

QUANTIFYING PATHOLOGY  
IN DIFFUSION WEIGHTED MRI

Matthan Caan

**About the cover**

Amyotrophic Lateral Sclerosis (ALS) is characterized by a loss of white matter tissue integrity, mainly along the corticospinal tract. The orange isosurface encapsulates the voxels in which tissue is affected. The color along the corticospinal tracts depicts the degree to which patients are affected at that particular location. Commissural fibers (in red) are reconstructed in a dual tensor atlas of this cohort, computed using our developed web portal to the grid.

QUANTIFYING PATHOLOGY  
IN DIFFUSION WEIGHTED MRI

PROEFSCHRIFT

ter verkrijging van de graad van doctor  
aan de Technische Universiteit Delft,  
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,  
voorzitter van het College voor Promoties,  
in het openbaar te verdedigen op maandag 11 oktober 2010 om 15:00 uur

door

Matthan Willem Abraham CAAN

natuurkundig ingenieur  
geboren te Kapelle.

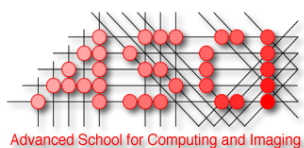
Dit proefschrift is goedgekeurd door de promotoren:  
Prof. dr. ir. L.J. van Vliet en prof. dr. ir. C.A. Grimbergen

Copromotor:  
Dr. F.M. Vos

Samenstelling promotiecommissie:

Rector Magnificus	voorzitter
Prof. dr. ir. L.J. van Vliet	Technische Universiteit Delft
Prof. dr. ir. C.A. Grimbergen	Universiteit van Amsterdam
Dr. F.M. Vos	Technische Universiteit Delft
Prof. dr. ir. F.C.T. van der Helm	Technische Universiteit Delft
Prof. dr. J. Sijbers	Universiteit Antwerpen (CDE)
Prof. dr. ir. B.M. ter Haar Romeny	Technische Universiteit Eindhoven
Prof. dr. S. Sunaert	Katholieke Universiteit Leuven
Prof. dr. W.J. Niessen (reservelid)	Erasmus Universiteit Rotterdam

This work is part of the programme of BiG Grid, the Dutch e-Science Grid, which is financially supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Organisation for Scientific Research, NWO).



This work was carried out in the ASCI graduate school.  
ASCI dissertation number 210.

Printed by: Ridderprint BV, Ridderkerk, The Netherlands.

ISBN 978-90-5335-309-7  
© 2010, Matthan Caan. All rights reserved.



# CONTENTS

---

1	INTRODUCTION	1
1.1	Brain structure	1
1.2	Measuring diffusion	3
1.2.1	Diffusion in brain tissue	3
1.2.2	Diffusion Weighted MRI	3
1.3	Pathology	6
1.4	Data analysis	6
1.5	The grid - a virtual laboratory	7
1.6	Thesis objectives	7
I	DISCRIMINANT ANALYSIS	11
2	SHAVING DTI IN DISCRIMINANT ANALYSIS	13
2.1	Introduction	14
2.2	Materials and Methods	15
2.2.1	Subjects	15
2.2.2	Data acquisition	15
2.2.3	Data preprocessing	15
2.2.4	Algorithm design	17
2.2.5	Algorithm evaluation	20
2.3	Results	21
2.3.1	Simulated data	21
2.3.2	Experimental data	22
2.4	Discussion and conclusion	25
3	PRUNING DATASETS IN DISCRIMINANT ANALYSIS	27
3.1	Introduction	28
3.2	Method	28
3.2.1	Mapping computation	28
3.2.2	Classification	29
3.2.3	Pruning	30
3.2.4	Thresholding	31
3.2.5	Subjects	31
3.2.6	Data acquisition and preprocessing	32
3.3	Results	33
3.4	Discussion	34
4	LINEAR AND KERNEL FISHER DISCRIMINANT ANALYSIS	35
4.1	Introduction	36
4.2	Method	36
4.2.1	Discrimination	36
4.2.2	Classification	39
4.2.3	Subjects	40

