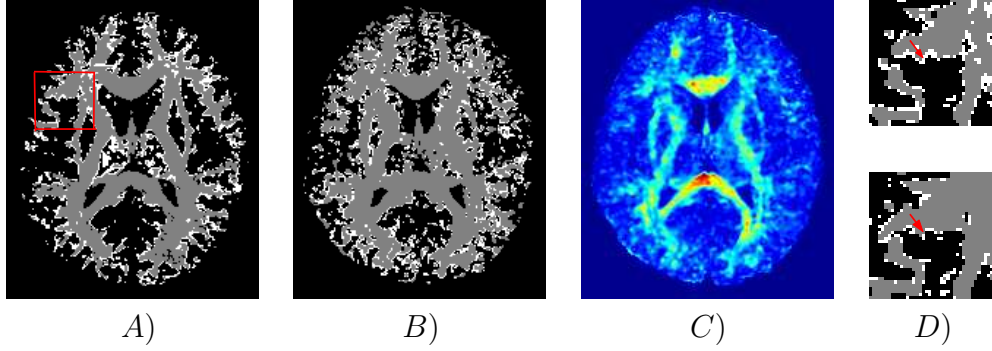


Figure 5-5. A)-B). Characterization: black, gray, and white regions represent the voxels with isotropic, one-fiber, and two-fiber diffusion, respectively. In A) we used 55 measurements, in B) we used 12 carefully selected measurements. C) Image of CRE calculated from 12 measurements. D) Characterization results of the region inside the red box in A) using CRE (top) and variance (bottom) based on 55 measurements. Red arrows point to a voxel that is wrongly characterized as one-fiber diffusion by using variance but correctly classified as two-fiber diffusion using CRE.

Figure 5-5A) shows a partition of a brain region into isotropic, one-fiber and two-fiber diffusion based on ADC calculated from 55 measurements. The black, gray, white voxels are identified as isotropic, one-fiber and two-fiber diffusion, respectively. The characterization is consistent with known fiber anatomy. Figure 5-5B) represents the characterization result based on the ADC estimated from 12 measurements. It is comparable with that from 55 measurements. CRE based on ADC estimated from

12 measurements (Figure5-5C)) is also comparable to that from 55 measurements (Figure5-4C)). Thus our characterization appears not to be sensitive to number of measurements as long as at least 12 measurements are used. Figure5-5D) illustrates a two-fiber diffusion voxel (indicated by red arrow) that is incorrectly characterized as one-fiber diffusion using variance (bottom image) but characterized as two-fiber correctly using CRE (top image). This further verifies the superiority of CRE over variance in characterizing diffusion anisotropy.

5.4 Conclusion

In this chapter, we presented a novel variational framework for simultaneous smoothing and estimation of ADC profiles depicted by two diffusion tensors. To our knowledge this was the first attempt to use the least amount of measurement to detect two-fiber diffusion from human brain HARD MRI data. We also demonstrated our algorithm for using CRE of e^{-bd} to characterize the diffusion anisotropy.

Our experiments on two sets of human brain HARD MRI data showed the effectiveness and robustness of the proposed model in the estimation of ADC profiles and the enhancement of the characterization of diffusion anisotropy. The characterization of diffusion from the proposed method was consistent with known neuroanatomy.

CHAPTER 6

RECONSTRUCTION OF INTRA-VOXEL STRUCTURE FROM DIFFUSION WEIGHTED IMAGES

Copyright [2006] IEEE [71]. Portions reprinted, with permission.

6.1 Determination of Fiber Directions

In our approach of determining fiber directions, the first step is to recover the ADC profiles d from noisy HARD data by using model (4-12) and (4-13). Then, from the SH representation of the recovered d we define

$$R_0(\mathbf{x}) = \frac{|A_{0,0}(\mathbf{x})|}{A(\mathbf{x})}, \quad R_2(\mathbf{x}) = \frac{\sum_{m=-2}^{m=2} |A_{2,m}(\mathbf{x})|}{A(\mathbf{x})},$$

where $A(\mathbf{x}) = \sum_{l=0,2,4} \sum_{m=-l}^l |A_{l,m}(\mathbf{x})|$. The voxels with significant R_0 and R_2 are identified as *strong* isotropic diffusion and one-fiber diffusion, respectively. The union of all these voxels is denoted by Ω_1 . On Ω_1 , f_1 in (1-9) should be close to 1.

Under the assumption of $p_t(\mathbf{r})$ being a mixture of two Gaussians, the diffusion is modelled by (1-9) with $n = 2$. The combination of (1-9) with $n = 2$ and (1-11) yields

$$e^{-bd(\mathbf{x},\theta,\phi)} = \sum_{i=1}^2 f_i e^{-b\mathbf{u}^T D_i(\mathbf{x}) \mathbf{u}}, \quad (6-1)$$

where $\mathbf{u}^T = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. For notation conciseness, denote f_1 by f . To estimate D_i and f in (1-10) we minimize the following function w.r.t L_1, L_2, f :

$$\begin{aligned} & \int_{\Omega} \left(\sum_{i=1}^2 |\nabla L_i|^{P_i(\mathbf{x})} + |\nabla f|^{p_f(\mathbf{x})} \right) d\mathbf{x} + \lambda_1 \int_{\Omega_1} (f - 1)^2 d\mathbf{x} \\ & + \lambda_2 \int_{\Omega} \int_0^{2\pi} \int_0^{\pi} \left| \sum_{i=1}^2 f_i e^{-b\mathbf{u}^T L_i L_i^T \mathbf{u}} - e^{-bd} \right|^2 \sin\theta d\theta d\phi d\mathbf{x}, \end{aligned} \quad (6-2)$$

with the constraints $L_i^{m,m} > 0$. In (6-2) for $i = 1, 2$, $\lambda_i > 0$ is a parameter, $p_i(\mathbf{x}) = 1 + \frac{1}{1+k|\nabla G_{\sigma} * \nabla L_i|^2}$, $p_f(\mathbf{x}) = 1 + \frac{1}{1+k|\nabla G_{\sigma} * \nabla f|^2}$, and L_i is a lower triangular matrix. D_i is recovered from L_i by $D_i = L_i L_i^T$ (i.e., $L_i L_i^T$ is the Cholesky factorization of D_i). Writing the cost function in terms of L_i rather than directly in terms of D_i

ensures that the positive-definite constraint on D_i is met (see [101]). Also we have

$$|\nabla L_i|^p = \sum_{1 \leq m, n \leq 3} |\nabla L_i^{m,n}|^p.$$

The first two terms in (6–2) are the regularization terms. By the choice of $p_i(\mathbf{x})$ (similarly for p_f), in the homogeneous region image gradients are close to zero and $p_i(\mathbf{x}) \approx 2$, so the smoothing is approximately isotropic. Along the edges, image gradient makes $p_i(\mathbf{x}) \approx 1$, so the smoothing is approximately total-variation-based and is almost only along the edges. At all other locations, the image gradient forces $1 < p < 2$, and the diffusion is between isotropic and total variation based, and varies depending on the local properties of the image. Therefore, the smoothing governed by this model well preserves relevant features in these images. The third term in (6–2) is forcing $f \approx 1$ on Ω_1 . The last term is the nonlinear data fidelity term based on (6–1).

The fiber orientations at each voxel are determined by the directions of the principal eigenvectors of D_1 and D_2 . For the voxels where f (or $1 - f$) is significantly large, we consider $1 - f$ (or f) as zero, and (1–10) with $n = 2$ reduces to the Gaussian diffusion model.

6.2 Validation and Application to HARD Weighted Images

We applied model (6–2) to a set of HARD MRI human data. The raw HARD MR images were obtained using a single shot spin-echo EPI sequence. The imaging parameters for the DW-MRI acquisition are repetition time (TR) = 1000ms, $echotime(TE)$ = 85ms. Diffusion-sensitizing gradient encoding is applied in fifty-five directions with $b = 1000s/mm^2$. Thus, a total of fifty-six DW images with the matrix size = 256 × 256 were obtained for each slice, and images through the entire brain are obtained by 24 slices. The slice is transversally oriented and the thickness is 3.8mm, and intersection gap between two contiguous slices is 1.2mm. The field of view (FOV) = 220mm × 220mm.

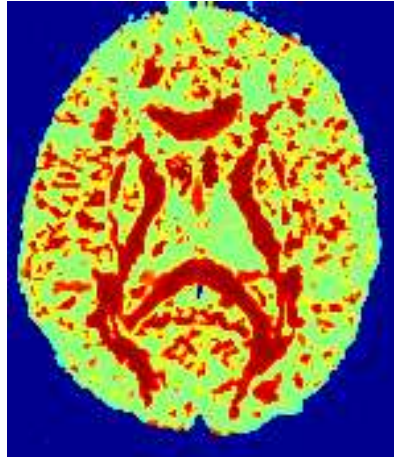
We will show our experimental results from a particular subject in a brain slice through the external capsule. In this experiment we first recovered the ADC profiles d using (4–12)-(4–13), and defined Ω_1 as the set of the voxels where $R_0 > 0.8416$ or

$R_2 > 0.1823$. These thresholds were selected using the histograms of R_0 and R_2 . Then, we solved the minimization problem (6–2) by the energy descent method. The information of $f \approx 1$ on Ω_1 was also incorporated into the selection of the initial f .

