University of Kentucky
UKnowledge

Theses and Dissertations--Computer Science | Computer Science

2013

RECOVERING LOCAL NEURAL TRACT DIRECTIONS AND RECONSTRUCTING NEURAL PATHWAYS IN HIGH ANGULAR RESOLUTION DIFFUSION MRI

Ning Cao
University of Kentucky, ningcao@yahoo.com

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Recommended Citation
Cao, Ning, "RECOVERING LOCAL NEURAL TRACT DIRECTIONS AND RECONSTRUCTING NEURAL PATHWAYS IN HIGH ANGULAR RESOLUTION DIFFUSION MRI" (2013). *Theses and Dissertations--Computer Science*. 13.
https://uknowledge.uky.edu/cs_etds/13

This Doctoral Dissertation is brought to you for free and open access by the Computer Science at UKnowledge. It has been accepted for inclusion in Theses and Dissertations--Computer Science by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

www.manaraa.com

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained and attached hereto needed written permission statements(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine).

I hereby grant to The University of Kentucky and its agents the non-exclusive license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless a preapproved embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's dissertation including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Ning Cao, Student

Dr. Jun Zhang, Major Professor

Dr. Miroslaw Truszczynski, Director of Graduate Studies

www.manaraa.com

University of Kentucky
UKnowledge

Theses and Dissertations--Computer Science | Computer Science

2013

RECOVERING LOCAL NEURAL TRACT DIRECTIONS AND RECONSTRUCTING NEURAL PATHWAYS IN HIGH ANGULAR RESOLUTION DIFFUSION MRI

Ning Cao
University of Kentucky, ningcao@yahoo.com

Recommended Citation
Cao, Ning, "RECOVERING LOCAL NEURAL TRACT DIRECTIONS AND RECONSTRUCTING NEURAL PATHWAYS IN HIGH ANGULAR RESOLUTION DIFFUSION MRI" (2013). *Theses and Dissertations--Computer Science.* Paper 13.
http://uknowledge.uky.edu/cs_etds/13

This Doctoral Dissertation is brought to you for free and open access by the Computer Science at UKnowledge. It has been accepted for inclusion in Theses and Dissertations--Computer Science by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

www.manaraa.com

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained and attached hereto needed written permission statements(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine).

I hereby grant to The University of Kentucky and its agents the non-exclusive license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless a preapproved embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's dissertation including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Ning Cao, Student

Dr. Jun Zhang, Major Professor

Dr. Miroslaw Truszczynski, Director of Graduate Studies

www.manaraa.com

RECOVERING LOCAL NEURAL TRACT DIRECTIONS AND RECONSTRUCTING NEURAL PATHWAYS IN HIGH ANGULAR RESOLUTION DIFFUSION MRI

DISSERTATION

A dissertation submitted in partial
fulfillment of the requirements for
the degree of Doctor of Philosophy
in the College of Engineering at the
University of Kentucky

By
Ning Cao
Lexington, Kentucky

Director: Dr. Jun Zhang, Professor of Computer Science
Lexington, Kentucky 2013

Copyright© Ning Cao 2013

www.manaraa.com

ABSTRACT OF DISSERTATION

RECOVERING LOCAL NEURAL TRACT DIRECTIONS AND RECONSTRUCTING NEURAL PATHWAYS IN HIGH ANGULAR RESOLUTION DIFFUSION MRI

Magnetic resonance imaging (MRI) is an imaging technique to visualize internal structures of the body. Diffusion MRI is an MRI modality that measures overall diffusion effect of molecules *in vivo* and non-invasively. Diffusion tensor imaging (DTI) is an extended technique of diffusion MRI. The major application of DTI is to measure the location, orientation and anisotropy of fiber tracts in white matter. It enables non-invasive investigation of major neural pathways of human brain, namely tractography. As spatial resolution of MRI is limited, it is possible that there are multiple fiber bundles within the same voxel. However, diffusion tensor model is only capable of resolving a single direction. The goal of this dissertation is to investigate complex anatomical structures using high angular resolution diffusion imaging (HARDI) data without any assumption on the parameters.

The dissertation starts with a study of the noise distribution of truncated MRI data. The noise is often not an issue in diffusion tensor model. However, in HARDI studies, with many more gradient directions being scanned, the number of repetitions of each gradient direction is often small to restrict total acquisition time, making signal-to-noise ratio (SNR) lower. Fitting complex diffusion models to data with reduced SNR is a major interest of this study. We focus on fitting diffusion models to data using maximum likelihood estimation (MLE) method, in which the noise distribution is used to maximize the likelihood. In addition to the parameters being estimated, we use likelihood values for model selection when multiple models are fit to the same data. The advantage of carrying out model selection after fitting the models is that both the quality of data and the quality of fitting results are taken into account. When it comes to tractography, we extend streamline method by using covariance of the estimated parameters to generate probabilistic tracts according to the uncertainty of local tract orientations.

KEYWORDS: diffusion-tensor MRI, high angular resolution MRI, tractography

www.manaraa.com

Author's signature: Ning Cao

Date: December 6, 2013

www.manaraa.com

RECOVERING LOCAL NEURAL TRACT DIRECTIONS AND RECONSTRUCTING NEURAL PATHWAYS IN HIGH ANGULAR RESOLUTION DIFFUSION MRI

By
Ning Cao

Director of Dissertation: Jun Zhang

Director of Graduate Studies: Miroslaw Truszczynski

Date: December 6, 2013

www.manaraa.com

This thesis is dedicated to Ruixue and Jeremy who have always stood by me
and dealt with my absence from many evenings and weekends
during the time when Jeremy was learning so many things.

www.manaraa.com

ACKNOWLEDGMENTS

I would first and foremost like to thank my advisor and graduate committee chair, Dr. Jun Zhang. My many discussions with Dr. Zhang provided invaluable advice on science and life.

I would also like to express my deep thanks to my committee members Fuhua (Frank) Cheng, Grzegorz W. Wasilkowski and Brian T. Gold for their support and helpful comments.

I also want to thank all that have provided datasets for us to evaluate our methods. Thanks to Jennifer Campbell of the McConnell Brain Imaging Center, Montreal Neurological Institute, McGill University, Canada for the phantom datasets. Thanks also to Ruiwang Huang of the MR Group, IME Research Center, Jülich, Germany, for the human brain datasets.

This work is supported in part by the U.S. National Science Foundation under grant CCF-0527967, in part by the Kentucky Science and Engineering Foundation under grant KSEF-148-502-06-186, and in part by Alzheimer's Association under grant NIGR-06-25460.

I would never have been able to finish this work without the support of my family.

www.manaraa.com

CONTENTS

Acknowledgments iii

Contents iv

List of Figures vi

List of Tables xii

List of Abbreviations xiii

Glossary xv

Chapter 1 Background 1
1.1 Diffusion-Weighted Imaging 1
1.2 Diffusion Tensor Imaging 4
1.3 High Angular Resolution Diffusion Imaging 8
1.4 Visualization 10
1.5 Motivation 11
1.6 Organization 12

Chapter 2 Noise Estimation 14
2.1 Introduction 14
2.2 Theory 15
2.3 Methods 21
2.4 Results 26
2.5 Discussion 35

Chapter 3 Apparent Diffusion Coefficient Estimation 37
3.1 Introduction 37
3.2 Methods 39
3.3 Results 46
3.4 Discussion 49

Chapter 4 Fitting Diffusion Models to Data 52
4.1 Introduction 52
4.2 Method 53
4.3 Results 63
4.4 Discussion 71

Chapter 5 Model Selection 76
5.1 Introduction 76
5.2 Method 77

www.manaraa.com

5.3 Results 81
5.4 Discussion 87

Chapter 6 Tractography 91
6.1 Introduction 91
6.2 Method 95
6.3 Results 98
6.4 Discussion 101

Chapter 7 Conclusion and Future Work 104
7.1 Research Accomplishments 104
7.2 Future Research Possibilities 105

Bibliography 108

Vita 117

www.manaraa.com

LIST OF FIGURES

2.1 The cumulative distribution function (CDF) of Rayleigh distribution with $\sigma = 5, 10,$ and 20. Horizontal dotted lines show the probability that a pixel intensity value will be found at a value not greater than 10, for different noise levels. 17
2.2 The discrete probability density function (PDF) Eq. 2.9 for different values of σ. The discrete PDF is shown using dots together with the Rayleigh PDF Eq. 2.3 for the same σ to show the difference between them. . . . 18
2.3 The first and second raw moments of the discrete distribution Eq. 2.9 for the same σ (solid line). The raw moments of the corresponding Rayleigh distribution Eq. 2.3 are shown as dashed line for comparison. 19
2.4 Difference between the true value and the noise parameters estimated by the mean estimator and the maximum likelihood (ML) estimator on rounded data. 20
2.5 Gradient directions used in producing the human brain data and the rat phantom data. (a) gradient directions derived from a truncated icosahedron with 60 vertices distributed on the unit sphere to acquire the human brain data; (b) gradient directions used to acquire the rat phantom data with 90 vertices scattered on a half hemisphere. The schema shown in (a) is also used in simulations. 22
2.6 Selected slices of the T_1-weighted image of the rat phantom showing the phantom configuration. Two of the spinal cords are put together to create a fiber crossing region. 23
2.7 Unweighted and diffusion-weighted images of the rat phantom of $b = 1300\,\mathrm{s/mm^2}$. Pixel intensities of the weighted image is much lower than that of the unweighted image. The two images are shown as grayscale images using different colormaps. 23
2.8 Underlying signal magnitude estimated from one noisy pixel intensity value, for different noise parameter σ and different noisy pixel intensity value. 26

www.manaraa.com

2.9 Multiple comparison result of the simulations of Rayleigh noise parameter estimation. The mean and 95% confidence interval are plotted for each simulation. Two means are considered different if their intervals are disjoint, and are considered not significantly different if the intervals overlap. The comparison lines help compare the interval of simulation (a) with that of other simulations. The input and estimator used for each simulation: (a) floating-point and the ML estimator, (b) floating-point and the mean estimator, (c) and (g) rounding down and the ML estimator, (d) and (h) rounding down and the mean estimator, (e) rounding to the nearest integer and the ML estimator, and (f) rounding to the nearest integer and the mean estimator. In (g) and (h), a correction of 0.5 is added to each rounded pixel intensity values before carrying out the computation. . . . 27
2.10 Noise parameters estimated from all diffusion-weighted images of the rat phantom. The dataset has four repetitions with each has 90 diffusion-weighted images, for both $b = 1300\,\mathrm{s/mm^2}$ and $b = 3000\,\mathrm{s/mm^2}$. For each gradient direction, noise parameters estimated from each repetition by assuming rounding down and rounding to the nearest integer are plotted using different markers. 28
2.11 Noise parameters estimated from four repetitions of the rat phantom shown as box plots for different b-values when rounding down is assumed. On each box of the box plots, the central mark is the median of values, the edges of the box are the 25th and 75th percentiles, the whiskers show the most extreme data points not considered outliers, and outliers are plotted individually using "+". 29
2.12 The PDF of the Rayleigh distribution with the noise parameter estimated from one diffusion-weighted image of the rat phantom for different rounding methods and different b-values together with the histograms of pixel intensity values. Both the PDF of the Rayleigh distribution and the corresponding discrete PDF for different rounding methods are plotted together for comparison. 30
2.13 Noise parameter estimated from all the images of the human brain data, for assuming both rounding methods. The first seven images are unweighted images and the rest are diffusion-weighted images. 31
2.14 The PDF with noise parameter estimated from one image of the human brain data together with the histogram of pixel intensity values for different rounding methods. 31
2.15 The SNR of diffusion-weighted images and T_1-weighted image of human brain data. The histograms of SNR of one unweighted image and that of one weighted image are plotted together in (a) for comparison. 32
2.16 Pixel intensity values and estimated signal intensity value of the human brain data. The estimated signal intensities as well as the 95% confidence intervals are shown using error bars. For the unweighted image, the signal intensity values are estimated from seven repetitions. While for the diffusion-weighted image, the signal intensity values are estimated from one measured intensity value. 34

www.manaraa.com

3.1 The second order and the third order tessellations of a regular icosahedron. (a) and (b) are the tessellation of a face; (c) and (d) are the tessellated regular icosahedrons. 43
3.2 The relative error in apparent diffusion coefficient (ADC) profiles when the spherical harmonics are truncated to different degrees to describe the biGaussian ADC profiles with different angle τ between the two tensor compartments. The errors on each gradient direction are used to create the boxes in the box plots. On each box of the box plots, the central mark is the median of values, the edges of the box are the 25th and 75th percentiles, the whiskers show the most extreme data points not considered outliers, and outliers are plotted individually using "+". (a) and (b) show the error in ADC profiles caused by Rician noise; (c) and (d) show the error in ADC profile caused by both noise and truncating the spherical harmonics to degree L. The figures shown here are for $b = 1000\,\mathrm{s/mm^2}$ and $S_0 = 500$. 48
3.3 The magnitude of the spherical harmonic expansion coefficients of an axial slice of the human brain data for $\ell = 0, 1,$ and 2, and $m = -\ell, \ldots, \ell$. The ranges of the coefficients of different ℓ are different, and the colormap for images of each ℓ is adjusted to get proper contract. (b) and (c) are made 5 and 10 times brighter so that the brightness is similar to that of (a). . 49
3.4 An axial slice of the root mean square error of the ADC profile reconstructed from spherical harmonics truncated to different L of the human brain data. 50

4.1 Comparison of different local optimization methods on fitting the biGaussian model to data from the same set of starting points. Local maximums returned by each method are plotted in a box plot. On each box of the box plot, the central mark is the median of values, the edges of the box are the 25th and 75th percentiles, the whiskers show the most extreme data points not considered outliers, and outliers are plotted individually using "+". All three methods returns multiple local maximums that have log-likelihood value similar to that of the global maximum. However, the active set method only returns a few good local maximums and the global maximum is returned by the interior point method. 58
4.2 Basins of attractions and local gradients of the negative log-likelihood function of fitting a diffusion tensor to synthetic data. 59
4.3 The crossing region of the rat phantom, marked with a rectangle on the fractional anisotropy (FA) map. 63
4.4 The crossing region of the human brain data, marked with a rectangle on a coronal slice of the FA map. Both the cingulum tract and the corpus callosum are included in the region of interest (ROI). 64

www.manaraa.com

4.5 Histograms of the diffusion coefficient D estimated from the isotropic diffusion model using MLE and the least squares (LS) method for estimating the variance of each estimator, obtained from bootstrapping. The normal distributions fitted to the histograms are shown using dotted line, and the normal distributions obtained from asymptotic normality for both methods are plotted using solid line. 65
4.6 Simulation result of fitting the diffusion tensor model to data using the MLE method and the LS method. For each b-value and σ combination, the same input is used to compare the two methods. 50 trials are used to create each box of the box plot. On each box of the box plots, the central mark is the median of values, the edges of the box are the 25th and 75th percentiles, the whiskers show the most extreme data points not considered outliers, and outliers are plotted individually using "+". (a) $\Lambda = (1200, 300, 300) \times 10^{-6}\,\mathrm{mm}^2$ /s, $b = 2000\,\mathrm{s/mm}^2$, and $\sigma = 4, 6,$ and 8; (b) $\Lambda = (1200, 400, 300) \times 10^{-6}\,\mathrm{mm}^2$ /s, $b = 1000, 2000,$ and $3000\,\mathrm{s/mm}^2$, and $\sigma = 6$. 66
4.7 Simulation results of fitting the biGaussian model to data using the MLE approach. Box plots are shown for different b-values, different tensor shapes, and different number of parameters. On each box of the box plots, the central mark is the median of values, the edges of the box are the 25th and 75th percentiles, the whiskers show the most extreme data points not considered outliers, and outliers are plotted individually using "+". 67
4.8 Noisy ADC profiles within the crossing region shown in Fig. 4.3 of the rat phantom to show the difference of different repetitions caused by noise and imaging artifacts. 68
4.9 Fitting different diffusion models to the crossing region of the rat phantom data. The crossing region is shown in Fig. 4.3. (a) Noisy ADC profiles of the crossing region; (b) ADC profiles recovered from the diffusion tensor model, together with the principal directions of the diffusion tensors; (c) ADC profiles recovered from biGaussian model with 10 parameters, together with the principal directions of each tensor compartments; (d) ADC profiles recovered from biGaussian model with 13 parameters, together with the principal directions of each tensor compartments. The ADC profiles of some voxels are scaled to show the details with good contrast. 69
4.10 Log-likelihood values of fitting the diffusion tensor model, the biGaussian model with 10 parameters and the biGaussian model with 13 parameters to the crossing region of the rat phantom shown in Fig. 4.3. Log-likelihood values of the three diffusion models are plotted for each voxel within the ROI as horizontal bars from top to bottom. 70

www.manaraa.com

4.11 The variance shown as PDF of the marginal distributions of parameters estimated from the diffusion tensor model. The parameters shown in (a) and (b) are estimated from a voxel within a single spinal cord of the rat phantom, and the parameters shown in (c) and (d) are estimated from a voxel within the crossing region of the rat phantom. The marginal distributions are obtained from the covariance matrix estimated from the asymptotic covariance of the ML estimator of the diffusion tensor model. 72
4.12 The variance shown as PDF of the marginal distributions of parameters estimated from a voxel of the rat phantom containing two tract crossing, using the biGaussian model with 10 parameters. The marginal distributions are obtained from the covariance matrix estimated from the asymptotic covariance of the ML estimator of the diffusion tensor model. . . . 73
4.13 Fitting different diffusion models to the crossing region of the human brain data. The crossing region is shown in Fig. 4.4. (a) Noisy ADC profiles of the crossing region; (b) ADC profiles recovered from the diffusion tensor model, together with the principal directions of the diffusion tensors; (c) ADC profiles recovered from biGaussian model with 10 parameters, together with the principal directions of each tensor compartments; (d) ADC profiles recovered from biGaussian model with 13 parameters, together with the principal directions of each tensor compartments. The ADC profiles of some voxels are scaled to show the details with good contrast. 74
4.14 Log-likelihood values of fitting the diffusion tensor model, the biGaussian model with 10 parameters and the biGaussian model with 13 parameters to the crossing region of the human brain data shown in Fig. 4.4. Log-likelihood values of the three diffusion models are plotted for each voxel within the ROI as horizontal bars from top to bottom. 75

5.1 Log-likelihood value obtained by fitting the isotropic model and the diffusion tensor model to synthetic data with different FA value. 82
5.2 Simulation results of likelihood ratio test between the isotropic model and the diffusion tensor model. The p-value of likelihood ratio test between the two models is plotted versus FA value. The log-likelihood values used to calculate the test statistic are shown in Fig. 5.1. 83
5.3 Log-likelihood value obtained by fitting the diffusion tensor model and the biGaussian model with 10 parameters to synthetic data with different angle τ between the two tensor compartments used to generate the biGaussian profiles. The log-likelihood value for each different τ are shown as box plots. On each box of the box plots, the central mark is the median of values, the edges of the box are the 25th and 75th percentiles, the whiskers show the most extreme data points not considered outliers, and outliers are plotted individually using "+". 84

www.manaraa.com

5.4 The p-values of likelihood ratio test between the biGaussian model with 10 parameters and the diffusion tensor model. The p-value of test statistic is plotted versus τ. Trials for $\tau = 30, 60,$ and 90° are plotted using different markers. 85
5.5 Simulation result of likelihood ratio test between the biGaussian models with 13 parameters and 10 parameters. The p-value of the test statistic is plotted versus FA value. Trials for $\tau = 30, 60,$ and 90° are plotted using different markers. 86
5.6 Likelihood ratio test between the diffusion tensor model and the isotropic diffusion model on a slice of the rat phantom data overlaid on FA map. Result of both likelihood ratio test with and without Bonferroni correction are shown. The white rectangles define the ROI where the test is carried out. Voxels where the isotropic diffusion model is rejected are marked with a white dot. 87
5.7 Model selection results of the rat phantom on the crossing region defined in Fig. 4.3. (a) shows the result of likelihood ratio test between the diffusion tensor model and the biGaussian model with 10 parameters; (b), (c), and (d) show model selection results on different criteria. 88
5.8 Model selection results of the human brain data on the crossing region defined in Fig. 4.4. (a) shows the result of likelihood ratio test between the diffusion tensor model and the biGaussian model with 10 parameters; (b), (c), and (d) show model selection results on different criteria. . . . 89

6.1 Fiber tracts reconstructed from the streamline method and the whole-brain tracking method. The two seed points used for the streamline method are marked as "×" and the ROI containing all seed points used for whole-brain tracking is shown as a rectangle. The tracts are overlaid with FA map and the two directions of a tract from the seed point are plotted in different colors. 95
6.2 Local direction uncertainty of an anisotropic voxel within a single fiber tract of the rat phantom. Local directions are plotted in spherical coordinates. 99
6.3 Local direction uncertainty of fitting a voxel containing fiber crossing structure to the diffusion tensor model and the biGaussian model. Local directions are plotted in spherical coordinates. 100
6.4 Fiber tracts of the rat phantom reconstructed by the streamline method with random sampling. The tracts are overlaid with FA map and the seed points are marked as "×". The two directions of a tract from the seed point are plotted in different colors. 101
6.5 Fiber tracts of the human brain data reconstructed by the streamline method with random sampling. The tracts are overlaid with FA map and the seed points are marked as "×". The corpus callosum tract is plotted on a coronal slice and the cingulum tracts are plotted on sagittal slices. The two directions of a tract from the seed point are plotted in different colors. 102

www.manaraa.com

LIST OF TABLES

2.1 Simulation result of noise reduction using MLE with the PDF of the Rician distribution Eq. 2.1 and discrete PDF Eq. 2.15 or Eq. 2.16. 34

3.1 Rank and condition number of $M^T M$ 41
3.2 The absolute error of $I_c(Y_5^m)$ 45

www.manaraa.com

List of Abbreviations

ADC apparent diffusion coefficient. 4, 10, 35, 37–39, 45, 46, 48, 49, 51, 54, 55, 64, 65, 68–71, 73, 75, 90–92, 110, 111

AIC Akaike information criterion. 82, 83, 90–92

AICc Akaike information criterion with correction. 82, 83, 90–92

ANOVA analysis of variance. 21

BIC Bayesian information criterion. 82, 90–92

CDF cumulative distribution function. 15–17, 20, 58

DOF degrees of freedom. 81, 88

DTI diffusion tensor imaging. iii, 1, 4, 5, 7, 9–11, 37

DWI diffusion-weighted imaging. 1–4, 14, 34, 35, 76, 109

FA fractional anisotropy. 7, 10, 65, 84–86, 88, 89, 96, 98, 99, 104–106

HARDI high angular resolution diffusion imaging. iii, 3, 4, 10, 11, 37–39, 41, 48, 52, 86, 109

HODT high order diffusion tensor. 8, 37

LS least squares. 38, 41, 45, 52, 55, 66–68

MAP maximum a posteriori probability. 55, 76

MD mean diffusivity. 4, 7, 9, 10

ML maximum likelihood. 16, 17, 19, 21, 23, 26, 34, 58, 70, 72–74, 101, 102

MLE maximum likelihood estimation. iii, 12, 20, 21, 25, 32, 33, 35, 55, 56, 62, 63, 66–68, 79, 80, 82, 99, 100

MR magnetic resonance. 14, 16, 17, 21, 96

MRI magnetic resonance imaging. iii, 1, 8–10, 14, 35, 38, 110

PDF probability density function. 14, 15, 17, 18, 24, 25, 27, 29, 31, 33, 57, 58, 72–74, 100

PVE partial volume effect. 11, 103, 110

www.manaraa.com

RA relative anisotropy. 7

ROI region of interest. 64, 65, 84, 89, 90, 96, 98

SHT spherical harmonic transform. 38, 39, 48, 52

SNR signal-to-noise ratio. iii, 7, 8, 10, 14, 15, 20, 30–32, 35, 45, 46, 48–50, 56, 57, 96, 97, 101, 110

www.manaraa.com

Glossary

A Signal amplitude. 14, 15, 20, 25, 33, 57

D Diffusion coefficient. 3–5, 56, 66, 67

I Integral of a function over a sphere. 39, 40, 42–45

K Number of tensor compartments in Gaussian Mixture model. 9, 57

L Degree to truncate spherical harmonics. 38–41, 45, 46, 48–52

S_0 Unweighted signal. 3, 4, 8, 9, 54, 55, 57, 58, 70

S Diffusion-weighted signal. 3, 4, 8, 9, 54, 55, 57, 58

Y Spherical harmonic series. 37–39, 44, 45

Λ Diagonal matrix of eigenvalues. 5

α Significant level. 32, 81, 84, 85

ℓ Degree of spherical harmonics. 37–39, 44, 45, 48, 51

η_{err} The angle error. 64, 67, 69

λ Eigenvalue of the diffusion tensor. 5, 7, 73, 83, 84, 102

$\mathcal{I}$ Fisher information matrix. 62

$\mathcal{L}$ Likelihood function. 57, 58, 62, 81, 82

D Test statistic of likelihood ratio test. 81

L Lower triangular matrix of the Cholesky decomposition. 101

$\mathsf{\Sigma}$ Covariance matrix. 62, 101

ν Distance between the reference point and the center of Rician distribution. 14, 15, 55, 56, 58

ϕ Azimuthal angle in the x y-plane from the x-axis. 2, 37–41, 45

ψ Euler Angle ψ. 6, 56

σ The parameter of Rayleigh distribution; the scale parameter of Rician distribution. 14–21, 23–27, 35, 55–58, 63, 68, 83

τ The angle between principal directions of the two tensors of biGaussian model. 9, 48–50, 64, 67, 69, 83, 86–88

www.manaraa.com

D Diffusion tensor. 4, 5, 7–9, 99

M A diffusion model. 54, 55, 57, 58

N Normal distribution. 62

P The set of all allowable values for the parameter. 54, 55, 57, 58, 76

R Rotation matrix. 5, 6

$\mathbf{S}_0$ Unweighted image. 3

S Diffusion-weighted image. 57

f The volume ratio of each tensor compartment in Gaussian Mixture model. 8, 9, 56

p The parameter vector of a model. 54–59, 62, 101

θ Polar angle from the z-axis. 2, 37–41, 45

φ Euler Angle φ. 6, 56

ϑ Euler Angle ϑ. 6, 56

$\vec{e}$ Eigenvector of the diffusion tensor. 5, 6, 9, 64

$\vec{g}$ Gradient. 2, 4, 5, 8, 9, 38, 54, 55, 57, 58, 63

c Expansion coefficient of spherical harmonic series. x, 37–39, 45

m Order of spherical harmonics. 37–39, 44, 45

n Number of gradient directions. 2, 39–41, 44, 45, 54–58

q A function used in calculating c. 39, 40, 42–45

r Number of acquisitions. 55–57

w Weight in quadrature formulas. 40, 42, 43, 45

d ADC function. 4, 38, 39

www.manaraa.com

Chapter 1 Background

In this chapter we give a literature review of diffusion-weighted imaging (DWI) and diffusion tensor model to introduce the background of the research work of this thesis.

Introduced in the mid-1980s (Le Bihan, 1986), diffusion magnetic resonance imaging (MRI) is a specific MRI modality that measures the overall diffusion effect of water molecules or some other metabolites *in vivo*. Diffusion anisotropy of water molecules in white matter of human brain (Chenevert et al., 1990) was reported in early 1990s.

Diffusion tensor imaging (DTI) is an extended technique of diffusion MRI. It is important because it enables the non-invasive investigation of major neural pathways of human brain, namely tractography. The tractography can be used for visualization and analyses of human brain data. And, diffusion anisotropy has proven useful when studying pathologies related to changes in white matter tissues (Pfefferbaum and Sullivan, 2003; Pagani et al., 2005; Chua et al., 2008). More complex diffusion models have been proposed as well. However, DTI is still important because it supports many important applications of human brain studies and the model only has six parameters, which can be solved by conventional mathematical methods.

1.1 Diffusion-Weighted Imaging

Diffusion MRI produces images of biological tissues weighted with local characteristics of water diffusion (Le Bihan et al., 2002). New techniques have been put forward to overcome the insufficiency of the old ones. The plain diffusion MRI can only characterize isotropic diffusion, and DTI can describe Gaussian diffusion (Basser et al., 1994b).

The signal attenuation observed in a voxel of diffusion MRI image reflects the

www.manaraa.com

overall effects of water diffusion within that voxel (Le Bihan et al., 2002). With anisotropic diffusion, the effects of water diffusion vary when the gradient is applied at different directions. The gradient direction, or the gradient, is represented by a unit vector $\vec{g}$, which is always written in Cartesian coordinates. In some studies (Frank, 2002; Özarslan and Mareci, 2003; Descoteaux et al., 2006), it is more convenient to express the unit vector $\vec{g}$ in spherical coordinate (ρ, θ, ϕ) with $\rho = 1$:

$$\vec{g} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}, \tag{1.1}$$

where θ and ϕ represent the polar angles ($0 \leq \theta \leq \pi$) from the z-axis and the azimuthal angle ($0 \leq \phi < 2\pi$) in the x y-plane from the x-axis. This notation is called physics convention, while in the mathematics convention as used in some of the references, θ and ϕ are reversed. An illustration of the spherical coordinate system in the physics convention is drawn by Frank (Frank, 2002).

We use $\vec{g}_i$, $i = 1, ..., n$, to denote the n gradients actually applied in DWI, and use $\vec{g}$ to represent an arbitrary direction on the unit sphere. The signal attenuation obtained along each gradient direction $\vec{g}_i$ can be viewed as samples of the diffusion profile.

Anisotropic Diffusion in the Human Brain

In human brain gray matter, the measured diffusivity is independent of the orientation of the tissue (Basser and Jones, 2002). But for tissues such as cardiac muscle and brain white matter, the measured diffusivity depends upon the orientation of the tissue (Basser and Jones, 2002). Water diffusion in the brain is not isotropic because axonal membranes and myelin sheaths of white matter fibers running in parallel prevent the motion of water molecules in directions other than along their own orientation (Le Bihan et al., 2001).

www.manaraa.com

Plain Diffusion Imaging

The amount of the diffusion signal loss with gradient applied is described by the Stejskal-Tanner formula (Stejskal and Tanner, 1965):

$$S = S_0 e^{-bD}, \tag{1.2}$$

where

$$b = \gamma^2 G^2 \delta^2 (\Delta - \delta/3) \tag{1.3}$$

is often referred to as the b-value. In Eq. 1.2, S is the signal attenuation with the diffusion gradient, S_0 is the signal attenuation without the diffusion gradient, γ is the gyromagnetic ratio, G is the strength of the gradient pulse, δ is the duration of the pulse, Δ is the time between the two pulses, and D is the diffusion constant (Mori and Barker, 1999).

When only isotropic diffusion is involved, diffusion constant D is independent of gradient direction. In plain diffusion imaging, diffusion is fully represented by diffusion constant D, which is a scalar. The diffusion constant D of each voxel can be mapped to create an apparent diffusion constant image (Mori and Barker, 1999).

Diffusion-Weighted Imaging

In anisotropic water diffusion, the amount of diffusion signal loss depends on the direction of the gradient applied. In DWI, the signal loss is measured with different gradients applied.

The images produced with gradient applied in DWI are called diffusion-weighted images, or weighted images, bearing the same acronym as the technique of DWI. In DWI, signals without gradient applied are also measured. The images of unweighted signals are called unweighted images, often referred to as the $\mathbf{S}_0$ images. When a large number of gradient directions are used to increase the angular resolution so that the complex detail of the diffusion profile can be captured, the technique is often called

www.manaraa.com

high angular resolution diffusion imaging (HARDI) (Tuch et al., 1999; Frank, 2001; Tuch et al., 2002).

Similar to the diffusion constant D in Eq. 1.2, we can calculate diffusion coefficient D_i at each gradient direction by

$$D_i = -\frac{1}{b} \ln \frac{S_i}{S_0}, \tag{1.4}$$

and S_i/S_0 is called spin-echo attenuation. The signal strength S meansured at any gradient direction and its corresponding diffusion coefficient D can be considered as functions of gradient direction $\vec{g}$. Therefore we have $S_i = S(\vec{g}_i)$, and we define apparent diffusion coefficient (ADC) profile (Alexander et al., 2002) as

$$d(\vec{g}) = -\frac{1}{b} \ln \frac{S(\vec{g})}{S_0}, \tag{1.5}$$

The estimation of $d(\vec{g})$ from D_i may be improved by using more gradient directions. The ADC profile can therefore be visualized in spherical coordinate (Alexander et al., 2002; Tuch et al., 2002; Özarslan and Mareci, 2003) and the surface may also be colored by function value (Frank, 2001; Descoteaux et al., 2006; Frank, 2002).

In traditional or "plain" DWI, the trace of $\mathbf{D}$, or the mean diffusivity (MD), is estimated by applying three gradients. It can be a useful technique for diagnosing vascular strokes in its acute stage (Moseley, 1990), but the MD does not make it possible to describe anisotropy. DWI is also the basis for more advanced diffusion models or imaging techniques, such as DTI and HARDI. At least six weighted images are needed in DTI studies in order to estimate the diffusion tensor of each voxel.

1.2 Diffusion Tensor Imaging

The plain diffusion imaging measures the diffusion constant, making it sufficient for isotropic diffusion. For general diffusion situations, the plain diffusion model is insufficient. Assuming that the anisotropic diffusion may be characterized with a Gaussian

www.manaraa.com

function, Basser et al. proposed to use diffusion tensor to describe the local diffusion and introduced DTI (Basser et al., 1994b), by using the following formula:

$$S(\vec{g}) = S_0 e^{-b\vec{g}^T \mathbf{D}\vec{g}}, \tag{1.6}$$

where $\mathbf{D}$ is a symmetric positive definite tensor,

$$\mathbf{D} = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{pmatrix}. \tag{1.7}$$

In Eq. 1.6, the term $\vec{g}^T \mathbf{D}\vec{g}$ replaced the diffusion coefficient D in Eq. 1.2.

Diffusion tensor can be computed from diffusion-weighted signals obtained from six different gradient directions and the unweighted signals, by solving the linear system defined by taking the natural logarithm on both sides of Eq. 1.6. Using more gradient directions to get a least squares (LS) solution will in general improve the accuracy of the tensor estimation (Jones et al., 1999; Papadakis et al., 2000; Hasan et al., 2001). When the noise level is high, non-linear LS is preferred to linear LS (Kingsley, 2006b), because the log transform would make the symmetric noise not symmetric on the log scale.

DTI provides information about both the extent of diffusion anisotropy and its orientation.