4.2.4	Data acquisition and preprocessing	40
4.3	Results	41
4.4	Discussion	42
II	RESOLVING CROSSING FIBERS	45
5	GENERATING FIBER CROSSING PHANTOMS OUT OF EXPERIMENTAL DWIS	47
5.1	Introduction	48
5.2	Method	49
5.2.1	Experimental phantom data	49
5.2.2	Two-tensor model	50
5.3	Results	51
5.4	Discussion	53
6	DUAL TENSOR ATLAS GENERATION	55
6.1	Introduction	56
6.2	Method	57
6.2.1	Dual Tensor Model	57
6.2.2	Simulating Variation in the Cohort	57
6.2.3	Building a Dual Tensor Atlas	58
6.3	Results	60
6.3.1	Synthetic Data	60
6.3.2	Clinical Data	60
6.4	Discussion	61
7	ESTIMATION OF DIFFUSION PROPERTIES IN CROSSING FIBER BUNDLES	63
7.1	Introduction	64
7.2	Background	65
7.2.1	Diffusion modeling	65
7.2.2	Optimization of diffusion imaging	65
7.3	Theory	66
7.3.1	Dual tensor model parameterization	66
7.3.2	Maximum likelihood estimation	67
7.3.3	Cramér-Rao analysis for model selection and optimizing $\mathbf{b}$	68
7.4	Experiments	70
7.4.1	Model selection	70
7.4.2	Optimizing $\mathbf{b}$ for different tissue configurations	73
7.4.3	Validation in brain data	77
7.5	Discussion	80
7.5.1	Model selection	80
7.5.2	Optimizing $\mathbf{b}$ for different tissue configurations	81
7.5.3	Validation in brain data	82
7.A	Appendix	83
7.A.1	Derivatives to estimated parameters	83
7.A.2	Cramér-Rao Lower Bounds on eigenvalues and FA	84

III	FIBER TRACT STATISTICS	87
8	NON-RIGID POINT SET MATCHING OF FIBER TRACTS	89
8.1	Introduction	90
8.2	Background	91
8.2.1	Data acquisition and reconstruction	91
8.2.2	Uncertainty in principal diffusion orientation	92
8.2.3	Dyadic coherence	93
8.3	Fiber tract matching	93
8.3.1	Feature selection	93
8.3.2	Classic point set registration	93
8.3.3	Clustering fiber tracts	96
8.3.4	Membership in fiber tracts	98
8.3.5	Parameters	99
8.4	Experiments	100
8.4.1	Synthetic data	100
8.4.2	Brain data	101
8.5	Conclusion	106
9	SPATIAL CONSISTENCY IN 3D TRACT-BASED CLUSTERING STATISTICS	109
9.1	Introduction	110
9.2	Method	111
9.2.1	Feature selection	111
9.2.2	Registration	111
9.2.3	Statistics	112
9.3	Results	113
9.4	Discussion	115
IV	CLINICAL STUDIES	117
10	WMFA CORRELATES WITH SPEED IN YOUNG CHILDHOOD CANCER SURVIVORS	119
10.1	Introduction	120
10.2	Methods and Materials	121
10.2.1	Patients	121
10.2.2	Measurements	122
10.2.3	MRI Data Analyses	122
10.2.4	Statistical Analyses	123
10.3	Results	124
10.3.1	Participants	124
10.3.2	Neurocognitive functioning	124
10.3.3	Structural MR Images	125
10.3.4	WMFA findings	126
10.3.5	Correlations of WMFA with cognitive functioning	127
10.4	Discussion	127
10.5	Conclusion	128
11	SPATIAL PROFILING OF THE CORTICOSPINAL TRACT IN RECENT ON-SET MND	131

11.1	Introduction	132
11.2	Materials and Methods	132
11.2.1	Subjects	132
11.2.2	Clinical parameters	133
11.2.3	Data acquisition	133
11.2.4	DTI processing	133
11.2.5	DTI postprocessing: Fibretracking	134
11.3	Statistical analysis	134
11.3.1	Clinical data	134
11.3.2	DTI data	135
11.3.3	Power analysis	135
11.4	Results	135
11.4.1	Clinical characteristics	135
11.4.2	DTI: mean FA value of the CST	135
11.4.3	DTT: profiles of FA along the CST	137
11.4.4	Correlation of clinical parameters with FA values	139
11.5	Discussion	139
11.5.1	Strengths and limitations of the study	139
11.5.2	PLS	139
11.5.3	ALS-L and ALS-B	140
11.5.4	PMA	141
11.5.5	Overall conclusion	141
V	GRID IMPLEMENTATION	143
12	USING DYNAMIC CONDOR-BASED SERVICES	145
12.1	Introduction and Related Work	146
12.2	Application	147
12.3	Infrastructure	149
12.3.1	Challenges in distributing the application	149
12.3.2	Running Compiled MATLAB Applications on a Condor grid	150
12.4	Implementation	150
12.4.1	Compiling the Image Processing Engine (IPE)	150
12.4.2	System Overview	151
12.4.3	Porting the application	152
12.4.4	Fault Tolerance	153
12.5	Results	154
12.5.1	Application output	154
12.5.2	Application Execution	155
12.5.3	Infrastructure Performance	156
12.6	Discussion and Conclusions	157
13	GRIDIFYING A DIFFUSION TENSOR IMAGING PIPELINE	159
13.1	Introduction	160
13.2	The Diffusion Tensor MRI Analysis Pipeline	160
13.3	Grid implementation	162
13.3.1	Command-line jobs on a cluster	164