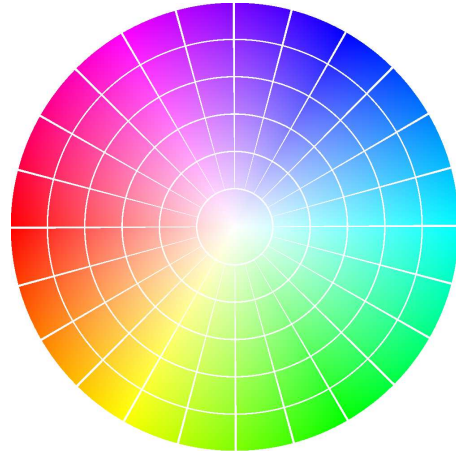
By solving (6–2) we obtained the solutions L_i and f , and consequently, $D_i = L_i L_i^T$ ($i = 1, 2$). Figure 6–1A) represents the model solution f . Function $f \approx 1$ on the dark red regions. The voxels in these regions are identified as isotropic or one-fiber diffusion. This is consistent with known neuroanatomy. Figures 6–1C) and 6–1D) show the color representation of the directions of the principal eigenvectors for $D_1(\mathbf{x})$ and $D_2(\mathbf{x})$, respectively. By comparing the color-coding in Figures 6–1C) and 6–1D) with the color pie shown in Figure 6–1B), the fiber directions are uniquely determined. The representation in Figure 6–1B) is implemented by relating the azimuthal angle (ϕ) of the vector to color hue (H) and the polar angle ($\theta \geq \pi/2$) to the color saturation (S). Slightly different from [113], we define $H = \phi/2\pi$, $S = 2(\pi - \theta)/\pi$, and Value $V = 1$. If the direction of the principal eigenvector is represented by (θ, ϕ) , the fiber orientation can be described by either (θ, ϕ) or $(\pi - \theta, \phi + \pi)$. To resolve this ambiguity, we choose the eigenvector to lie in the lower hemisphere, i.e. $\theta \geq \pi/2$. The upper hemisphere is just an antipodally symmetric copy of the lower one. The xy plane is the plane of discontinuity.

Figure 6–2 shows the shapes of $d(\mathbf{x}, \theta, \phi)$ together with the fiber directions at 4 particular voxels. The blue and red arrows indicate the orientations of the fibers determined from the principal eigenvectors of D_1 and D_2 respectively. The last shape corresponds to isotropic diffusion. Figs. 6–1 and Figure 6–2 indicate that our model (6–2) is effective in recovering the intra-voxel structure.

To examine the accuracy of the model in recovering fiber directions, we selected a region inside the corpus callosum where the diffusion is known to be of one-fiber. For each voxel in this region we computed the direction in which d is maximized. This direction vector field is shown in Figure 6–3A). On the other hand we solved (6–2) and obtained the model solution $f \approx 1$ on this region. The direction field generated from the principal



A)



B)



C)



D)

Figure 6-1. Fiber direction color map. A) Model solution f . B) color pie. C) Color-coding of the 1st fiber direction mapping. D) Color-coding of the 2nd fiber direction mapping.

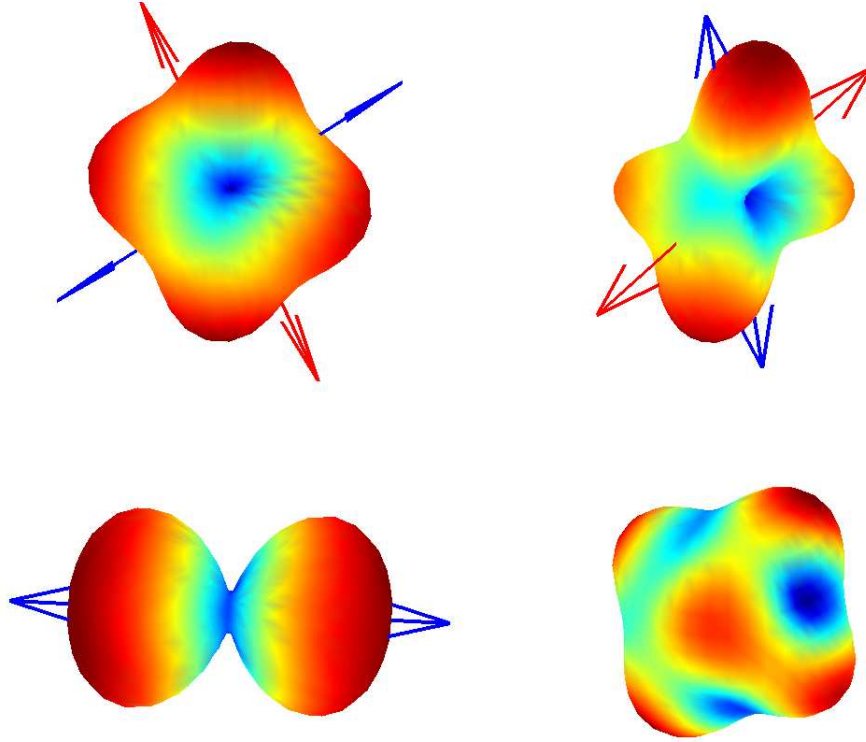


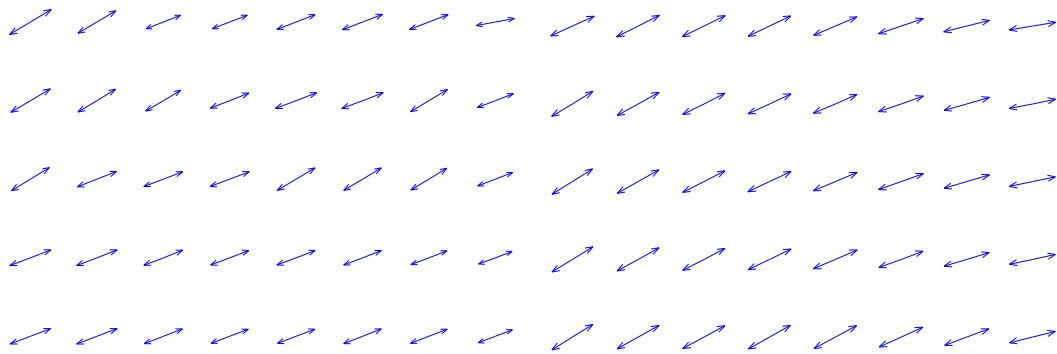
Figure 6-2. Shape of d , and the orientations of the principal eigenvectors of D_1 (blue) and D_2 (red) at 4 particular voxels

eigenvector of D_1 is shown in Figure 6-3B), in which the vector field is not only well preserved but also more regularized due to the regularization terms in the model.

6.3 Conclusion

A novel model for determination of fiber directions for biGaussian diffusion from HARD MRI was presented. In this model two tensor fields were recovered by simultaneous smoothing and data fitting, as well as incorporating the information on voxel classification of diffusion anisotropy. The numerical results indicated the effectiveness of the model in recovering intra-voxel structure.

The choice of the parameters in (6-2) and the determination of the region Ω_1 influence the results. Our choice was made based on the principle that the one-fiber direction from the model agrees with the direction in which d was maximized.



A)

B)

Figure 6-3. Fiber direction field. A) That obtained by maximizing d . B) The principal eigenvector of D (solution of (6-2))

CHAPTER 7

RECONSTRUCT OF WHITE MATTER FIBER TRACES USING MULTI-TENSOR DEFLECTION IN DWI

Copy right [2006] IEEE [114]. Portions reprinted, with permission.

7.1 Introduction

We will assume p_t is a mixture of 2 Gaussians($n = 2$ in (1-9)):

$$s(\mathbf{u}) = s_0(fe^{-b\mathbf{u}^T D_1 \mathbf{u}} + (1 - f)e^{-b\mathbf{u}^T D_2 \mathbf{u}}), \quad (7-1)$$

where $f \geq 0, 1 - f \geq 0$ are considered as the apparent volume fractions of diffusion tensor D_1, D_2 respectively. Recently, Parker et al. [69] and Tuch et al. [70] used a mixture of two Gaussian densities to model the diffusion for the voxels where the Gaussian model fits the data poorly.

A prime problem in recovering multi tensor field $D_i(\mathbf{x}), i = 1, 2$, and $f(\mathbf{x})$ is the acquisition noise which corrupts the data measurement. Neither Parker nor Tuch considered removing noise. In this note we present a new variational method which differs from the existing methods in the following aspects. First, we recover tensor field $D_i(\mathbf{x}), i = 1, 2$, and $f(\mathbf{x})$ globally by simultaneous smoothing and data fitting, rather than estimating them from (7-1) at each isolated voxel, which leads to an ill-posed problem and is impossible to get a smooth multi-tensor field. Second, we applied the biGaussian model to all the voxels in the field while Parker et al [69] and Tuch et al. [70] only applied biGaussian model to the voxels where the Gaussian model fits the data badly. So they need to do preprocessing to distinguish voxels at which Gaussian model fits the data poorly from those at which Gaussian model fits data well. In our approach, this kind of voxel classification is not required and thus avoids the errors coming from it. Section 7.2 will explain in detail how to recover smooth multi-diffusion tensor field.