Diffusion Tensor Diagonalization

Because $\mathbf{D}$ is symmetric and positive definite, it can be rewritten in terms of its eigenvalues λ_1, λ_2, and λ_3 ($\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$), and the corresponding eigenvectors $\vec{e}_1$, $\vec{e}_2$, and $\vec{e}_3$ as (Basser et al., 1994b; Basser and Pierpaoli, 1996):

$$\mathbf{D} = \mathbf{R}\Lambda\mathbf{R}^{-1}, \tag{1.8}$$

with orthogonal matrix

$$\mathbf{R} = \begin{pmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \end{pmatrix} \tag{1.9}$$

www.manaraa.com

and

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}. \tag{1.10}$$

The eigenvector associated with the largest eigenvalue, $\vec{e}_1$, is often called the principal eigenvector. It is generally believed that local fiber orientation is parallel to eigenvector $\vec{e}_1$.

According to Euler's rotation theorem, any rotation may be described by only three angles. In the so called "x-convention", the rotation is given by Euler angles $(\varphi, \vartheta, \psi)$, where φ is a rotation about the z-axis, $\vartheta \in [0, \pi]$ is a rotation about the former x-axis, and ψ is a rotation about the former z-axis again. The rotation $\mathbf{R}$ can be written as

$$\mathbf{R} = \mathbf{R}_\psi \cdot \mathbf{R}_\vartheta \cdot \mathbf{R}_\varphi, \tag{1.11}$$

where each matrix contains a single Euler angle and the component rotations are given by

$$\mathbf{R}_\varphi = \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{1.12}$$

$$\mathbf{R}_\vartheta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\vartheta & \sin\vartheta \\ 0 & -\sin\vartheta & \cos\vartheta \end{pmatrix}, \tag{1.13}$$

and

$$\mathbf{R}_\psi = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1.14}$$

The Euler angles that decide a rotation matrix $\mathbf{R}$ is not unique. Two groups of Euler angles can be extracted from the rotation matrix $\mathbf{R}$, if $\mathbf{R}_{3,3} \equiv \cos\vartheta \neq \pm 1$. Otherwise,

www.manaraa.com

the rotation matrix $\mathbf{R}$ reduces to a 2D rotation, from which only $\varphi + \psi$ or $\varphi - \psi$ can be decided.

Anisotropy Measures

Diffusion anisotropy indices can be used to quantify the amount of anisotropic diffusion. Most popular anisotropy indices are defined in the diffusion tensor model, although more complex diffusion models have also been used (Frank, 2002; Alexander et al., 2002; Özarslan and Mareci, 2003; Chen et al., 2004, 2005).

MD (Basser et al., 1994b) is the earliest anisotropy measure, which is defined as

$$MD = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3}. \tag{1.15}$$

Fractional anisotropy (FA) (Basser and Pierpaoli, 1996) is the most widely reported diffusion anisotropy index. Relative anisotropy (RA) (Basser and Pierpaoli, 1996) is also used in studies regarding anisotropy. They are defined as follows:

$$FA = \sqrt{\frac{3}{2}\frac{(\lambda_1 - \bar{\lambda})^2 + (\lambda_2 - \bar{\lambda})^2 + (\lambda_3 - \bar{\lambda})^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}, \tag{1.16}$$

$$RA = \frac{\sqrt{(\lambda_1 - \bar{\lambda})^2 + (\lambda_2 - \bar{\lambda})^2 + (\lambda_3 - \bar{\lambda})^2}}{\sqrt{6}\bar{\lambda}}, \tag{1.17}$$

where

$$\bar{\lambda} = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3}. \tag{1.18}$$

For an anisotropy index to characterize the degree of anisotropy in DTI, it should be a monotonic function of some physical quantities, and should be a translation and rotation invariant (Basser and Pierpaoli, 1996). It can be calculated from the eigenvalues of diffusion tensor $\mathbf{D}$ (Kingsley, 2006a; Skare et al., 2000).

www.manaraa.com

Limitations of DTI

DTI is an inherently low-signal-to-noise ratio (SNR) technique (Taylor et al., 2004; Parker, 2004) with low spatial resolution, where each voxel may contain tens of thousands of axons.

DTI tractography corresponds well to major fiber pathways such as the corpus callosum, but it does not work well in some other fibers (Johansen-Berg and Behrens, 2006). It suffers from two types of errors: the first type of error is caused by spatial resolution limitation, image noise, artifacts and accumulation of numerical errors; the second type of error is due to the fact that diffusion tensor model is incapable of describing non-Gaussian diffusion (Tuch et al., 1999, 2002).

1.3 High Angular Resolution Diffusion Imaging

The current practical voxel size of diffusion MRI is around 2 mm, which is much larger than the diameter of neural axons. Therefore it is normal that there are multiple fiber tract orientations within a given voxel. Non-Gaussian diffusion is also observed at white matter structures, like the corona radiata lateral to the lateral ventricle (Tuch et al., 1999). The diffusion tensor model is not capable of describing these non-Gaussian diffusion. As it is much harder to reduce the voxel size than to increase the number of gradients due to SNR issue, sampling with increased angular resolution of the acquisition schema is therefore proposed to fully capture the details of non-Gaussian diffusion (Tuch et al., 1999).

Spherical Harmonics and Higher Order Tensor

Methods to describe the non-Gaussian diffusion profile include spherical harmonics (Frank, 2002; Alexander et al., 2002; Chen et al., 2004) and high order diffusion tensor (HODT) (Özarslan and Mareci, 2003). Actually, there is an analytical relationship between the spherical harmonics coefficients and the elements of the

www.manaraa.com

HODT (Özarslan and Mareci, 2003; Descoteaux et al., 2006). These model-free methods are useful for quantifying anisotropy (Frank, 2002; Özarslan and Mareci, 2003; Özarslan et al., 2005), however fiber orientations cannot be recovered directly, given that there is no simple correspondence between fiber tract orientations and the maxima of the diffusion profile in non-Gaussian diffusion.

BiGaussian and Gaussian Mixture Model

The biGaussian model is proposed to assess two tract orientations within the same voxel (Alexander et al., 2001; Tuch et al., 2002; Parker and Alexander, 2003) with little or no exchange between the compartments:

$$S(\vec{g}) = S_0(\mathbf{f}e^{-b\vec{g}^T\mathbf{D}_1\vec{g}} + (1-\mathbf{f})e^{-b\vec{g}^T\mathbf{D}_2\vec{g}}), \tag{1.19}$$

where $\mathbf{f} > 0$ is the volume ratio. As each compartment of Eq. 1.19 is a Gaussian diffusion, the model is a weighted mixture of two Gaussian compartments. When with exchange between the two compartments, the corresponding model (Alexander et al., 2001) is

$$S(\vec{g}) = S_0 e^{-b\vec{g}^T(\mathbf{f}\mathbf{D}_1+(1-\mathbf{f})\mathbf{D}_2)\vec{g}}. \tag{1.20}$$

However, for two distinct tissues within a single voxel, Eq. 1.19 is more reasonable to describe the effects than Eq. 1.20 (Alexander et al., 2001), which does not work well with large b-value (Frank, 2001).

The angle between the principal directions of the two tensor compartments in biGaussian model is given by

$$\tau = \arccos(|\vec{e}_1 \cdot \vec{e}_1'|), \tag{1.21}$$

where $\vec{e}_1$ and $\vec{e}_1'$ are the two principal eigenvectors of the two tensor compartments.

www.manaraa.com

The generalization of biGaussian model Eq. 1.19 is Gaussian mixture model with K Gaussian compartments:

$$S(\vec{g}) = S_0 \sum_{i=1}^{K} \mathbf{f}_i e^{-b\vec{g}^T \mathbf{D}_i \vec{g}}, \tag{1.22}$$

where volume ratio $\mathbf{f}_i > 0$ and is subject to

$$\sum_{i=1}^{K} \mathbf{f}_i = 1. \tag{1.23}$$

1.4 Visualization

Visualization of diffusion-weighted data is both the objective of image processing and a useful tool to explore the three-dimensional images.

MRI images can be visualized as scalar images. T_1-weighted images are usually displayed using signal magnitude directly. Anisotropy indices are often used to visualize diffusion-weighted images or diffusion tensor images. Traditionally, diffusion-weighted image is shown as an MD map. In DTI studies, diffusion tensor image is often displayed as an FA map. When displaying scalar images, we can use different colormaps to scale or enhance the data.

Vector field can be used to visualize principal eigenvectors of diffusion tensors. A small arrow, or just a small line segment, is sometimes used to show the fiber orientation recovered from each voxel (Peled et al., 1998). When Gaussian mixture model is used, multiple glyphs can be used in the same voxel to present multiple orientations. The length of the arrow or line segment can also be used to encode a scalar, such as FA. Similarly, a diffusion colormap combines the scalar image and a vector field by encoding the orientation of principal eigenvector and FA of each voxel in RGB space, providing more information about white matter fiber tracts (Douek et al., 1991; Pajevic et al., 1999). Some major nerve pathways can even be easily identified in diffusion colormap. When accurate diffusion colormap is desired, each

www.manaraa.com

individual's diffusion tensor data should be registered to standard space, whereas MD and FA are rotationally invariants and can be used for comparison directly.

More complicated glyphs are also used to visualize diffusion profiles (Westin et al., 2002). For diffusion tensor model, diffusion ellipsoid is a standard method to visualize tensors (Basser et al., 1994a,b; Pierpaoli et al., 1996), as a tensor can be fully encoded by the ellipsoid. As mentioned above, for HARDI data, a peanut-like plot is also used to visualize the ADC profile.

1.5 Motivation

Diffusion MRI is a low-SNR modality therefore we need to pay attention to the noise. Some methods used in DTI and HARDI studies assume that the noise is normally distributed in diffusion-weighted images. The Gaussian approximation of Rician noise is good for high SNR value (Gudbjartsson and Patz, 1995). However, Rician noise introduces a significant bias into diffusion-weighted signals if the noise level is high. The noise in diffusion-weighted signals can be regarded as Gaussian noise, if Rician noise removal is carried out and the number of repetitions is large. But the number of repetitions in typical DTI studies and most HARDI studies is small. Therefore we need investigate the impact of Rician noise. We also want to verify the noise distribution of diffusion-weighted images because it is reported that the distribution is hard to verify (Gudbjartsson and Patz, 1995).

Better tractography results help neurosurgical planning and provide better understanding of the brain (Wakana et al., 2004; Mori and Wakana, 2005). However, partial volume effect (PVE) has negative impact on both tractography algorithms and anisotropy measurements. Diffusion-weighted images often have relative large voxel sizes compared to the diameters of some fiber tracts. Therefore voxel-based group studies on diffusion-weighted data are often affected by PVE (Cao et al., 2008). Tractography algorithms developed for diffusion-tensor model often fail to pass regions

www.manaraa.com

containing crossing fiber tracts. Therefore, we would like to fit Gaussian mixture model to diffusion-weighted data so that anisotropy measurements defined for diffusion tensor model can be estimated more accurately and crossing fiber tracts can be reconstructed.

1.6 Organization

The dissertation is composed of seven chapters. The following chapters are organized as follows:

- In Ch. 2, the research focuses on Rician noise of diffusion-weighted images. We study the noise distribution of image intensity values stored as 12-bit integers, which are truncated or rounded from the image magnitudes. Noise distribution verification is carried out using phantom data as the noise distribution is the basis of remaining chapters.

- In Ch. 3, we present a fast method computing spherical harmonic for several widely used gradient schema. Spherical harmonic is an important representation of functions on the sphere, which are widely used in HARDI.

- In Ch. 4, we fit different diffusion models to data using maximum likelihood estimation (MLE) method. This chapter is based on noise distribution discussed in Ch. 2. When fitting models to data, we do not use any feedback from tracking algorithms or any prior distribution of the parameters.

- Ch. 5 focuses on choosing a better diffusion model from two candidate models, or selecting the best model from a set of candidate models. Instead of selecting a complex model when a simple model appears to fail, fitting result of all candidate models are evaluated and a best model, in the sense of both accuracy and simplicity, is selected from two or more models.

www.manaraa.com

- In Ch. 6, we use the result of Chs. 4 and 5 to reconstruct fiber pathways. We extend the streamline method by considering the covariance matrix of parameters estimated in Ch. 4. This method allows us to reconstruct possible fiber pathways based on the quality of diffusion-weighted data.

- Ch. 7 provides the conclusion of this dissertation and some possible directions for future research work.

Copyright© Ning Cao, 2013.

www.manaraa.com

Chapter 2 Noise Estimation

In DWI, SNR is always a major concern, as diffusion weighting makes the signal magnitude lower than that of a typical T_1-weighted image. High diffusion weighting factor and high spatial resolution improve the performance of tractography algorithms of DWI, but they also decrease the SNR.

Noise estimation and reduction in MRI data has been studied intensively (Gudbjartsson and Patz, 1995; Andersen, 1996; Sijbers et al., 1998a,b, 1999; Karlsen et al., 1999; Sijbers and Den Dekker, 2004; Basu et al., 2006; Kristoffersen, 2007). But most of the studies overlook the effect of the quantization process in producing MRI data. In this chapter, we focus on the effects of rounding error introduced by the quantization process.

2.1 Introduction

It is known that the pixel intensity values measured in magnetic resonance (MR) images are Rician distributed (Edelstein et al., 1984; Bernstein et al., 1989; Brummer et al., 1993). The probability density function (PDF) of the Rician distribution or Rice distribution with parameters ν and σ is

$$f(x;\nu,\sigma) = \frac{x}{\sigma^2} e^{-\frac{x^2+\nu^2}{2\sigma^2}} I_0\left(\frac{x\nu}{\sigma^2}\right), \tag{2.1}$$

where $\nu \geq 0$ is the underlying noise-free signal amplitude, the scalar parameter $\sigma > 0$ is the standard deviation of normal distribution in both the real and imaginary signals, and $I_0(x)$ is the modified Bessel function of the first kind of order zero. The scale parameter σ is decided by the imaging hardware and protocol, and can be considered as a constant for the whole image.

In the air background of an MR image, the signal strength ν is 0. The Rician distribution reduces to the Rayleigh distribution with the scale parameter σ.

www.manaraa.com

The common method to estimate the "true" signal amplitude A of each pixel in an MR image is to first estimate the scale parameter σ of the whole image from some pixels of the air background, then use measured values of the same pixel to estimate A with the scale paraemter σ being fixed.

To simplify the calculation, we define the local SNR of a pixel as the ratio of signal amplitude and the scale parameter σ of the Rician distribution (Gudbjartsson and Patz, 1995):

$$\text{SNR} = \frac{A}{\sigma}. \tag{2.2}$$

2.2 Theory

Parameter Estimation of the Rayleigh Distribution

The PDF of the Rayleigh distribution is

$$f(x;\sigma) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, \tag{2.3}$$

for $x \geq 0$ and $\sigma > 0$. It is the special case of the Rician distribution Eq. 2.1 given by $\nu = 0$. The cumulative distribution function (CDF) of the Rayleigh distribution is

$$F(x;\sigma) = 1 - e^{-\frac{x^2}{2\sigma^2}}. \tag{2.4}$$

The mean or the first raw moment of the population is

$$\int_0^\infty x f(x;\sigma)\ dx = \sigma\sqrt{\frac{\pi}{2}}, \tag{2.5}$$

and the second raw moment of the population is

$$\int_0^\infty x^2 f(x;\sigma)\ dx = 2\sigma^2. \tag{2.6}$$

The population mean can be estimated from the sample mean of a set of N observations $\{m_1, \ldots, m_N\}$. The parameter σ can therefore be estimated using Eq. 2.5 from the sample mean (Gudbjartsson and Patz, 1995) as

$$\hat{\sigma} = \sqrt{\frac{2}{\pi}} \frac{1}{N} \sum_{i=1}^{N} m_i, \tag{2.7}$$

www.manaraa.com

or using Eq. 2.6 from the second moment (Sijbers et al., 1999; Sijbers and Den Dekker, 2004; He and Greenshields, 2009) as

$$\hat{\sigma} = \sqrt{\frac{1}{2N}\sum_{i=1}^{N} m_i^2}, \tag{2.8}$$

which is the maximum likelihood (ML) estimator of the parameter σ.

The Quantization of MR Data

A good estimation of the parameter σ of the Rayleigh distribution is prerequisite to recovering the underlying signal magnitudes from the noise occupied pixel intensity values.

Although the floating-point representation is preferred for storing the signal magnitudes in terms of precision, the pixel intensity values are often stored in 12-bit integers. In this chapter, we consider two ways of rounding a non-negative floating-point number to an integer: rounding down and rounding to the nearest integer. Rounding down, or taking the floor, is denoted as $\lfloor x \rfloor$, and rounding to the nearest integer, or taking the nearest integer function is denoted as $[x]$. For half-integers, rounding half up, rounding half down and rounding to the nearest even integer do not make much difference in MR images.

The quantization error is large for small intensity values, especially when rounding down is used. The CDF of the Rayleigh distribution with $\sigma = 5, 10,$ and 20, plotted in Fig. 2.1 show that, about 39% pixels from the Rayleigh distribution with $\sigma = 10$ have an intensity value of 10 or less, and 86% pixels have an intensity value of 10 or less for $\sigma = 5$. Therefore we investigate the effect of the quantization error in pixel intensity values.

When the floating-point numbers are rounded down to integers, the continuous Rayleigh distribution Eq. 2.3 changes to the following discrete distribution

$$f(m;\sigma) = \int_{m}^{m+1} f(x;\sigma)\, dx, \tag{2.9}$$

www.manaraa.com

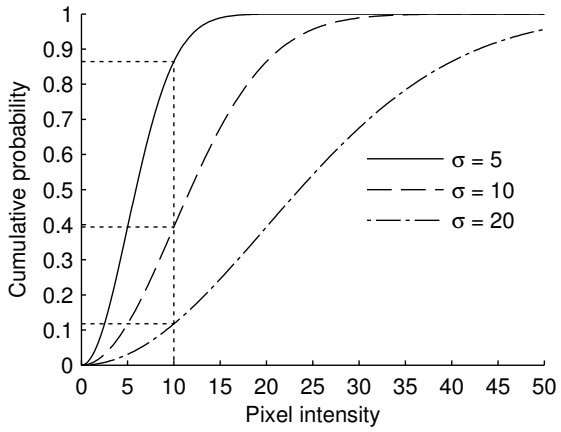


Figure 2.1: The CDF of Rayleigh distribution with $\sigma = 5, 10,$ and 20. Horizontal dotted lines show the probability that a pixel intensity value will be found at a value not greater than 10, for different noise levels.

where $m = 0, 1, \ldots$ are intensity values stored as integers in raw MR data; if the floating-point numbers are rounded to the nearest integer, the discrete distribution is

$$f(m;\sigma) = \int_{\max(m-1/2,0)}^{m+1/2} f(x;\sigma)\, dx. \tag{2.10}$$

Discrete distributions Eqs. 2.9 and 2.10 can be computed by the CDF of the Rayleigh distribution Eq. 2.4 or numerical integration of the PDF Eq. 2.3. The midpoint rule of Eq. 2.9 equals evaluating $f(m+0.5;\sigma)$. The PDF of the discrete distribution Eq. 2.9 is plotted in Fig. 2.2 to show that the difference between the Rayleigh distribution and the discrete distribution can be large.

Noise Parameter Estimation

Although the ML estimator Eq. 2.8 is often considered to be the gold-standard approach (Sijbers et al., 1999), we notice that the mean estimator Eq. 2.7 performs well when the floating-point numbers are rounded integers. This can be seen from the expectation of the first and second raw moments of the discrete Rayleigh distribution Eqs. 2.9 and 2.10, which are used to estimate the parameter σ.

www.manaraa.com

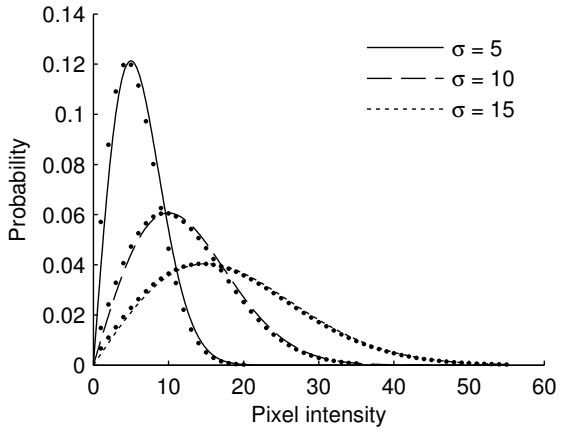


Figure 2.2: The discrete PDF Eq. 2.9 for different values of σ. The discrete PDF is shown using dots together with the Rayleigh PDF Eq. 2.3 for the same σ to show the difference between them.

When rounding down is used, the first and second raw moments are

$$\begin{aligned}\int_0^\infty \lfloor x \rfloor f(x;\sigma)\ dx &= \sum_{k=1}^{\infty}(k \int_k^{k+1} f(x;\sigma)\ dx) \\ &= \sum_{k=1}^{\infty} e^{-\frac{k^2}{2\sigma^2}},\end{aligned} \tag{2.11}$$

and

$$\begin{aligned}\int_0^\infty \lfloor x \rfloor^2 f(x;\sigma)\ dx &= \sum_{k=1}^{\infty}(k^2 \int_k^{k+1} f(x;\sigma)\ dx) \\ &= \sum_{k=1}^{\infty}(2k-1)e^{-\frac{k^2}{2\sigma^2}}.\end{aligned} \tag{2.12}$$

When rounding to the nearest integer is used, the first and second raw moments are

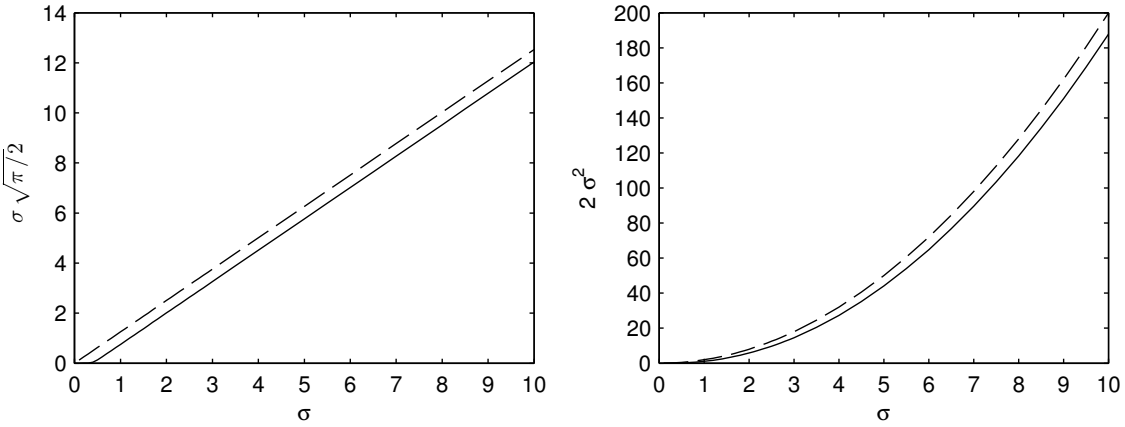
$$\begin{aligned}\int_0^\infty [x] f(x;\sigma)\ dx &= \sum_{k=1}^{\infty}(k \int_{k-1/2}^{k+1/2} f(x;\sigma)\ dx) \\ &= \sum_{k=1}^{\infty} e^{-\frac{(k-1/2)^2}{2\sigma^2}},\end{aligned} \tag{2.13}$$

and

$$\begin{aligned}\int_0^\infty [x]^2 f(x;\sigma)\ dx &= \sum_{k=1}^{\infty}(k^2 \int_{k-1/2}^{k+1/2} f(x;\sigma)\ dx) \\ &= \sum_{k=1}^{\infty}(2k-1)e^{-\frac{(k-1/2)^2}{2\sigma^2}}.\end{aligned} \tag{2.14}$$

www.manaraa.com

There are no closed forms for the sums of the above series. They can be approximated by finite sum.



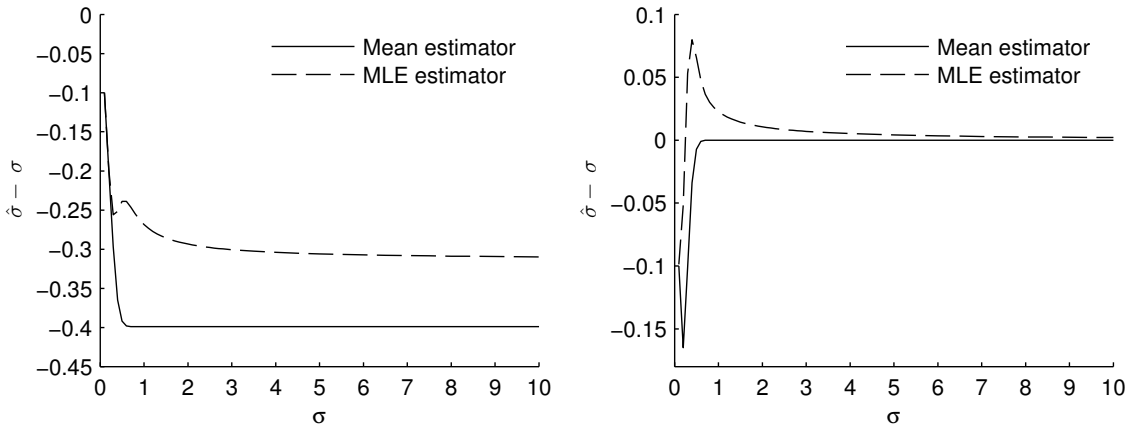
(a) the first raw moment (b) the second raw moment

Figure 2.3: The first and second raw moments of the discrete distribution Eq. 2.9 for the same σ (solid line). The raw moments of the corresponding Rayleigh distribution Eq. 2.3 are shown as dashed line for comparison.

In Fig. 2.3 we compare the raw moments of the discrete distribution Eq. 2.9 using Eqs. 2.11 and 2.12 to those of the Rayleigh distribution Eq. 2.3. The difference between the mean of the Rayleigh distribution and the mean of the discrete distribution Eq. 2.9 approximately equals 0.5 for $\sigma > 1$, making it straightforward to estimate the population mean from the sample mean. The difference between the raw moments obtained from the rounded distribution Eq. 2.10 and those of the Rayleigh distribution Eq. 2.3 is much smaller thus not plotted.

Fig. 2.4 compares the expectation of the two estimators on rounded data, for different σ. Note that the differences of the mean estimator Eq. 2.7 between the Rayleigh distribution and the discrete distributions Eqs. 2.9 and 2.10 are approximately constants, for $\sigma > 1$. So, we can estimate the parameter σ by correcting the sample means. But for the ML estimator Eq. 2.8, the correction is more difficult, although the bias is much smaller.

www.manaraa.com



(a) rounding down (b) rounding to the nearest integer

Figure 2.4: Difference between the true value and the noise parameters estimated by the mean estimator and the ML estimator on rounded data.

Rician Noise Reduction

For recovering the underlying signal magnitudes, the method of MLE is preferred to averaging the repetitions, especially for low SNR area.

Similar to the case of the Rayleigh distribution, the Rician distribution changes to the following discrete distribution when the floating-point numbers are rounded down:

$$f(m; A, \sigma) = \int_{m}^{m+1} f(x; A, \sigma)\ dx, \tag{2.15}$$

and the following discrete distribution when rounding to the nearest integer is used:

$$f(m; A, \sigma) = \int_{\max(m-1/2,0)}^{m+1/2} f(x; A, \sigma)\ dx. \tag{2.16}$$

Evaluating those functions by subtracting two values of the CDF directly is not always stable, because the CDF of the Rician distribution involves summation. Numerical integration can be used when $f(x; A, \sigma)$ is small.

When the MLE is used, the difference between the Rician distribution Eq. 2.1 and the discrete distribution Eq. 2.15 or Eq. 2.16 is significant when the underlying signal A is small. Another advantage of Eqs. 2.15 and Eq. 2.16 is that we do not

www.manaraa.com

have to discard zero pixel intensity values when computing the joint density function in MLE.

2.3 Methods

Noise Parameter Estimation

Estimating the parameter σ of a given MR images using the mean estimator Eq. 2.7 or the ML estimator Eq. 2.8 is done by computing the moment of the pixel values obtained from a large region in the air background, which contains only noise (Gudbjartsson and Patz, 1995).

Simulation

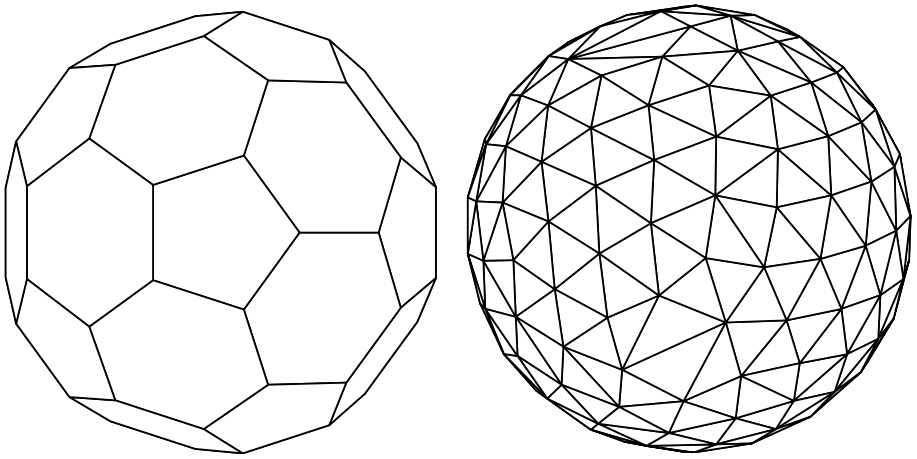
Besides the analytical result shown in Fig. 2.4, we carry out simulations by generating 50000 Rayleigh distributed random numbers with predefined parameter $\sigma = 5$. The floating-point numbers are then rounded to integers to simulate the quantization process. The parameter is then estimated from the floating-point numbers and the rounded integers, to compare the performance of the two estimators. The same simulation is repeated 15 times. We use the method of analysis of variance (ANOVA) to test if the parameter estimated from different method differs.

Phantom Experiments

We use the rat phantom constructed for a diffusion-weighted study (Campbell et al., 2005) to test our method. The rat phantom was constructed by embedding rat spinal cords in 2% agar. The cords were scanned one hour after they were surgically excised. Totally 90 gradient directions are used to acquire the rat phantom data, and the gradient directions are shown in Fig. 2.5(b). The unweighted image was produced by averaging ten unweighted images and the original unweighted images were not saved. Four co-registered datasets were acquired for $b = 1300\,\mathrm{s/mm^2}$ and another four for

www.manaraa.com

$b = 3000\,\mathrm{s/mm^2}$. The spinal cords are 5 mm in diameter. The image dimension is $128 \times 96 \times 40$, and the voxel size is $2.5\,\mathrm{mm} \times 2.5\,\mathrm{mm} \times 2.5\,\mathrm{mm}$. Water diffusion in agar is isotropic while water diffusion within the spinal cords is anisotropic. Curved, straight and crossing fiber tracts are all included in this phantom. T_1-weighted image of this phantom is shown in Fig. 2.6 to illustrate the configuration. The voxel size of the T_1-weighted image is $1\,\mathrm{mm} \times 1\,\mathrm{mm} \times 1\,\mathrm{mm}$.



(a) truncated icosahedron (b) gradient directions of rat phantom

Figure 2.5: Gradient directions used in producing the human brain data and the rat phantom data. (a) gradient directions derived from a truncated icosahedron with 60 vertices distributed on the unit sphere to acquire the human brain data; (b) gradient directions used to acquire the rat phantom data with 90 vertices scattered on a half hemisphere. The schema shown in (a) is also used in simulations.

Diffusion-weighted image of the rat phantom is shown in Fig. 2.7. In this chapter, we only use the air background of the phantom. Pixels from two separate regions in the air background are marked manually. The air background can be identified as the black region in Fig. 2.7(a). For the images of $b = 1300\,\mathrm{s/mm^2}$, the sizes of the two regions are $9 \times 76 \times 19$ and $17 \times 76 \times 19$; for the images of $b = 3000\,\mathrm{s/mm^2}$, the size of the two regions are $18 \times 76 \times 19$ and $21 \times 76 \times 19$. As the rounding method used in quantization is not known to us, the parameter $\hat{\sigma}$ is estimated in both situations

www.manaraa.com

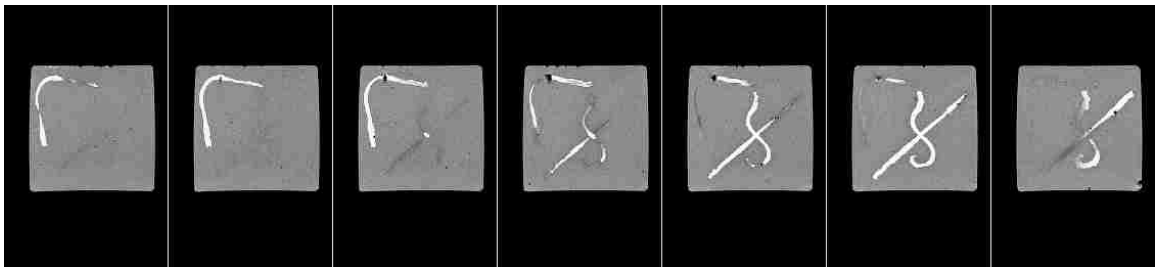
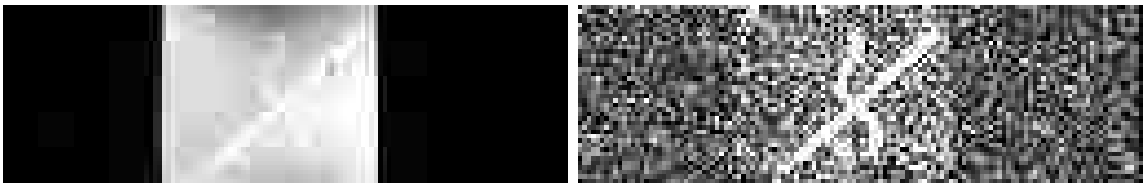


Figure 2.6: Selected slices of the T_1-weighted image of the rat phantom showing the phantom configuration. Two of the spinal cords are put together to create a fiber crossing region.

by both the mean estimator Eq. 2.7 and the ML estimator Eq. 2.8.



(a) unweighted image (b) diffusion-weighted image

Figure 2.7: Unweighted and diffusion-weighted images of the rat phantom of $b = 1300\,\mathrm{s/mm^2}$. Pixel intensities of the weighted image is much lower than that of the unweighted image. The two images are shown as grayscale images using different colormaps.