13.3.2	Condor pool of desktop machines	164
13.3.3	MOTEUR workflows on the Dutch Grid	164
13.3.4	Web-based service with nested fault-tolerant workflows	165
13.4	Results	167
13.4.1	Parameter sweep for pattern recognition on Dutch grid and on Condor pool	167
13.4.2	Coregistration as a MOTEUR workflow	169
13.4.3	Atlas building using a web-based service	169
13.4.4	Performance Summary	169
13.4.5	Clinical results	169
13.5	Discussion and Conclusions	171
14	CONCLUSIONS	173
14.1	Part I - Discriminant analysis	173
14.2	Part II - Resolving crossing fibers	173
14.3	Part III - Fiber tract statistics	174
14.4	Part IV - Clinical studies	175
14.5	Part V - Grid implementation	175
14.6	Future research	176
	BIBLIOGRAPHY	179
	LIST OF ABBREVIATIONS	189
	LIST OF PUBLICATIONS	191
	SUMMARY	193
	SAMENVATTING	195
	ACKNOWLEDGEMENTS	197
	CURRICULUM VITAE	199



## INTRODUCTION

---

Early exploration of the human brain can be dated back to 1600 B.C., when the Egyptians prepared the dead for embalming. This led to the discovery of e.g. cerebral spinal fluid (figure 1.1) [104]. Curiously enough, the exact working of the brain is still only marginally understood. Until approximately 30 years ago, measurements on brain tissue composition could only be performed on dead tissue outside the body (*ex vivo*). In the past decades Magnetic Resonance Imaging (MRI) techniques allowed for *in vivo* inspection of living brain tissue. More recent advances led to Diffusion Weighted MRI (DW-MRI), an MR protocol that measures the orientation and shape of the diffusion profile of water in white matter. Employing this imaging modality allows for studying pathologies that affect the white matter.

### 1.1 BRAIN STRUCTURE

The central nervous system (CNS) maintains both control over the human body and communication with the surroundings. Incoming or afferent signals from the periphery traverse through the spinal chord to the brain. After processing, an outgoing or efferent signal may initiate remote action. Afferent and efferent fibers are arranged in bundles called nerves, as shown in figure 1.2(a).

A nerve cell, or neuron (see figure 1.2(b)), transmits impulses from one part of the body to the other. It consists of a cell body, to which dendrites are connected that conduct an impulse towards the cell body. These cell bodies color the gray matter of the brain. Attached to the cell body is a nerve fiber, which is built up by an axon, a single and potentially long (up to 1 meter) cell that conducts impulses to other neurons or tissue. A microtubular structure of neurofibrils gives structure to the axon, while it may be surrounded by a myelin sheath to enhance conductivity. This myelin sheath is known to give white matter its characteristic white color [234].

On macroscopic scale, several brain parts can be recognized, as shown in figure 1.3. Fibers enter the brain in the brainstem. Above the brainstem lies the diencephalon, which includes the thalamus and hypothalamus. These gray matter structures are important for vital functions and are regarded as ‘connecting stations’. The cerebrum spreads out over the diencephalon and forms the majority of the brain. The cerebrum (or telencephalon) is divided into left and right hemispheres. The cerebrum consists of the cerebral cortex: frontal, parietal, temporal and occipital lobes as indicated in figure 1.3,