7.2 Recovery of Multi-Tensor Field in HARD MRI

We assume the data acquisition noise is additive and the distribution is modeled as a mixture of two Gaussians(7-1). The goal of this section is to recover a smooth

multi-tensor field from noisy HARD MRI data. By Cholesky factorization theorem, a symmetric matrix D is positive definite if and only if $D = LL^T$, for some lower triangular matrix L with positive diagonal entries. Uniqueness of the factorization is ensured by positiveness of the diagonal entries of L . To guarantee diffusion tensor D_1, D_2 are semi-definite, we will therefore let $D_i = L_i L_i^T$, for $i = 1, 2$, with L_i a lower triangular matrix with positive diagonal entries. The constraint $0 \leq f \leq 1$ is fulfilled through the variable relaxation method. For example, if we let $f(\mathbf{x}) = .5 + \frac{\arctan(\omega(\mathbf{x}))}{\pi}$, which is an increasing smooth function of ω , it is obvious that the values of f lie in $[0, 1]$. To get a regularized estimation of functions $f(\mathbf{x}), D_1(\mathbf{x}), D_2(\mathbf{x})$, we minimize the following functional w.r.t $L_1(\mathbf{x}), L_2(\mathbf{x}), f(\mathbf{x})$:

$$\int_{\Omega} (\sum_{i=1}^2 \sum_{m=1}^3 \sum_{n=1}^m \alpha |\nabla L_i^{mn}(\mathbf{x})| + \beta |\nabla f(\mathbf{x})|) d\mathbf{x} + \int_{\Omega} \int_0^{2\pi} \int_0^{\pi} |s_0(\mathbf{x})(f e^{-b\mathbf{u}^T L_1 L_1^T \mathbf{u}} + (1-f) e^{-b\mathbf{u}^T L_2 L_2^T \mathbf{u}}) - s(\mathbf{x}, \theta, \phi)|^2 \sin\theta d\theta d\phi d\mathbf{x} \quad (7-2)$$

with the constraint $L_i^{mm} > 0$, for $i = 1, 2, m = 1, 2, 3$. Here we have denote the mn^{th} entry of L_i by L_i^{mn} ; α, β are weighting factors; and $\mathbf{u} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)^T$ with $0 \leq \theta < \pi$ and $0 \leq \phi < 2\pi$. The first two terms in the first integral are the regularization terms, without which model(7-2) would be an ill-posed problem, and the second integral is the nonlinear data fidelity term based on (7-1).

We employ a gradient descent scheme to solve the minimization problem(7-2). Our initial f, L_1, L_2 are carefully chosen to avoid getting caught in local minima. For conciseness, we show the gradient-descent only for f, L_1^{11} as following:

$$\begin{aligned} \frac{\partial f}{\partial t} = & -2s_0(\mathbf{x}) \int_0^{2\pi} \int_0^\pi [s_0(\mathbf{x})(fe^{-\mathbf{b}\mathbf{u}^T L_1 L_1^T \mathbf{u}} + (1-f)e^{-\mathbf{b}\mathbf{u}^T L_2 L_2^T \mathbf{u}}) - s(\mathbf{x}, \theta, \phi)] \cdot \\ & (e^{-\mathbf{b}\mathbf{u}^T L_1 L_1^T \mathbf{u}} - e^{-\mathbf{b}\mathbf{u}^T L_2 L_2^T \mathbf{u}}) \sin(\theta) d\theta d\varphi + \beta \operatorname{div}\left(\frac{\nabla f}{|\nabla f|}\right) \end{aligned} \quad (7-3)$$

$$\begin{aligned} \frac{\partial L_1^{11}}{\partial t} = & 4bf s_0(\mathbf{x}) \int_0^{2\pi} \int_0^\pi [s_0(\mathbf{x})(fe^{-\mathbf{b}\mathbf{u}^T L_1 L_1^T \mathbf{u}} + (1-f)e^{-\mathbf{b}\mathbf{u}^T L_2 L_2^T \mathbf{u}}) - s(\mathbf{x}, \theta, \phi)] \cdot \\ & e^{-\mathbf{b}\mathbf{u}^T L_1^T \mathbf{u}} (\mathbf{u}_1 L_1^{11} + \mathbf{u}_1 \mathbf{u}_2 L_1^{21} + \mathbf{u}_1 \mathbf{u}_3 L_1^{31}) \sin(\theta) d\theta d\varphi + \alpha \operatorname{div}\left(\frac{\nabla L_1^{11}}{|\nabla L_1^{11}|}\right) \end{aligned} \quad (7-4)$$

7.3 White Matter Fiber Tractography

Results of the above section provide a smooth multi-tensor vector field and a smooth volume fraction field f , the fiber tractography based on which is almost not sensitive to noise and thus is more accurate. In this section, we will provide an improved line propagation algorithm for reconstructing white matter fiber traces. A line propagation scheme is defined by: $\mathbf{x}(t+1) = \mathbf{x}(t) + \mathbf{v}(t+1)\delta$, where $\mathbf{x}(t)$ is the position vector in R^3 of the streamline at time t , $\mathbf{v}(t+1)$ is a unit vector that leads the particle from the position $\mathbf{x}(t)$ to the next step $\mathbf{x}(t+1)$, and δ is the step size.

In DTI data, Westin et al [76] used the entire tensor D at location $\mathbf{x}(t)$ to determine $\mathbf{v}(t+1)$ as $D \cdot \mathbf{v}(t)$. They also provided a scheme which dynamically modulates the principal eigenvector(PE) \mathbf{e}_1 of tensor D and the tensor deflection contribution $D \cdot \mathbf{v}(t)$ to trace steering:

$$\mathbf{v}(t+1) = \alpha \mathbf{e}_1 + (1-\alpha)((1-\beta)\mathbf{v}(t) + \beta D \cdot \mathbf{v}(t)) \quad (7-5)$$

where α and β are user-defined weighting factors that vary between 0 and 1, and \mathbf{e}_1 and $\mathbf{v}(t)$ are normalized before used. However, (7-5) cannot handle voxels with degenerate anisotropy, since PE PE does not make sense there. Therefore we define $\mathbf{v}(t+1)$ as

$$\mathbf{v}(t+1) = (1-\beta)\mathbf{v}(t) + \beta D \cdot \mathbf{v}(t) \quad (7-6)$$

Here $D \cdot \mathbf{v}(t)$ is also normalized before used. Normalization of it was ignored in [76], but this is essentially necessary as for human brain HARD MRI data, norm of $D \cdot \mathbf{v}(t)$ is

usually in the order of 10^{-3} , tensor deflection without normalization would not contribute as much as expected. To make following explanation concise, let $f_1 = f, f_2 = 1 - f$. At each voxel, for $i = 1, 2$, apply (7-6) to D_i to get $\mathbf{v}_i(t + 1)$. Define corresponding step size as $\delta_i = cf_i(\mathbf{x}(t))FA_i(\mathbf{x}(t))\mathbf{v}(t) \cdot \mathbf{v}_i(t + 1)$ with c a fixed constant, $FA_i(\mathbf{x}(t))$ the fractional anisotropy (defined in 1-8) corresponding to tensor D_i . As we know, if f_i is very close to 0, channel D_i could be ignored; if FA_i is very low, anisotropy of D_i is low; if $\mathbf{v}(t) \cdot \mathbf{v}_i(t + 1)$ is low, there is too much bending between $\mathbf{v}(t)$ and $\mathbf{v}_i(t + 1)$. So fiber tracking should be terminated at channel D_i when anyone of the above quantities is low. This could simply be done by setting a threshold to step size so that channels with step size less than this threshold are terminated. The threshold is a statistical value obtained through a large size of experiments. This self adapting step size decreases propagation speed in regions with high curvature of fiber tracks and low diffusion anisotropy, increases speed in regions with low curvature and high diffusion anisotropy, and it also automatically terminates fiber tracking at channel(s) with extremely low step size(s).

Our scheme thus generalizes the usual tensor line propagation algorithm based on a single tensor. We call it multi-tensor line propagation(MTEND). The challenging aspect of this method is the estimation of $\mathbf{v}(t + 1)$'s at non-grid points. We linearly interpolate f and 6 entries of D_1, D_2 respectively, $\mathbf{v}(t + 1)$'s are then calculated using (7-6) based on the interpolated f, D_i 's.

7.4 Experimental Results

In this section we present synthetic as well as real data experiments. We did experiments on a set of subjects, but only list results of one subject for demonstration.