Human Brain Data

Four sets of diffusion-weighted images (referred as DIFF_30, DIFF_32, DIFF_33 and DIFF_64) with different gradient schemes and T_1-weighted (MP-RAGE) images of one 43 years old healthy male adult are acquired using Siemens Sonata 1.5 T MRI system in the same session, and at the same position. DIFF_30 consisted of one unweighted and 30 diffusion-weighted images; DIFF_32 consisted of three unweighted and 30 diffusion-weighted images; DIFF_33 consisted of four unweighted and 30 diffusion-weighted images; DIFF_64 consisted of one unweighted and 64 diffusion-weighted images. Different datasets can be combined to obtain datasets with more

www.manaraa.com

gradient directions. Diffusion-weighted images are acquired with the following parameters: slice thickness, 1.8 mm; 75 slices; field of view, 230.4 mm × 201.6 mm, 230.4 mm × 216 mm, 230.4 mm × 216 mm, and 230.4 mm × 201.6 mm; in-plane resolution, 1.8 mm × 1.8 mm; matrix, 128 × 112, 128 × 120, 128 × 120, and 128 × 112; $b = 800\,\mathrm{s/mm^2}$; TR and TE of the four sets are 9146 ms/86 ms, 9600 ms/90 ms, 9600 ms/90 ms, and 9146 ms/86 ms. T_1-weighted image is acquired with the following parameters: slice thickness, 1.0 mm; field of view, 256 mm × 256 mm; matrix 256 × 256; 176 slices; TR 2250 ms, TE 3.03 ms.

Images of DIFF_32 and DIFF_33 together make a set of 60 diffusion-weighted images with icosahedral scheme, named as DIFF_60. Images of DIFF_30 and DIFF_64 make a set of 94 diffusion-weighted images named DIFF_94. The icosahedral scheme used by DIFF_60 is shown in Fig. 2.5(a).

A 17 × 18 × 68 region is manually selected from the air background of the DIFF_32 dataset. The same estimation and χ^2 goodness-of-fit test as the phantom experiment are carried out.

Noise Parameter Verification

It is reported that the Rayleigh distribution of the noise in the air background cannot to be verified using the χ^2 goodness-of-fit test (Gudbjartsson and Patz, 1995). One of the debates in the discussions (Andersen, 1996; Gudbjartsson and Patz, 1996) on the possible causes is whether quantization error is a reason of the failure.

As shown above, the quantization process changes the continuous distribution to a discrete one. So, it is more reasonable to carry out the χ^2 goodness-of-fit test using the discrete PDF.

The null hypothesis of the χ^2 goodness-of-fit test is that "the data is a random sample from a Rayleigh distribution with the estimated parameter $\hat{\sigma}$". The test is performed by first grouping the pixel intensity values into bins, each of which con-

www.manaraa.com

tains only one integer. Bins in the tails with a count less than a threshold are pooled with neighboring bins to make sure that the count in each bin is large enough. The frequency counted from each bin is then compared with the expected frequency computed by Eq. 2.9 for rounding down and Eq. 2.10 for rounding to the nearest integer. Suppose we get n bins at last, the test statistic is compared to a χ^2 distribution with $n-2$ degrees of freedom, as we have one estimated parameter. The power of the test for 55000 pixels is 0.95 if the effect size index w is 0.02.

Rician Noise Reduction

We estimate the underlying pixel intensity value using the MLE method with the discrete distributions Eq. 2.15 or Eq. 2.16.

For simulation, Rician distributed random numbers are generated with predefined A and σ. Assuming $\hat{\sigma} = \sigma$, $\hat{A}$ is estimated from floating-point numbers using the PDF of the Rayleigh distrituion Eq. 2.1, from rounded integers using Eqs. 2.15 and 2.10, with four repetitions. The simulation is repeated 10000 times. We simply skip all zero values and use the non-zero pixel intensity values to estimate $\hat{A}$, as zeros make MLE using Eq. 2.1 not feasible.

We look into combinations of different rounding methods and different PDFs in simulations. We compare the estimated values from rounded data using discrete PDFs Eq. 2.15 or Eq. 2.16 to the values estimated using the Rician PDF Eq. 2.1 from floating-point numbers without any rounding.

For the human brain data, there are seven repetitions for the unweighted image and only one image for each gradient direction. When only one measured pixel intensity is available, estimating the underlying signal magnitude by maximizing the probability of the Rician distribution Eq. 2.1 or discrete distribution Eq. 2.15 or Eq. 2.16 with fixed σ is a simple correction to the measured intensity. We plot the estimated signal magnitude from one measured pixel intensity value for different noise

www.manaraa.com

levels in Fig. 2.8.

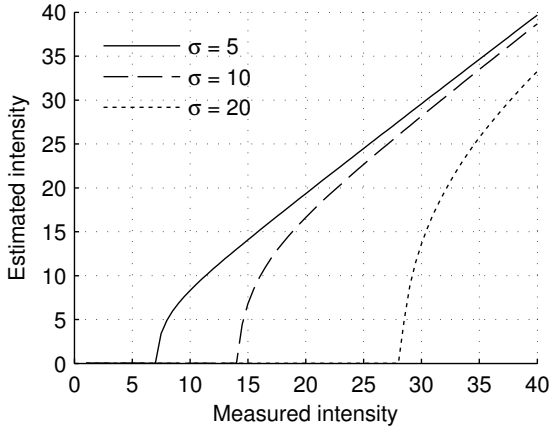


Figure 2.8: Underlying signal magnitude estimated from one noisy pixel intensity value, for different noise parameter σ and different noisy pixel intensity value.

2.4 Results

Noise Parameter Estimation

Simulation

No significant difference is found between the two estimators derived from Eq. 2.5 and Eq. 2.6. Multiple comparisons of the noise parameter estimation using the mean estimator Eq. 2.7 and the ML estimator Eq. 2.8 are shown in Fig. 2.9. To compare the mean estimator and the ML estimator on different rounding methods, 50000 floating-point Rayleigh distributed random numbers are generated. The floating-point numbers are then rounded to integers using both rounding down and rounding to the nearest integer methods. Simulations (a) and (b) use the floating-point numbers as input, (c), (d), (g) and (h) use the integers which are rounded down from floating-point numbers, and (e) and (f) use the integers obtained using rounding to the nearest integer. Simulations (a), (c), (e), and (g) use the ML estimator, while (b), (d), (f) and (h) use the mean estimator. Simulations (g) and (h) are the same as (c) and (d)

www.manaraa.com

except that 0.5 is added to each input integer. The same procedure is repeated 15 times to calcuate the mean and confidence interval of each simulation. The difference between the two estimators is not significant except for integers that are rounded down. Simply adding 0.5 to the intensity value of each pixel helps both estimators get better results if rounding down is used.

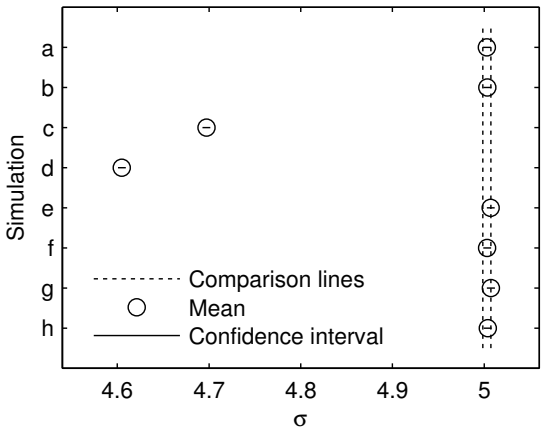


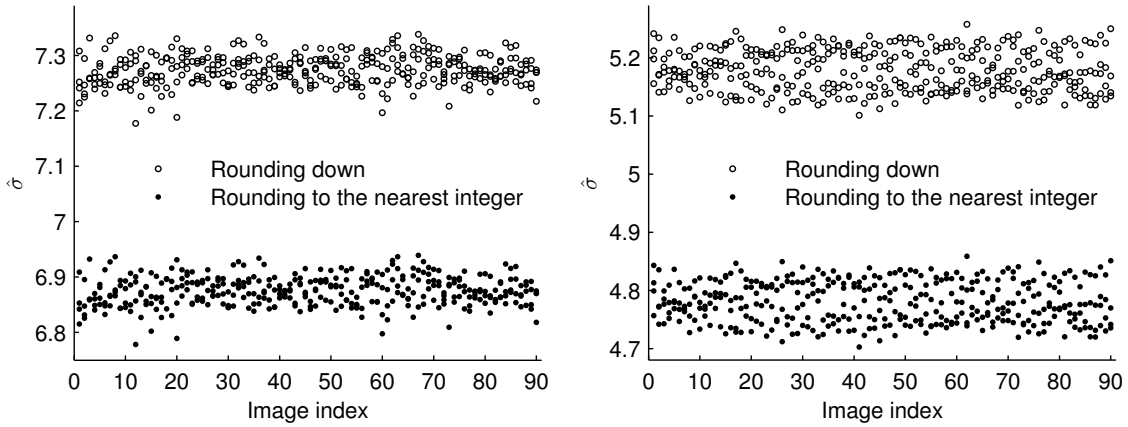
Figure 2.9: Multiple comparison result of the simulations of Rayleigh noise parameter estimation. The mean and 95% confidence interval are plotted for each simulation. Two means are considered different if their intervals are disjoint, and are considered not significantly different if the intervals overlap. The comparison lines help compare the interval of simulation (a) with that of other simulations. The input and estimator used for each simulation: (a) floating-point and the ML estimator, (b) floating-point and the mean estimator, (c) and (g) rounding down and the ML estimator, (d) and (h) rounding down and the mean estimator, (e) rounding to the nearest integer and the ML estimator, and (f) rounding to the nearest integer and the mean estimator. In (g) and (h), a correction of 0.5 is added to each rounded pixel intensity values before carrying out the computation.

Phantom Experiments

In Fig. 2.10, we plot the noise parameter σ estimated from all the 360 diffusion-weighted images of the four repetitions. The noise parameter is estimated for both rounding down and rounding to the nearest integer. The noise level estimated by assuming rounding down as the rounding method is higher than that estimated by

www.manaraa.com

assuming rounding to the nearest integer. We carry out paired samples t-test and the p-value is close to zeros, indicating the difference is significant.



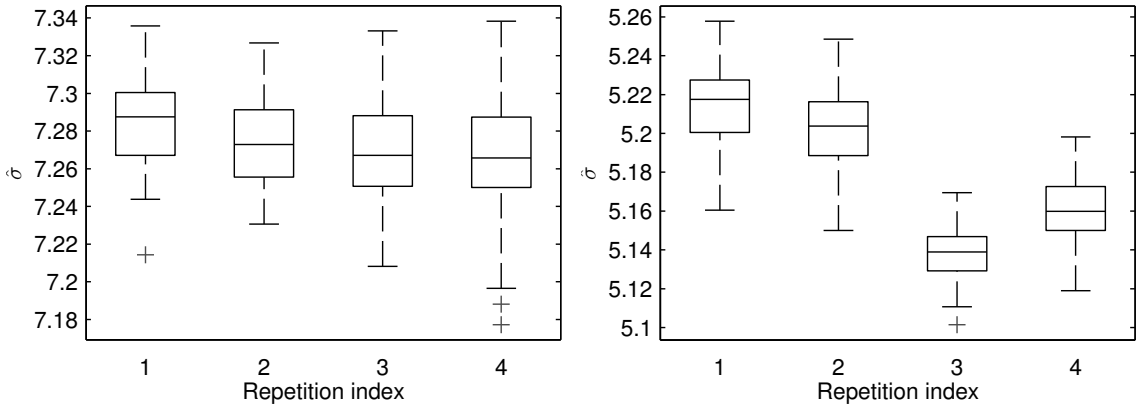
(a) $b = 1300\,\mathrm{s/mm^2}$ (b) $b = 3000\,\mathrm{s/mm^2}$

Figure 2.10: Noise parameters estimated from all diffusion-weighted images of the rat phantom. The dataset has four repetitions with each has 90 diffusion-weighted images, for both $b = 1300\,\mathrm{s/mm^2}$ and $b = 3000\,\mathrm{s/mm^2}$. For each gradient direction, noise parameters estimated from each repetition by assuming rounding down and rounding to the nearest integer are plotted using different markers.

Box plots are shown in Fig. 2.11 to compare the noise parameter estimated from the four repetitions when rounding down is assumed. The box plots are produced by grouping the results in Fig. 2.10 by repetition. The plots show that the variation of estimated noise parameter is small across different repetitions or different gradient directions. The box plots of assuming rounding to the nearest integer are similar and are not shown.

The PDF of the Rayleigh distribution with the parameter estimated from one diffusion-weighted image is plotted for different b-values and different rounding methods in Fig. 2.12. The discrete PDF is evaluated by Eq. 2.9 for rounding down and Eq. 2.10 for rounding to the nearest integer using the estimated noise parameter. The value of the discrete PDF Eq. 2.9 for rounding down is different from that of the Rayleigh PDF Eq. 2.3 with the same noise parameter. The histogram in Figs. 2.12(a)

www.manaraa.com



(a) $b = 1300\,\mathrm{s/mm^2}$ (b) $b = 3000\,\mathrm{s/mm^2}$

Figure 2.11: Noise parameters estimated from four repetitions of the rat phantom shown as box plots for different b-values when rounding down is assumed. On each box of the box plots, the central mark is the median of values, the edges of the box are the 25th and 75th percentiles, the whiskers show the most extreme data points not considered outliers, and outliers are plotted individually using "+".

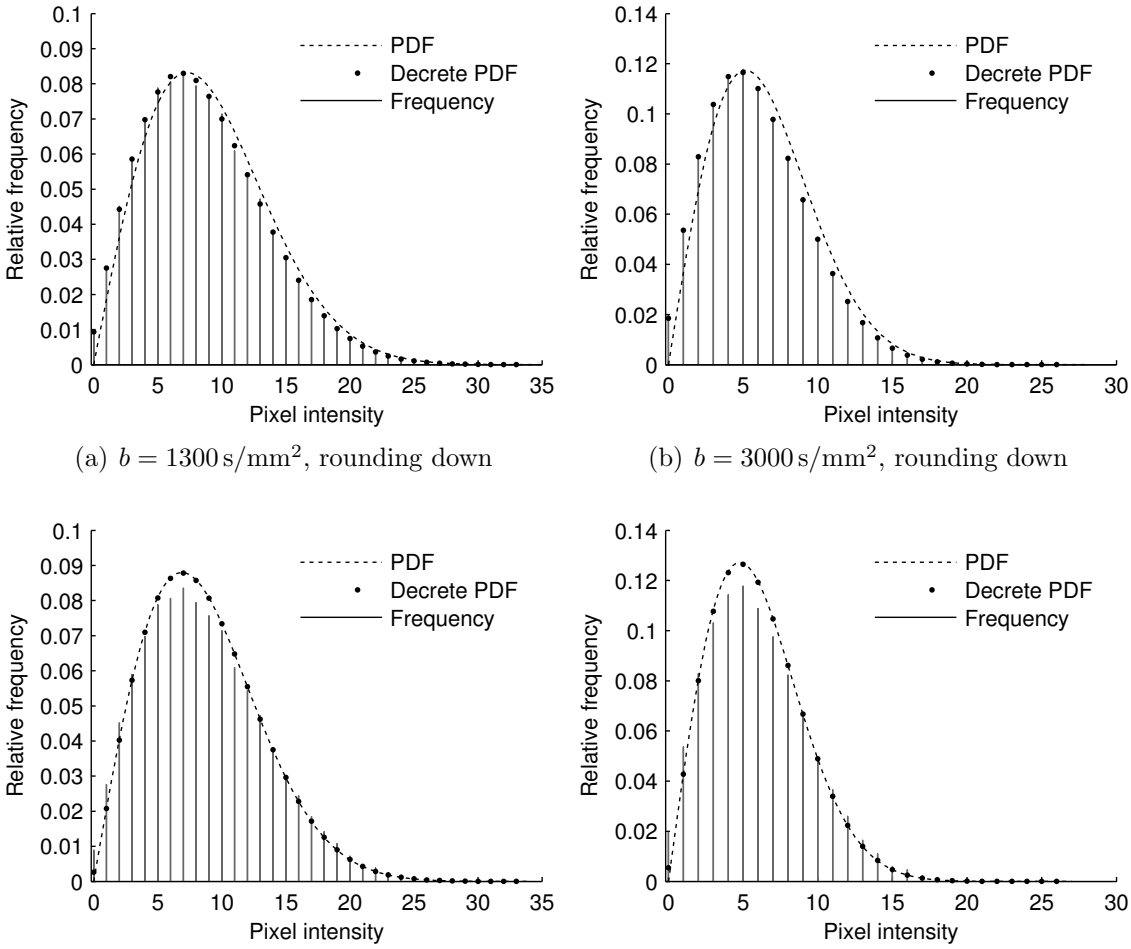
and 2.12(b) matches to the discrete PDF better when rounding down is assumed.

Human Brain Data

Noise parameter estimated from each image is plotted in Fig. 2.13. The value of the noise parameter estimated assuming rounding down as the rounding method is higher than that estimated assuming rounding to the nearest integer. Paired samples t-test is carried out and the p-value is close to zero. The result of not considering quantization error is identical to that of assuming rounding to the nearest integer. The estimated noise parameter of the diffusion-weighted images is between 19 and 20, which is higher than that of the rat phantom, making the quantization error less noticeable. The noise parameter estimated from the T_1-weighted image is 5.6.

Histogram of the local SNR calculated from pixels within the brain is shown in Fig. 2.15. The SNR of each pixel is calculated by Eq. 2.2 using the intensity value of that pixel and the noise parameter estimated from the whole image. The brain is extracted using Brain Extraction Tool (Battaglini et al., 2008). The local SNR of the

www.manaraa.com



(a) $b = 1300\,\mathrm{s/mm^2}$, rounding down

(b) $b = 3000\,\mathrm{s/mm^2}$, rounding down

(c) $b = 1300\,\mathrm{s/mm^2}$, rounding to the nearest integer

(d) $b = 3000\,\mathrm{s/mm^2}$, rounding to the nearest integer

Figure 2.12: The PDF of the Rayleigh distribution with the noise parameter estimated from one diffusion-weighted image of the rat phantom for different rounding methods and different b-values together with the histograms of pixel intensity values. Both the PDF of the Rayleigh distribution and the corresponding discrete PDF for different rounding methods are plotted together for comparison.

www.manaraa.com

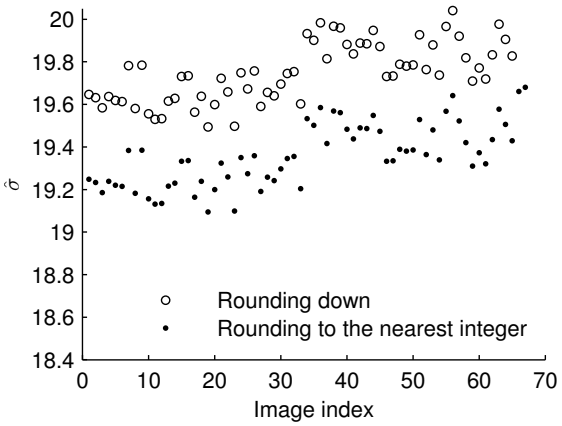
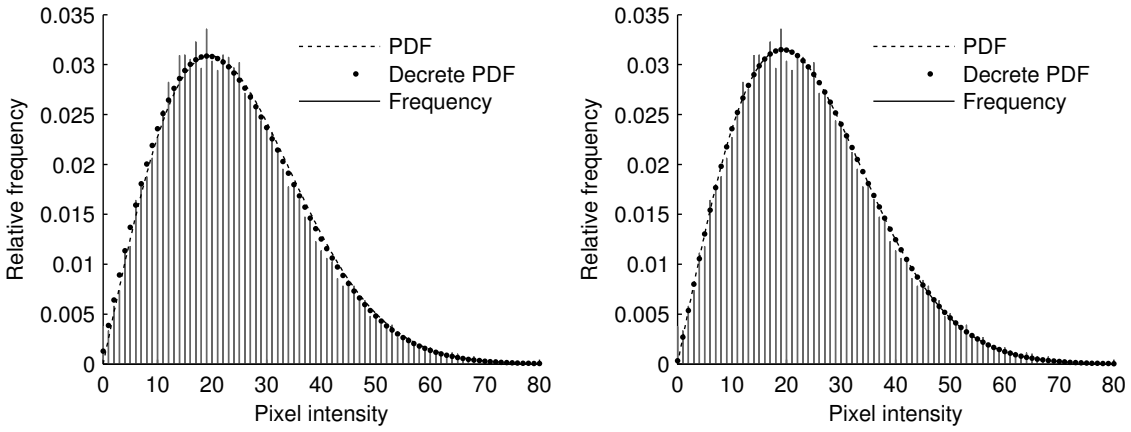


Figure 2.13: Noise parameter estimated from all the images of the human brain data, for assuming both rounding methods. The first seven images are unweighted images and the rest are diffusion-weighted images.



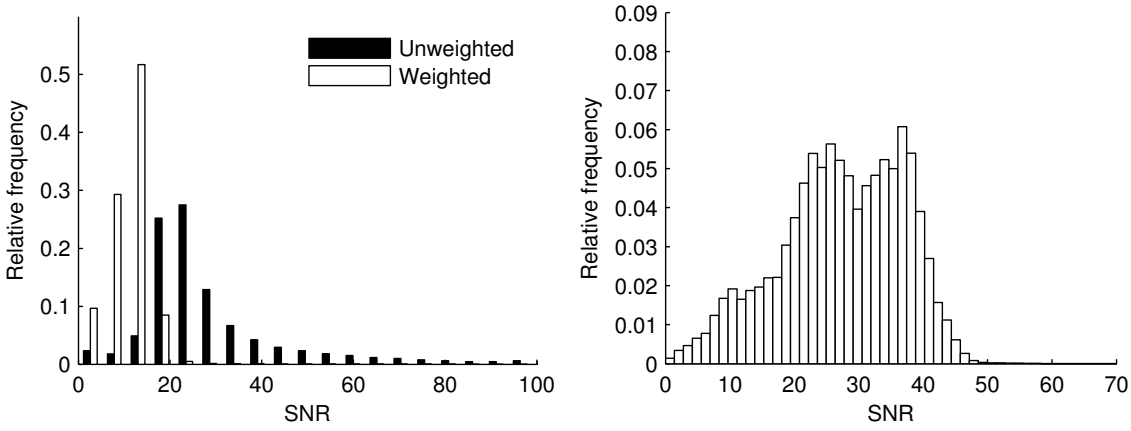
(a) rounding down

(b) rounding to the nearest integer

Figure 2.14: The PDF with noise parameter estimated from one image of the human brain data together with the histogram of pixel intensity values for different rounding methods.

www.manaraa.com

T_1-weighted image is between 20 and 40, which is higher than that of the diffusion-weighted images, for more than 70% voxels within the brain. The local SNR of the unweighted image is around 25 and that of the diffusion-weighted image is below 20.



(a) diffusion-weighted images (b) T_1-weighted image

Figure 2.15: The SNR of diffusion-weighted images and T_1-weighted image of human brain data. The histograms of SNR of one unweighted image and that of one weighted image are plotted together in (a) for comparison.

Model Verification

Phantom Experiment

Although the rounding method is not provided with the images, rounding down is more likely than rounding to the nearest integer to have been used in producing the data. We carry out χ^2 goodness-of-fit test of the noise parameters estimated for both rounding methods. If rounding to the nearest integer has been used, the null hypnosis is rejected for all 360 diffusion-weighted images of both $b = 1300\,\mathrm{s/mm^2}$ and $b = 3000\,\mathrm{s/mm^2}$, at significant level $\alpha = 0.05/360$. When rounding down is assumed, the null hypnosis is not rejected for any image at the same significant level. Therefore we conclude that rounding down has been used to store the raw images. It can also

www.manaraa.com

be seen from Fig. 2.12 that the estimated noise parameters fit the frequencies better for rounding down.

Human Brain Data

We carry out the same procedures for the human brain data. However, there is not much difference between rounding down and rounding to the nearest integer, because the noise level of the human brain data is high and truncation does not introduce large error compared to the noise. When assuming rounding down as rounding method, χ^2 goodness-of-fit test is rejected for 2 out of 67 images at significant level $\alpha = 0.05/67$, while the same test is rejected for 10 images when rounding to the nearest integer is assumed. Although we are not able to judge the rounding method used in producing the data, the noise distribution is verified for majority of the images.

Rician Noise Reduction

Simulations show that the standard deviation of the MLE decreases when the number of repetition increases, or the SNR increases.

The result of using Eq. 2.1 to conduct the MLE with rounding down data is significantly different from that of other estimators, when the number of repetitions is large enough.

Although there is no statistical difference between those estimators when the number of repetitions is small, the result of using Eq. 2.1 is always 0.5 lower than that of other estimators.

Table 2.1 lists simulation results of the noise reduction using MLE, for $A = 12, 18, \text{and } 60$.

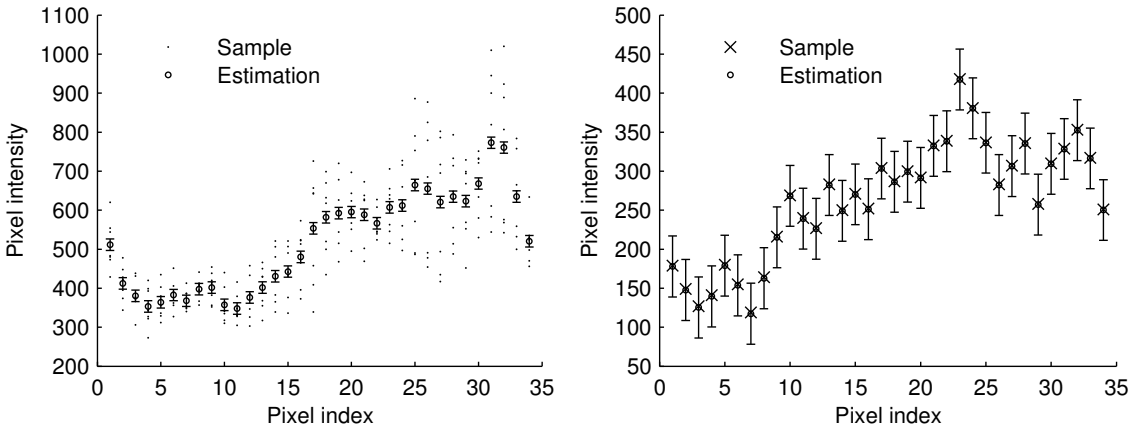
The underlying pixel intensity value estimated from human brain data is shown in Fig. 2.16. For the unweighted image, the pixel intensity values are estimated from seven repetitions, while for the diffusion-weighted image, the pixel intensity values

www.manaraa.com

Data type or rounding method	PDF	$A = 12$	$A = 18$	$A = 60$
floating-point	Eq. 2.1	11.8 ± 3.5	18.0 ± 3.1	60.0 ± 3.0
rounding down	Eq. 2.1	11.2 ± 3.6	17.5 ± 3.1	59.5 ± 3.0
rounding down	Eq. 2.15	11.8 ± 3.5	18.0 ± 3.1	60.0 ± 3.0
rounding to the nearest integer	Eq. 2.1	11.8 ± 3.5	18.0 ± 3.1	60.0 ± 3.0
rounding to the nearest integer	Eq. 2.16	11.8 ± 3.5	18.0 ± 3.1	60.0 ± 3.0

Table 2.1: Simulation result of noise reduction using MLE with the PDF of the Rician distribution Eq. 2.1 and discrete PDF Eq. 2.15 or Eq. 2.16.

are estimated from one image. Fig. 2.16(a) shows that the variation of pixel intensity values of the same pixel from different repetitions may be large. In our experiment, the underlying pixel intensity value estimated using MLE is actually close to the sample mean. For the diffusion-weighted image, as we do not have repetitions, the estimated underlying value is just a small correction to the measured value.



(a) unweighted image (b) diffusion-weighted image

Figure 2.16: Pixel intensity values and estimated signal intensity value of the human brain data. The estimated signal intensities as well as the 95% confidence intervals are shown using error bars. For the unweighted image, the signal intensity values are estimated from seven repetitions. While for the diffusion-weighted image, the signal intensity values are estimated from one measured intensity value.

www.manaraa.com

2.5 Discussion

We propose a simple correction of adding 0.5 to pixel intensity values if rounding down is used to save the data. Both of the two estimators Eq. 2.7 and Eq. 2.8 are derived without taking the quantization error into account. Although the ML estimator Eq. 2.8 is preferred in some studies, the results from both methods are similar in our simulation if the correction is carried out. However, using integers rounded down from floating-point number as if they are floating-point numbers introduces a bias in the result if the noise level is low. Although the difference is only 5% in the rat phantom, it causes the goodness-of-fit test to fail if another rounding method is assumed. When noise level is high, the effect of quantization error is less noticeable.

Noise parameter estimation and noise reduction should be integrated into the DWI processing pipeline. In order to employ the Rician noise reduction, it is better not to average multiple acquisitions by the scanner. The noise parameter may be estimated from a single acquisition before or after the eddy current correction process, by selecting a large region in the air background manually. Rician noise reduction is then carried out after registering multiple acquisitions. In both eddy current correction and registration, we recommend to use the nearest-neighbor interpolation because it does not change the noise distribution.

The MLE approach of Rician noise reduction does not provide much improvement when b-value is large, especially when the number of acquisitions is small. The limitation is due to the fact that the diffusion-weighted signals are much weaker with high b-values. The quantization error can sometimes be ignored here because the Rician noise is the major concern. In Ch. 4 we avoid doing noise reduction on each single diffusion-weighted signal by carrying out MLE on the diffusion model.

The noise level is often simply reported as an averaged SNR. This practice may be sufficient for structural MRI, but it is not appropriate for diffusion MRI. It is because the local SNR can differ a lot from location to location or image to image.

www.manaraa.com

The noise level σ can be considered fixed for the whole image, while the ADC may differ greatly (Kristoffersen, 2007). Therefore, it is preferred to report the parameter σ in DWI studies, because the local SNR is not a constant.

Copyright© Ning Cao, 2013.

www.manaraa.com

Chapter 3 Apparent Diffusion Coefficient Estimation

Diffusion tensor model is insufficient for complex tissue structures within a single voxel. Although complex diffusion models (Alexander et al., 2001; Tuch et al., 2002; Parker and Alexander, 2003) are proposed, none of them is able to fully capture the complex details of local diffusion. Therefore, we look into methods that describe the ADC profile without a model.

When the diffusion tensor model is sufficient to describe a diffusion profile, we are able to use the diffusion tensor to fully assess the diffusion profile. It means that in addition to the measured values for each gradient direction, we have a continuous function defined on the unit sphere to describe the diffusion profile. In HARDI, diffusion-weighted signal is recorded for many more gradient directions than in a typical DTI study. We need a continuous function to describe the shape of the diffusion profile when the diffusion tensor model is not sufficient. HODT and spherical harmonic series are proposed to describe the shape of local diffusion profiles (Frank, 2002; Alexander et al., 2002; Özarslan and Mareci, 2003; Chen et al., 2004; Özarslan et al., 2005; Descoteaux et al., 2006) without assigning a particular diffusion model. In the studies using spherical harmonic series, LS minimization is often used to obtain the coefficients of spherical harmonic series. We show that LS method is not always stable and provide alternative methods to calculate the coefficients of spherical harmonic series.

3.1 Introduction

Any square-integrable function $f(\theta, \phi)$ defined on the unit sphere can be expanded using spherical harmonics (Byerly, 1895; MacRobert and Sneddon, 1967), as:

$$f(\theta, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} c_\ell^m Y_\ell^m(\theta, \phi), \tag{3.1}$$

www.manaraa.com

where Y_ℓ^m is the spherical harmonic function of degree ℓ and order m. It is a standard method in physical sciences and computer graphics. The notational convention of θ and ϕ is defined in Eq. 1.1. It follows the one commonly used in physics and some HARDI studies (Frank, 2002; Özarslan and Mareci, 2003; Chen et al., 2005). The expansion coefficient c_ℓ^m of degree ℓ and order m is given by the following integral:

$$\begin{aligned} c_\ell^m &= \int_\Omega d\Omega f(\theta,\phi) Y_\ell^{m*}(\theta,\phi) \\ &= \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta f(\theta,\phi) Y_\ell^{m*}(\theta,\phi), \end{aligned} \tag{3.2}$$

where $Y_\ell^{m*}(\theta,\phi)$ denotes complex conjugation of $Y_\ell^m(\theta,\phi)$, and the process of obtaining c_ℓ^m is known as spherical harmonic transform (SHT). In addition to the spherical harmonics, the real form of spherical harmonics can be used to simplify the computation.

In diffusion-weighted MRI studies, ADC profile $d(\vec{g})$ defined in Eq. 1.5 is the underlying diffusion profile of water molecules within a voxel. It is often convenient to write $d(\vec{g})$ in spherical coordination as $d(\theta,\phi)$. Studies regarding the ADC profiles approximate $d(\theta,\phi)$ using spherical harmonic series truncated to a certain degree L (Alexander et al., 2002; Frank, 2002; Chen et al., 2004, 2005) as:

$$d(\theta,\phi) = \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} c_\ell^m Y_\ell^m(\theta,\phi). \tag{3.3}$$

Odd-degrees of the expansion coefficients c_ℓ^m can be omitted because the diffusion-weighted signal is antipodally symmetric (Frank, 2002; Alexander et al., 2002). Choosing $L = 2$ in Eq. 3.3 equals the diffusion tensor model (Frank, 2002; Özarslan and Mareci, 2003; Descoteaux et al., 2006), and a higher degree of L allows more details of the ADC profile to be captured.

The commonly used approach treats the problem as curve fitting. The straightforward LS minimization is used in some studies (Alexander et al., 2002). The coefficients c_ℓ^m in Eq. 3.3 can be obtained by solving the linear system

$$\mathbf{Mc} = \mathbf{d} \tag{3.4}$$

www.manaraa.com

using linear LS method, where $\mathbf{M}$ is the matrix containing all even order spherical harmonic basis for all the n sampling directions as

$$\mathbf{M} = \begin{pmatrix} Y_0^0(\theta_1,\phi_1) & Y_2^{-2}(\theta_1,\phi_1) & \cdots & Y_L^L(\theta_1,\phi_1) \\ \vdots & \vdots & \ddots & \vdots \\ Y_0^0(\theta_n,\phi_n) & Y_2^{-2}(\theta_n,\phi_n) & \cdots & Y_L^L(\theta_n,\phi_n) \end{pmatrix}, \tag{3.5}$$

$\mathbf{c} = [c_0^0, \ldots, c_L^L]^T$ is the vector of unknown coefficients to be decided, and $\mathbf{d} = [d(\theta_1,\phi_1), \ldots, d(\theta_n,\phi_n)]^T$ is the vector of the ADC values at each sampling direction.