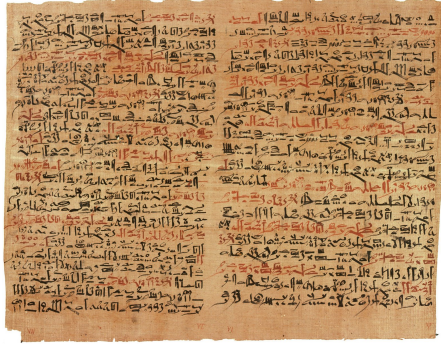


Figure 1.1: The Edwin Smith papyrus, the world's oldest surgical document, mentioning the brain and cerebrospinal fluid: "If thou examinest a man having a gaping wound in his head (...) rending open the brain of his skull, thou shouldst palpate his wound. Shouldst thou find that smash which is in his skull [like] those corrugations which form in molten copper, (and) something therein throbbing (and) fluttering under thy fingers (...): "An ailment not to be treated." Thou shouldst anoint that wound with grease." [49, 104]

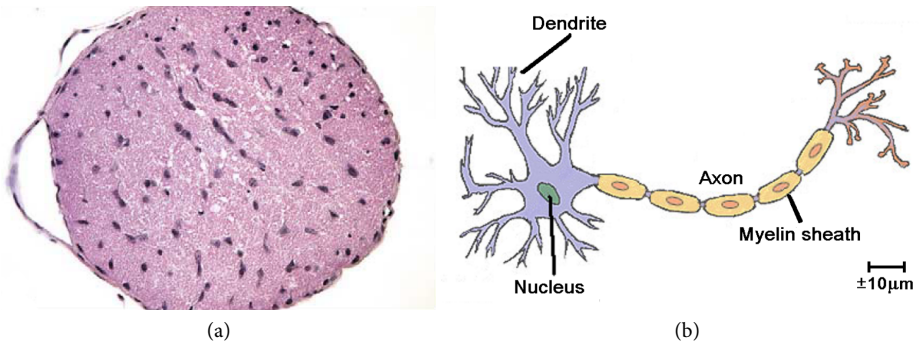


Figure 1.2: (a) Staining of the mouse optic nerve [227]. (b) An illustration of a neuron [255].

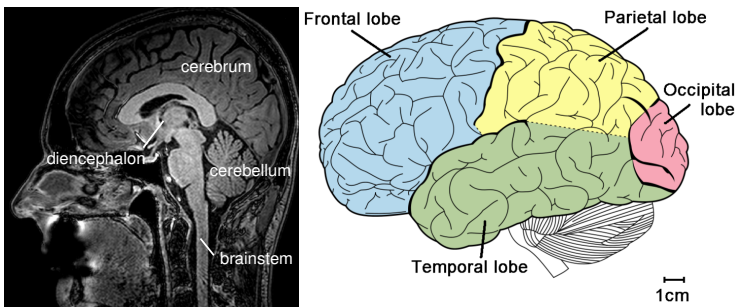


Figure 1.3: (a) Overview of the Central Nervous System, annotated on a ( $T_1$ -weighted) structural MR mid-sagittal slice. (b) The cerebrum may be subdivided in several lobes [20].



and the basal ganglia. At the cerebral surface, gray matter develops more rapidly than the underlying white matter, such that the cerebral surface folds into gyri. The sulci, observed as downfolds, are filled with cerebrospinal fluid (CSF). The cerebellum, or small brain, is situated beneath the cerebrum, posterior to the brain stem.

Communication between these different brain regions takes place on different scales, over three types of white matter fibers [138], as is illustrated in figure 1.4. *Association fibers* connect gyri within one hemisphere. Connections between both hemispheres are called *commissural fibers*. The largest commissural fiber is the corpus callosum, connecting the hemispheres. *Projection fibers* connect the cortex of the cerebrum to lower brain parts, as well as the spinal chord. An example is the corticospinal tract, mainly related to motor functions. Half of its fibers arise in the precentral gyrus (the primary motor cortex) and travel through brainstem and spinal chord towards feet and hands.