Our first experiment compare the proposed model (7-2) to Parker et. al's method([69]) in recovering smooth multi-tensor field and the volume fraction f using human brain HARD MRI data. The data set consists of 33 diffusion weighted images as well as one image in the absence of a diffusion-sensitizing field gradient. 27 evenly spaced axial planes with 128×128 voxels in each slice are obtained using a 3T MRI scanner

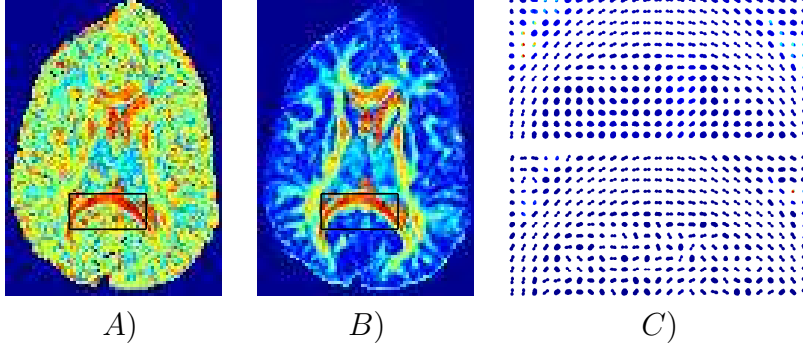


Figure 7-1. FA and ADC profile of the first channel. A)-B) FA maps of the first tensor field D_1 obtained using Parker et. al's([69]) and proposed model(7-2) respectively. C) Images of ADC profile $\mathbf{u}^T D_1 \mathbf{u}$ of voxels inside the selected regions in A) and B), the top one and the bottom one are obtained using D_1 calculated using proposed model(7-2) and Parker's method respectively.

with a single shot spin-echo EPI sequence. Slice thickness is $3.8mm$, gap is 0 between two consecutive slices, repetition time (TR) = $1000ms$, $echotime(TE)$ = $85ms$ and $b = 1000s/mm^2$, and the field of view (FOV) = $200mm \times 200mm$. Figure 7-1A) – C) compare maps of FA of the first tensor field D_1 and apparent diffusion coefficient(ADC) profiles of D_1 , i.e. $\mathbf{u}^T D_1 \mathbf{u}$ using solutions obtained from Parker et al's([69]) and proposed model(7-2). It is clear that all the results obtained from our model are much more smooth and seem more reasonable than those obtained by Parker's method. Specifically, Figure 7-1B)(result of proposed model(7-2)) gives a reasonable FA image from which we are able to distinguish voxels with high anisotropy(red region) from that with low anisotropy(blue region), while in Figure 7-1A)(result of Parker's method) FA collapses except in regions around the corpus callosum. In Figure 7-1C), at each voxel we show shape of ADC profile $\mathbf{u}^T D_1 \mathbf{u}$ in 41×21 directions, i.e. $41 \times 21 \mathbf{u}$'s. In the top image of Figure 7-1C)(result of proposed model(7-2)), shapes of ADC profile change smoothly from voxel to voxel, and in the region below corpus callosum, voxels which are mostly likely to be of isotropic diffusion have sphere-shaped ADC. But in the bottom image of Figure 7-1C)(result of Parker's method) shapes of ADC profile jump a lot from voxel to

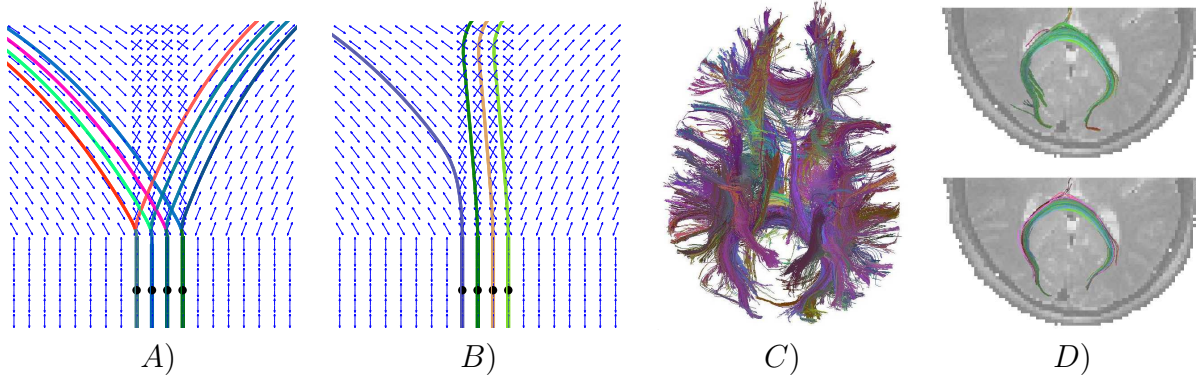


Figure 7-2. Comparison between TEND and MTEND. A) Traces recovered by using MTEND based on simulated multi-tensor field, the black points at the bottom are seeds. B) Traces recovered by using TEND based on simulated DTI data. C) Axial view of fiber tracking result of the whole brain by MTEND algorithm. D) Axial view of fiber tracking results using MTEND(top), TEND(bottom) algorithm .

voxel, especially in regions outside the corpus callosum. Thus our model appears superior to Parker et al.'s in these aspects.

Our second experiment is to compare the MTEND algorithm and the TEND algorithm in reconstructing fiber traces that include bifurcation. This is done on simulated data. We first simulate a $20 \times 20 \times 3$ multi-vector field shown by blue arrows in Figure 7-2A)B), where each voxel with one arrow owns only one tensor that has the direction shown by this blue arrow as the principal eigenvector, while each voxel with two arrows owns two tensors that have the directions shown by the two blue arrows as the principal eigenvectors. Second, we construct a multi-tensor field so that the multi-vector field is the corresponding principal eigenvector field. Raw DTI data are finally simulated based on the simulated multi-tensor field using (7-1) with $s_0 = 400$, $b = 1000$, $f = 1$ at voxels with one vector, $f = .5$ at voxels with two vectors, and 6 \mathbf{u} 's which are uniformly distributed on a sphere. Applying MTEND algorithm to the multi-tensor field while applying TEND algorithm to the single-tensor field obtained from DTI data, we obtain the results shown in Figure 7-2A), B) respectively. The four black points are the initiations, i.e. seeds of the fiber tracking. Nice bifurcations are observed in Figure 7-2A), and they

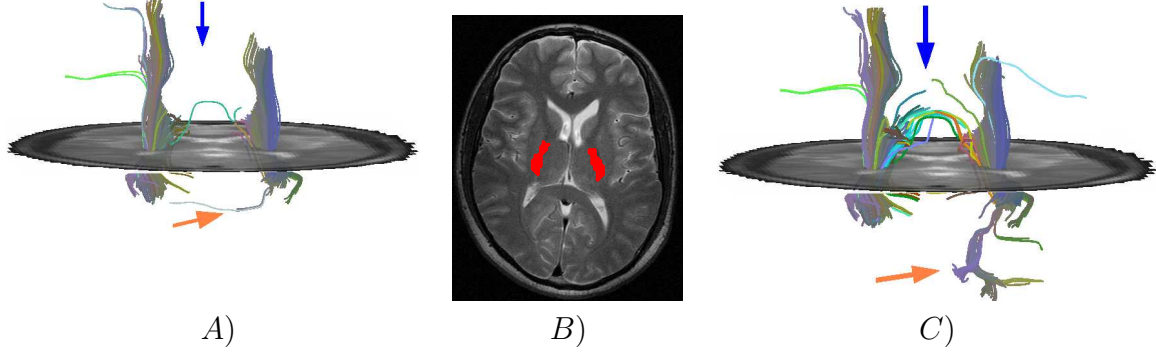


Figure 7-3. Comparison between TEND and MTEND at internal capsule. A)C) are tracking results using TEND, MTEND method respectively. B) Anatomic image of one slice with red regions as the seeds of tracking.

take place in voxels with two diffusion tensors as expected. In comparison, no bifurcation is visualized from Figure 7-2B) and only the left-most fiber trace goes almost along the vector field, while the other 3 fiber traces do not make sense at all. This indicated that the MTEND algorithm outdoes the TEND algorithm in recovering fibers with branching.

Results of experiments on human brain HARD MRI data are also shown in Figure 7-2 and Figure 7-3. These show that MTEND and TEND work similarly in the corpus callosum region, where Gaussian diffusion is dominant, but they differ in regions with non-Gaussian diffusion. Figure 7-2C) shows an axial view of the whole brain's fiber traces. The initiations of tracking are set at all the anisotropic voxels in the whole brain volume. Different strong fiber bundles and branches are clearly visualized, and are consistent with known neuroanatomy.

Figure 7-2D) shows axial view of tracking results around the corpus callosum region using the MTEND(top), and TEND(bottom) algorithms. Tracking results are embedded on a 2D anatomic image. Tracking starts from a small portion inside the corpus callosum. No significant difference is observed because in the corpus callosum region, Gaussian distribution is dominant, and therefore biGaussian model with $f \simeq 1$ and the Gaussian model work just as well in recovering a single-tensor field.