3.2 Methods

The sampling schemas used in HARDI studies are either obtained from the regular polyhedrons (Tuch et al., 1999; Frank, 2001, 2002; Tuch et al., 2002; Özarslan and Mareci, 2003; Özarslan et al., 2005; Descoteaux et al., 2006) or from numerical optimizations (Papadakis et al., 2000; Hasan et al., 2001; Jansons and Alexander, 2003).

When considering the problem as sampling, we follow the notations of mathematic references to use the word "node" to refer to each sampling direction, and use "sampling grid" for the point grid on the unit sphere formed by all the sampling directions of a sampling schema. The coefficients can be obtained by evaluating the integrals Eq. 3.2 using samples $f_i = f(\theta_i, \phi_i)$.

Let $q_\ell^m(\theta,\phi)$ denote the integrand in Eq. 3.2:

$$q_\ell^m(\theta,\phi) = f(\theta,\phi) Y_\ell^{m*}(\theta,\phi). \tag{3.6}$$

As the integral Eq. 3.2 is needed for all ℓ and m, we simply write $q_\ell^m(\theta,\phi)$ as $q(\theta,\phi)$. To get c_ℓ^m, we need to compute the integral Eq. 3.2 using samples $q_i = q(\theta_i,\phi_i)$. We use $I[q]$ to denote such formulas.

www.manaraa.com

Computation of SHT

As a continuous function $\tilde{q}$ can be obtained from the samples q_i by interpolation, the integral Eq. 3.2 can be approximated by numerically integrating the interpolated function. This approach is not adopted here because it introduces artificial values that are not from the input data.

Previous studies (Wang and Carrington, 2003; Stevensson and Edén, 2006) using Lebedev quadrature (Lebedev, 1976) and Gaussian quadrature show that Eq. 3.2 can be exactly evaluated by the following format with some restrictions (Lebedev, 1977; Koch and Becker, 2004):

$$\tilde{I}[q] = \sum_{i=1}^{n} w_i q_i, \tag{3.7}$$

where w_i are weights to be decided. In these methods, the function q_i is known and the sampling grid is carefully chosen to find formulae that are accurate to a degree as high as possible with as few nodes as possible. In our case, only the samples q_i are available.

Although a higher degree of L in Eq. 3.3 may be demanded for better approximation, we show that the value L that can be chosen is limited by the sampling schema.

Least Squares Minimization

Let Y_ℓ^m denote $Y_\ell^m(\theta_i, \phi_i)$ and the n sampling points satisfy the orthogonality relationships Eq. 3.12. To find the best fit in the LS sense of f_i, we minimize

$$J = \sum_{i=1}^{n} [(f_i - \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} c_\ell^m Y_\ell^m)(f_i - \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} c_\ell^m Y_\ell^m)^*]. \tag{3.8}$$

By setting the partial derivatives of J to zero,

$$\frac{\partial J}{\partial c_\ell^m} = \sum_{i=1}^{n} [(f_i - \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} c_\ell^m Y_\ell^m)(-Y_\ell^{m*}) + (f_i - \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} c_\ell^m Y_\ell^m)^*(-Y_\ell^m)] \equiv 0 \tag{3.9}$$

www.manaraa.com

and utilizing the orthogonality relationships Eq. 3.12, we have

$$c_\ell^m + c_\ell^{m*} = \frac{1}{n}\sum_{i=1}^{n} f_i(Y_\ell^m + Y_\ell^{m*}). \tag{3.10}$$

So a formula in the format of

$$c_\ell^m = \frac{1}{n}\sum_{i=1}^{n} f_i Y_\ell^m \tag{3.11}$$

is also the LS solution.

Therefore, the LS minimization is exactly the same as the quadrature formulae, like Eqs. 3.14, 3.15, and 3.16, if the set of basis used is orthogonal on the sampling grid, e.g.,

$$\sum_{i=1}^{n} Y_\ell^m(\theta,\phi)Y_{\ell'}^{m'*}(\theta,\phi) = \delta_{\ell\ell'}\delta_{mm'}. \tag{3.12}$$

The LS method by solving Eq. 3.4 only works for some combinations of degree and sampling schema. For other combinations, there is a rank deficiency issue. Rank and condition numbers of the normal matrix M^TM for choosing $L = 2, 4, 6,$ and 8 in Eq. 3.3 for the sampling grids mentioned above are listed in Table 3.1. Note that full rank of M^TM for $L = 2, 4, 6,$ and 8 are 6, 15, 28, and 45. Therefore, we need alternative approaches to calculate the coefficients.

Table 3.1: Rank and condition number of M^TM

Schema (n)	Rank				Condition number			
	L=2	4	6	8	L=2	4	6	8
icosahedron (20)	6	6	6	6	1	9.9×10^{16}	2.8×10^{18}	1.5×10^{18}
2nd tessellation (42)	6	15	22	22	20	1.4×10^{3}	8.4×10^{17}	6.0×10^{17}
3rd tessellation (92)	6	15	28	45	17	490	4.28×10^{5}	1.7×10^{7}
4th tessellation (162)	6	15	28	45	16	520	2.5×10^{5}	2.4×10^{6}
5th tessellation (252)	6	15	28	45	16	466	2.1×10^{5}	1.3×10^{6}
truncated icosahedron (60)	6	15	27	30	1	1.2	1.8×10^{16}	1.0×10^{17}
Elec061 (61)	6	15	28	45	1	1.0	1.1	1.2

www.manaraa.com

Directions from the Icosahedron

Sampling grids obtained from the regular icosahedron are often used in HARDI studies (Tuch et al., 1999; Frank, 2001, 2002; Tuch et al., 2002; Özarslan and Mareci, 2003; Özarslan et al., 2005; Descoteaux et al., 2006). The regular icosahedron has 12 vertices, 20 faces and 30 edges. One or more groups of the vertices (denoted as a^v), the midpoints of each edge (denoted as a^e) and the centers of each face (denoted as a^f) can be used as gradient directions. Elements of each of the above set are equivalent under icosahedronal symmetry, so their corresponding weights are the same. Tessellations of the regular icosahedron (Tuch et al., 1999; Frank, 2001) generate more nodes thus increase the angular resolution. See Fig. 3.1 for the tessellation process and the nodes a^v, a^e, and a^f.

The quadrature formula is in the format of

$$\tilde{I}[q] = w^v \sum_{i=1}^{12} q(a_i^v) + w^e \sum_{i=1}^{30} q(a_i^e) + w^f \sum_{i=1}^{20} q(a_i^f) + \ldots, \tag{3.13}$$

where $w^v, w^e, w^f, \ldots$ are constants to be decided. The method of deciding the constants $w^v, w^e, w^f, \ldots$ have been studied elsewhere (McLaren, 1963) and the constants are tabulated (Stroud, 1971).

For example, the quadrature of the 12 vertices (McLaren, 1963; Stroud, 1971; Bažant and Oh, 1986) is

$$\tilde{I}_{12}[q] = \frac{\pi}{3} \sum_{i=1}^{12} q(a_i^v). \tag{3.14}$$

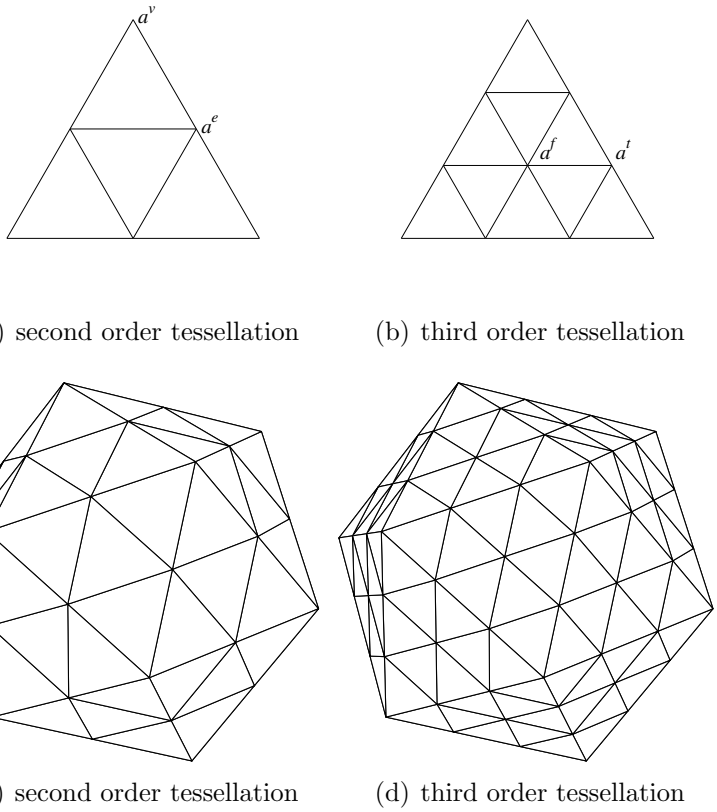
The alternative one is that of the 20 centers of each faces (McLaren, 1963; Bažant and Oh, 1986)

$$\tilde{I}_{20}[q] = \frac{\pi}{5} \sum_{i=1}^{20} q(a_i^f), \tag{3.15}$$

and that of the 30 midpoints of each edge (McLaren, 1963)

$$\tilde{I}_{30}[q] = \frac{2\pi}{15} \sum_{i=1}^{30} q(a_i^e). \tag{3.16}$$

www.manaraa.com



(a) second order tessellation (b) third order tessellation

(c) second order tessellation (d) third order tessellation

Figure 3.1: The second order and the third order tessellations of a regular icosahedron. (a) and (b) are the tessellation of a face; (c) and (d) are the tessellated regular icosahedrons.

They are accurate to the degree of 5 (McLaren, 1963). Using the 60 vertices (denoted as a^t, see Fig. 3.1(d)) of the truncated icosahedron also yields a quadrature accurate to degree of 5:

$$\tilde{I}_{60}[q] = \frac{\pi}{15} \sum_{i=1}^{60} q(a_i^t). \tag{3.17}$$

The quadrature for the second order tessellation of a regular icosahedron, which is shown in Fig. 3.1(c), is the 9th order with 42 points, with 12 points of a_i^v and 30 points of a_i^e (McLaren, 1963; Bažant and Oh, 1986):

$$\tilde{I}_{42}[q] = \frac{25\pi}{315} \sum_{i=1}^{12} q(a_i^v) + \frac{32\pi}{315} \sum_{i=1}^{30} q(a_i^e). \tag{3.18}$$

We obtain the quadrature for the third order tessellation of a regular icosahedron

www.manaraa.com

(see Fig. 3.1(d)) numerically. It is a combination of a_i^v, a_i^f and the 60 vertices of the truncated icosahedron:

$$\tilde{I}_{92}[q] = w^v \sum_{i=1}^{12} q(a_i^v) + w^f \sum_{i=1}^{20} q(a_i^f) + w^t \sum_{i=1}^{60} q(a_i^t), \tag{3.19}$$

with $w^v = 0.08480585664765$, $w^f = 0.15198893252538$ and $w^t = 0.14181536140133$. It is accurate to the degree of 11 (Sobolev, 1962).

Finite Element Integration

In general, formulas like Eqs. 3.18 and 3.19 do not exist for the directions obtained from numerical optimizations.

When the number of nodes is small, the sampling grid provides triangulation of the sphere surface. The tessellation process provides a triangular grid of the icosahedron; for sampling grids obtained from numerical optimization, we may use the Delaunay triangulation of the vertices. With more points added to the sampling grid, the triangular grids become good approximations of the sphere surface, so that a slightly revised centroid rule can be used.

Let $T_{k_1}, \ldots, T_{k_m}$ be the m triangles that share the vertex v_k, $A_{k_1}, \ldots, A_{k_m}$ be their area, and $A'_{k_1}, \ldots, A'_{k_m}$ be the area of their projections on the sphere surface, the quadrature can be approximated (Atkinson, 1982) as

$$\tilde{I}[q] = \frac{1}{3} \sum_{k=1}^{n} [q(v_k) \sum_{j=1}^{m} A'_{k_j}], \tag{3.20}$$

or

$$\tilde{I}[q] = \frac{1}{3} \sum_{k=1}^{n} [q(v_k) \sum_{j=1}^{m} A_{k_j}]. \tag{3.21}$$

The degree of precision of the centroid rule on the tessellations of a regular icosahedron is 5 (Atkinson, 1982). The error of using it on the tessellations of a regular icosahedron can be found by integrating Y_ℓ^m, $\ell = 1, \ldots, 5$. The maximum absolute error of $I_c(Y_5^m)$, $m = -5, \ldots, 5$ is shown in Table 3.2.

www.manaraa.com

Table 3.2: The absolute error of $I_c(Y_5^m)$

sampling schema	n	error of $I_c(Y5^m)$
icosahedron	12	2.66×10^{-15}
2nd tessellation	42	9.71×10^{-16}
3rd tessellation	92	4.16×10^{-16}
4th tessellation	162	3.97×10^{-16}
5th tessellation	252	2.75×10^{-16}

Applying Eq. 3.20 to the 12 vertices or the 20 face centers of the regular icosahedron yields the same formula as Eqs. 3.14 and 3.15.

Monte Carlo Integration

If the nodes are fairly distributed on the sphere, it is possible to use the Monte Carlo integration (Caflisch, 1998) on the unit sphere to calculate the integral of Eq. 3.2 with discrete samples:

$$\tilde{I}_{MC}[q] = \frac{4\pi}{n} \sum_{i=1}^{n} q_i. \tag{3.22}$$

The weights of Monte Carlo integration Eq. 3.22 are exactly the same as the quadrature Eqs. 3.14, 3.15 and 3.16. And, the weights of Eq. 3.20 on the Elec061 schema and the 5th order tessellation of the regular icosahedron are from 0.08446 to 0.1058, and 0.0314 to 0.0562.

Complexity

The complexity of both the quadratures in this section and LS minimization is $O(n)$. In Eq. 3.7, both the weights w_i and the samples $Y_\ell^m(\theta_i, \phi_i)$ are decided by the sampling schema, thus are computed only once for the whole image volume. For Eq. 3.4, the matrix $(M^T M)^{-1} M^T$ can be computed in advance, if the normal matrix $M^T M$ has full rank. The ADC values $\tilde{D}_i$ are obtained directly from the diffusion-weighted measurement. So, computing a single c_ℓ^m using Eq. 3.7 or Eq. 3.4 is the same as

www.manaraa.com

computing the inner product of two vectors, requiring n multiplications and $n-1$ additions.

Choosing L for Gaussian Mixture model

Spherical harmonics truncated to a certain degree L are used to describe the ADC profiles of the biGaussian model (Alexander et al., 2002; Frank, 2002; Chen et al., 2005; Cao et al., 2009). However, the degree L is decided arbitrarily or by the model residual only. We carry out simulations to compare the relative error in ADC profiles caused by truncating spherical harmonics with that caused by Rician noise to choose the degree L so that the truncation error is similar to the error caused by Rician noise.

Two tensor compartments with eigenvalues being $(1200, 200, 200) \times 10^{-6}\,\mathrm{mm}^2$ /s are used to generate the biGaussian ADC profiles. In the simulations, the angle τ between the two tensor compartments covers the range of 0° to 90° because it is reported that the choice of L depends on the relative orientation of the two tensor compartments (Frank, 2002). The same simulation is repeated for different σ and different b-values. To ensure accuracy, we use the 5810-node Lebedev grid, which is accurate to the degree of 131, to compute the integrals needed by SHT.

3.3 Results

Simulation

The relative error in biGaussian ADC profiles caused by Rician noise are plotted in Figs. 3.2(a) and 3.2(b) for different noise level σ. The relative error caused by both the noise and truncating the spherical harmonics to L are shown in Figs. 3.2(c) and 3.2(d). The SHT is computed using the Lebedev grid and the error is checked on the directions obtained from the 4th tessellation of the icosahedron. The Rician noise is decided by σ only, therefore the absolute ADC error caused by Rician noise does

www.manaraa.com

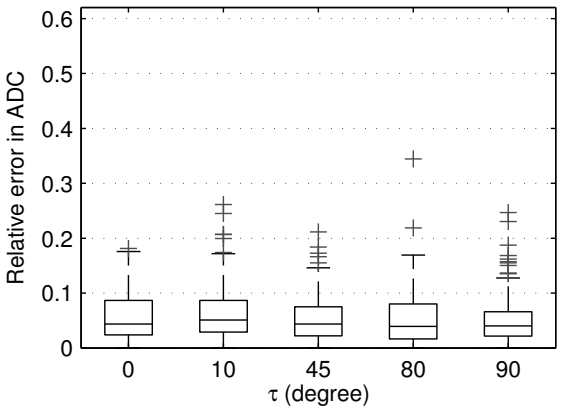
not depend on gradient direction, b-value, and τ. The difference of the relative error for different τ is mainly because the biGaussian ADC values depend on τ. The error caused by truncating the spherical harmonics to L, or the model residual, depends greatly on the angle τ. Although it is reported that higher L is required when the two tensor orientations are more closely aligned (Frank, 2002), the peak of ADC error is observed when τ is around 45° in Figs. 3.2(c) and 3.2(d) for $L \geq 2$ in our simulations. It also indicates that the power spectrum of ADC profiles of the biGaussian model is not generally limited to a small number, however, a small L can be used when τ is towards 0° or 90°. Therefore, the choice of L depends on the application and other imaging parameters. However, for the biGaussian model and relative low noise level, a large L is not necessary.

It can also be seen from Figs. 3.2(c) and 3.2(d) that using $L = 2$ to truncate Eq. 3.3, which equals the diffusion tensor model, maintains a moderate low relative ADC error when τ is around 0° or 90°. This phenomena indicates that the diffusion tensor model is a good approximation of the biGaussian model in these situations. When τ is small, it is difficult to distinguish the two-fiber case from the single-fiber case, making it not feasible to reconstruct the curving fibers using biGaussian model. When τ is towards 90°, a tensor with $\lambda_1 \approx \lambda_2 > \lambda_3$ approximates the biGaussian ADC profile well. Therefore, planer-shaped diffusion tensors should be checked carefully in diffusion tensor model.

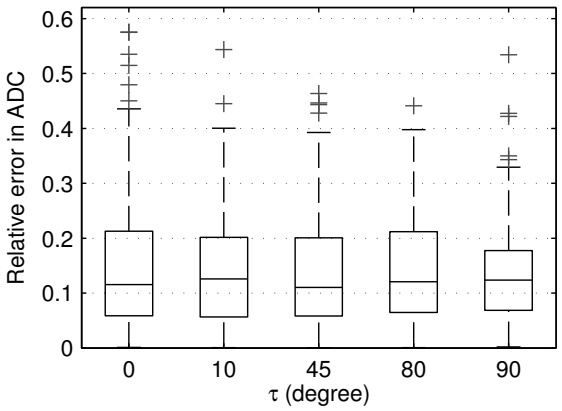
Human Brain Data

The same human brain data mentioned in Ch. 2 is used to test the quadrature formulae. The spherical harmonics coefficients for $\ell = 0, 1,$ and 2 of the human brain data are shown in Fig. 3.3. The air background area is kept in the figures to show the effect of pure noise. The coefficients are computed using the 60 point quadrature formula Eq. 3.17. The coefficients of $\ell = 1$ are mainly caused by imaging artifacts and Rician

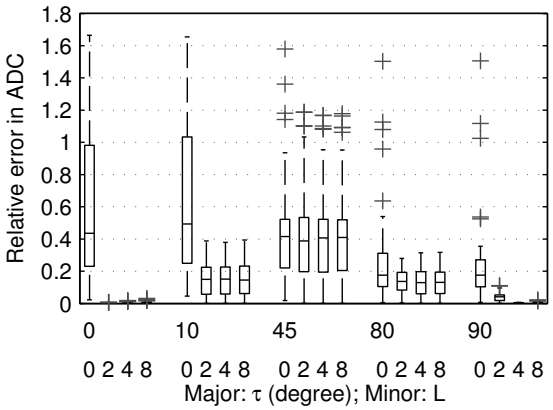
www.manaraa.com



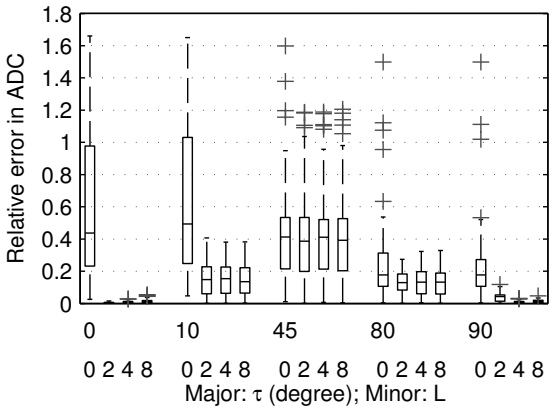
(a) $\sigma = 10$, noise only



(b) $\sigma = 25$, noise only



(c) $\sigma = 10$, truncated spherical harmonics



(d) $\sigma = 25$, truncated spherical harmonics

Figure 3.2: The relative error in ADC profiles when the spherical harmonics are truncated to different degrees to describe the biGaussian ADC profiles with different angle τ between the two tensor compartments. The errors on each gradient direction are used to create the boxes in the box plots. On each box of the box plots, the central mark is the median of values, the edges of the box are the 25th and 75th percentiles, the whiskers show the most extreme data points not considered outliers, and outliers are plotted individually using "+". (a) and (b) show the error in ADC profiles caused by Rician noise; (c) and (d) show the error in ADC profile caused by both noise and truncating the spherical harmonics to degree L. The figures shown here are for $b = 1000\,\mathrm{s/mm^2}$ and $S_0 = 500$.

www.manaraa.com

noise, which appear in all degrees. The model residual for truncating the spherical harmonics to $L = 0, 2,$ and 4 is shown Fig. 3.4. Truncating the spherical harmonics to $L = 0$ equals the isotropic model, and truncating to $L = 2$ is the same as the diffusion tensor model. Some major fiber bundles can be seen in Fig. 3.4(a) because the isotropic model is not able to describe anisotropic diffusion. From Figs. 3.4(b) and 3.4(c), we can see that the error in ADC profiles for different brain structure is different, indicating different L can be chosen for different applications.

(a) $\ell = 0$

(b) $\ell = 1$

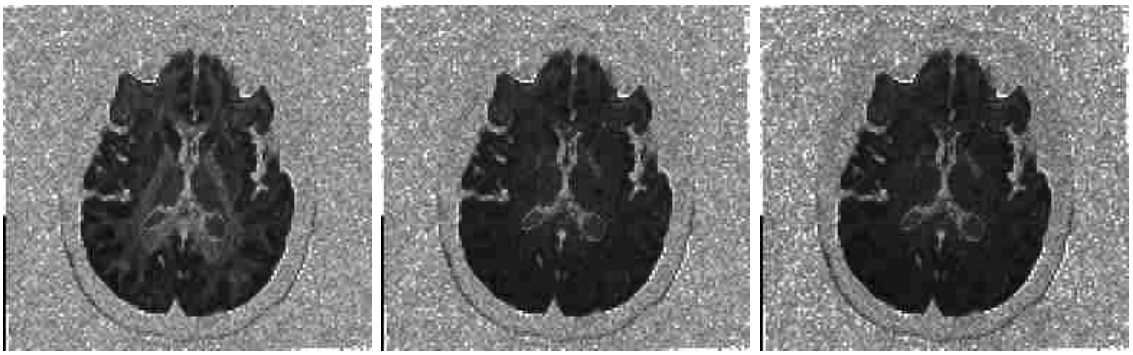
(c) $\ell = 2$

Figure 3.3: The magnitude of the spherical harmonic expansion coefficients of an axial slice of the human brain data for $\ell = 0, 1,$ and 2, and $m = -\ell, \ldots, \ell$. The ranges of the coefficients of different ℓ are different, and the colormap for images of each ℓ is adjusted to get proper contract. (b) and (c) are made 5 and 10 times brighter so that the brightness is similar to that of (a).

3.4 Discussion

The development of Fourier series is an important tool in rectangular domains. The spherical harmonics represent functions on a sphere in the same way as Fourier se-

www.manaraa.com



(a) $L = 0$ (b) $L = 2$ (c) $L = 4$

Figure 3.4: An axial slice of the root mean square error of the ADC profile reconstructed from spherical harmonics truncated to different L of the human brain data.

ries do in rectangular domains. Many methods in Fourier analysis can be used in spherical harmonic analysis by taking expansions in spherical harmonics rather than trigonometric functions. Therefore, it worth to investigate fast computing algorithms analogous to the discrete Fourier transform.

Gradient schemas from icosahedron tessellations or truncated icosahedron are often used in HARDI studies to acquire the data. We propose to use the quadrature formulas developed in numerical integration on spheres to carry out SHT for these schemas. We also show that some other methods are exactly the same as the quadrature formulas on these schemas. LS minimization is used in some HARDI studies. It is shown in this chapter that the method of inverting the matrix $M^T M$ in LS minimization should be avoided on these schemas, because the matrix $M^T M$ is ill-conditioned when $L > 4$. The gradient schema in HARDI studies should be chosen carefully to maintain numerical stability in calculating spherical harmonic coefficients.

Using spherical harmonics truncated to a degree L to present the non-Gaussian ADC profiles is a popular approach in HARDI studies. However, recovering Gaussian mixture models is not always supported by the relatively low b-values and relatively low angular resolution in some studies. This is one of the reasons why recovering Gaussian mixture models with more than two tensor compartments is reported not

www.manaraa.com

numerically stable (Tuch et al., 2002; Kreher et al., 2005).

Attention should also be paid to some simply model selection approaches. It is expected that it is difficult to recover Gaussian mixture models with closely aligned tensor compartments. A less intuitive fact is that when tensor compartments are perpendicular to each other, the diffusion tensor model may also fit the ADC profile well.

Copyright© Ning Cao, 2013.

www.manaraa.com

Chapter 4 Fitting Diffusion Models to Data

In Ch. 3, we look for spherical functions that describe the shape of ADC profile using diffusion-weighted signals measured at different gradient directions. In this chapter, we describe how we fit different diffusion models to data. By fitting a diffusion model to data, useful information, for example, local anisotropy, local fiber orientation, can be derived from the model. We focus on finding the value of parameter vector $\mathbf{p}$ that "best fits" a given diffusion model $\mathbf{M}(\mathbf{p})$. Sometimes it is not straightforward to choose one diffusion model over the other. We discuss how we choose a suitable diffusion model from two candidate models or a set of candidate models based on the result of this chapter in Ch. 5.

4.1 Introduction

For a given parametric model $\mathbf{M}(\mathbf{p})$ with parameter vector $\mathbf{p}$, we may either take the descriptive approach which looks for a parameter vector $\mathbf{p}$ so that the difference between the measured data and the values predicted by the model $\mathbf{M}$ is minimized, or look for the parameter that "most likely" to produce the data, with or without a prior distribution. When discussing in the context of diffusion models, $\mathbf{M}$ can be any diffusion model, for example, diffusion tensor model and biGaussian model. These techniques calculate a point estimate.

The descriptive approach of fitting a diffusion model often looks for parameter vector $\mathbf{p}$ that minimizes the mean squared error between the diffusion-weighted signal predicted by the model and the measured diffusion-weighted signal S_i:

$$\mathbf{p} = \arg\min_{\mathbf{p}\in\mathbf{P}} \frac{1}{n}\sum_{i=1}^{n}(\mathbf{M}(\mathbf{p}|\vec{g}_i, b, S_0) - S_i)^2, \tag{4.1}$$

where $\mathbf{M}(\mathbf{p}|\vec{g}_i, b, S_0)$ is the diffusion-weighted signal predicted by model $\mathbf{M}$ with $\mathbf{p}$ being the variable of $\mathbf{M}$. Note that the widely used linear LS method (Basser et al.,

www.manaraa.com

1994a) in solving Eq. 1.6 minimizes mean squared error of ADC, instead of diffusion-weighted signal:

$$\mathbf{p} = \arg\min_{\mathbf{p}\in\mathbf{P}} \frac{1}{n}\sum_{i=1}^{n}(\log \mathbf{M}(\mathbf{p}|\vec{g}_i, b, S_0) - \log S_i)^2. \tag{4.2}$$

Linear LS is identical to assuming the distribution of the error terms to belong to a normal distribution with mean zero. Besides linear LS, weighted linear LS, and nonlinear fits are also presented in previous studies (Kingsley, 2006b).

The MLE method maximizes the likelihood function without using any prior distribution, while a maximum a posteriori probability (MAP) estimation (Behrens et al., 2003; Friman et al., 2006) uses the prior distribution of the parameter being estimated. The major argument of the MAP estimation is that the optimization objective incorporates a prior distribution of the parameters being estimated. Landman et al. use the MLE approach to fit the diffusion tensor model (Landman et al., 2007), but they do not use the confidence interval to indicate the reliability of the estimate. Behrens et al. present a probabilistic framework based on the MAP estimation and report the marginal posterior distributions of the parameters regarding the fiber orientations to illustrate how well the fit is (Behrens et al., 2003). However, there was no practical guideline for deciding the prior distributions. Therefore, the prior distributions are somehow decided arbitrarily.

In addition to diffusion tensor model, we use the MLE approach on biGaussian models with different number of parameters. We also compute the covariance matrix of the parameters to get a confidence interval and for model selection.

4.2 Method

For a given voxel, suppose we have diffusion-weighted signals measured on n different gradient directions with r acquisitions on each gradient direction. It is straightforward to treat the r measurements on each gradient direction as samples of a Rician

www.manaraa.com

distribution, and the parameters ν and σ of the Rician distribution can be estimated together using the r repetitions. It is also possible to estimate σ of the Rayleigh distribution for the whole image using Eq. 2.3 first, so that ν can be estimated with σ being fixed (Sijbers et al., 1998a; Kristoffersen, 2007). However, this approach does not always work properly if r is small, especially when b-value is large and SNR is low. Instead, we consider the r repetitions as r samples of the diffusion model, and each sample has n measurements on different gradient directions.

Local Diffusion Models

In order to evaluate the MLE method, we choose the following parametric models:

1. The plain diffusion model, or the isotropic model, defined in Eq. 1.2. It only has one parameter, the diffusion constant D.

2. The diffusion tensor model defined in Eq. 1.6. It has three parameters for the shape and another three parameters for rotations. The parameter vector is

$$\mathbf{p}_{dt} = (\lambda_1, \lambda_2, \lambda_3, \varphi, \vartheta, \psi). \tag{4.3}$$

3. The biGaussian model defined in Eq. 1.19. If the two tensor compartments are not related, there are totally 13 parameters, with six for each tensor compartment and one for volume ratio $\mathbf{f}$. If the shapes of the two tensor compartments are exactly the same, the number of parameters reduces to 10. The parameter vector of the biGaussian model with 13 parameters is

$$\mathbf{p}_{bg13} = (\lambda_1^{(1)}, \lambda_2^{(1)}, \lambda_3^{(1)}, \varphi^{(1)}, \vartheta^{(1)}, \psi^{(1)}, \lambda_1^{(2)}, \lambda_2^{(2)}, \lambda_3^{(2)}, \varphi^{(2)}, \vartheta^{(2)}, \psi^{(2)}, \mathbf{f}), \tag{4.4}$$

and the parameter vector of the model with 10 parameters is

$$\mathbf{p}_{bg10} = (\lambda_1, \lambda_2, \lambda_3, \varphi^{(1)}, \vartheta^{(1)}, \psi^{(1)}, \varphi^{(2)}, \vartheta^{(2)}, \psi^{(2)}, \mathbf{f}). \tag{4.5}$$

www.manaraa.com

It is possible to further restrict or even fix the anisotropy of tensor compartments (Tuch et al., 2002) to reduce the number of parameters to make numerical optimization more stable. Also, some simplified diffusion models (Frank, 2001; Alexander et al., 2001) make numerical optimization much easier.

We want to try other complex diffusion models as well, for example, the general Gaussian mixture model. However, the numerical optimization is slow and is not stable. It is also reported in other studies that Gaussian mixture model with $K = 3$ is unstable (Tuch et al., 2002; Kreher et al., 2005).

Maximum Likelihood Estimation

We regard the diffusion-weighted signals S_i measured at a gradient direction $\vec{g}_i$ as observations of the Rician distribution with PDF $f(x|A_i, \sigma_i)$, where A_i is the diffusion-weighted signal predicted by the diffusion model $\mathbf{M}(\mathbf{p})$ at gradient direction $\vec{g}_i$, and the scale parameter σ_i of diffusion-weighted image $\mathbf{S}_i$ can be estimated using the mean estimator of the Rayleigh distribution Eq. 2.7.

We look for the parameter $\mathbf{p} \in \mathbf{P}$, where $\mathbf{P}$ is the allowable set of parameter $\mathbf{p}$, that "best fits" the diffusion model $\mathbf{M}$ for a given voxel, by maximizing the joint PDF

$$f(S_1, \ldots, S_n|\mathbf{p}) = \prod_{i=1}^{n} f(S_i|A_i, \sigma_i) \tag{4.6}$$

of these n individual Rician distributions. The joint PDF Eq. 4.6 is an extension of the one we used in Ch. 2, allowing PDFs with different parameters in the joint PDF. When there are r repetitions for each gradient direction, the term of $f(S_i|A_i, \sigma_i)$ is replaced by $\prod_{j=1}^{r} f(S_i^{(j)}|A_i, \sigma_i)$, where $S_i^{(j)}$ represents the j-th repetition. Although the unweighted signal S_0 can also be modeled using Rician distribution, we treat it as a parameter of the diffusion model to simplify the computation. We use the value recovered from the Rician distribution as S_0, in order to simplify the calculation of the likelihood function. We do so without much concern because the SNR of

www.manaraa.com

the unweighted signal is always much higher, and it is often acquired with multiple repetitions. The likelihood function $\mathcal{L}$ is therefore defined as

$$\mathcal{L}(\mathbf{p}|S_1,\ldots,S_n) = \prod_{i=1}^{n} f(S_i|\mathbf{M}(\mathbf{p}|\vec{g}_i,b,S_0),\sigma_i), \tag{4.7}$$

and we find the value of $\mathbf{p}$ that maximizes the natural logarithm of the likelihood function, which is called the log-likelihood function,

$$\log \mathcal{L}(\mathbf{p}|S_1,\ldots,S_n) = \sum_{i=1}^{n} \log f(S_i|M(\mathbf{p}|\vec{g}_i,b,S_0),\sigma_i), \tag{4.8}$$

that is

$$\hat{\mathbf{p}}_{mle} = \arg\max_{\mathbf{p}\in\mathbf{P}} \log \mathcal{L}(\mathbf{p}|S_1,\ldots,S_n). \tag{4.9}$$

The logarithm function is monotonically increasing, thus it achieves the maximum value at the same point as the original function. And, the logarithm makes calculating the gradient of the log-likelihood function Eq. 4.8 simpler. The derivative of the logarithm of the Rician PDF $\log f(x)$ is

$$(\log(f(x)))' = \frac{\nu}{\sigma^2} - \frac{x}{\sigma^2}\cdot\frac{I_1(\frac{\nu x}{\sigma^2})}{I_0(\frac{\nu x}{\sigma^2})}, \tag{4.10}$$

where $I_1(x)$ is modified Bessel function of the first kind of order one. The gradient functions of diffusion model $\mathbf{M}$ can then be used together with the derivative of the logarithm of the PDF Eq. 4.10 by applying the chain rule to compute the gradient of log-likelihood function Eq. 4.8.