## 1.2 MEASURING DIFFUSION

### 1.2.1 *Diffusion in brain tissue*

Water is abundant within brain tissue. Fiber bundles can be characterized by studying diffusion properties of water. Diffusion of water molecules in a free medium due to thermal energy, called Brownian motion, was derived by Einstein to be  $\langle r^2 \rangle = 2D\tau$ , where  $\langle r^2 \rangle$  is the average squared distance, and  $\tau$  a specific diffusion time. Molecular displacement is Gaussian distributed and characterized by the diffusion coefficient  $D$ . For a representative value of  $D = 2 \cdot 10^{-3} \text{ mm}^2\text{s}^{-1}$  and a diffusion time of  $\tau = 50 \text{ ms}$  this implies an average diffusion distance of  $r = 17 \mu\text{m}$ , that is 32% of all molecules have traveled at least this distance [44]. Within white matter tissue, however, axonal boundaries form geometric restrictions which cause an apparent anisotropic diffusion profile. As such, in diffusion times between 50 and 100 ms water molecules diffuse over distances ranging from 1 to 15  $\mu\text{m}$ . The anisotropy of the diffusion profile provides key insights into the geometry of local tissue structure.

### 1.2.2 *Diffusion Weighted MRI*

Diffusion processes can be measured and visualized by means of Diffusion Weighted Magnetic Resonance Imaging (DW-MRI). This MR-modality specifically allows to measure the orientation of anisotropic tissues composing the brain. In addition, the amount of diffusivity can be precisely estimated, such that potentially subtle changes in tissue composition can be measured. A common voxel size is  $2.0^3 \text{ mm}^3$  (at a field strength of 3.0 Tesla).

Classically, a single Gaussian diffusion process is assumed within a voxel, described by a symmetric positive definite rank-two diffusion tensor  $\mathbf{D}$ .

$$\mathbf{D} = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{pmatrix}. \quad (1.1)$$

The Stejskal and Tanner equation Stejskal and Tanner [228] describes the diffusion weighted magnitude signal attenuation  $S(\mathbf{g})/S_0$  in the case of a Gaussian diffusion process:

$$\frac{S(\mathbf{g})}{S_0} = e^{-b\mathbf{g}^T \mathbf{D} \mathbf{g}} \quad (1.2)$$

where  $\mathbf{g}$  is the gradient direction and  $S_0$  is the  $T_2$ -signal without diffusion weighting. The diffusion weighting coefficient  $b$  is

$$b \equiv \gamma^2 \delta^2 \left( \Delta - \frac{\delta}{3} \right) G^2, \quad (1.3)$$

with  $\gamma$  is the gyro-magnetic ratio,  $\delta$  is the diffusion weighting gradient duration,  $\Delta$  is the time between both gradients and  $G$  is the gradient strength. Also, it is assumed that  $\delta \ll \Delta$ . In figure 1.5 the diffusion weighting part of an MR pulse sequence [228] is plotted. Two gradient pulses are centered around the  $180^\circ$  refocusing pulse, inducing and inverting a phase shift. Spins that have moved due to Brownian motion will experience different phase shifts, resulting in signal attenuation at echo time TE.

The Apparent Diffusion Coefficient (ADC)-profile consists of at least six independently measured  $S(\mathbf{g})$ -values, from which  $\mathbf{D}$  is computed via a linear system of equations [256]. An eigen-decomposition of the tensor results in three eigenvectors  $\mathbf{e}_{1..3}$  and their corresponding eigenvalues  $\lambda_{1..3}$ , which ease the analysis of the data. Parameters that are clinically studied are the Mean Diffusivity (MD)

$$\text{MD} = \frac{1}{3} (\lambda_1 + \lambda_2 + \lambda_3) \quad (1.4)$$

and the Fractional Anisotropy (FA) [32],

$$\text{FA} = \sqrt{\frac{3}{2}} \sqrt{\frac{(\lambda_1 - \bar{\lambda})^2 + (\lambda_2 - \bar{\lambda})^2 + (\lambda_3 - \bar{\lambda})^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}, \quad (1.5)$$

which provides a distance measure to the nearest isotropic tensor.

Fiber tracking aims at reconstructing global white matter trajectories, by generating streamlines through the principal eigenvector-field [39]. Note that the resolution of the

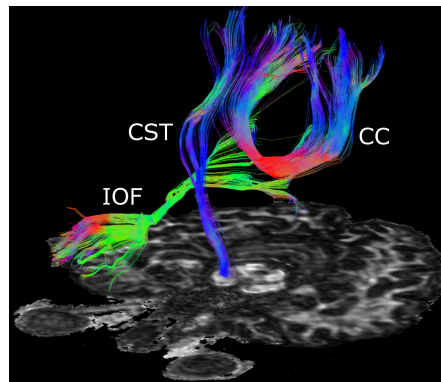


Figure 1.4: An axial slice of the Fractional Anisotropy (FA), with exemplar white matter fibers, in the author's brain. Association fibers: inferior occipitofrontal fasciculus (IOF). Projection fibers: corticospinal tract (CST). Commissural fibers: body of the corpus callosum (CC).