Finally, we select two regions of interest (ROI) from the internal capsule (red region in Figure 7-3B)) for another set of comparisons. We set all the anisotropic voxels in the whole brain volume as seeds, then apply MTEND based on multi-tensor field recovered using model (7-2) and TEND based on single-tensor field recovered using Gaussian diffusion model to reconstruct fiber traces separately. In MTEND we set $\beta = 0.9$ and the threshold value of step size δ to be 0.1; these are obtained from a large size of experiments. Only those fibers passing through the ROIs are retained and shown in Figure 7-3A)C) for TEND and MTEND respectively. Clearly, MTEND method recovers more branching fibers than TEND method does. Specifically, this happens in three different locations. One is at the lower right position and is indicated by orange arrows. Bunches of fiber traces with several branches are nicely shown in Figure 7-3C), but do not appear in Figure 7-3A). The second location is at the middle and is indicated by blue arrows. A strong bundle connecting the left portion and the right portion is clearly visualized in Figure 7-3C) but only one fiber trace is shown in Figure 7-3A). The third location lies in the most upper left position: Figure 7-3C) looks thicker and includes more fibers in each branching than Figure 7-3A) does. The main reason for the difference is that voxels involving branching in MTEND method are characterized as isotropic, so TEND algorithm terminates at these voxels.

7.5 Conclusion

A new variational framework for simultaneous reconstruction and regularization of multi-diffusion tensor field and a new fiber tractography algorithm based on multi-tensor field was provided. The performance of the proposed model was evaluated on synthetic data and several human subjects' HARD MR images. The experimental results indicated that proposed model (7-2) for recovering multi-tensor field together with MTEND for reconstruction of white matter fiber traces work more accurately than the Gaussian diffusion model together with TEND.

The proposed model assumes that the probability density function of diffusion is of linear combination of two Gaussians. This results in 13 unknowns at each voxel, and hence

the acquisition of at least 13 diffusion weighted images was required to solve for the 13 unknowns accurately.

CHAPTER 8

FAST SEGMENTATION OF WHITE MATTER FIBER TRACTS BASED ON GEOMETRIC FLOWS

8.1 Introduction

Geometric flows, especially curvature shortening flows, are becoming more and more important regularization tools in computer vision. A curvature flow is a curve or surface that evolves at each point along its normal with a velocity depending on the curvature at that point. A geometric flow is a curve or surface that evolves with a velocity depending on external properties determined by the image features.

A general flow for a 3D closed surface can be described as

$$\frac{\partial S}{\partial t} = (F + H)N, \quad (8-1)$$

where F is an image based speed function, H is an intrinsic speed depending on the curvature of the surface S , N is the normal of the surface and t is time.

To solve this time dependent partial differential equation, we use the level set method introduced by Osher and Sethian (1988), where the evolving surface is considered as 0 level set of a function of one dimension higher(ϕ). By doing this, a numerically stable algorithm that easily handles topology changes of the evolving surface is obtained. Finally, (8-1) becomes

$$\frac{\partial \phi}{\partial t} = (F + H)|\nabla \phi| \quad (8-2)$$

The key point of a new model is to design an efficient and senseful external force or speed function F .

8.2 Model

L. Jonasson et al. [115] introduced a similarity based front propagation algorithm. They defined

$$F = \text{mean}(NTSP(D_i, D_{i-1}), NTSP(D_i, D_{i-2})), \quad (8-3)$$

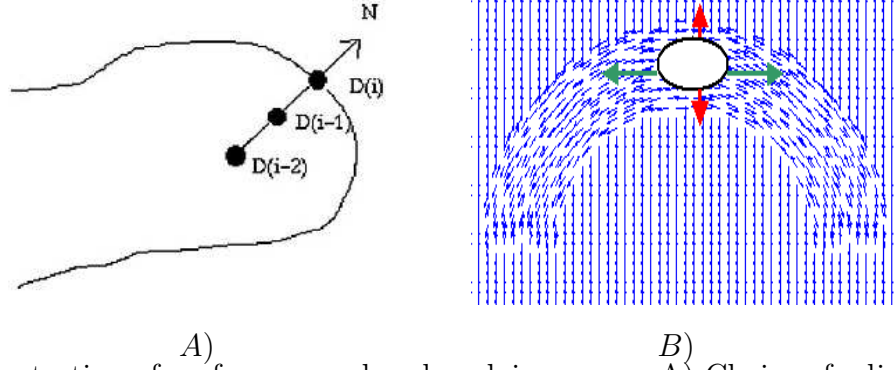


Figure 8-1. Illustration of surface normal and evolving curve. A) Choice of adjacent voxels with respect to the normal of the surface, courtesy of [115]. B) Evolving curve(black) superimposed inside a semi-circle shaped 2D tensor field.

where the notation is as follows. D_i is the diffusion tensor in the current voxel. For $p = 1, 2$, D_{i-p} is the diffusion tensor in the grid element found by following the normal to the surface p voxels backwards from the original voxel i , as illustrated in Figure8-1A), and then selecting the nearest neighbor. $NTSP(D_i, D_{i-1})$ [116] is the Normalized Tensor Scalar Product defined as

$$NTSP(D_i, D_{i-1}) = \frac{D_i : D_{i-1}}{Trace(D_i)Trace(D_{i-1})}, \quad (8-4)$$

with $D_i : D_{i-1} = Trace(D_i D_{i-1})$, $NTSP(D_i, D_{i-2})$ is defined similarly. The fundamental assumption of the segmentation technique in [115] is that adjacent voxels in a tract have similar diffusion property. The proposed model maintains this assumption, but the goal is to obtain a faster and more accurate geometric flow.

Figure8-1D) shows an example of 2D tensor field to demonstrate motivation of the proposed model. In Figure8-1D) the black ellipse depicts the evolving curve, red arrows and green arrows show normal directions at corresponding locations. Intuitively, evolution speed along green arrows should be larger than that along the red arrows, this is because the evolving contour is close to the boundary at locations with red arrows, while pretty far away from the boundary at locations with red arrows. Mathematically, evolution should be faster at location where the normal direction is more coincident with the tensor field.

To enforce this behavior, our model replaces (8-3)

$$F = \text{mean}(NTSP(D_i, D_{i+1}), NTSP(D_i, D_{i+2})) + \beta COIN, \quad (8-5)$$

where $COIN$ is a measure of coincidence of normal direction to the evolving curve (surface in 3D) and the tensor field. We have tried two candidates, the first of which is:

$$COIN_1 = |N \cdot PE| \quad (8-6)$$

with PE the principal eigenvector of the current diffusion tensor D_i , and N is the normal direction. The absolute value is to guarantee the positivity of COIN measure.

The second candidate is

$$COIN_2 = (FA) * |N \cdot \frac{D_i * N}{\|D_i * N\|}| \quad (8-7)$$

with FA the fractional anisotropy value of tensor D_i . The definition of $COIN_1$ is quite intuitive: if normal direction is coincident with the principle direction of tensor, the evolution speed is higher. The definition of $COIN_2$ is motivated by Mariana Lazar et al [117]. $COIN_2$ is less sensitive to image noise and experimental results show that it is more efficient than $COIN_1$.

Note that in definition (8-5), similarity is defined using D_{i+1} and D_{i+2} , which are neighbors of D_i along the positive normal direction instead of negative normal direction used in (8-3). This is to avoid overshooting.

We terminate evolution of the curve/surface at locations where the similarity measure $NTSP$ is less than a threshold or the norm of the gradient of $NTSP$ is less than a threshold. Our experiments indicate that these two criteria combined together can accurately catch the boundary. In [115] only the first one is used, but this does not work for segmentation of a vector/tensor field composed of homogeneous regions. In this case, the similarity map of the vector/tensor field is a piecewise constant function, and the boundary between one homogeneous region and another is signaled by a high gradient

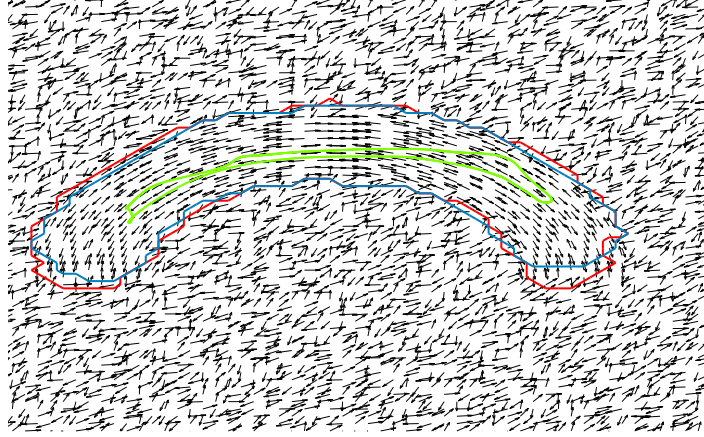


Figure 8-2. Segmentation of a 2-D tensor field using two methods. Red contour: Segmentation result of the proposed model; Blue contour: segmentation result of L. Jonasson's model in [115]; Green curve: initial contour.

of the similarity measure; while a threshold of similarity alone could not tell where the boundary is.

8.3 Experimental Results

For all the results shown in this section, we use $COIN_2$ as it is better than $COIN_1$.

8.3.1 Synthetic results

Figure 8-2 compares segmentation quality and speed of proposed method and method in [115]. The black arrows show the principal eigenvectors of a 2D diffusion tensor field. The green curve is the initial curve, and the blue and the red one show the result of model in [115] and the proposed model, respectively. In the proposed model, we set $\beta = 10$, it costs 25 iterations and 53.45 seconds in a computer with PIV 2.8GHZ CPU and 2G RAM using Matlab script code, while for model in [115], it costs 50 iterations and 105.55 seconds. It shows that the proposed model is more than twice as fast and the result is even more accurate.