For the isotropic model and the diffusion tensor model, we use the solution obtained from Eq. 1.2 and tensor diagonalization Eq. 1.8 as the starting point. For other models, we use multiple restarts with starting points chosen from the whole allowable parameter set $\mathbf{P}$ randomly because there is no easy way to find good starting points. We do not use feedbacks from fiber tracking algorithm in generating starting points.

The confidence interval of $\mathbf{p}_{mle}$ is calculated by assuming that $\mathbf{p}_{mle}$ is normally distributed. The asymptotic covariance of ML estimator at $\mathbf{p}_{mle}$ is approximated by finite difference. The confidence interval is then decided by the CDF of the normal distribution with corresponding standard deviation.

www.manaraa.com

Optimization

The optimization problem Eq. 4.9 is generally solved by iterative methods because there is no known direct method. The optimization problem of available solvers is often stated in terms of minimization. Therefore, instead of maximizing the log-likelihood Eq. 4.8, we minimize the negative log-likelihood function. When the local diffusion model is more complex than the one-parameter isotropic model, the objective function often has more than one local minimum. We use both local optimization and global optimization.

For local minimums, we use constrained nonlinear optimization with linear constraints

$$A_{ineq} \cdot \mathbf{p} \leq b_{ineq}, \tag{4.11}$$

as well as lower bound lb and upper bound ub, such that $lb \leq \mathbf{p} \leq ub$. We compare different gradient-based algorithms (Byrd et al., 1999, 2000, 2006) available in MATLAB and KNITRO (Byrd et al., 2006) on our optimization problems and use interior point methods. In order to compare different available local optimization methods, we run different solvers from the same set of starting points to find local minimums. In Fig. 4.1, we plots all the local maximums of the log-likelihood function found from a set of random starting points by each method on fitting the biGaussian model to data. In the comparison, the interior point algorithm reaches local minimums similar to the global minimum for more starting points.

We use lower bound lb and upper bound ub to limit the tensor size, as well as the range of Euler angles. The linear constraints Eq. 4.11 are used to assure proper ordering of eigenvalues. The Hessian matrix is computed numerically by central difference of the gradient. The solvers scale the objective function and constraints.

It is common for the objective function Eq. 4.8 to have multiple local minimums. Even for a diffusion tensor model with fixed eigenvalues, the objective function may have multiple local minimums. Running the interior point search from equal-spaced

www.manaraa.com

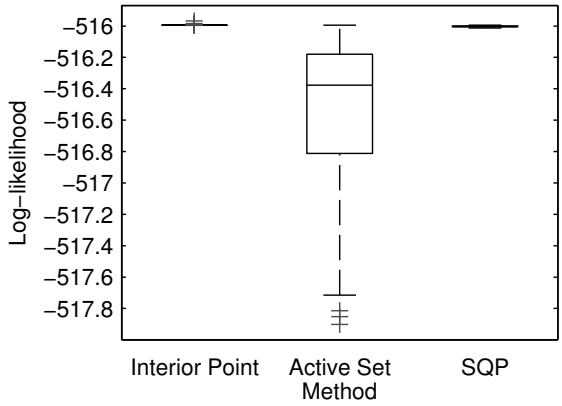


Figure 4.1: Comparison of different local optimization methods on fitting the biGaussian model to data from the same set of starting points. Local maximums returned by each method are plotted in a box plot. On each box of the box plot, the central mark is the median of values, the edges of the box are the 25th and 75th percentiles, the whiskers show the most extreme data points not considered outliers, and outliers are plotted individually using "+". All three methods returns multiple local maximums that have log-likelihood value similar to that of the global maximum. However, the active set method only returns a few good local maximums and the global maximum is returned by the interior point method.

grid points in a subspace of the parameter space leads to multiple local minimums. By clustering on the vectors rotated by the local minimums into three clusters, we got the approximate basins of attraction of the specific optimization algorithm on the data. The clustering is based on the cosine of the included angle between rotated vectors, and the number of clusters is decided visually. We plot the basins of attraction and the local gradients in Fig. 4.2. When working on the biGaussian model, we notice that the solvers return multiple local minimums.

We use multiple restart on simpler models and global optimization (Ugray et al., 2007) for biGaussian models. For using multiple restarts, we run the local solver from randomly picked starting points within the bounds given by lb and ub, and choose the local minimum with lowest objective function value. For the global optimization approach, a large set of trail points are evaluated and only some of them are used as starting points (Ugray et al., 2007).

www.manaraa.com

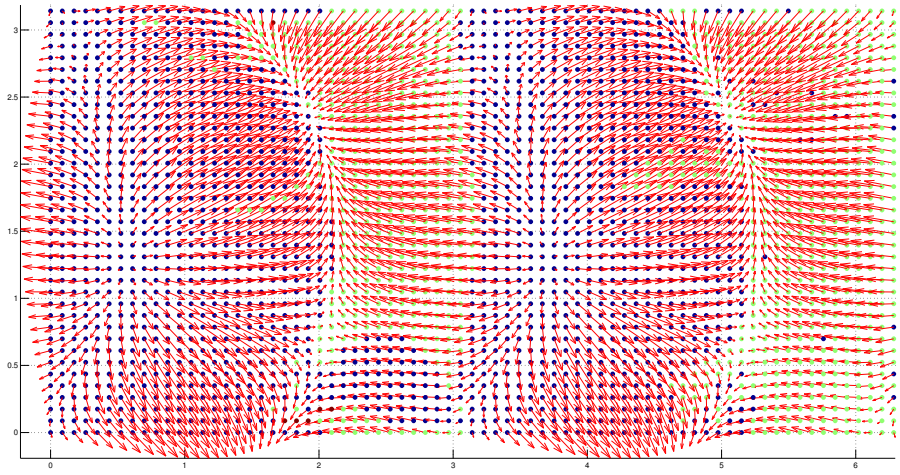


Figure 4.2: Basins of attractions and local gradients of the negative log-likelihood function of fitting a diffusion tensor to synthetic data.

Local Maxima and Multiple Global Maxima

Optimization algorithms cannot always reach the global maxima. When the objective function has local maxima, we choose multiple starting points randomly and pick the solution that maximizes the likelihood. For simple models, e.g., the diffusion tensor model, we can analytically compute the LS solutions as starting points, which help avoid local maxima problem. However, when working on more complex models, e.g., the biGaussian model, optimization algorithms suffer from local maxima. We had no practical method to find good starting points for biGaussian model. Therefore we choose enough starting points so that we have some confidence that the global maxima can be reached. More than 300 starting points are used with real data.

There may be multiple local maxima with likelihood values similar to the global maxima. Our simulations on the biGaussian model show that this phenomenon is common. It indicates that we should not use the likelihood value as the only criterion for complex models. One of the possible reasons is that the diffusion profiles of some

www.manaraa.com

configurations do not show obvious biGaussian structure. Instead of picking the global maxima or picking one of multiple global maxima, we report all solutions with likelihood value similar to the global maxima as candidates. When a voxel of biGaussian model is encountered, tractography algorithm can use neighborhood information to pick one candidate from those reported by model fitting process.

Variance of Estimator

The variance of parameter can be used to measure the reliability of the estimates. Larger variance indicates the estimates are less precise. Generally, the estimates from MLE can be considered following a normal distribution. The confidence interval can then be derived from the variance.

Asymptotic normality

The distribution of an estimate obtained from MLE tends to a normal distribution when the sample size increases to infinity. The mean of this distribution is the parameter $\mathbf{p}$ being estimated, and the covariance matrix Σ is the inverse of Fisher information matrix $\mathcal{I}(\mathbf{p})$. The Fisher information matrix is defined as the Hessian matrix of the log-likelihood function Eq. 4.8. As we only have the estimation of $\mathbf{p}$, which is $\hat{\mathbf{p}}_{mle}$, we use

$$\mathbf{N}(\hat{\mathbf{p}}_{mle}, \mathcal{I}^{-1}(\hat{\mathbf{p}}_{mle})) \tag{4.12}$$

to approximate $\mathbf{N}(\mathbf{p}, \mathcal{I}^{-1}(\mathbf{p}))$.

Covariance matrix Σ can also be estimated from outer product of gradient of likelihood function $\mathcal{L}$. Let J be the Jacobian matrix of $\mathcal{L}$ when $\mathcal{L}$ is maximized, $(J'J)^{-1}$ is the estimate of covariance matrix.

Estimating the covariance matrix Σ using Fisher information matrix or outer product of gradient is straightforward and does not require much computation. The Hessian matrix can be computed from the Jacobian matrix, which is often computed

www.manaraa.com

analytically for gradient based optimization. However, the variance of the estimator also depends on the solver used in optimization. For example, the variance of the estimate will be greater if the step size or tolerance is too large.

Bootstrapping

Bootstrapping method can be used to estimate properties of an estimator by measuring them from an approximating distribution. A standard way of getting an approximating distribution is to resample the data with replacement to get the same number of samples.

For diffusion-weighted data, we have one or more repetitions on each gradient direction. We treat a measurement on gradient direction $\vec{g}_i$ as a sample. As the number of gradient directions is generally large for HARDI data, we use the Monte Carlo approach to perform the resampling. The procedure of bootstrapping is simple but it requires a lot of computation.

Simulation

We evaluate the MLE approach by simulation. Parameters estimated from the models are compared to the true values used to generate the data. We use the 60 vertices of the truncated icosahedron showed in Fig. 2.5(a) as gradient directions to generate diffusion-weighted data.

Isotropic and Diffusion Tensor Models

For both isotropic model and diffusion tensor model, we generate synthetic data from diffusion tensor model. Diffusion-weighted images are generated with different eigenvalues and different b-values. Rician noise with predefined parameter σ is then added to the diffusion-weighted image. For the isotropic model, we use $(1200,\ 1190,\ 1180) \times 10^{-6}\,\mathrm{mm}^2\ /\mathrm{s}$ as eigenvalues to generate nearly isotropic diffu-

www.manaraa.com

sion profiles; for the diffusion tensor model, we use $(800,\ 300,\ 300) \times 10^{-6}\,\mathrm{mm}^2$ /s, $(800,\ 400,\ 400) \times 10^{-6}\,\mathrm{mm}^2$ /s, $(1200,\ 400,\ 300) \times 10^{-6}\,\mathrm{mm}^2$ /s, and $(1200,\ 300,\ 300) \times 10^{-6}\,\mathrm{mm}^2$ /s as eigenvalues to generate diffusion tensors of different shapes. We also add rotations to the diffusion profiles to test the parameterization of the diffusion tensor model.

BiGaussian Model

We use the same eigenvalues for both tensor compartments to generate biGaussian profile. We change the angle τ between the two tensor compartments to simulate fiber crossing at different angles.

When evaluating the fitting results, we mainly focus on tensor orientations. The angle difference η_{err} (Tuch et al., 2002) is defined as

$$\eta_{err} = \frac{1}{2}\min\{\arccos(\hat{\vec{e}}_1 \cdot \vec{e}_1) + \arccos(\hat{\vec{e}}_2 \cdot \vec{e}_2), \arccos(\hat{\vec{e}}_1 \cdot \vec{e}_2) + \arccos(\hat{\vec{e}}_2 \cdot \vec{e}_1)\}, \quad (4.13)$$

where $\vec{e}_1$ and $\vec{e}_2$ are the true principal eigenvectors used to generate the synthetic data, and $\hat{\vec{e}}_1$ and $\hat{\vec{e}}_2$ are the principal eigenvectors recovered by optimization.

We create biGaussian profiles by mixing two tensors of different shape: the "thin" tensor has eigenvalues $(1200,\ 300,\ 300) \times 10^{-6}\,\mathrm{mm}^2$ /s and the "wide" tensor has eigenvalues $(1200,\ 400,\ 400) \times 10^{-6}\,\mathrm{mm}^2$ /s. Simulation is carried out for different b-values and for models with different numbers of parameters. We use 0.48 as the volume ratio so that the weights of the two tensor compartments are similar.

Phantom Experiments

We use the rat phantom data described in Ch. 2 to conduct phantom experiments. We compare the recovered ADC profile with the noisy ADC profile to assess the quality of fitting, because the true parameter values are unknown. The goodness of the fit is measured by the log-likelihood function Eq. 4.8. A region of interest (ROI)

www.manaraa.com

in an $x - z$ slice containing two fiber crossing is identified manually and shown in Fig. 4.3.

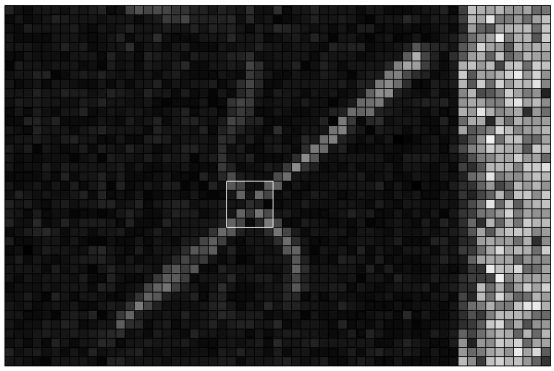


Figure 4.3: The crossing region of the rat phantom, marked with a rectangle on the FA map.

Human Brain Data

We fit different diffusion models to the human brain data described in Ch. 2. A 5×3 ROI shown in Fig. 4.4 including part of the cingulum tract and part of the corpus callosum is manually identified on a coronal slice. The two rows from top are within the cingulum tract, and two rows from bottom are within the corpus callosum. The third row contains a mixture of the above two. The crossing region can also be identified visually by the shape of the ADC profiles in Fig. 4.13(a).

4.3 Results

Simulation

Isotropic Model

We gather the results of the following approaches

1. recovering the diffusion coefficient using the MLE method;

www.manaraa.com

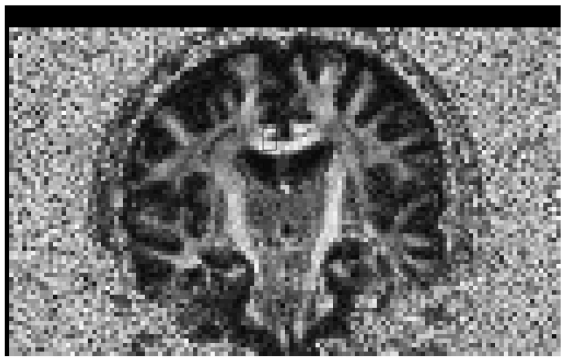


Figure 4.4: The crossing region of the human brain data, marked with a rectangle on a coronal slice of the FA map. Both the cingulum tract and the corpus callosum are included in the ROI.

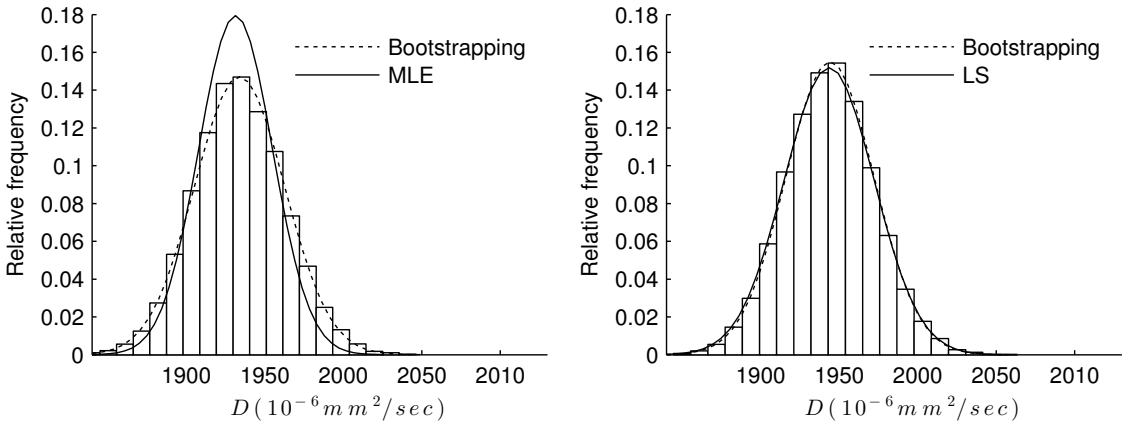
2. solving Eq. 1.2 using linear LS directly without processing the Rician noise;

3. recovering the underlying diffusion-weighted signal assuming Rician noise and then solving Eq. 1.2 using linear LS.

Recovering the underlying diffusion-weighted signal with high noise level, high b-values and no repetition does not work well, because the diffusion-weighted signals are often weak. So we only compare the result of the MLE method with that of solving Eq. 1.2 directly.

The result of the MLE method is similar to that of solving Eq. 1.2 using linear LS directly. We carry out paired samples t-test and no significant difference between the results is found when N is small, i.e., $N = 100$. To compare the MLE and the LS method, the diffusion coefficient from a voxel of the rat phantom is estimated using both methods. The variance estimated from both methods using asymptotic normality and bootstrapping using 10000 samples are plotted in Fig. 4.5. We plot the histograms of the diffusion coefficient D estimated from each sample of the bootstrapping, as well as the normal distribution fitted to the histogram.

www.manaraa.com



(a) MLE (b) LS

Figure 4.5: Histograms of the diffusion coefficient D estimated from the isotropic diffusion model using MLE and the LS method for estimating the variance of each estimator, obtained from bootstrapping. The normal distributions fitted to the histograms are shown using dotted line, and the normal distributions obtained from asymptotic normality for both methods are plotted using solid line.

Diffusion Tensor Model

In fiber tracking, we mainly use the principal eigenvector of the diffusion tensor as the local orientation. So, we compare the angle difference between the ground truth value used to generate the synthetic data with the result of the MLE method and that of the LS method.

We carry out one-tailed paired samples t-test to compare the results of the two methods. The angle difference of the MLE method is smaller for some configurations with higher b-value, higher noise level and more anisotropic tensor shapes. This is expected because the MLE approach takes the Rician noise into consideration, even if there is no repetition. Simulation result of some test settings are shown in Fig. 4.6 as box plots.

www.manaraa.com

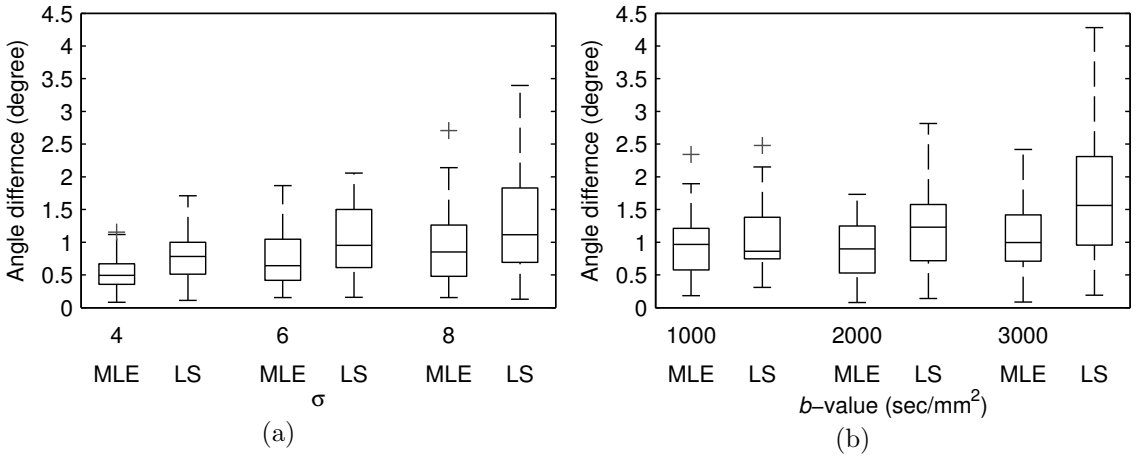


Figure 4.6: Simulation result of fitting the diffusion tensor model to data using the MLE method and the LS method. For each b-value and σ combination, the same input is used to compare the two methods. 50 trials are used to create each box of the box plot. On each box of the box plots, the central mark is the median of values, the edges of the box are the 25th and 75th percentiles, the whiskers show the most extreme data points not considered outliers, and outliers are plotted individually using "+". (a) $\Lambda = (1200, 300, 300) \times 10^{-6}\,\mathrm{mm}^2$ /s, $b = 2000\,\mathrm{s/mm}^2$, and $\sigma = 4, 6,$ and 8; (b) $\Lambda = (1200, 400, 300) \times 10^{-6}\,\mathrm{mm}^2$ /s, $b = 1000, 2000,$ and $3000\,\mathrm{s/mm}^2$, and $\sigma = 6$.

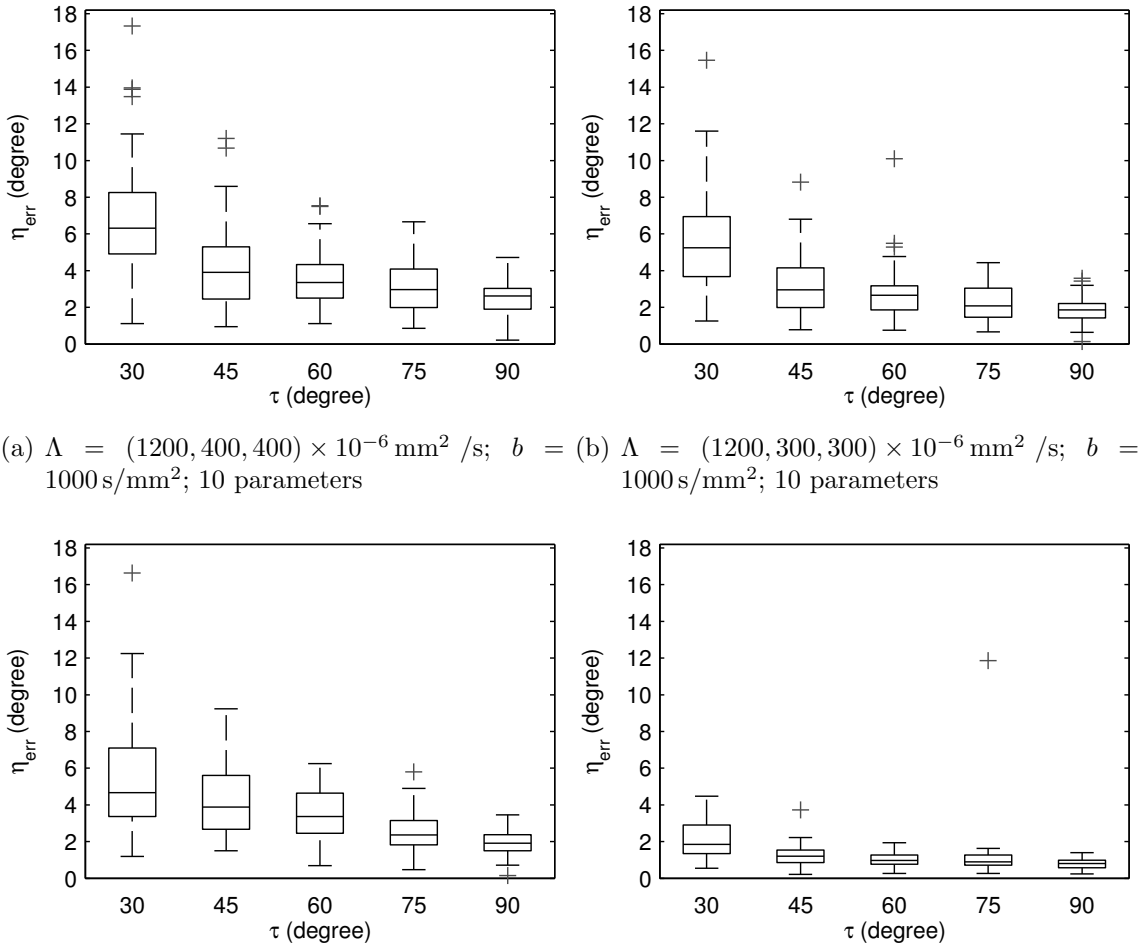
BiGaussian Model

The angle error η_{err} of fitting the biGaussian models to data with Rician noise added are shown in Fig. 4.7, for τ between 30° and 90°. We carry out global optimization for 50 trails to create the box plots for each test configuration. The simulations show that the angle error is smaller for larger b-values, more anisotropic tensor compartments and fewer numbers of parameters.

Phantom Experiments

Four repetitions of the rat phantom were scanned. Ideally we can use these four repetitions as a larger sample. However, from the diffusion-weighted images of the four repetitions, we notice that the repetitions are not well aligned. Therefore, we fit these four repetitions individually without registering them, to avoid blurring the thin

www.manaraa.com



(a) $\Lambda = (1200, 400, 400) \times 10^{-6}\,\mathrm{mm}^2\,/\mathrm{s}$; $b = 1000\,\mathrm{s/mm}^2$; 10 parameters

(b) $\Lambda = (1200, 300, 300) \times 10^{-6}\,\mathrm{mm}^2\,/\mathrm{s}$; $b = 1000\,\mathrm{s/mm}^2$; 10 parameters

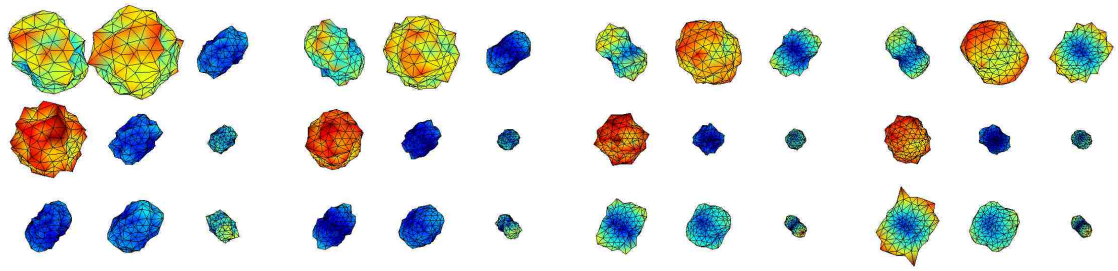
(c) $\Lambda = (1200, 300, 300) \times 10^{-6}\,\mathrm{mm}^2\,/\mathrm{s}$; $b = 1000\,\mathrm{s/mm}^2$; 13 parameters

(d) $\Lambda = (1200, 400, 400) \times 10^{-6}\,\mathrm{mm}^2\,/\mathrm{s}$; $b = 3000\,\mathrm{s/mm}^2$; 10 parameters

Figure 4.7: Simulation results of fitting the biGaussian model to data using the MLE approach. Box plots are shown for different b-values, different tensor shapes, and different number of parameters. On each box of the box plots, the central mark is the median of values, the edges of the box are the 25th and 75th percentiles, the whiskers show the most extreme data points not considered outliers, and outliers are plotted individually using "+".

www.manaraa.com

fibers with resampling. ADC profiles from the four repetitions of the same crossing region are shown in Fig. 4.8 to illustrate the difference between them, especially the voxel in the center of the region.



(a) repetition 1 (b) repetition 2 (c) repetition 3 (d) repetition 4

Figure 4.8: Noisy ADC profiles within the crossing region shown in Fig. 4.3 of the rat phantom to show the difference of different repetitions caused by noise and imaging artifacts.

We apply MLE method on the rat phantom using the isotropic model, the diffusion tensor model and the biGaussian models. The noise level of each individual diffusion-weighted image is estimated in Ch. 2. In our experiment, more complex models have higher likelihood values, indicating they fit the data better.

Since the individual unweighted images were not saved by the scanner and only the average is available, we use the averaged value for S_0.

ADC profiles of the region shown in Fig. 4.3 recovered from different diffusion models are shown in Fig. 4.9, and the log-likelihood values are plotted in Fig. 4.10 as horizontal bar graphs. The shapes of the ADC profiles show that the diffusion tensor model fits most of the voxels well. For voxels exhibiting crossing structures, the biGaussian models fit the noisy ADC profiles better than the diffusion tensor model does. However, from the shapes of the ADC profiles and the log-likelihood values, we find that the biGaussian model with 13 parameters is only slightly better than the biGaussian model with 10 parameters for these voxels.

Local fiber orientations recovered from the diffusion tensor model and the bi-

www.manaraa.com

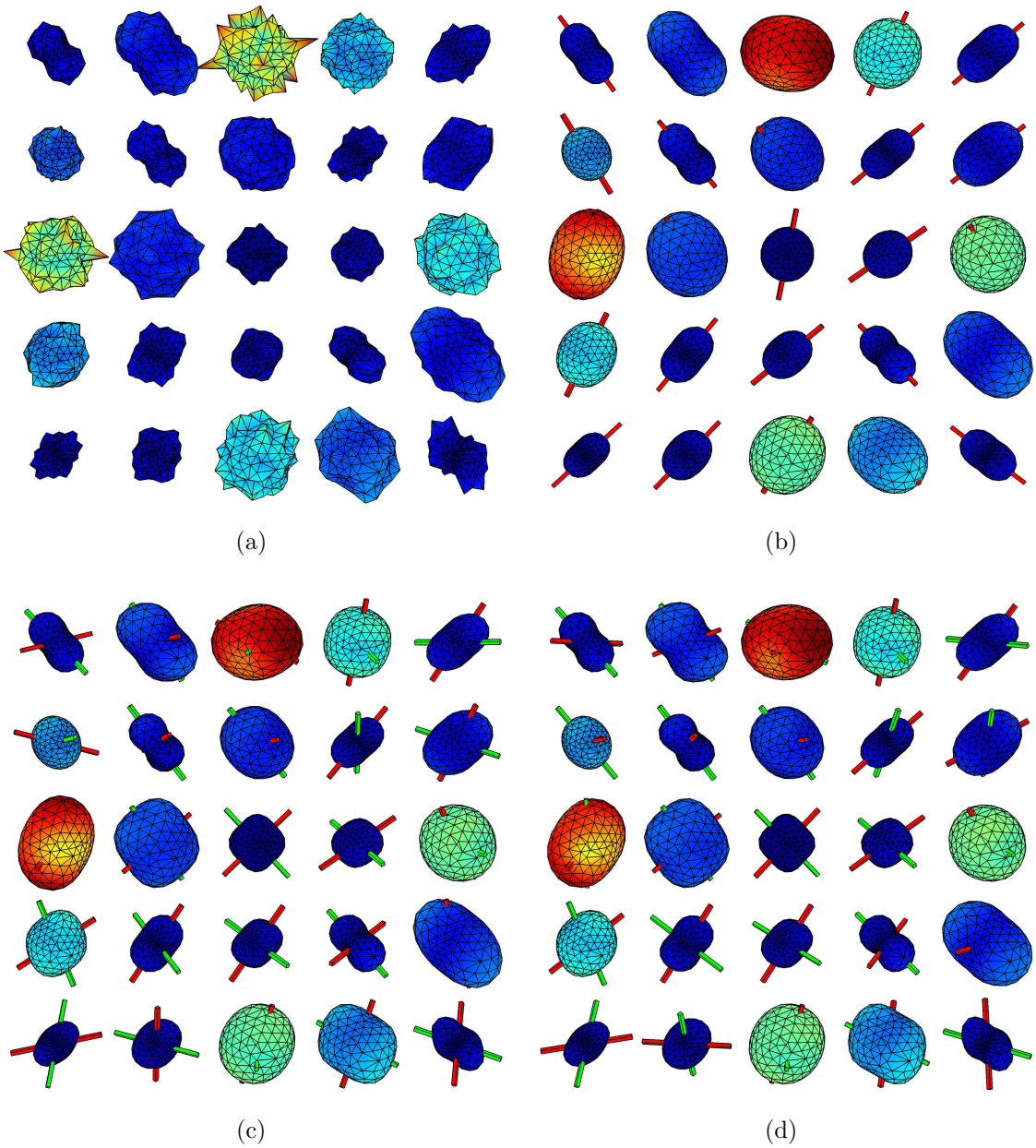


Figure 4.9: Fitting different diffusion models to the crossing region of the rat phantom data. The crossing region is shown in Fig. 4.3. (a) Noisy ADC profiles of the crossing region; (b) ADC profiles recovered from the diffusion tensor model, together with the principal directions of the diffusion tensors; (c) ADC profiles recovered from biGaussian model with 10 parameters, together with the principal directions of each tensor compartments; (d) ADC profiles recovered from biGaussian model with 13 parameters, together with the principal directions of each tensor compartments. The ADC profiles of some voxels are scaled to show the details with good contrast.

www.manaraa.com

Gaussian models of the same crossing region are also plotted in Figs. 4.9(b), 4.9(c), and 4.9(d). The fiber orientations recovered from the biGaussian models are more reasonable than those recovered from the diffusion tensor model for voxels within the crossing area. However, the difference between the results of the biGaussian models with different numbers of parameters is not significant. For the diffusion tensor model, fiber orientation recovered from a more anisotropic voxel is more reliable than that recovered from a more isotropic voxel.

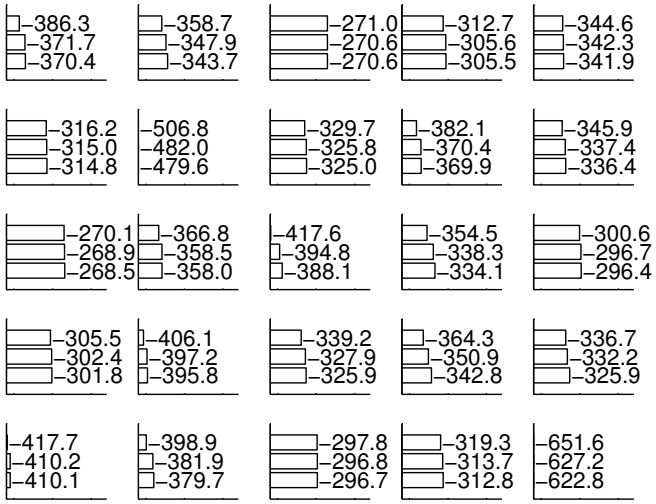


Figure 4.10: Log-likelihood values of fitting the diffusion tensor model, the biGaussian model with 10 parameters and the biGaussian model with 13 parameters to the crossing region of the rat phantom shown in Fig. 4.3. Log-likelihood values of the three diffusion models are plotted for each voxel within the ROI as horizontal bars from top to bottom.

We estimate the asymptotic covariance matrix of the ML estimator using Eq. 4.12. The covariance matrix as well as the marginal distributions of the eigenvalues and Euler angles are plotted separately.

We fit the diffusion tensor model to the diffusion profile of a voxel within a single fiber tract. Figs. 4.11(a) and 4.11(b) show the covariance matrix of the ML estimator. The diffusion tensor model describes the anisotropic diffusion within that voxel well and the covariance is small.

www.manaraa.com

The diffusion tensor model is fit to the diffusion profile of a voxel containing crossing fiber tracts, and the covariance matrix of the estimator is plotted in Figs. 4.11(c) and 4.11(d). Both the planar-shape of the tensor characterized by $\lambda_1 \approx \lambda_2 > \lambda_3$ in Fig. 4.11(c) and the covariance matrix of the Euler angles shown in Fig. 4.11(d) indicate that the uncertainty of the principal eigenvector is large.

We also fit the biGaussian model with 10 parameters to the diffusion profile of the same voxel and the covariance matrix is plotted in Fig. 4.12. The distribution of the estimated Euler angles fitted to the biGaussian model is narrower than that of fitting to the diffusion tensor model, because the biGaussian model describes the data of a fiber crossing voxel better. However, the relatively wide distribution in Fig. 4.12 indicates that the quality of fitting is not as good as that of Figs. 4.11(a) and 4.11(b).

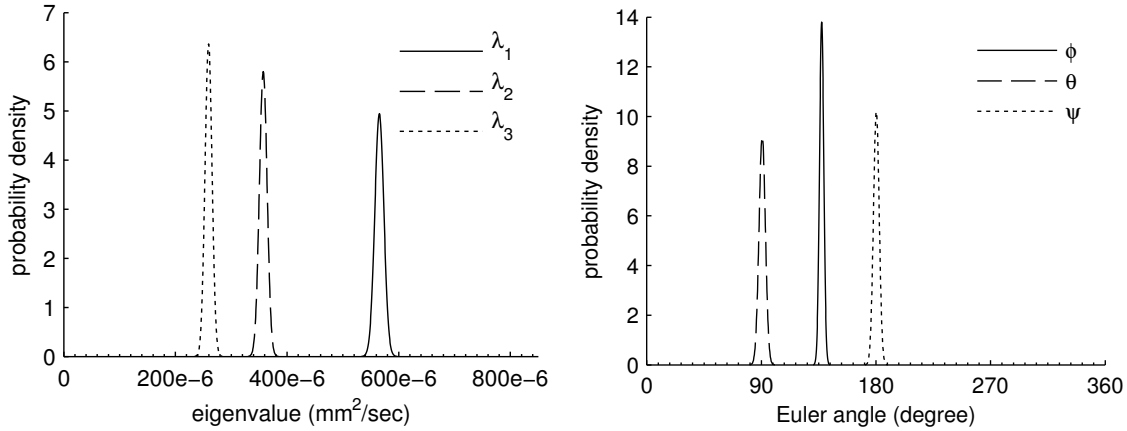
Human Brain Data

Result of fitting different diffusion models to the human brain data is shown in Fig. 4.13 and the log-likelihood values are shown in Fig. 4.14. Original noisy ADC profile as well as ADC profiles recovered from different diffusion models are plotted together with principal directions of tensor compartments. For the voxels except the third row, which is the crossing region, ADC profiles recovered by different diffusion models are similar, but the principal directions recovered from diffusion tensor model are better aligned with the tract orientations. For the crossing region, the directions recovered from biGaussian models exhibit expected crossing pattern.

4.4 Discussion

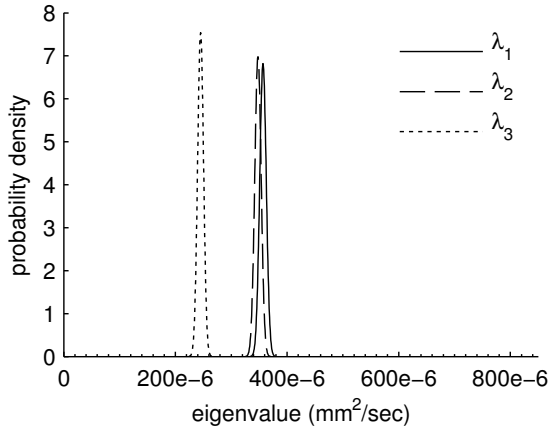
The major advantage of our method over the MAP approach is that we do not need the prior distribution and provide information based on the diffusion profile only. Feedback from the fiber tracking algorithms may improve the result, but it is more likely to introduce artifacts. Another benefit is that we are able to use all the

www.manaraa.com

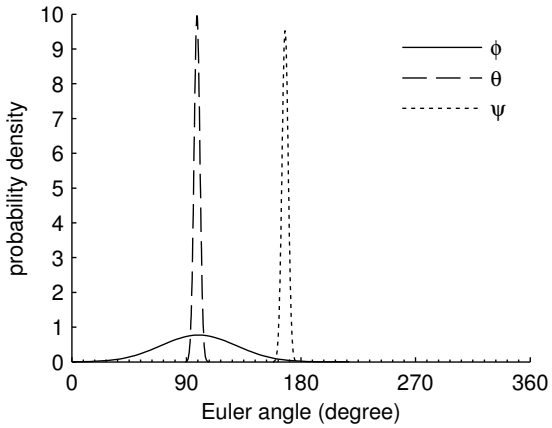


(a) single tract, eigenvalues

(b) single tract, Euler angles



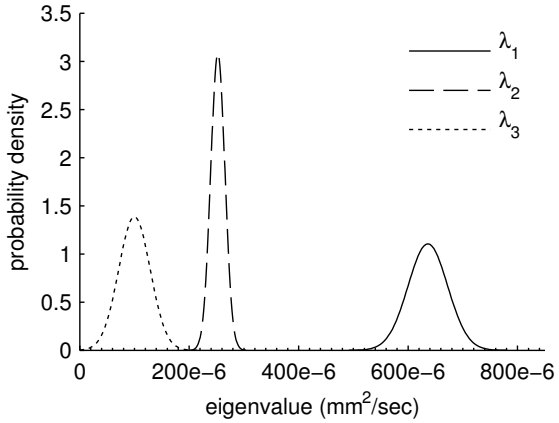
(c) two tract crossing, eigenvalues



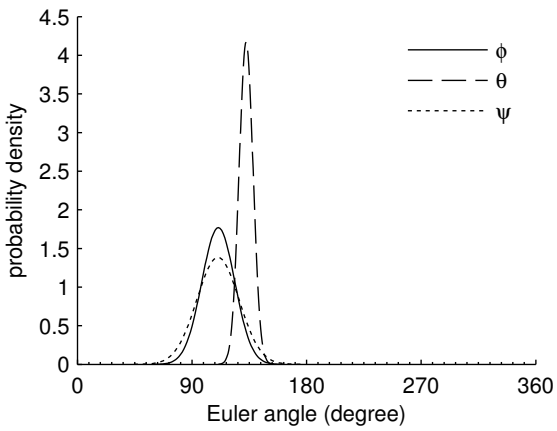
(d) two tract crossing, Euler angles

Figure 4.11: The variance shown as PDF of the marginal distributions of parameters estimated from the diffusion tensor model. The parameters shown in (a) and (b) are estimated from a voxel within a single spinal cord of the rat phantom, and the parameters shown in (c) and (d) are estimated from a voxel within the crossing region of the rat phantom. The marginal distributions are obtained from the covariance matrix estimated from the asymptotic covariance of the ML estimator of the diffusion tensor model.

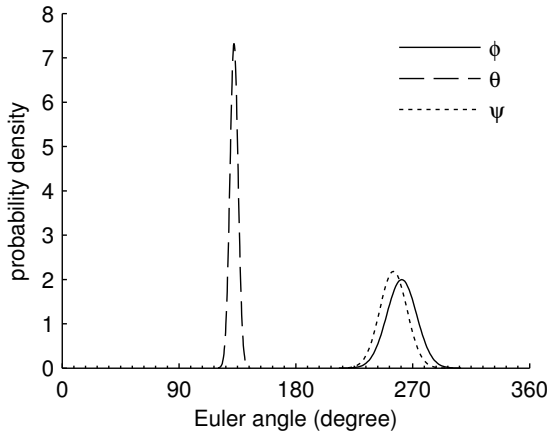
www.manaraa.com



(a) eigenvalues



(b) Euler angles of the first tensor



(c) Euler angles of the second tensor

Figure 4.12: The variance shown as PDF of the marginal distributions of parameters estimated from a voxel of the rat phantom containing two tract crossing, using the biGaussian model with 10 parameters. The marginal distributions are obtained from the covariance matrix estimated from the asymptotic covariance of the ML estimator of the diffusion tensor model.

www.manaraa.com

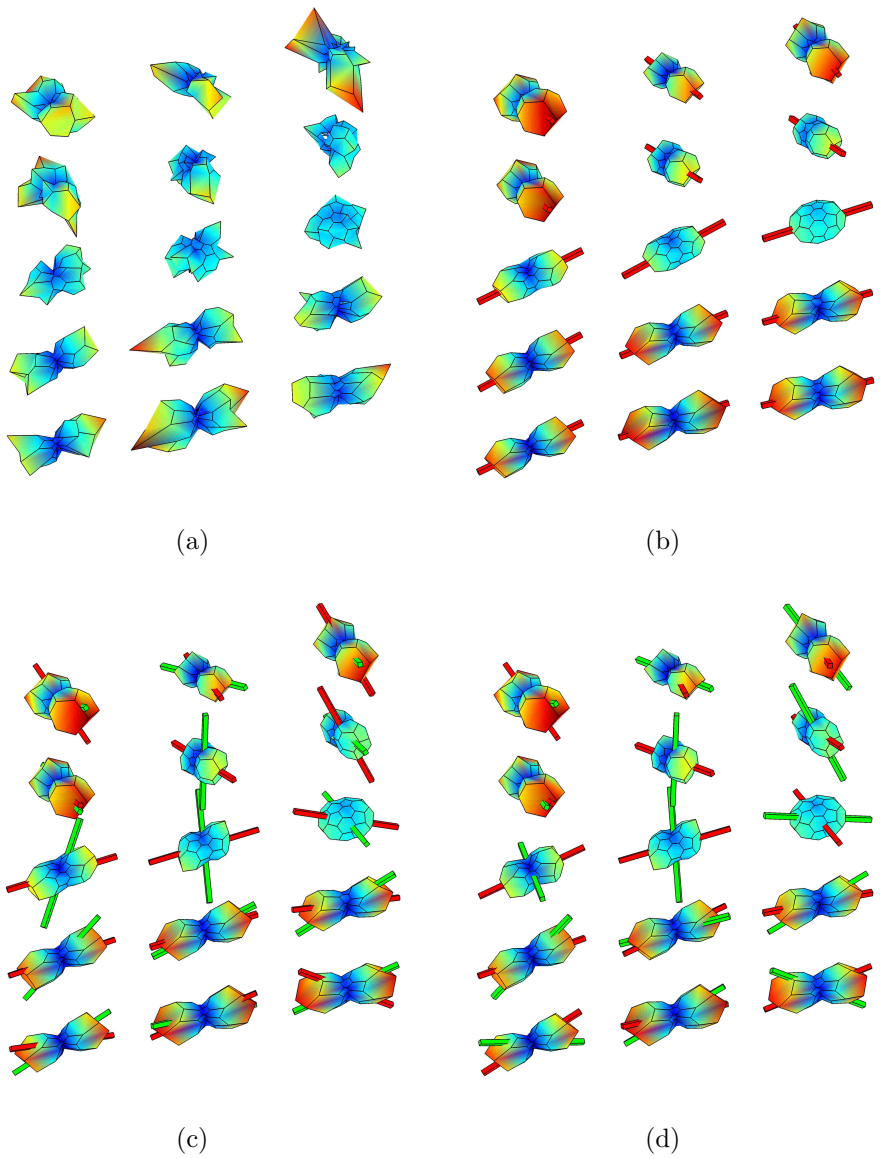


Figure 4.13: Fitting different diffusion models to the crossing region of the human brain data. The crossing region is shown in Fig. 4.4. (a) Noisy ADC profiles of the crossing region; (b) ADC profiles recovered from the diffusion tensor model, together with the principal directions of the diffusion tensors; (c) ADC profiles recovered from biGaussian model with 10 parameters, together with the principal directions of each tensor compartments; (d) ADC profiles recovered from biGaussian model with 13 parameters, together with the principal directions of each tensor compartments. The ADC profiles of some voxels are scaled to show the details with good contrast.

www.manaraa.com

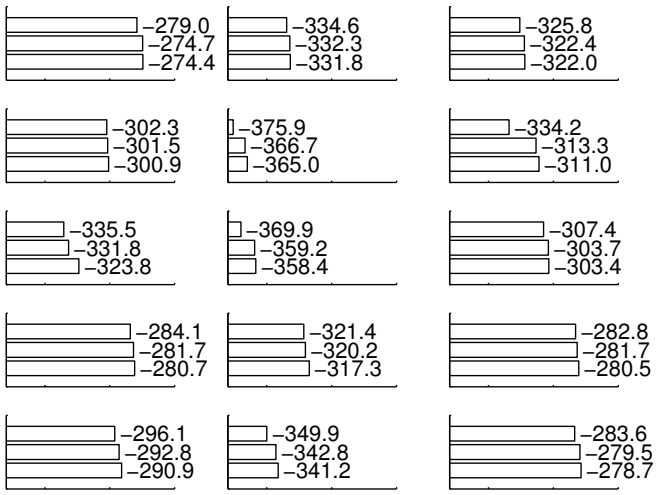


Figure 4.14: Log-likelihood values of fitting the diffusion tensor model, the biGaussian model with 10 parameters and the biGaussian model with 13 parameters to the crossing region of the human brain data shown in Fig. 4.4. Log-likelihood values of the three diffusion models are plotted for each voxel within the ROI as horizontal bars from top to bottom.

measured DWI data, instead of rejecting the measurement as outliers when they are far from the predicted values.

On the other hand, the method we proposed suffers from the same problems as those search-based descriptive methods (Cao et al., 2007). The quality of solution depends on both the starting points and the optimization algorithm. And, calculating the likelihood functions is computationally intensive, preventing us from using more starting points when computation resource is limited. Another problem that is also found in the MAP method is that the allowable parameter set $\mathbf{P}$ is chosen arbitrarily. However, we can overcome those issues by carrying out more computation, because the optimization of different voxels can be done independently which is suitable for more advanced high performance computer architectures.

Also, we need to find a strategy to deal with the scenario that different candidate solutions have similar likelihood values.

Copyright© Ning Cao, 2013.

www.manaraa.com

Chapter 5 Model Selection

In Ch. 4, we focus on fitting diffusion models to data. Models with more parameters generally fit the data better. However, the models with more parameters sometimes fit the data only slightly better. The additional parameters that do not present anything useful are not desired. Overfitting caused by random error or noise sometimes leads to artifacts in fiber tracking thus should be avoided by assessing the quality of the fitting. Model selection is the process of choosing the best model from a set of candidate models for the given data. A good model selection algorithm balances the goodness of fit with the number of parameters. In this chapter, we discuss model selection algorithms based on likelihood values obtained from MLE described in Ch. 4.

5.1 Introduction

Most people believe that the diffusion tensor model only derives a single fiber orientation. When more complex models that support two or more fiber orientations are also supported by the diffusion profile, model selection is always involved to justify the choice of a particular model. The underlying idea is to check if the diffusion tensor model is capable of describing the observed data. It is straightforward to define some measurements of goodness of fit, and then use empirical thresholds to decide which model to choose. For example, Pearson's correlation of 0.95 are used as a threshold to decide if the diffusion tensor model agrees with the observed diffusion profile well (Tuch et al., 2002). The drawback of this approach is that only the data is considered. To be more specific, the goodness of fit of the models other than the diffusion tensor model is not checked. When a model does not fit the data well, we should evaluate the goodness of fit of other candidate models together with the increased complexity, to avoid the risk of selecting a more complex model, which is not

www.manaraa.com

significantly better or even worse.

On the other hand, fiber orientation density function (Tournier et al., 2004) is a model-free approach that provides a continuous distribution of fiber orientation.

5.2 Method

Candidate Models

Choosing the set of candidate models is a critical task in determining the best model (Burnham and Anderson, 2002). It always involves developing a global model and deriving some special cases of the global model by reducing the number of parameters. We use the Gaussian mixture model Eq. 1.22 as the global model. The isotropic model Eq. 1.2, the diffusion tensor model Eq. 1.6, and the biGaussian model Eq. 1.19 are special cases derived from the Gaussian mixture model. The isotropic model is the simplest diffusion-weighted model; the diffusion tensor model is widely accepted and has been validated in many studies; the biGaussian model is also used in many studies and some promising results have been presented.

There are straightforward methods to fit the isotropic model and the diffusion tensor model. But, fitting of more than two Gaussian compartments was often not numerically stable (Tuch et al., 2002; Kreher et al., 2005). Therefore, our set of candidate models includes the isotropic model, the diffusion tensor model, and the biGaussian model.

Hypothesis Testing

One advantage of the MLE approach is that the value of the log-likelihood function Eq. 4.8, along with the estimated variance, can be used by statistical tests of model selection.

Likelihood ratio test, Wald test and Lagrange multiplier test are statistical tests that compare the fit of two models, one of which is nested within the other. The

www.manaraa.com

nested model is often called the null model or the restricted model, while the other is often called the alternative model or the unrestricted model. Likelihood ratio test requires the estimate of both the restricted and the unrestricted models, while Wald test and Lagrange multiplier test need the estimate of one of the models as well as the covariance matrix. The test statistic of likelihood ratio test is given by

$$\mathsf{D} = -2\log(\frac{\mathcal{L}_0}{\mathcal{L}_A}), \tag{5.1}$$

where $\mathcal{L}_0$ is the likelihood of the null model, and $\mathcal{L}_A$ is the likelihood of the alternative model. Let k_0 denote the number of free parameters of the null model and k_A be that of the alternative model. The distribution of the test statistic is approximately a χ^2 distribution with degrees of freedom (DOF) being $k_A - k_0$.

The isotropic model is a special case of the diffusion tensor model, and the diffusion tensor model is a special case of the biGaussian model. For biGaussian models with different numbers of parameters, the 10-parameter model is a special case of the 13-parameter one. In these cases, the model with more parameters generally fits the data better, unless there are more constraints on the more complex model.

Multiple Comparisons

Multiple comparisons arise when we carry out many statistical tests simultaneously. Diffusion-weighted images always have a large number of voxels, because the voxel size is relatively small compared to the physical size of the subject. In order to select a local diffusion model for each voxel, we use the same test repeatedly for each voxel. Correction of the significant value should be done to avoid increasing type I error, which is the incorrect rejection of a true null model.

One of the most commonly used approaches is Bonferroni correction. In Bonferroni correction, the experiment-wide significance level $\bar{\alpha}$ is given by α/n, where n is the number of independent comparisons to be performed, and α is the per comparison significance level. It is simple to calculate but it is very conservative.

www.manaraa.com

Model Selection Criteria

Hypothesis tests are only defined for the situation that the null model is nested within the alternative model. However, we often need to compare non-nested models or a set of candidate models.

Akaike information criterion (AIC), derived from information entropy, is a measure of relative goodness of fit of a model (Akaike, 1974). AIC is defined as

$$AIC = 2k - 2\log(\mathcal{L}), \tag{5.2}$$

where k is the number of parameters of the model, and $\mathcal{L}$ is the maximum value of likelihood function found by MLE. When we have AIC values for a set of candidate models, and let AIC_{min} denote the AIC of the minimum of them, the quantity

$$\exp(\frac{AIC_{min} - AIC_i}{2}) \tag{5.3}$$

is the relative likelihood of model i.

While the statistical tests are only valid for nested models, AIC can be used in other scenarios as well.

Similar measures can also be used. Akaike information criterion with correction (AICc), which is AIC with a correction for finite sample sizes, has a greater penalty for introducing extra parameters (Sugiura, 1978). It is defined as

$$AIC_c = AIC + \frac{2k(k+1)}{n-k-1}, \tag{5.4}$$

where n is the sample size. As we generally assume the estimator is normally distributed, we can also use Bayesian information criterion (BIC) (Schwarz, 1978), which is defined as

$$BIC = -2\log\mathcal{L} + k\log(n). \tag{5.5}$$

Note that it is possible and common to get different results by using different model selection criteria. In our phantom and human brain data experiments, the

www.manaraa.com

model selection results of three criteria are mostly consistent. However, for some low anisotropy voxels, BIC prefers simpler models, which cohere with the author's intuitive choice.

Simulation

We carry out simulation for both statistical test of nested models and comparison of multiple models.

We create two sets of synthetic data. One dataset is generated from diffusion tensors with low anisotropy. The data is fit to the isotropic diffusion model and the diffusion tensor model. The other dataset, which is generated from the mixture of two diffusion tensors with different angle between them, is fit to the diffusion tensor model and the biGaussian model with different numbers of parameters. We do likelihood ratio test (1) between the isotropic diffusion model and the diffusion tensor model, (2) between the diffusion tensor model and the biGaussian model, and (3) between the biGaussian model with different numbers of parameters. We also calculate AIC and AICc to compare multiple models of the second dataset. The first dataset is created from diffusion tensor with eigenvalues in the format of $(\lambda_a, \lambda_b, \lambda_b)$, where $\lambda_a = 1200 \times 10^{-6}\,\mathrm{mm}^2$ /s and λ_b changes from $840 \times 10^{-6}\,\mathrm{mm}^2$ /s to $1200 \times 10^{-6}\,\mathrm{mm}^2$ /s, to include diffusion tensors with different degree of anisotropy. The second dataset is created by mixing two diffusion tensors with the same shape. The eigenvalues of the tensor compartments are also in the format of $(\lambda_a, \lambda_b, \lambda_b)$, with value $(1200, 300, 300) \times 10^{-6}\,\mathrm{mm}^2$ /s. The angle τ between the two tensors compartments varies from 30° to 90°. The test data has 60 diffusion-weighed images as shown in Fig. 2.5(a) and three unweighed images. Rician noise of $\sigma = 6$ is added to both diffusion-weighted and unweighted images to simulate the real noise level we have in the phantom data.

www.manaraa.com

Phantom Experiments

We carry out likelihood ratio test on the rat phantom data described in Ch. 2.

Likelihood ratio test between the diffusion tensor model and the isotropic diffusion model are used to segment voxels into agar and spinal cords. We fit diffusion-weighted data to both diffusion models and record the log-likelihood values at each voxel. For each voxel, the test statistic is computed to derive the p-value. The likelihood ratio test is carried out on a ROI whose dimension is 35×32. The experiment-wide significance level $\bar{\alpha}$ is set to 5%, thus the per comparison significance level α with Bonferroni correction is given by $0.05/1120 \approx 4.5 \times 10^{-5}$.

Likelihood ratio test is performed between the diffusion tensor model and the bi-Gaussian model on the 5×5 region shown in Fig. 4.3. We use the same experiment-wide significance level $\bar{\alpha} = 5\%$, and the per comparison significance level with Bonferroni correction is $0.05/9 \approx 5.6 \times 10^{-3}$.

We also use model selection criteria on the same region to choose a model from the isotropic diffusion model, the diffusion tensor model and the biGaussian models.

5.3 Results

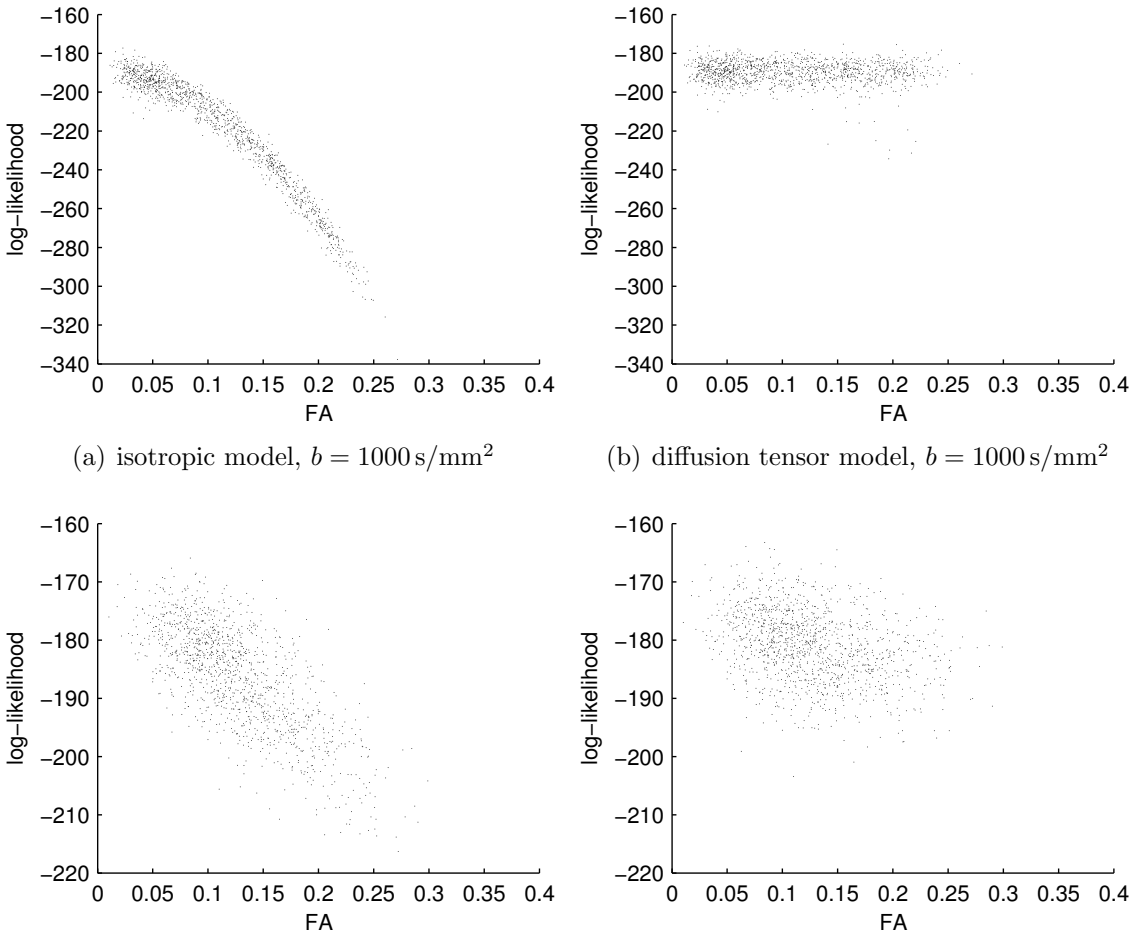
Simulation

Hypothesis Testing

Isotropic model and diffusion tensor model Log-likelihood values obtained by fitting the isotropic diffusion model, the restricted model, and the diffusion tensor model, the unrestricted model, are plotted in Fig. 5.1. For each configuration of λ_b/λ_a, we draw 100 trials. For each trail, FA value is calculated from the diffusion tensors. Although only a few discrete values of λ_b are used in the simulation, the FA value of all trials seems to be continuous because of the effect of added noise. A decrease in log-likelihood value indicates that the model is not likely to describe the data. In

www.manaraa.com

Figs. 5.1(a) and 5.1(c), log-likelihood value of the isotropic model decreases as the FA value increases, indicating that the isotropic diffusion model does not describe the diffusion profile well when anisotropy becomes stronger. However, Figs. 5.1(b) and 5.1(d) show that log-likelihood value almost remains the same when the FA value increases, because the diffusion tensor model is capable of describing anisotropic Gaussian diffusion.



(a) isotropic model, $b = 1000\,\mathrm{s/mm^2}$

(b) diffusion tensor model, $b = 1000\,\mathrm{s/mm^2}$

(c) isotropic model, $b = 3000\,\mathrm{s/mm^2}$

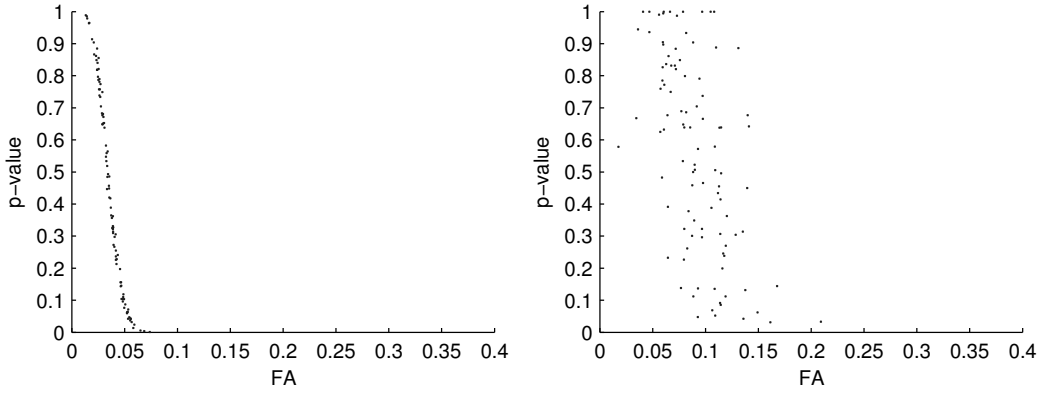
(d) diffusion tensor model, $b = 3000\,\mathrm{s/mm^2}$

Figure 5.1: Log-likelihood value obtained by fitting the isotropic model and the diffusion tensor model to synthetic data with different FA value.

The p-value of likelihood ratio test between the isotropic diffusion model and the diffusion tensor model for different b-values are shown in Fig. 5.2. We plot the p-value

www.manaraa.com

versus the FA value. The isotropic diffusion model is rejected at significance level α if the p-value is below α. From the plot we can see that the statistical test does not always equal choosing a single FA value as threshold. An FA value may be used as threshold to get the same effect as likelihood ratio test for $b = 1000\,\mathrm{s/mm^2}$ as shown in Fig. 5.2(a), however, it can be seen from Fig. 5.2(b) that it is not feasible to use a single FA value as threshold for $b = 3000\,\mathrm{s/mm^2}$.

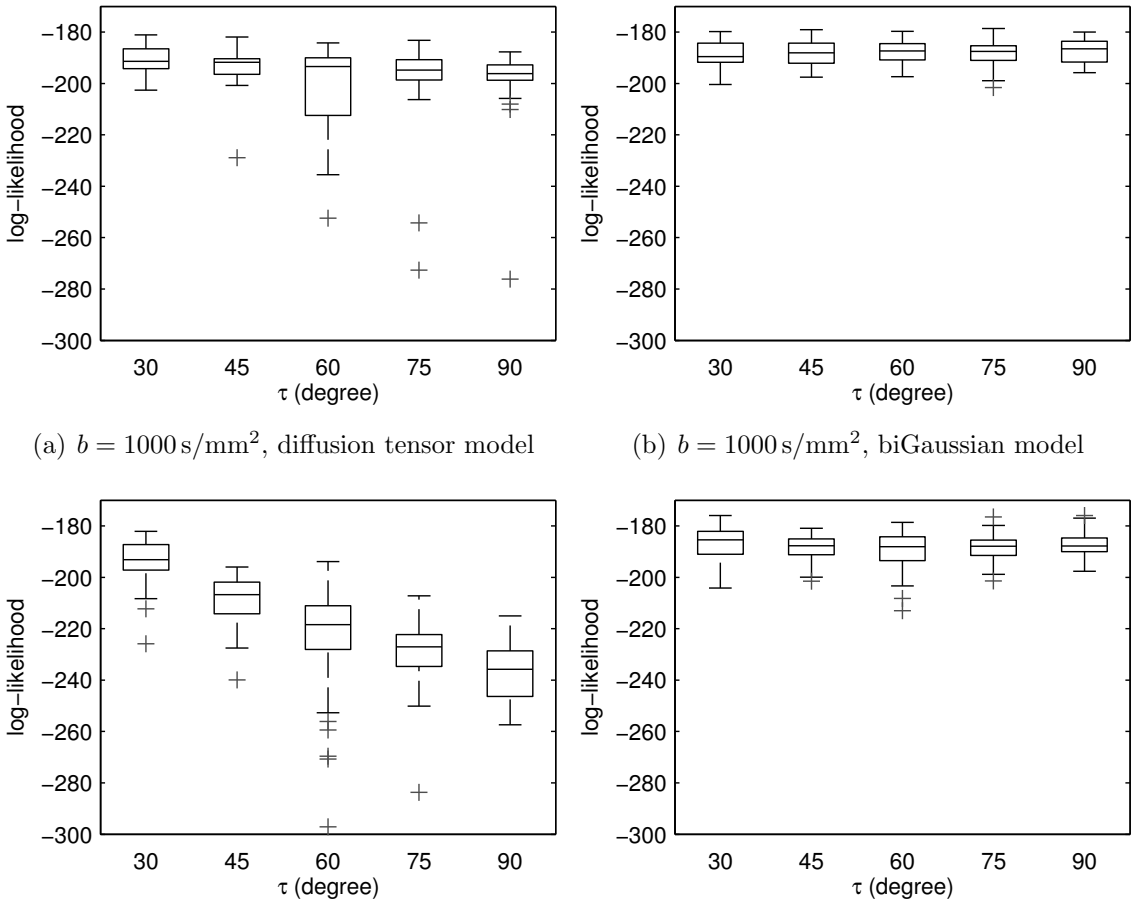


(a) $b = 1000\,\mathrm{s/mm^2}$ (b) $b = 3000\,\mathrm{s/mm^2}$

Figure 5.2: Simulation results of likelihood ratio test between the isotropic model and the diffusion tensor model. The p-value of likelihood ratio test between the two models is plotted versus FA value. The log-likelihood values used to calculate the test statistic are shown in Fig. 5.1.