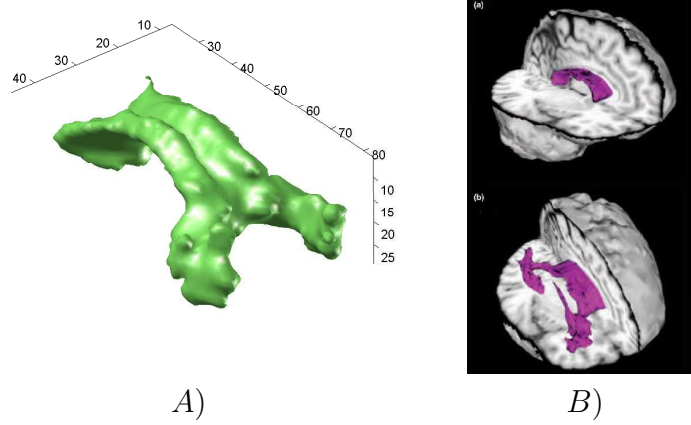


Figure 8-3. Comparison of corpus callosum segmentation results from two models.
A)-B) Segmentation of corpus callosum obtained by using the proposed model and the one in [115] respectively.)

8.3.2 Human brain results

Figure 8-3 show comparison of segmentation of corpus callosum obtained using the proposed model and the model in [115]. Figure 8-3B) shows two sets of segmentation results obtained from two threshold values for NTSP. The top one corresponds to a higher threshold, while the bottom one corresponds to a lower threshold. It is clear that when a higher threshold is chosen, only part of the corpus callosum is segmented, but when a lower threshold is chosen, regions other than corpus callosum are also segmented. In comparison, the proposed model (Figure 8-3A)) captures more accurate corpus callosum details.

8.4 Conclusion

The proposed model designed a new external force that depended on two magnitudes: similarity of diffusion tensors and the coincidence level of the evolving surface normal with the tensor field. The new model based on this new external force was able to segment white matter fiber bundles in DT-MRI more accurately and efficiently than the one provided by L. Jonasson et al. in [115].

REFERENCES

- [1] M. Kass, A. Witkin, and D. Terzopoulos, “Snakes: Active contour models,” *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.
- [2] V. Caselles, F. Catté, T. Coll, and F. Dibos, “A geometric model for active contours in image processing,” *Numerische Mathematik*, vol. 66, pp. 1–31, 1993.
- [3] R. Malladi, J. Sethian, and B. Vemuri, “Shape modeling with front propagation: A level set approach,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 158–175, 1995.
- [4] V. Caselles, R. Kimmel., and G. Sapiro, “On geodesic active contours,” *Intel. Journal of Computer Vision*, vol. 22, no. 1, pp. 61–79, 1997.
- [5] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. J. Yezzi, “Gradient flows and geometric active contour models,” *Proc. ICCV’95*, pp. 810–815, 1995.
- [6] A. Yezzi, S. Kichenassamy, A. Kumar, P. J. Olver, and A. Tannenbaum, “A geometric snake model for segmentation of medical imagery,” *IEEE Trans. Medical Imaging*, vol. 16, pp. 199–209, 1997.
- [7] S. C. Zhu and A. Yuille, “Region competition: unifying snakes, region growing, and Bayes/MDL for multiband image segmentation,” *IEEE PAMI*, vol. 18, pp. 884–90, 1996.
- [8] D. Mumford and J. Shah, “Optimal approximation by piecewise smooth functions and associated variational problems,” *Comm. Pure Appl. Math.*, vol. 42, pp. 557–685, 1989.
- [9] D. Cremers, F. Tischhuser, J. Weickert, and C. Schnrr, “Diffusion-snakes: Introducing statistical shape knowledge into the Mumford-Shah functional,” *International Journal of Computer Vision*, vol. 50, no. 3, pp. 295–315, 2002.
- [10] A. Tsai, A. Yezzi Jr., , and A. S. Willsky, “Curve evolution implementation of the Mumford-Shah functional for image segmentation, denoising, interpolation, and magnification,” *IEEE Transactions on Image Processing*, vol. 10, no. 8, pp. 1169–1186, 2001.
- [11] A. Chakraborty, H. Staib, and J. Duncan, “Deformable boundary finding in medical images by integrating gradient and region information,” *IEEE Transactions on Medical Imaging*, vol. 15, no. 6, pp. 859–870, 1996.
- [12] N. Paragios and R. Deriche, “Geodesic active regions for supervised texture segmentation,” *ICCV, Cofu, Greece*, pp. 926–932, 1999.
- [13] N. Paragios, “Geodesic active regions and level set methods: contributions and applications in artificial vision,” *Ph.D. thesis, School of Computer Engineering, University of Nice/Sophia Antipolis*, 2000.

- [14] N. Paragios, “A variational approach for the segmentation of the left ventricle in MR cardiac images,” *Proceedings 1st IEEE Workshop on Variational and Level Set methods in Computer Vision, Vancouver, B.C., Canada*, pp. 153–160, 13 July 2001.
- [15] L. Ambrosio, N. Fusco, and D. Pallara, “Partial regularity of free discontinuity sets,” *Ann. Scuola Norm. Sup. Pisa Cl. Sci.*, vol. 24, pp. 39–62, 1997.
- [16] T. F. Chan and L. A. Vese, “Active contours without edges,” *IEEE Trans. Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.
- [17] L. A. Vese and T. F. Chan, “A multiphase level set framework for image segmentation using the mumford and shand model,” *Int. Journal of Computer Vision*, vol. 50(3), pp. 271–293, 2002.
- [18] S. Osher and J. A. Sethian, “Fronts propagating with curvature-dependent speed: algorithm based on Hamilton-Jacobi formulation,” *Journal of Computational Physics*, vol. 70, pp. 12–49, 1988.
- [19] A. Yezzi, S. Soatto, A. Tsai, and A. Willsky, “The Mumford-Shah functional: From segmentation to stereo,” *Mathematics and Multimedia*, 2002.
- [20] T. F. Chan and S. Esedoglu, “Aspects of total variation regularized L^1 function approximation,” *SIAM J. Appl. Math.*, vol. 65(5), pp. 1817–1837, 2005.
- [21] T. F. Chan, S. Esedoglu, and M. Nikolova, “Algorithms for finding global minimizers of image segmentation and denoising models,” *SIAM J. Appl. Math.*, vol. 66(5), pp. 1632–1648, 2006.
- [22] J. Lie, M. Lysaker, and X.C.Tai, “A variant of the level set method and applications to image segmentation,” *Math. Comp.*, vol. 75, pp. 1155–1174, 2006.
- [23] X.C.Tai and C. H. Yao, “Image segmentation by piecewise constant Mumford-Shah without estimating the constants,” *J. Comput. Math.*, vol. 24(3), pp. 435–443, 2006.
- [24] S. Gao and T.D. Bui, “A new image segmentation and smoothing model,” *Proc. of Int. Sym. on Biomed. Img.*, pp. 137–140, 2004.
- [25] Tony Chan, Berta Sandberg, and Mark Moelich, “Some recent developments in variational image segmentation,” *UCLA CAM Reports*, , no. 06-52, Sep. 2006.
- [26] M. Rousson and R. Deriche, “A variational framework for active and adaptive segmentation of vector values images,” *Proceedings of workshop on motion and video computing*, pp. 56–61, Dec. 2002.
- [27] J. Kim, J. W. Fisher III, A. Yezzi, M. Cetin, and A. S. Willsky, “A nonparametric statistical method for image segmentation using information theory and curve evolution,” *IEEE Transactions on Image Processing*, vol. 14:10, pp. 1482–1502, Oct. 2005.

- [28] X. Huang, D. Metaxas, and T. Chen, "Metamorphs: Deformable shape and texture models," *Proceedings of IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, vol. I, pp. 496–503, 2004.
- [29] W. Abd-Almageed, C. E. Smith, and S. Ramadan, "Kernel snakes: non-parametric active contour models," *IEEE Int. Conf. on Systems, Man and Cybernetics*, vol. 1, pp. 240–244, 2003.
- [30] T. Cootes, C. Beeston, G. Edwards, and C. Taylor, "Unified framework for atlas matching using active appearance models," in *Int'l Conf. Inf. Proc. in Med. Imaging*, Springer-Verlag, 99, pp. 322–333.
- [31] Y. Wang and L. Staib, "Boundary finding with corresponding using statistical shape models," *Proc. IEEE Conf. Comp. Vision and Patt. Recog.*, pp. 338–345, 1998.
- [32] T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active shape model - their training and application," *Computer Vision and Image Understanding*, vol. 61, pp. 38–59, 1995.
- [33] L. Staib and J. Duncan, "Boundary finding with parametrically deformable contour methods," *IEEE Trans. Patt. Analysis and Mach. Intell.*, vol. 14, no. 11, pp. 1061–1075, 1992.
- [34] H. D. Tagare, "Deformable 2-d template matching using implementation of the Mumford-Shah functional for image segmentation, denoising, interpolation, and magnification," *IEEE Transactions on Image Processing*, vol. 10, no. 8, pp. 1169–1186, 2001.
- [35] A. Yuille, P. W. Hallinan, and D. S. Cohen, "Feature extraction from faces using deformable templates," *Int. J. Computer Vision*, vol. 8, pp. 99–111, 1992.
- [36] T. Cootes, A. Hill, C. Taylor, and J. Haslam, "The use of active shape models for locating structures in medical images," *Image Vision Comput.*, vol. 13, no. 6, pp. 255–266, 1994.
- [37] G. Szekel, A. Kelemen, Ch. Brechbuhler, and G. Gerig, "Segmentation of 2-D and 3-D objects from MRI volume data using constrained elastic deformation of flexible fourier surface models," *Medical Image Analysis*, vol. 1, no. 1, pp. 19–34, 1996.
- [38] D. Metaxas and D. Terzopoulos, "Shape and nonrigid motion estimation through physics-based synthesis," *IEEE trans. Pattern Anal. Machine Intelligence*, vol. 15, no. 6, pp. 580–591, 1993.
- [39] B. C. Vemuri and A. Radisavljevic, "Multiresolution stochastic hybrid shape models with fractal priors," *ACM Trans. on Graphics*, vol. 13, no. 2, pp. 177–207, 1994.
- [40] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis: a survey," *Medical Image Analysis*, vol. 1, no. 2, pp. 91–108, 1996.