Diffusion tensor model and biGaussian model Log-likelihood values from the diffusion tensor model and the biGaussian model with 10 parameters are plotted in Fig. 5.3. We draw 40 trials for each test configuration. In Fig. 5.3(a), the log-likelihood values do not change significantly when the angle τ increases. The log-likelihood value drops in Fig. 5.3(c) when the angle τ increases, indicating that the goodness of fit of the diffusion tensor model decreases when τ increases for high b-values. This explains why higher b-value is preferred over lower b-values in HARDI studies.

www.manaraa.com



(a) $b = 1000\,\mathrm{s/mm^2}$, diffusion tensor model

(b) $b = 1000\,\mathrm{s/mm^2}$, biGaussian model

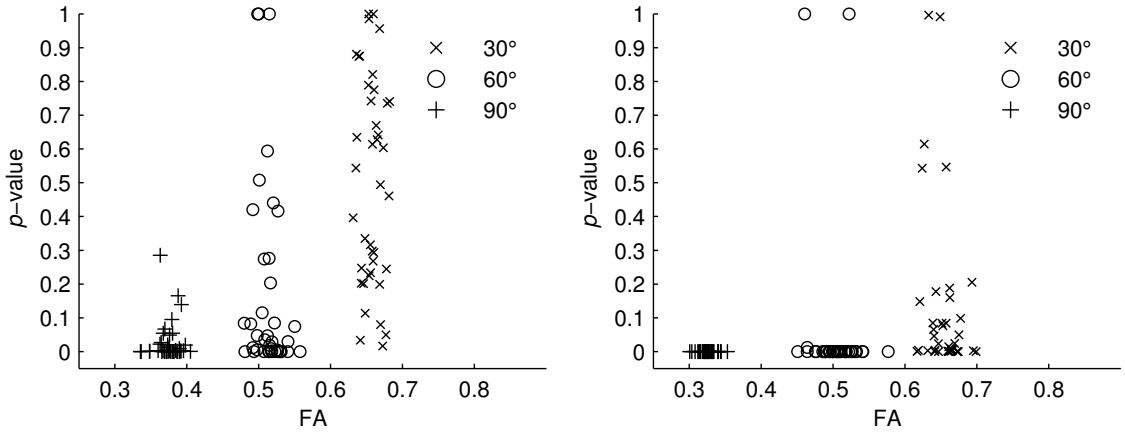
(c) $b = 3000\,\mathrm{s/mm^2}$, diffusion tensor model

(d) $b = 3000\,\mathrm{s/mm^2}$, biGaussian model

Figure 5.3: Log-likelihood value obtained by fitting the diffusion tensor model and the biGaussian model with 10 parameters to synthetic data with different angle τ between the two tensor compartments used to generate the biGaussian profiles. The log-likelihood value for each different τ are shown as box plots. On each box of the box plots, the central mark is the median of values, the edges of the box are the 25th and 75th percentiles, the whiskers show the most extreme data points not considered outliers, and outliers are plotted individually using "+".

www.manaraa.com

Simulations of likelihood ratio test between the diffusion tensor model and the biGaussian model are shown in Fig. 5.4. We follow the same method as above to calculate the p-value of each trial. The figure shows that the biGaussian model is preferred to the diffusion tensor model for most cases. However, diffusion tensor model is not always rejected when the angle between the two tensor compartments is small. Compared to the simulation result of $b = 1000\,\mathrm{s/mm^2}$, the diffusion tensor model is more often rejected at a small significant level than in the simulation of $b = 3000\,\mathrm{s/mm^2}$.

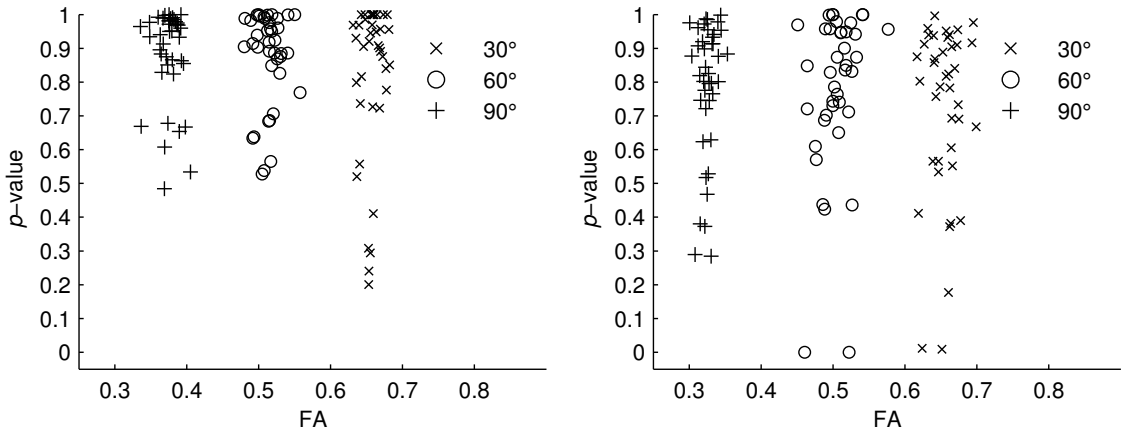


(a) $b = 1000\,\mathrm{s/mm^2}$ (b) $b = 3000\,\mathrm{s/mm^2}$

Figure 5.4: The p-values of likelihood ratio test between the biGaussian model with 10 parameters and the diffusion tensor model. The p-value of test statistic is plotted versus τ. Trials for $\tau = 30, 60,$ and 90° are plotted using different markers.

BiGaussian models Because the synthetic data is generated with two tensor compartments of the same shape, the log-likelihood values obtained from fitting the two biGaussian models do not differ much. Therefore, we only plot the p-value of the likelihood ratio test between the two biGaussian models in Fig. 5.5. The biGaussian model with fewer parameters is not likely to be rejected because it is simpler and is able to describe the data as good as the more complex model.

www.manaraa.com



(a) $b = 1000\,\mathrm{s/mm^2}$ (b) $b = 3000\,\mathrm{s/mm^2}$

Figure 5.5: Simulation result of likelihood ratio test between the biGaussian models with 13 parameters and 10 parameters. The p-value of the test statistic is plotted versus FA value. Trials for $\tau = 30, 60,$ and 90° are plotted using different markers.

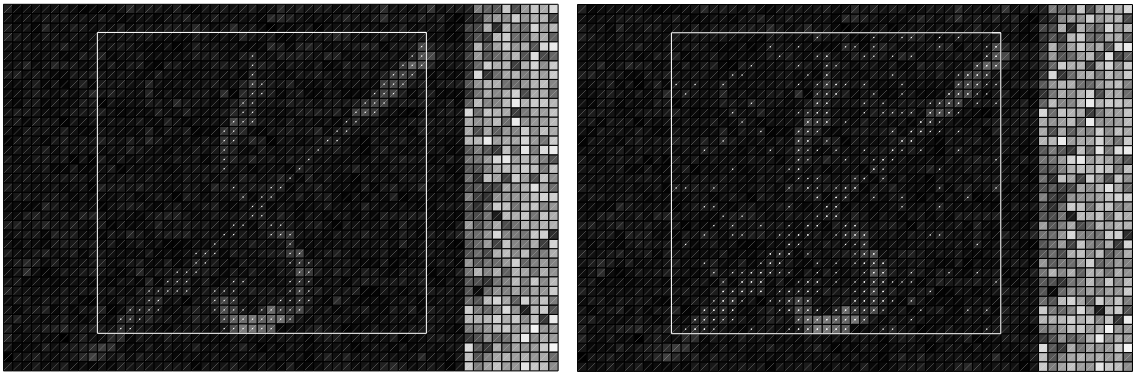
Phantom Experiments

Hypothesis Testing

Model selection between the diffusion tensor model and the isotropic diffusion model with per comparison significance level adjusted by Bonferroni correction is shown in Fig. 5.6(a). It shows that up to three voxels along the spinal cords are considered anisotropic by the statistical test, which complies with the fact that the diameter of rat spinal cords is 5 mm and the voxel is 2.5 mm isotropic. The result of using experiment-wide significance level without Bonferroni correction is shown in Fig. 5.6(b) for comparison. The isotropic diffusion model is rejected by likelihood ratio test on some voxels in the agar background if Bonferroni correction is not used.

Likelihood ratio test between the diffusion tensor model and the biGaussian model of the crossing region described in Fig. 4.3 is plotted in Fig. 5.7(a).

www.manaraa.com



(a) with Bonferroni correction (b) without Bonferroni correction

Figure 5.6: Likelihood ratio test between the diffusion tensor model and the isotropic diffusion model on a slice of the rat phantom data overlaid on FA map. Result of both likelihood ratio test with and without Bonferroni correction are shown. The white rectangles define the ROI where the test is carried out. Voxels where the isotropic diffusion model is rejected are marked with a white dot.

Model Selection Criteria

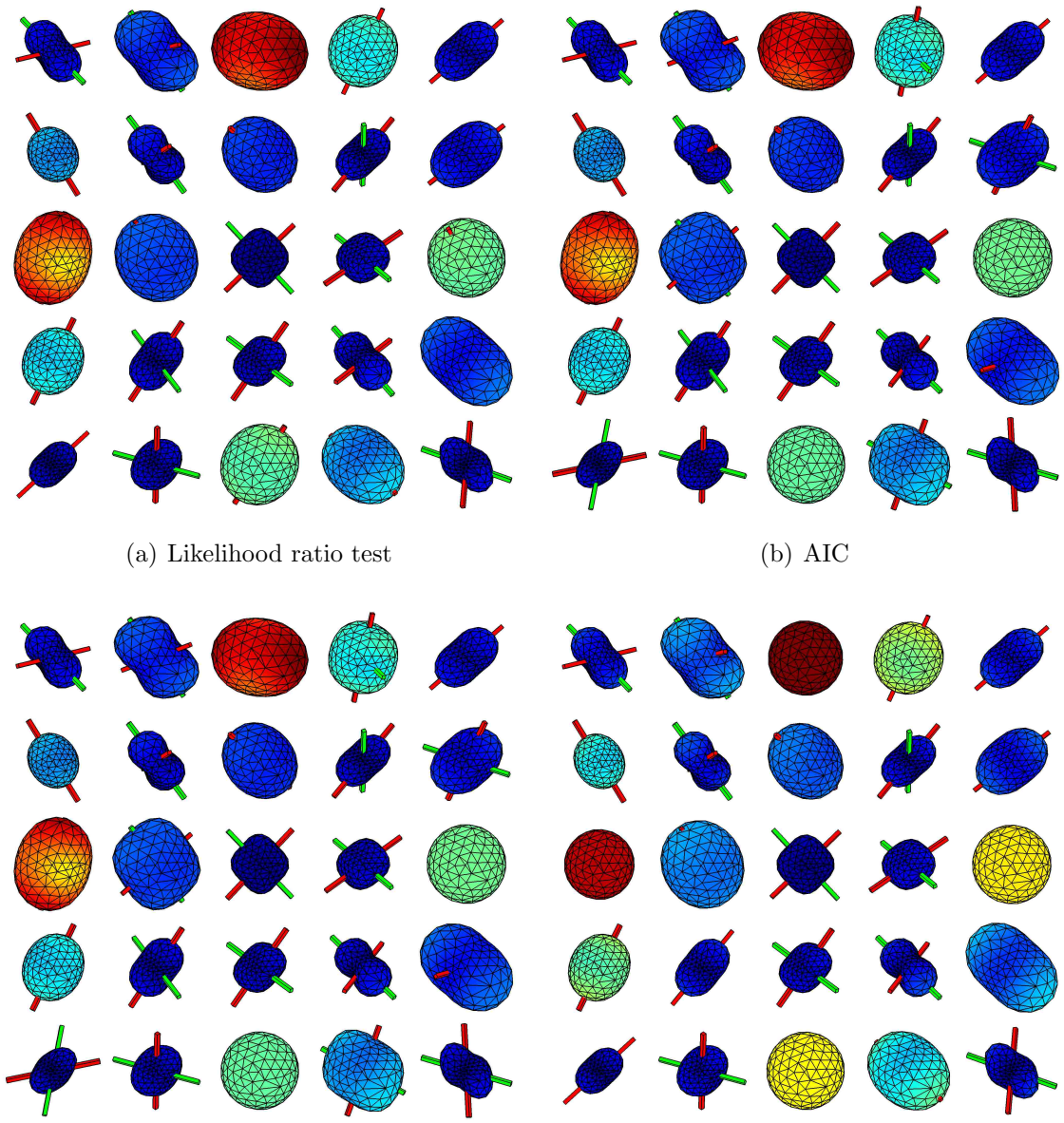
Model selection result on the crossing region described in Fig. 4.3 is plotted in Fig. 5.7. For each voxel, estimated ADC profile of the model picked by the model selection criteria is plotted to compare with the original noisy ADC profile. Fitting result of candidate models is already shown in Fig. 4.9. AIC, AICc and BIC give similar result in this region.

5.4 Discussion

Hypothesis tests and model selection criteria are developed with different consideration. Hypothesis tests assume the alternative model is true and test if the additional parameters have any significant effect, while model selection criteria check the relative quality of each model in a set of candidate models.

Model selection also depends on the sample size. While a simpler model may underfit the data, a more complex model tends to overfit. Underfitting induces bias and overfitting produces variability. With a larger sample size, the bias is the major

www.manaraa.com



(a) Likelihood ratio test

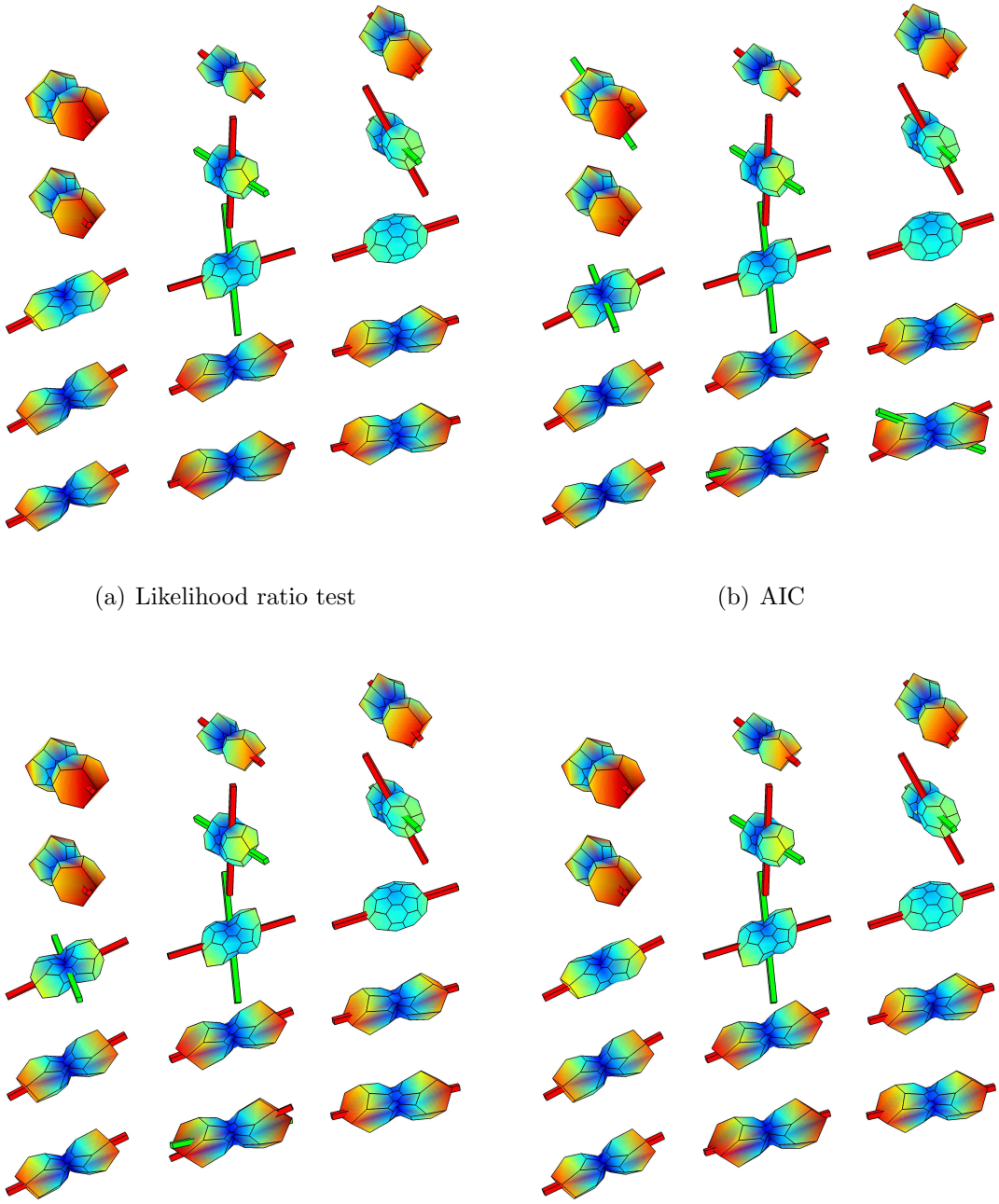
(b) AIC

(c) AICc

(d) BIC

Figure 5.7: Model selection results of the rat phantom on the crossing region defined in Fig. 4.3. (a) shows the result of likelihood ratio test between the diffusion tensor model and the biGaussian model with 10 parameters; (b), (c), and (d) show model selection results on different criteria.

www.manaraa.com



(a) Likelihood ratio test

(b) AIC

(c) AICc

(d) BIC

Figure 5.8: Model selection results of the human brain data on the crossing region defined in Fig. 4.4. (a) shows the result of likelihood ratio test between the diffusion tensor model and the biGaussian model with 10 parameters; (b), (c), and (d) show model selection results on different criteria.

www.manaraa.com

concern, and for a smaller sample size, attention should be paid to avoiding the variability caused by overfitting.

In terms of selecting a local diffusion model for the given data, the choice depends on the true model, which is the underlying structure within the voxel, the sample size, and the noise level.

It is hard to validate model selection result using real data. Although people prefer a biGaussian model for the voxels with fiber crossing or branching, the data may not support the complexity. As mentioned in Ch. 3, the diffusion tensor model may fit a voxel containing crossing tracts well enough. For the rightmost voxel of the third row in Figs. 4.13 and 5.8, although the biGaussian model is able to recover two tract orientations, as shown in Figs. 4.13(c) and 4.13(d), the preferred model selected by likelihood ratio test and the three model selection criteria is the diffusion tensor model, as shown in Fig. 5.8, because it is relatively simple and fits the ADC profile well, which can be seen from Fig. 4.14. Note that the human brain dataset has a relatively low b-value, which is $800\,\mathrm{s/mm^2}$, and relatively low angular resolution.

Although the model selection methods mentioned in this chapter always return a preferred model from two models a set of candidate models, please note that the preferred model may be only a little better than another model. Other criteria should be used when the values of the relative likelihood Eq. 5.3 of two models are similar.

Copyright© Ning Cao, 2013.

www.manaraa.com

Chapter 6 Tractography

6.1 Introduction

A neural tract, or neural pathway, connects relatively distant parts of the brain or nervous system. It usually consists of bundles of neurons, which are covered by the myelin sheath. Tractography, or fiber tracking, is the modeling process to reconstruct neural tracts using diffusion-weighted data. Tractography algorithms can be roughly classified into deterministic tractography and probabilistic tractography. Most tractography methods are based on the streamline method.

In this chapter, we describe how we improve tractography by taking local orientation uncertainty into consideration.

Streamline Method

Streamlines are widely used in the study of fluid flow, which is characterized by a three-dimensional velocity vector field. Streamlines are curves formed by tangent lines to the velocity vector of the flow, and are used to visualize vector fields. Local directions in three-dimensional grid recovered from diffusion-weighted data is similar to steady flow, in which the velocity field does not change over time. The streamline fiber tracking method can be described as finding the solution of the following ordinary differential equation with a given initial value (Basser et al., 2002)

$$\begin{cases} \vec{y}'(t) = f(\vec{y}(t)) \\ \vec{y}(t_0) = \vec{y}_0, \end{cases} \tag{6.1}$$

where $f(\vec{y})$ is the local direction at position $\vec{y}$, $\vec{y}(t)$ is the path of the tract, which is parameterized by the arc length t, and $\vec{y}_0$ is the seed point. The initial value problem can be solved by the Euler method, which is a first-order numerical method to solve

www.manaraa.com

an initial value problem. The global error of the Euler method is proportional to the step size. The Euler method is the simplest Runge-Kutta method, while Runge-Kutta methods of higher order give better result with the same step size.

The streamline algorithm and its variants often have several aspects. Seed point selection is often performed by an operator on the co-registered T_1-weighted image or the FA map. Both the centers of each voxel (Mori et al., 1999) within the ROI and multiple points within each voxel (Basser et al., 2002) are used. Local tract orientation of each voxel can be obtained from a diffusion model or without a particular diffusion model. Propagation is the process of solving the initial value problem or simply the method of deciding the next point of the tract. Termination criteria, or stopping criteria, are the standards by which the propagation should stop. Many implementations stop the propagation when the anisotropy is lower than an anisotropy threshold or the angle between directions of neighboring segments of the tract is larger than an angle threshold.

Regarding the quality of tracts, many factors contribute to the accumulated error of the tracts (Tournier et al., 2002): (1) the spatial resolution is relatively low compared to tract size; (2) the diffusion-weighted images are noisy thus uncertainty of local direction can be large; (3) total accumulated error of solving initial value problem can be large for long tracts. The total accumulated error can be improved by using higher order methods, but the spatial resolution and SNR cannot be improved easily.

Although other tractography methods have been proposed in recent years (Parker et al., 2002, 2003; Tournier et al., 2003; Behrens et al., 2003; Tuch et al., 2003; Campbell et al., 2005; Parker and Alexander, 2005; Kang et al., 2005a,b; Zhang et al., 2005b,a; Behrens et al., 2007; Descoteaux et al., 2009), the streamline method is still popular and is available through major MR hardware manufacturers (Mukherjee et al., 2008) and third parties.

www.manaraa.com

Deterministic Tractography

Most deterministic tractography algorithms are derived from the streamline method directly. They differ from the streamline method on one or more aspects discussed above.

Because the distance between two neighboring points on a tract is often smaller than the size of a voxel, which is often around 2 mm, interpolation or approximation are used to get local direction or anisotropy value at sub-voxel accuracy when a diffusion model is fit to data. Fiber assignment by continuous tracking algorithm (Mori et al., 1999) uses nearest-neighbor interpolation, while tri-linear interpolation (Conturo et al., 1999), B-spline interpolation (Basser et al., 2000; Pajevic et al., 2002), and B-spline approximation (Basser et al., 2000; Pajevic et al., 2002) are also used. Interpolation and approximation can be done on diffusion-weighted signals (Conturo et al., 1999) or computed diffusion tensor components (Basser et al., 2000; Pajevic et al., 2002).

Besides recovering local direction based on a diffusion model, local direction can also be recovered using q-space methods (Tuch et al., 2003; Behrens et al., 2007; Descoteaux et al., 2009).

For solving the initial value problem Eq. 6.1, both the Euler method (Conturo et al., 1999) and the higher-order Runge-Kutta methods (Basser et al., 2000) are used. Tensor deflection is also proposed to allow fiber tracts to pass through crossing regions (Lazar et al., 2003).

Deterministic methods are relatively easy to understand and implement, but they share the same limitations as the streamline method.

Limitations

There is no indicator on reliability of the tracking result on a given input. When it comes to local directions, deterministic method only takes one direction, which does

www.manaraa.com

not consider the data quality. As a result of low SNR, one may find that the tracts from two nearby seed points differ a lot. Another phenomenon is that sometimes many individual tracts are short, although they follow the major trend of a fiber bundle. This can commonly be seen in thin fiber tracts.

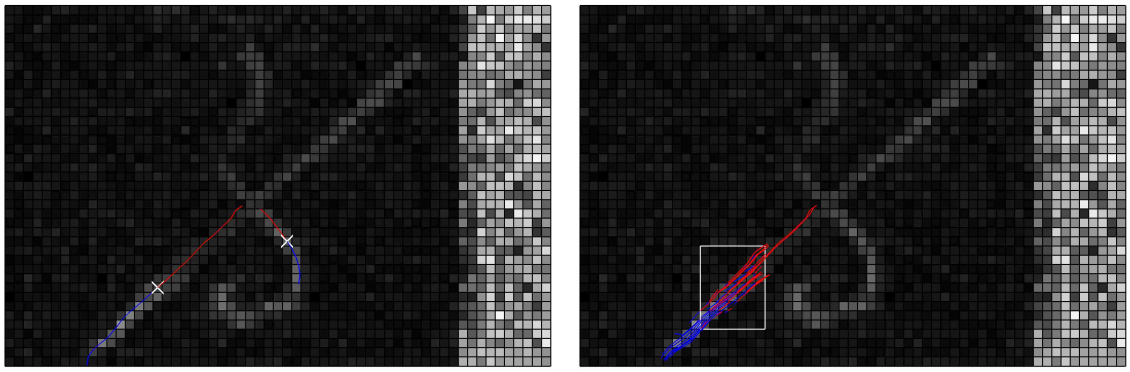
Sometimes it is difficult to choose seed points. Because of the low SNR nature of diffusion-weighted data (Taylor et al., 2004; Parker, 2004) and the accumulated error of propagation, most tracts recovered by deterministic methods are relatively short. As mentioned above, moving a seed point by a short distance may create a very different tract. Sometimes an operator has to spend long time to move the seed points to get a satisfying tract. Full-brain tracking is proposed to solve this problem (Conturo et al., 1999; Stieltjes et al., 2001). Full-brain tracking method uses seed points all over the whole data volume, often with several seed points within each voxel. Full-brain tracking makes recovering satisfying tracts easier, but we still need select good fiber tracts from many candidates. Using ROIs at both ends of the tracts works when the tracts are long enough; otherwise fiber clustering algorithms (Liang et al., 2009) may help.

Fiber tracts of the rat phantom obtained from the streamline algorithm and the whole-brain tracking method are shown in Fig. 6.1. Configuration of the phantom is shown in Fig. 2.6. Fourth-order Runge-Kutta method with step size being 1.0 is used to generate both figures. Although the configuration is simple and the seed points are located at the center of each tract, the recovered tracts are much shorter than the length of the neural cords. For the whole-brain tracking, most of the fiber tracts are also short.

Probabilistic Tractography

Deterministic tractography algorithms reconstruct a single tract for each seed point, while probabilistic tractography algorithms provide multiple tracts for the same seed

www.manaraa.com



(a) streamline method (b) whole-brain tracking

Figure 6.1: Fiber tracts reconstructed from the streamline method and the whole-brain tracking method. The two seed points used for the streamline method are marked as "×" and the ROI containing all seed points used for whole-brain tracking is shown as a rectangle. The tracts are overlaid with FA map and the two directions of a tract from the seed point are plotted in different colors.

point or a probability distribution of connectivity.

Most probabilistic tractography methods are derived from deterministic methods. The difference between a probabilistic tractography algorithm and its deterministic counterpart is that the local direction is chosen randomly at each voxel. The local direction can be drawn from diffusion tensor randomly based on FA value (Parker et al., 2003), using bootstrap methods (Jones and Pierpaoli, 2005), or from Bayesian inference (Behrens et al., 2003). Probabilistic tractography can also be achieved by using different propagation methods (Parker et al., 2002; Tournier et al., 2003; Kang et al., 2005a). In a similar way to deterministic tractography, q-space methods (Campbell et al., 2005; Parker and Alexander, 2005; Descoteaux et al., 2009) can also be used to decide a probability distribution of the local direction.

6.2 Method

We propose a probabilistic tractography method based on the MLE described in Ch. 4, which extends the streamline method.

www.manaraa.com

Local tract orientation as a three-dimensional vector is often not in the parameters being estimated thus it has to be computed from the parameters. In diffusion tensor model Eq. 1.6, local direction is obtained by diagonalization of the diffusion tensor $\mathbf{D}$ by Eq. 1.8. In the MLE approach described in Ch. 4, local directions can be computed from the Euler angles by Eq. 1.11.

From the software engineering point of view, we factor the streamline algorithm into a driver program and a callback function. The driver program solves the initial value problem using the Euler method or the Runge-Kutta method, while the callback function takes the location and a reference direction as input and returns a local direction. Global stopping criteria are checked in the driver program, and local stopping criteria, like the FA threshold and the maximum angle difference, can be handled in either the callback function or the driver program. When we modify the streamline method, the driver program does not change. For the simplest streamline method, the callback function returns a single direction for each voxel. When multiple local directions are supported, the callback function chooses one direction from all available directions using the reference direction as a hint. For probabilistic tracking, the callback function returns a random local direction each time it is called by the driver program.

Multiple Local Directions

Crossing and branching in fiber tracts can be recovered by considering multiple local directions. By using Gaussian mixture models Eq. 1.22 together with model selection techniques discussed in Ch. 5, it is possible to recover multiple local directions from a single voxel when the diffusion profile supports crossing or branching. When more than one direction is available, we choose the direction which has the minimum angle difference between itself and the reference direction so that the reconstructed tract does not turn shapely.

www.manaraa.com

Random Sampling of Local Direction

We draw random samples of local direction using covariance of the parameters instead of deriving a PDF of the local direction. The covariance matrix of the parameters can be obtained from MLE as described in Ch. 4. Random samples of local direction are then calculated from the random sample of Euler angles. Using covariance of parameters to characterize the uncertainty is more philosophic than methods based on simple factors (Parker et al., 2003). One approach to obtain the covariance is the bootstrapping method described in Ch. 4. However, our method is different from similar methods which use bootstrapping to draw samples of local direction (Jones and Pierpaoli, 2005) in that we draw random samples from the covariance matrix, instead of calculating from diffusion-weighted data directly. We estimate the covariance matrix by asymptotic normality, which requires less computation than the bootstrapping method does. Our approach follows frequentist inference, which is an alternative approach of Bayesian inference (Behrens et al., 2003).

We draw a random sample of local direction by drawing a sample of the parameter vector from multivariate normal distribution. We take the matrix L from the covariance matrix $\mathbf{\Sigma}$ of the ML estimator described in Ch. 4, such that

$$\mathbf{\Sigma} = \mathsf{L}\mathsf{L}^T, \tag{6.2}$$

and draw a sample of the parameter vector from multivariate normal distribution by

$$x = \hat{\mathbf{p}} + \mathsf{L} * z, \tag{6.3}$$

where $\hat{\mathbf{p}}$ is the estimator of the parameters, and the vector z is multivariate normally distributed. The lower triangular matrix L can be obtained from $\mathbf{\Sigma}$ by the Cholesky decomposition. Local direction is then calculated from the random samples x in the same way as from the estimated parameter vector $\hat{\mathbf{p}}$.

www.manaraa.com

The driver program is executed multiple times to generate multiple streamline samples from the same seed point.

The uncertainty of the local direction is decided by both the model and the data. The variance will be smaller if a model fits the data better than another model does. Once the model has been decided, the variance is affected by the diffusion-weighted data. The uncertainty is smaller for data with higher overall SNR, leading to less spreading tracts with higher precision.

6.3 Results

Multiple Local Directions

Multiple local orientations are recovered using the biGaussian model from the crossing region of the rat phantom. However, the quality of these multiple orientations is often not good enough to support the streamline method in our test.

Uncertainty

We illustrate the uncertainty of local orientations by drawing random samples from the multivariate normal distribution decided by the asymptotic covariance matrix of the ML estimators.

The uncertainty of local orientations of an anisotropic voxel within a single fiber tract is shown in Fig. 6.2. From the samples we can see that the distribution is narrow and using the direction obtained from the estimated parameter does not introduce much error to the fiber tract.

Uncertainty of the local direction in a voxel within the crossing area is illustrated in Fig. 6.3. The same diffusion profile is fit to both the diffusion tensor model and the biGaussian model. Uncertainty of the local direction recovered from diffusion tensor model is shown in Fig. 6.3(a). The uncertainty of the local direction is large, because the planar-shaped tensor has eigenvalues satisfying $\lambda_1 \approx \lambda_2 > \lambda_3$, which make the

www.manaraa.com

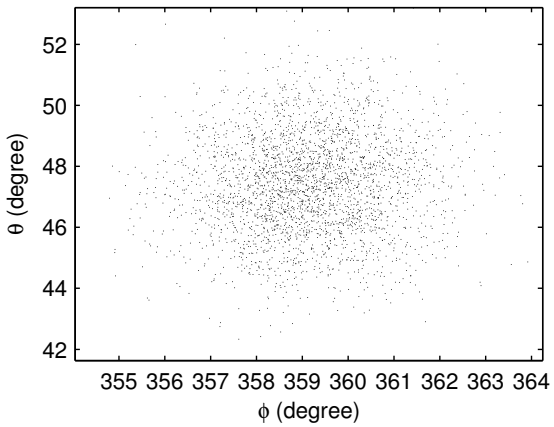
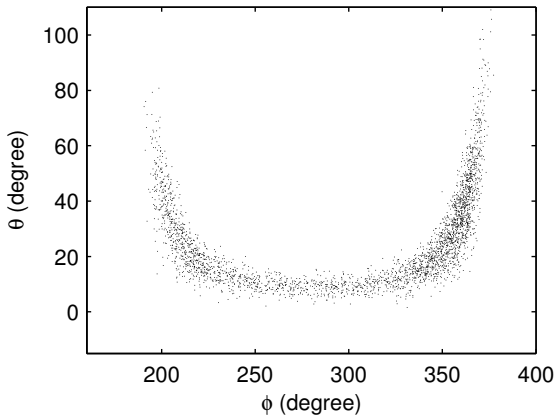


Figure 6.2: Local direction uncertainty of an anisotropic voxel within a single fiber tract of the rat phantom. Local directions are plotted in spherical coordinates.

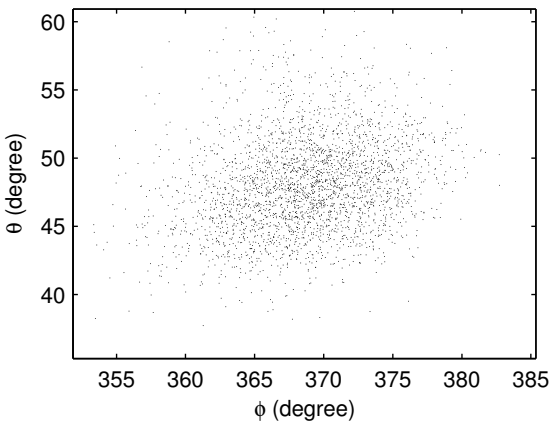
uncertainty of the principal eigenvector large. However, it does not indicate that the diffusion tensor model does not describe the diffusion profile of a crossing structure well. When the same data is fit to the biGaussian model, the two orientations have smaller uncertainty, as shown in Figs. 6.3(b) and 6.3(c). If the uncertainty is not considered in tractography, the directions from biGaussian model is acceptable because the distribution of recovered directions is approximately as narrow as that of the anisotropic voxel shown in Fig. 6.2. If we consider the uncertainty, local direction from the diffusion tensor model should not prevent a more sophisticated tracking algorithm from overcoming the crossing voxel because a large range of direction is returned. However, the local model does not provide much information in this case as the distribution covers a large potion of the unit sphere.

Generally, the uncertainty of local direction is small in the center of fiber tracts., while it is large within the area of fiber crossing or fiber curving. The uncertainty is also large on the boundary of the fiber tracts where anisotropy is lower than that of the center of the tracts because of the PVE.

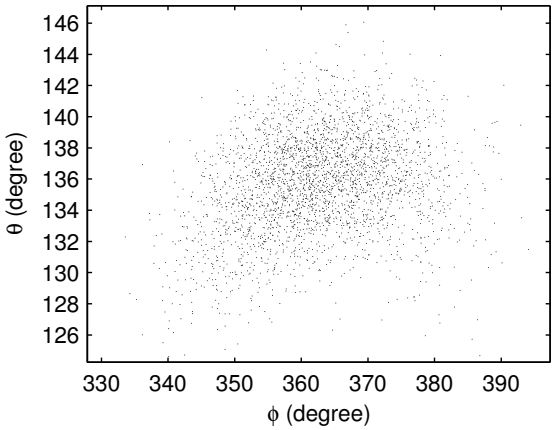
www.manaraa.com



(a) fit to diffusion tensor model



(b) the first orientation of biGaussian model



(c) the second orientation of biGaussian model

Figure 6.3: Local direction uncertainty of fitting a voxel containing fiber crossing structure to the diffusion tensor model and the biGaussian model. Local directions are plotted in spherical coordinates.

Tracking

With random sampling of streamlines, we are able to recover fiber crossing from the rat phantom data. Fiber tracking results of the rat phantom is shown in Fig. 6.4. The parameters of streamline tracking are the same as those used in Fig. 6.1. The two seed points are the same as the ones used in Fig. 6.1 and 200 random samples are drawn.

The reconstructed tracts cover most part of the rat spinal cords except the lower part of the curving one. Drawing random samples of local orientation is better than

www.manaraa.com

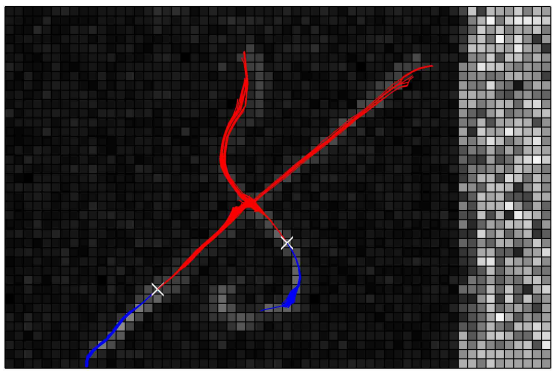


Figure 6.4: Fiber tracts of the rat phantom reconstructed by the streamline method with random sampling. The tracts are overlaid with FA map and the seed points are marked as "×". The two directions of a tract from the seed point are plotted in different colors.

just randomly perturbing the recovered local orientation. When the tract is within higher anisotropy regions, the uncertainty of the local orientations is relatively small; the uncertainty is larger near the crossing or the curving areas, because the covariance of the fitting result is larger within these areas.

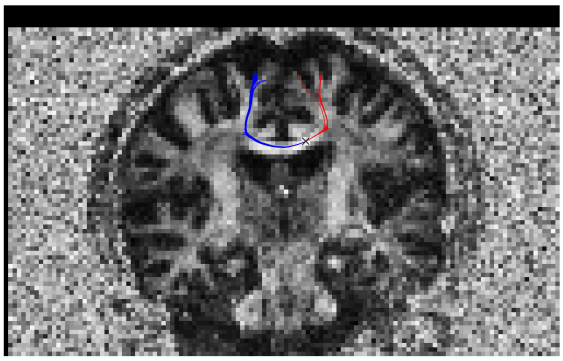
The cingulum tracts and a tract within the corpus callosum constructed form the human brain data are shown in Fig. 6.5. The three seed points are manually selected on the FA map. The cingulum tracts are difficult to reconstruct because they are relatively thin and surround the corpus callosum, which is strong in anisotropy. The cingulum tracts reconstructed by our method follow the expected paths.

6.4 Discussion

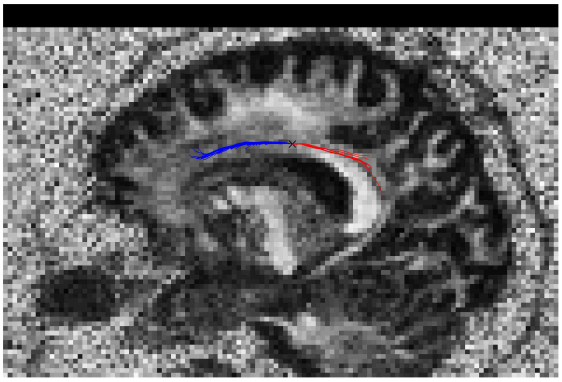
The tracking algorithm presented in this chapter is a direct application of the modeling approach described in Ch. 4.

Similar to other probabilistic tractography algorithms (Behrens et al., 2003; Parker et al., 2003), samples of local fiber orientations are used to create streamlines and these streamlines are regarded as samples of global connectivity. The global connectivity presented by these streamlines includes possible fiber tracts from the seed point,

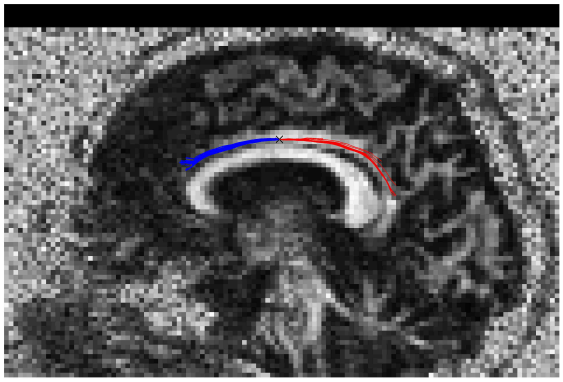
www.manaraa.com



(a) a tract within the corpus callosum



(b) the cingulum tract (left)



(c) the cingulum tract (right)

Figure 6.5: Fiber tracts of the human brain data reconstructed by the streamline method with random sampling. The tracts are overlaid with FA map and the seed points are marked as "×". The corpus callosum tract is plotted on a coronal slice and the cingulum tracts are plotted on sagittal slices. The two directions of a tract from the seed point are plotted in different colors.

considering the uncertainty caused by imaging noise in the given data. Imaging artifacts, inadequacy of the diffusion models caused by partial volume averaging of fiber orientations, and limited angular and spatial resolution also contribute to the uncertainty of local orientation. Therefore, care should be taken when interpreting the global connectivity.

The result shows that our tractography algorithm together with the biGaussian model is capable of reconstructing fiber tracts with crossing structure. Although the deterministic streamline method has difficulty passing through the crossing region in the phantom data, it is possible for our probabilistic algorithm to overcome the

www.manaraa.com

crossing structure even with the diffusion tensor model. It is because the uncertainty of the fiber orientations recovered by the diffusion tensor model is so large that it covers a large portion of all the directions. However, using an inadequate diffusion model to recover local tract orientations may result in more artifacts in the global connectivity. Therefore, another way of interpreting the local orientation uncertainty of a voxel is that the amount of uncertainty indicates how reliable the reconstructed tracts are when passing this voxel.

The Gaussian mixture model is not able to describe the continuity of the orientation of a curving tract within a voxel. It is also a limitation for the diffusion orientation distribution function presentation used in q-ball imaging (Tuch, 2004).

Copyright© Ning Cao, 2013.

www.manaraa.com

Chapter 7 Conclusion and Future Work

This dissertation presents my research work on using HARDI data to recover local tract orientations and reconstruct fiber pathways in human brain. It generally follows the frequentist approach. The work involves estimating noise parameter in diffusion-weighted images, model selection, and utilizing local tract orientations to reconstruct fiber pathways. In this chapter, I provide a summary of the dissertation work and future research possibilities.

7.1 Research Accomplishments

Reconstructing fiber pathways non-invasively is an important application of DWI. It generally includes recovering local tract orientations from each voxel and reconstructing fiber pathways from local orientations. The performance of the diffusion tensor model is good when there is only one tract orientation within the voxels. Recovering multiple tract orientations from the same voxel is the topic of ongoing research. The dissertation starts with the study of noise distribution of diffusion-weighted images. Then it focuses on recovering multiple tract orientations from a single voxel by fitting Gaussian mixture models to data. In addition to fitting diffusion models to data, we collect the likelihood values for model selection and estimate the covariance matrix of the estimated parameters for probabilistic tractography. When more than one model is fitted to data, we carry out model selection to choose a model from two nested models or from a set of candidate models to avoid overfitting. We also extend the streamline method to enable probabilistic tracking by sampling the covariance matrix of the estimated parameters so that we provide many possible fiber tracts instead of a single tract.

We evaluate the methods proposed in this work with simulations. For phantom

www.manaraa.com

experiments, we mainly use the rat phantom data because it is acquired with a large number of gradient directions, achieving a higher angular resolution and a greater overall SNR. Another advantage of the rat phantom is that motion artifacts are minimized because the rat spinal cords are embedded in agar. As a result, we can focus on evaluating our algorithms with known fiber tracts.

7.2 Future Research Possibilities

Noise Estimation

Determining the optimum scanning parameters is not a straightforward task in diffusion MRI (Tournier et al., 2011). Images acquired from a large number of gradient directions are required to capture the complex diffusion profile caused by multiple fiber orientations within the same voxel. On the other hand, we also want to decrease voxel size to minimize PVE so that simpler models are sufficient. Both increasing gradient directions and decreasing voxel size require longer total acquisition time to maintain a fair SNR. However, longer acquisition time allows more motion artifacts. Therefore, we need make trade-offs between these parameters by studying the noise characteristics in diffusion-weighted images and their effect in different applications to determine better scanning parameters for different purposes.

Apparent Diffusion Coefficient Estimation

Multiple repeated scans are preferred to more gradient directions in the sense that motion-induced mis-registration between scans can be easily identified (Tournier et al., 2011). These two approaches do not make much difference in the diffusion tensor model. However, when fitting more complex models, a large number of gradient directions are often considered essential. It is possible to assess image quality by the power spectrum of ADC profile.

www.manaraa.com

Fitting Diffusion Models to Data and Model Selection

The diffusion tensor model is considered sufficient when the voxel is within a single tract. It is also used where the fiber tract only occupies part of the voxel. We should carefully investigate the effect on anisotropy indices caused by anisotropic diffusion and isotropic diffusion mixed in the same voxel, because anisotropy indices are often used as features in group studies.

The method proposed in Ch. 4 depends on numerical optimization. Better optimization algorithms and implementations are always desired. As mentioned in Ch. 4, fitting the Gaussian mixture model is not always numerically stable. We should determine the angular resolution and the level of image quality at which Gaussian mixture model with different number of compartments can be recovered.

The techniques described in Ch. 5 use goodness-of-fit of the diffusion-weighted signals as the model selection criteria. When local tract orientation is the major interest, it is desired to develop techniques that use the parameters being estimated as the major criteria. For example, a planar-shaped tensor may describe the ADC profile of a voxel containing fiber crossing well, especially when the angular resolution is low. However, the diffusion tensor model is unable to recover the underlying local orientations.

Tractography

Tractography techniques are promising in reconstructing major white matter fiber tracts. However, improvements in tractography algorithms are always desired, especially when the fiber tract of interest is relatively thin. It is possible to incorporate the method proposed in Ch. 6 into shape models of tract templates, which can be obtained from brain atlases, to improve the tracking accuracy of tracts of interest in group studies, because the anatomy of important fiber tracts are generally known and the major concern is to locate the same fiber tracts across different individuals.

www.manaraa.com

Copyright© Ning Cao, 2013.

www.manaraa.com

Bibliography

Akaike, H. (1974). A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6):716–723.

Alexander, A., Hasan, K., Lazar, M., Tsuruda, J., and Parker, D. (2001). Analysis of partial volume effects in diffusion-tensor MRI. *Magnetic Resonance in Medicine*, 45(5):770–780.

Alexander, D., Barker, G., and Arridge, S. (2002). Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data. *Magnetic Resonance in Medicine*, 48(2):331–340.

Andersen, A. (1996). On the Rician distribution of noisy MRI data. *Magnetic Resonance in Medicine*, 36(2):331–332.

Atkinson, K. (1982). Numerical integration on the sphere. *The Journal of the Australian Mathematical Society. Series B. Applied Mathematics*, 23(03):332–347.

Basser, P. and Jones, D. (2002). Diffusion-tensor MRI: theory, experimental design and data analysis–a technical review. *NMR in Biomedicine*, 15(7-8):456–467.

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

Basser, P., Mattiello, J., and Le Bihan, D. (1994b). MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66(1):259–267.

Basser, P., Pajevic, S., Pierpaoli, C., and Aldroubi, A. (2002). Fiber tract following in the human brain using DT-MRI data. *IEICE Transactions on Information and Systems*, 85(1):15–21.

Basser, P., Pajevic, S., Pierpaoli, C., Duda, J., and Aldroubi, A. (2000). In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine*, 44(4):625–632.

Basser, P. and Pierpaoli, C. (1996). Microstructural and physiological features of tissues Elucidated by quantitative-diffusion-Tensor MRI. *Journal of Magnetic Resonance, Series B*, 111(3):209–219.

Basu, S., Fletcher, P., and Whitaker, R. (2006). Rician noise removal in diffusion tensor MRI. *Lecture Notes in Computer Science*, 4190:117.

Battaglini, M., Smith, S. M., Brogi, S., and De Stefano, N. (2008). Enhanced brain extraction improves the accuracy of brain atrophy estimation. *NeuroImage*, 40(2):583–589.

www.manaraa.com

Bažant, P. and Oh, B. (1986). Efficient numerical integration on the surface of a sphere. *ZAMM-Journal of Applied Mathematics and Mechanics*, 66(1).

Behrens, T., Berg, H., Jbabdi, S., Rushworth, M., and Woolrich, M. (2007). Probabilistic diffusion tractography with multiple fibre orientations: What can we gain. *NeuroImage*, 34:144–55.

Behrens, T., Woolrich, M., Jenkinson, M., Johansen-Berg, H., Nunes, R., Clare, S., Matthews, P., Brady, J., and Smith, S. (2003). Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine*, 50(5):1077–1088.

Bernstein, M., Thomasson, D. M., and Perman, W. H. (1989). Improved detectability in low signal-to-noise ratio magnetic resonance images by means of a phase-corrected real reconstruction. *Medical Physics*, 16:813.

Brummer, M. E., Mersereau, R. M., Eisner, R. L., and Lewine, R. R. (1993). Automatic detection of brain contours in MRI data sets. *Medical Imaging, IEEE Transactions on*, 12(2):153–166.

Burnham, K. and Anderson, D. (2002). *Model Selection and Multimodel Inference: A Practical Information-theoretic Approach*. Springer Verlag.

Byerly, W. (1895). *An Elementary Treatise on Fourier's Series, and Spherical, Cylindrical, and Ellipsoidal Harmonics, with Applications to Problems in Mathematical Physics*. Ginn And Company, Boston, New York, Chicago, London, Atlanta, Dallas, Columbus, San Franciso.

Byrd, R. H., Gilbert, J. C., and Nocedal, J. (2000). A trust region method based on interior point techniques for nonlinear programming. *Mathematical Programming*, 89(1):149–185.

Byrd, R. H., Hribar, M. E., and Nocedal, J. (1999). An interior point algorithm for large-scale nonlinear programming. *SIAM Journal on Optimization*, 9(4):877–900.

Byrd, R. H., Nocedal, J., and Waltz, R. A. (2006). KNITRO: an integrated package for nonlinear optimization. In *Large-scale nonlinear optimization*, pages 35–59. Springer.

Caflisch, R. (1998). Monte Carlo and quasi-Monte Carlo methods. *Acta Numerica*, 7(1-49):553.

Campbell, J., Siddiqi, K., Rymar, V., Sadikot, A., and Pike, G. (2005). Flow-based fiber tracking with diffusion tensor and q-ball data: validation and comparison to principal diffusion direction techniques. *NeuroImage*, 27(4):725–736.

Cao, N., Gold, B. T., and Zhang, J. (2008). Partial volume effect of cingulum tract in diffusion-tensor MRI. In *Medical Imaging*, pages 69161U–69161U. International Society for Optics and Photonics.

www.manaraa.com

Cao, N., Liang, X., Zhuang, Q., and Zhang, J. (2009). Approximating high angular resolution apparent diffusion coefficient profiles using spherical harmonics under bigaussian assumption. In *SPIE Medical Imaging*, pages 726204–726204. International Society for Optics and Photonics.

Cao, N., Zhuang, Q., Liang, X., Huang, R., and Zhang, J. (2007). Computing white matter fiber orientations in high angular resolution diffusion-weighted MRI. In *Bioinformatics and Biomedical Engineering, 2007. ICBBE 2007. The 1st International Conference on*, pages 740–743. IEEE.

Chen, Y., Guo, W., Zeng, Q., Yan, X., Huang, F., Zhang, H., He, G., Vemuri, B. C., and Liu, Y. (2004). Estimation, smoothing, and characterization of apparent diffusion coefficient profiles from high angular resolution DWI. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 1, pages 588–593. IEEE.

Chen, Y., Guo, W., Zeng, Q., Yan, X., Rao, M., and Liu, Y. (2005). Apparent diffusion coefficient approximation and diffusion anisotropy characterization in DWI. *Proceedings of the 19th International Conference on Information Processing in Medical Imaging, Glenwood Springs, CO, USA*, pages 246–257.

Chenevert, T., Brunberg, J., and Pipe, J. (1990). Anisotropic diffusion within human white matter: demonstration with NMR techniques in vivo. *Radiology*, 177(2):401–405.

Chua, T., Wen, W., Slavin, M., and Sachdev, P. (2008). Diffusion tensor imaging in mild cognitive impairment and Alzheimer's disease: a review. *Current Opinion in Neurology*, 21(1):83.

Conturo, T., Lori, N., Cull, T., Akbudak, E., Snyder, A., Shimony, J., McKinstry, R., Burton, H., and Raichle, M. (1999). Tracking neuronal fiber pathways in the living human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 96(18):10422–10427.

Descoteaux, M., Angelino, E., Fitzgibbons, S., and Deriche, R. (2006). Apparent diffusion coefficients from high angular resolution diffusion imaging: estimation and applications. *Magnetic Resonance in Medicine*, 56:395–410.

Descoteaux, M., Deriche, R., Knosche, T., and Anwander, A. (2009). Deterministic and probabilistic tractography based on complex fibre orientation distributions. *Medical Imaging, IEEE Transactions on*, 28(2):269–286.

Douek, P., Turner, R., Pekar, J., Patronas, N., and Le Bihan, D. (1991). MR color mapping of myelin fiber orientation. *Journal of Computer Assisted Tomography*, 15(6):923–929.

Edelstein, W., Bottomley, P. A., and Pfeifer, L. M. (1984). A signal-to-noise calibration procedure for NMR imaging systems. *Medical Physics*, 11:180.

www.manaraa.com

Frank, L. (2001). Anisotropy in high angular resolution diffusion-weighted MRI. *Magnetic Resonance in Medicine*, 45(6):935–939.

Frank, L. (2002). Characterization of anisotropy in high angular resolution diffusion-weighted MRI. *Magnetic Resonance in Medicine*, 47(6):1083–1099.

Friman, O., Farneback, G., and Westin, C. (2006). A Bayesian approach for stochastic white matter tractography. *IEEE Transactions on Medical Imaging*, 25(8):965–978.

Gudbjartsson, H. and Patz, S. (1995). The Rician distribution of noisy MRI data. *Magnetic Resonance in Medicine*, 34(6):910–914.

Gudbjartsson, H. and Patz, S. (1996). Response. *Magnetic Resonance in Medicine*, 36(2):332–333.

Hasan, K., Parker, D., and Alexander, A. (2001). Comparison of gradient encoding schemes for diffusion-tensor MRI. *Journal of Magnetic Resonance Imaging*, 13(5):769–780.

He, L. and Greenshields, I. (2009). A nonlocal maximum likelihood estimation method for Rician noise reduction in MR images. *Medical Imaging, IEEE Transactions on*, 28(2):165–172.

Jansons, K. and Alexander, D. (2003). Persistent angular structure: new insights from diffusion magnetic resonance imaging data. *Inverse Problems*, 19(5):1031–1046.

Johansen-Berg, H. and Behrens, T. (2006). Just pretty pictures? What diffusion tractography can add in clinical neuroscience. *Current Opinion in Neurology*, 19(4):379–85.

Jones, D., Horsfield, M., and Simmons, A. (1999). Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Optimization*, 525.

Jones, D. K. and Pierpaoli, C. (2005). Confidence mapping in diffusion tensor magnetic resonance imaging tractography using a bootstrap approach. *Magnetic Resonance in Medicine*, 53(5):1143–1149.

Kang, N., Zhang, J., and Carlson, E. (2005a). Fiber tracking by simulating diffusion process with diffusion kernels in human brain with DT-MRI data. *Proceedings of SPIE*, 5746:126–137.

Kang, N., Zhang, J., Carlson, E., and Gembris, D. (2005b). White matter fiber tractography via anisotropic diffusion simulation in the human brain. *Medical Imaging, IEEE Transactions on*, 24(9):1127–1137.

Karlsen, O., Verhagen, R., and Bovee, W. (1999). Parameter estimation from Rician-distributed data sets using a maximum likelihood estimator: application to T1, and perfusion measurements. *Magnetic Resonance in Medicine*, 41(3):614–623.

www.manaraa.com

Kingsley, P. (2006a). Introduction to diffusion tensor imaging mathematics: Part II. Anisotropy, diffusion-weighting factors, and gradient encoding schemes. *Concepts in Magnetic Resonance Part A*, 28(2):123–154.

Kingsley, P. (2006b). Introduction to diffusion tensor imaging mathematics: Part III. Tensor calculation, noise, simulations, and optimization. *Concepts in Magnetic Resonance Part A*, 28(2):155–179.

Koch, R. and Becker, R. (2004). Evaluation of quadrature schemes for the discrete ordinates method. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 84(4):423–435.

Kreher, B., Schneider, J., Mader, I., Martin, E., Hennig, J., and Il'yasov, K. (2005). Multitensor approach for analysis and tracking of complex fiber configurations. *Magnetic Resonance in Medicine*, 54(5):1216.

Kristoffersen, A. (2007). Optimal estimation of the diffusion coefficient from non-averaged and averaged noisy magnitude data. *Journal of Magnetic Resonance*, 187(2):293–305.

Landman, B., Bazin, P.-L., and Prince, J. (2007). Diffusion tensor estimation by maximizing rician likelihood. *Computer Vision, IEEE International Conference on*, 0:1–8.

Lazar, M., Weinstein, D., Tsuruda, J., Hasan, K., Arfanakis, K., Meyerand, M., Badie, B., Rowley, H., Haughton, V., and Field, A. (2003). White matter tractography using diffusion tensor deflection. *Human Brain Mapping*, 18(4):306–321.

Le Bihan, D. (1986). MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology*, 161(2):401–407.

Le Bihan, D., Mangin, J., Poupon, C., Clark, C., Pappata, S., Molko, N., and Chabriat, H. (2001). Diffusion tensor imaging: concepts and applications. *Journal of Magetic Resonance Imaging*, 13:534–546.

Le Bihan, D., van Zijl, P., et al. (2002). From the diffusion coefficient to the diffusion tensor. *NMR in Biomedicine*, 15(7-8):431–434.

Lebedev, V. (1977). Spherical quadrature formulas exact to orders 25–29. *Siberian Mathematical Journal*, 18(1):99–107.

Lebedev, V. I. (1976). Quadratures on a sphere. *USSR Computational Mathematics and Mathematical Physics*, 16(2):10–24.

Liang, X., Zhuang, Q., Cao, N., and Zhang, J. (2009). Shape modeling and clustering of white matter fiber tracts using fourier descriptors. In *Computational Intelligence in Bioinformatics and Computational Biology, 2009. CIBCB'09. IEEE Symposium on*, pages 292–297. IEEE.

www.manaraa.com

MacRobert, T. and Sneddon, I. (1967). *Spherical Harmonics: An Elementary Treatise on Harmonic Functions with Applications*. Pergamon Press, 3 edition.

McLaren, A. (1963). Optimal numerical integration on a sphere. *Mathematics of Computation*, 17(84):361–383.

Mori, S. and Barker, P. (1999). Diffusion magnetic resonance imaging: its principle and applications. *The Anatomical Record*, 257:102–109.

Mori, S., Crain, B., Chacko, V., and van Zijl, P. (1999). Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Annals of Neurology*, 45(2):265–269.

Mori, S. and Wakana, S. (2005). *MRI Atlas of Human White Matter*. American Society of Neuroradiology.

Moseley, M. (1990). Diffusion-weighted MR imaging of acute stroke: correlation with T2-weighted and magnetic susceptibility-enhanced MR imaging in cats. *American Journal of Neuroradiology*, 11(3):423–429.

Mukherjee, P., Chung, S., Berman, J., Hess, C., and Henry, R. (2008). Diffusion tensor MR imaging and fiber tractography: technical considerations. *American Journal of Neuroradiology*, 29(5):843–852.

Özarslan, E. and Mareci, T. (2003). Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 50(5):955–965.

Özarslan, E., Vemuri, B., and Mareci, T. (2005). Generalized scalar measures for diffusion MRI using trace, variance, and entropy. *Magnetic Resonance in Medicine*, 53(4).

Pagani, E., Filippi, M., Rocca, M., and Horsfield, M. (2005). A method for obtaining tract-specific diffusion tensor MRI measurements in the presence of disease: application to patients with clinically isolated syndromes suggestive of multiple sclerosis. *NeuroImage*, 26(1):258–265.

Pajevic, S., Aldroubi, A., and Basser, P. (2002). A continuous tensor field approximation of discrete DT-MRI data for extracting microstructural and architectural features of tissue. *Journal of Magnetic Resonance*, 154(1):85–100.

Pajevic, S., Pierpaoli, C., et al. (1999). Color schemes to represent the orientation of anisotropic tissues from diffusion tensor data: application to white matter fiber tract mapping in the human brain. *Magnetic Resonance in Medicine*, 42(3):526–540.

Papadakis, N., Murrills, C., Hall, L., Huang, C., and Adrian Carpenter, T. (2000). Minimal gradient encoding for robust estimation of diffusion anisotropy. *Magnetic Resonance Imaging*, 18(6):671–679.

www.manaraa.com

Parker, G. (2004). Analysis of MR diffusion weighted images. *British Journal of Radiology*, 77(suppl_2):S176.

Parker, G. and Alexander, D. (2003). Probabilistic Monte Carlo based mapping of cerebral connections utilizing whole-brain crossing fibre information. *Information Processing in Medical Imaging*, 18:684–695.

Parker, G., Haroon, H., and Wheeler-Kingshott, C. (2003). A framework for a streamline-based probabilistic index of connectivity using a structural interpretation of MRI diffusion measurements. *Journal of Magnetic Resonance Imaging*, 18(2):242–254.

Parker, G. J. and Alexander, D. C. (2005). Probabilistic anatomical connectivity derived from the microscopic persistent angular structure of cerebral tissue. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1457):893–902.

Parker, G. J., Wheeler-Kingshott, C. A., and Barker, G. J. (2002). Estimating distributed anatomical connectivity using fast marching methods and diffusion tensor imaging. *Medical Imaging, IEEE Transactions on*, 21(5):505–512.

Peled, S., Gudbjartsson, H., Westin, C.-F., Kikinis, R., and Jolesz, F. A. (1998). Magnetic resonance imaging shows orientation and asymmetry of white matter fiber tracts. *Brain Research*, 780(1):27–33.

Pfefferbaum, A. and Sullivan, E. (2003). Increased brain white matter diffusivity in normal adult aging: Relationship to anisotropy and partial voluming. *Magnetic Resonance in Medicine*, 49(5):953–961.

Pierpaoli, C., Jezzard, P., Basser, P., Barnett, A., and Di Chiro, G. (1996). Diffusion tensor MR imaging of the human brain. *Radiology*, 201(3):637–48.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464.

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

Sijbers, J., Den Dekker, A., Raman, E., and Van Dyck, D. (1999). Parameter estimation from magnitude MR images. *International Journal of Imaging Systems and Technology*, 10:11–4.

Sijbers, J., Den Dekker, A., Scheunders, P., and Van Dyck, D. (1998a). Maximum-likelihood estimation of Rician distribution parameters. *IEEE Transactions on Medical Imaging*, 17(3):357–361.

Sijbers, J., Den Dekker, A., Van Audekerke, J., Verhoye, M., and Van Dyck, D. (1998b). Estimation of the noise in magnitude MR images. *Magnetic Resonance Imaging*, 16(1):87–90.

www.manaraa.com

Skare, S., Li, T., Nordell, B., and Ingvar, M. (2000). Noise considerations in the determination of diffusion tensor anisotropy. *Magnetic Resonance Imaging*, 18(6):659–669.

Sobolev, S. (1962). Formulas of mechanical cubature on the surface of the sphere. *Siberian Mathematical Journal*, 3:769–796.

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

Stevensson, B. and Edén, M. (2006). Efficient orientational averaging by the extension of Lebedev grids via regularized octahedral symmetry expansion. *Journal of Magnetic Resonance*, 181(1):162–176.

Stieltjes, B., Kaufmann, W., van Zijl, P., Fredericksen, K., Pearlson, G., Solaiyappan, M., and Mori, S. (2001). Diffusion tensor imaging and axonal tracking in the human brainstem. *NeuroImage*, 14(3):723–735.

Stroud, A. (1971). *Approximate Calculation of Multiple Integrals*. Prentice Hall, Englewood Cliffs, New Jersey.

Sugiura, N. (1978). Further analysts of the data by akaike's information criterion and the finite corrections: further analysts of the data by akaike's. *Communications in Statistics–Theory and Methods*, 7(1):13–26.

Taylor, W., Hsu, E., Krishnan, K., and MacFall, J. (2004). Diffusion tensor imaging: background, potential, and utility in psychiatric research. *Biological Psychiatry*, 55(3):201–207.

Tournier, J., Calamante, F., Gadian, D., and Connelly, A. (2003). Diffusion-weighted magnetic resonance imaging fibre tracking using a front evolution algorithm. *NeuroImage*, 20(1):276–88.

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

Tournier, J., Calamante, F., King, M., Gadian, D., and Connelly, A. (2002). Limitations and requirements of diffusion tensor fiber tracking: An assessment using simulations. *Magnetic Resonance in Medicine*, 47(4):701–708.

Tournier, J.-D., Mori, S., and Leemans, A. (2011). Diffusion tensor imaging and beyond. *Magnetic Resonance in Medicine*, 65(6):1532–1556.

Tuch, D., Reese, T., Wiegell, M., Makris, N., Belliveau, J., and Wedeen, V. (2002). High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine*, 48(4):577–582.

www.manaraa.com

Tuch, D., Reese, T., Wiegell, M., and Wedeen, V. (2003). Diffusion MRI of complex neural architecture. *Neuron*, 40(5):885–895.

Tuch, D., Weisskoff, R., Belliveau, J., and Wedeen, V. (1999). High angular resolution diffusion imaging of the human brain. *Proceedings of the 7th Annual Meeting of ISMRM, Philadelphia*.

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

Ugray, Z., Lasdon, L., Plummer, J., Glover, F., Kelly, J., and Martí, R. (2007). Scatter search and local NLP solvers: a multistart framework for global optimization. *INFORMS Journal on Computing*, 19(3):328–340.

Wakana, S., Jiang, H., Nagae-Poetscher, L., van Zijl, P., and Mori, S. (2004). Fiber tract–based atlas of human white matter anatomy. *Radiology*, 230:77.

Wang, X. and Carrington, T. (2003). Using Lebedev grids, sine spherical harmonics, and monomer contracted basis functions to calculate bending energy levels of HF trimer. *Journal of Theoretical and Computational Chemistry*, 2(4):599–608.

Westin, C.-F., Maier, S. E., Mamata, H., Nabavi, A., Jolesz, F. A., and Kikinis, R. (2002). Processing and visualization for diffusion tensor MRI. *Medical image analysis*, 6(2):93–108.

Zhang, J., Ji, H., Kang, N., and Cao, N. (2005a). Fiber tractography in diffusion tensor magnetic resonance imaging: a survey and beyond. In *Proceedings of the 2005 International Symposium on Medical Imaging and Computing*, pages 39–56.

Zhang, J., Kang, N., and Rose, S. (2005b). Approximating anatomical brain connectivity with diffusion tensor MRI using kernel-based diffusion simulations. *Information Processing in Medical Imaging*, 19:64–75.

www.manaraa.com

Vita

Personal Data

Name Ning Cao

Birth Place Beijing, China

Educational Background

- Master of Science in Computer Science, Tsinghua University, China, 2002.
- Bachelor of Science in Computer Science, Tsinghua University, China, 2000.

Professional Experience

- Research Assistant, 08/2004 - 05/2007, Department of Computer Science, University of Kentucky, Lexington, Kentucky.
- Teaching Assistant, 08/2007 - 12/2009, Department of Computer Science, University of Kentucky, Lexington, Kentucky.

Publications

- Ning Cao, Brian T. Gold, and Jun Zhang. Partial volume effect of cingulum tract in diffusion-tensor mri. In *Medical Imaging*, pages 69161U–69161U. International Society for Optics and Photonics, 2008.
- Ning Cao, Xuwei Liang, Qi Zhuang, and Jun Zhang. Approximating high angular resolution apparent diffusion coefficient profiles using spherical harmonics under bigaussian assumption. In *SPIE Medical Imaging*, pages 726204–726204. International Society for Optics and Photonics, 2009.
- Ning Cao, Qi Zhuang, Xuwei Liang, Ruiwang Huang, and Jun Zhang. Computing white matter fiber orientations in high angular resolution diffusion-weighted mri. In *Bioinformatics and Biomedical Engineering, 2007. ICBBE 2007. The 1st International Conference on*, pages 740–743. IEEE, 2007.
- Xuwei Liang, Qi Zhuang, Ning Cao, and Jun Zhang. Quantitative and visual analysis of white matter integrity using diffusion tensor imaging. In *SPIE Medical Imaging*, pages 726131–726131. International Society for Optics and Photonics, 2009.

www.manaraa.com

- Xuwei Liang, Qi Zhuang, Ning Cao, and Jun Zhang. Shape modeling and clustering of white matter fiber tracts using fourier descriptors. In *Computational Intelligence in Bioinformatics and Computational Biology, 2009. CIBCB'09. IEEE Symposium on*, pages 292–297. IEEE, 2009.

- Cheng Qin, Ning Kang, and Ning Cao. Performance evaluation of anisotropic diffusion simulation based tractography on phantom images. In *Proceedings of the 45th annual southeast regional conference*, pages 521–522. ACM, 2007.

- Jun Zhang, Hao Ji, Ning Kang, and Ning Cao. Fiber tractography in diffusion tensor magnetic resonance imaging: A survey and beyond. In *Proceedings of the 2005 International Symposium on Medical Imaging and Computing*, pages 39–56, 2005.

- Qi Zhuang, Xuwei Liang, Ning Cao, and Jun Zhang. Generalized analytic expressions for the b matrix of twice-refocused spin echo pulse sequence. In *SPIE Medical Imaging*, pages 725920–725920. International Society for Optics and Photonics, 2009.

www.manaraa.com