- [41] M. E. Leventon, W. E. L. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," *Proc. IEEE Conf. Comp. Vision and Patt. Recog.*, pp. 316–323, 2000.
- [42] Y. Chen, H. Tagare, S.R. Thiruvankadam, F. Huang, D.Wilson, A. Geiser, K. Gopinath, and R. Briggs, "Using prior shapes in geometric active contours in a variational framework," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 315–328, 2002.
- [43] N. Paragios and M. Rousson, "Shape prior for level set representations," *Proceeding of ECCV*, 2002.
- [44] B. C. Vemuri, Y. Chen, and Z. Wang, "Registration associated image smoothing and segmentation," *Proceeding of ECCV 2002*, 2002.
- [45] F. Richard and L. Cohen, "A new image registration technique with free boundary constraints: Application to mammography," *ECCV 2002, LNCS*, pp. 531–545, 2002.
- [46] J. Dias and J. Leita, "Wall position and thickness estimation from sequences of echocardiograms images," *IEEE Trans. Med. Imag.*, vol. 15, pp. 25:38, Feb. 1996.
- [47] N. Paragios, M. Rousson, and V. Ramesh, "Marching distance functions: a shape-to-area variational approach for global-to-local registration," *Computer Vision-ECCV*, pp. 775–789, May 2002.
- [48] S. Soatto and A. Yezzi, "Deformation: deforming motion, shape average and joint registration and segmentation of images," *ECCV 2002*, 2002.
- [49] N. Paragios, "User-interactive level set method for image segmentation," *ICCV*, 2003.
- [50] D. LeBihan and P. J. Basser, "Molecular diffusion and nuclear magnetic resonance," *Diffusion and perfusion magnetic resonance imaging*, 1995.
- [51] M. E. Moseley, Y. Cohen, J. Mintorovitch, J. Suruda, L. Chieffitt, D. Norman, and P. Weinstein, "Evidence of anisotropic self-diffusion in cat brain," *Proc. of the 8th ISMRM*, pp. 136–136, 1989.
- [52] M. E. Moseley, J. Kucharczyk, H. S. Asgari, and D. Norman, "Anisotropy in diffusion weighted MRI," *Magn. Reson. Med.*, vol. 19, pp. 321–326, 1991.
- [53] P. J. Basser and C. Pierpaoli, "Microstructural and physiological features of tissues elucidated by quantitative diffusion tensor MRI," *Magn. Reson. Med.*, vol. 111(B), pp. 209–219, 1996.
- [54] R. Gupta, S. Saksena, A. Agarwal, K. Hasan, M. Husain, and V. Gupta et al., "Diffusion tensor imaging in late posttraumatic epilepsy," *Epilepsia*, vol. 46, no. 9, pp. 1465–1471, 2005.

- [55] J. Roh, D. Wang, S. Lee, B. Yoon, and K. Chang, "Significance of acute multiple brain infarction on Diffusion-Weighted Imaging," *Stroke*, vol. 31, no. 3, pp. 688:694, 2000.
- [56] D. S. Tuch, R. M. Weisskoff, J. W. Belliveau, and V. J. Wedeen, "High angular resolution diffusion imaging of the human brain," in *Proc. of the 7th ISMRM*, Philadelphia, 1999, p. 321.
- [57] V. J. Wedeen, T. G. Reese, D. S. Tuch, M. R. Weigel, J-G Dou, R. M. Weisskoff, and D. Chesler, "Mapping fiber orientation spectra in cerebral white matter with fourier transform diffusion MRI," *Proc. of the 8th ISMRM*, pp. 82–82, 2000.
- [58] E. O. Stejskal and J. E. Tanner, "Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient," *Chem. Phys.*, vol. 42, pp. 288–292, 1965.
- [59] P. J. Basser, J. Mattiello, and D. LeBihan, "Estimation of the effective self-diffusion tensor from the NMR," *Spin Echo. J. Magn. Reson.*, vol. series B 103, pp. 247–254, 1994.
- [60] P. J. Basser, J. Mattiello, and D. LeBihan, "MR diffusion tensor spectroscopy and imaging," *Biophys*, vol. 66, pp. 259–267, 1994.
- [61] T. L. Chenevert, J. A. Brunberg, and J. G. Pipe, "Anisotropic diffusion in human white matter: demonstration with MR techniques in vivo," *Radiology*, vol. 177, pp. 401–405, 1990.
- [62] E. W. Hsu and S. Mori, "Analytical expression for the NMR apparent diffusion coefficients in an anisotropy system and a simplified method for determining fiber orientation," *Magn Reson Med*, vol. 34, pp. 194–200, 1995.
- [63] R. Deriche, D. Tschumperle, and C. Lenglet, "DT-MRI estimation, regularization and fiber tractography," in *Proc. of 2nd ISBI*, Washington D.C, 2004, pp. 9–12.
- [64] P. J. Basser, J. Mattiello, and D. LeBihan, "MR diffusion tensor spectroscopy and imaging," *Biophys*, vol. 66, pp. 259–267, 1994.
- [65] A. L. Alexander, K. M. Hasan, M. Lazar, J.S. Tsuruda, and D. L. Parker, "Analysis of partial volume effects in diffusion-tensor MRI," *Magn. Reson. Med.*, vol. 45, pp. 770–780, 2001.
- [66] L. Frank, "Characterization of anisotropy in high angular resolution diffusion weighted mri," in *Proc. of 9th ISMRM*, Glasgow, Scotland, 2001, p. 1531.
- [67] L. Frank, "Anisotropy in high angular resolution diffusion-weighted MRI," *Magn Reson Med*, vol. 45, pp. 935–939, 2001.

- [68] D. S. Tuch, R. M. Weisskoff, J. W. Belliveau, and V. J. Wedeen, "High angular resolution diffusion imaging of the human brain," *Proc. of the 7th ISMRM*, pp. 321–321, 1999.
- [69] G. J. M. Parker and D. C. Alexander, "Probabilistic Monte Carlo based mapping of cerebral connections utilising whole-brain crossing fiber information," in *Information Processing in Medical Imaging*, Ambleside UK, 07 2003, pp. 684–696.
- [70] D. S. Tuch, T. G. Reese, M. R. Wiegell, N. Makris, J. W. Belliveau, and V. J. Wedeen, "High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity," *Magn Reson Med*, vol. 48, pp. 577–582, 2002.
- [71] Y. Chen, W. Guo, Q. Zeng, G. He, Y. Liu, and B.C Vemuri, "Recovery of intra-voxel structure from HARD MRI," *Proc. of 2nd ISBI*, pp. 1028–1031, 2004.
- [72] D.C.Alexander, G.J.Barker, and S.R.Arridge, "Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data," *Magn.Reson.Med.*, vol. 48, pp. 331–340, 2002.
- [73] D. Le Bihan, E. Brethon, D. Lallemand P. Grenier, E. Cabanis, and M. Laval-Jeantet, "MR imaging of intravoxel incoherent motions: Application to diffusion and perfusion in neurologic disorders," *Radiology*, vol. 161, pp. 401–407, 1986.
- [74] B. C. Vemuri, Y. Chen, M. Rao, Z. Wang, T McGraw, T. Mareci, S. J. Blackband, and P. Reier, "Automatic fiber tractography from DTI and its validation," in *Proc. of 1st IEEE International Symposium on Biomedical Imaging*, 2002, pp. 505–508.
- [75] C. Pierpaoli, P. Jezzard, P. J. Basser, A. Barnett, and G. Di Chiro, "Diffusion tensor MR imaging of the human brain," *Radiology*, vol. 201(3), pp. 637–648, 1996.
- [76] C. F. Westin, S. E. Maier, B. Khidir, P. Everett, F.A. Jolesz, and R. Kikinis, "Image processing for diffusion tensor magnetic resonance imaging," in *Lecture notes in computer science:MICCAI*, Cambridge:Springer, 1999, pp. 441–452.
- [77] M. Lazar, D. M. Weinstein, J. S. Tsuruda, K. M. Hasan, K. Arfanakis, M. Meyerand, B. Badie, H. Rowley, V. Haughton, A. Field, and A. Alexander, "White matter tractography using diffusion tensor deflection," *Human Brain Mapping*, vol. 18, pp. 306–321, 2003.
- [78] Y. Chen, W. Guo, F. Huang, and D. Wilson, "Using prior shape and points in medical image segmentation," *Lecture Notes in Computer Science, Proceedings of International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, vol. 44, pp. 625–632, 2000.
- [79] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithm," *Physica D*, vol. 60, pp. 259–268, 1992.

- [80] G. Aubert and L. Vese, “A variational method in image recovery,” *SIAM J. Num. Anal.*, vol. 5, no. 34, pp. 1948–1979, 1997.
- [81] H. K. Zhao, T. Chan, B. Merriman, and S. Osher, “A variational level set approach to multiphase motion,” *J. Comput. Phys.*, vol. 127, pp. 179–195, 1996.
- [82] Randall L. Rubank, *Nonparametric Regression and Spline Smoothing, 2nd Edition*, Marcel Dekker, Inc., 1999.
- [83] P.J. Green and Bernard. W. Silverman, *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*, Chapman and Hall/CRC, 1994.
- [84] W. Guo and Y. Chen, “Using non-parametric kernel to segment and smooth images simultaneously,” *Proceedings of International Conference in Image Processing*, pp. 217–220, 2006.
- [85] X. Tai, “Global extrapolation with a parallel splitting method,” *Numerical Algorithm*, vol. 3, pp. 527–440, 1991.
- [86] J. Weickert, B. Romeny, and M. Viergever, “Efficient and reliable schemes for nonlinear diffusion filtering,” *IEEE Trans. on Img. Proc.*, vol. 7, no. 3, pp. 398–410, March 1998.
- [87] G. Kühne, J. Weickert, M. Beier, and W. Effelsberg, “Fast implicit active contour models,” *Lecture Notes on Computer Science*, vol. 2449, pp. 133–140, 2002.
- [88] H. Federer, “Geometric measure theory,” *Berlin, Heidelberg, New-York: Springer-Verlag*, 1969.
- [89] L. C. Evans and R. R. Gariepy, “Measure theory and fine properties of functions,” *CRC Press*, 1992.
- [90] E. Guisti, “Minimal surfaces and functions of bounded variation,” *Basel, Boston, Stuttgart: Birkhäuser Verlag*, 1984.
- [91] G. J. M. Parker, J. A. Schnabel, M. R. Symms, D. J. Werring, and G. J. Baker, “Nonlinear smoothing for reduction of systematic and random errors in diffusion tensor imaging,” *Magn. Reson. Med.*, vol. 11, pp. 702–710, 2000.
- [92] B. C. Vemuri, Y. Chen, M. Rao, T. McGraw, Z. Wang, and T. Mareci, “Fiber tract mapping from diffusion tensor MRI,” in *Proc. of IEEE Workshop on VLISM*, Philadelphia, 2001, pp. 81–88.
- [93] C. Feddern, J. Weickert, and B. Burgeth, “Level-set methods for tensor-valued images,” in *Proceedings of the 9th Annual Meeting of ISMRM*, Nice, 2003, pp. 65–72.
- [94] C. Poupon, J. F. Mangin, C. A. Clark, V. Frouin, J. Regis, D. LeBihan, and I. Block, “Towards inference of human brain connectivity from MR diffusion tensor data,” *Med. Image Anal.*, vol. 5, pp. 1–15, 2001.

- [95] D. Tschumperle and R. Deriche, "Regularization of orthonormal vector sets regularization with PDE's and applications," *IJCV*, vol. 50(3), pp. 237–252, 2002.
- [96] C. Ched'hotel, D. Tschumperle, and Olivier D. Faugeras, "Constrained flows of matrix-valued functions: Application to diffusion tensor regularization," in *ECCV*, 2002, vol. 1, pp. 251–265.
- [97] Sinisa Pajevic and Carlo Pierpaoli, "Color schemes to represent the orientation of anisotropic tissues from diffusion tensor data : Application to white matter fiber tract mappying in the human brain," *Magn Reson Med*, vol. 42, pp. 526–540, 1999.
- [98] D. Tschumperle and R. Deriche, "Tensor field visualization with pde's and application to DT-MRI fiber visualization," in *Proc. of IEEE Workshop on VLSM*, Nice, 2003, pp. 256–26.
- [99] J. Weickert and T. Brox, "Diffusion and regularization of vector- and matrix-valued images," *Contemporary Mathematics*, vol. 313, pp. 251–268, 2002.
- [100] Z. Wang, B. C. Vemuri, Y. Chen, and T. Mareci, "Simultaneous smoothing and estimation of the tensor field from diffusion tensor MRI," in *Proc. of CVPR*, Wisconsin, 2003, vol. 2.
- [101] Z. Wang, B.C.Vemuri, Y.Chen, and T. Mareci, "A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from DWI," in *Proc. of IPMI*, Ambleside, UK, 2003, pp. 660–671.
- [102] P. Blomgren, T. Chan, P. Mulet, and C. K. Wong, "Total variation image restoration: Numerical methods and extensions," *Proceeding of IEEE Int'l Conference on Image Processing*, vol. 3, pp. 384–387, 1997.
- [103] A.Chambolle and P-L.Lions, "Image recovery via total variation minimization and related problems," *Numerische Mathematik*, vol. 1(76), pp. 167–188, 1997.
- [104] Y. Chen, S. Levine, and M. Rao, "Functionals with $p(x)$ -growth in image restoration," *SIAM Journal of App. Math.*, vol. 66(4), pp. 1383–1406, 2006.
- [105] P. Blomgren and T. Chan, "Color TV: total variation methods for restoration of vector-valued images," *IEEE Trans. on Image Processing*, vol. 7(3), pp. 304–309, 1998.
- [106] Z. Wang, B. C. Vemuri, Y. Chen, and T. Mareci, "A constrained variational principle for direct estimation and smoothing of the tensor field from complex DWI," *IEEE TMI*, vol. 23:8, pp. 930–939, 2004.
- [107] T. Lu, P. Neittaanm, and X. Tai, "A parallel splitting up method and its application to Navier-Stokes equations," *Applied Mathematics Letters*, vol. 4(2), pp. 25–29, 1991.
- [108] S. D. Conte and C. DeBoor, "Book," *Elementary Numerical Analysis. McGraw-Hill, New York*, 1972.

- [109] Y. Chen, W. Guo, Q. Zeng, X. Yan, and Y. Liu, "Apparent diffusion coefficient approximation and diffusion anisotropy characterization in DWI," *Lecture Notes in Computer Science, Proceedings of international conference on information processing in medical imaging*, pp. 246–257, 2005.
- [110] M. Rao, Y. Chen, B. C. Vemuri, and F. Wang, "Cumulative residual entropy: A new measure of information," *IEEE Trans. on Info. Theory*, vol. 50, pp. 1220–1228, 2004.
- [111] Y. Chen, W. Guo, Q. Zeng, X. Yan, F. Huang, H. Zhang, G. He, B.C Vemuri, and Y. Liu, "Estimation, smoothing, and charaterization of apparent diffusion coefficient profiles from high angular resolution DWI," *Proc. of CVPR*, pp. 588–593, 2004.
- [112] M. Rao, "More on a new concept of entropy and information," *Jour. of Theo. Prob.*, vol. 18(4), pp. 967–981, 2004.
- [113] D. Strong and T. Chan, "Spatial and scale adaptive total variation based regularization and anisotropic diffusion in image processing," *UCLA-CAM Report*, vol. 46, 1996.
- [114] W. Guo, Q. Zeng, Y. Chen, and Y. Liu, "Reconstruct white matter fiber traces using multi-tensor deflection in DWI," *Proceedings of International Symposium on Biomedical Image*, pp. 69–72, 2006.
- [115] L. Jonasoon, X. Bresson, P. Hagmann, O. Cuisenaire, R. Meuli, and J. Thiran, "White matter fiber trace segmentation in DT-MRI using geometric flows," *Med. Img. Ana.*, vol. 9, pp. 223–236, 2005.
- [116] D. Alexander, J. Gee, and R. Bajcsy, "Similarity measures for matching diffusion tensor images," *Proc. BMCV*, pp. 93–102, 1999.
- [117] M. Lazar, D. Weinstein, and J. Tsuruda et. al, "White matter tractography using diffusion tensor deflection," *Human Brain Mapping*, vol. 18, pp. 306–321, 2003.

BIOGRAPHICAL SKETCH

Weihong Guo was born in Hei Fei, Anhui Province, P. R. China. She got her bachelor of science degree from Central University for Nationalities, P.R. China, in 1999. She then was admitted to Beijing Normal University graduate program without taking an admission exam. She earned her Doctor of Philosophy degree in Applied Mathematics and Master of Statistics degree from University of Florida in May 2007. Her research interests include mathematical modeling and algorithm developing in medical image processing, statistical image processing, medical image analysis and numerical partial differential equation.