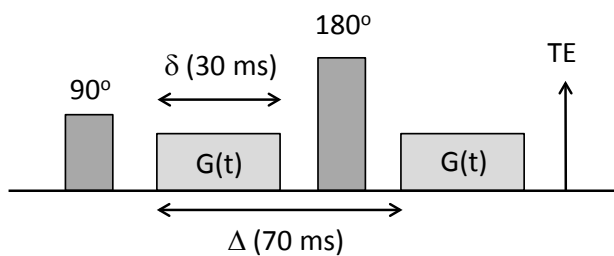


Figure 1.5: Diffusion weighting of an MR pulse sequence with typical values for a clinical setting.

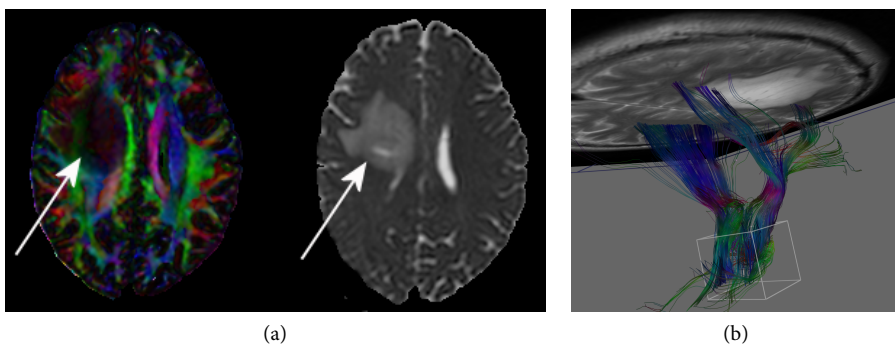


Figure 1.6: MR imaging of a brain tumor. (a) An axial slices of the FA, color coded by the principal orientation and the MD, with annotated tumor. (b) Tracked bilateral pyramidal tracts, superimposed on an axial  $T_2$ -weighted slice. The tumor has infiltrated into healthy white matter tissue. Note the low FA-value within the tumor.

data is two orders of magnitude lower than the axonal diameter. As such, DW-MRI only describes macroscopic brain tissue properties.

### 1.3 PATHOLOGY

Pathological processes may affect the CNS in different ways. Aggressive infiltration by a brain tumor is easily identified by the human eye in figure 1.6. Other diseases may not be identified in individual MR scans, while the clinical observations are distinct. Schizophrenia is a cognitive disorder occurring in about 1% of the world population. The first symptoms usually occur at early adolescence, with signs of abnormal social behavior and hallucinations. What is generally assumed, is that schizophrenia affects white matter structure in the brain [224]. Amyotrophic lateral sclerosis (ALS) is a progressive motor neuron disease with features of both upper motor neuron and lower motor neuron degeneration and poor prognosis [91, 267]. The cause of sporadic ALS is currently unknown. Functional evidence for brain diseases can be explained by specific structural loss in the white matter of the brain. That is, certain *biomarkers* may exist where the disease inhibits proper functioning.

DW-MRI is expected to be able to identify such biomarkers by capturing subtle white matter changes in an early stage [44]. Axonal and myelin sheath damage are found to respectively decrease the parallel and increase the perpendicular diffusion [227]. Consequently, a decreased FA may be found, correlated with an increased MD, such as in a study into Multiple Sclerosis [77]. In this retrospective study, patients are scanned after the disease manifests itself and compared to carefully matched healthy controls, to minimize negative side-effects of confounding factors such as age, education level and handedness. Statistics are performed on group level to quantify differences that exceed normal variations, as well as measurement noise, as is described below.

### 1.4 DATA ANALYSIS

Based on prior knowledge, different strategies have been developed for analyzing the DW-MRI data. In case no clear hypothesis for the disease is available, Voxel Based Analysis (VBA) is to be performed, which tests for significant differences in all voxels between the patient and control group. VBA is preceded by spatial normalization or coregistration of the data, aiming at finding spatial correspondence at each voxel between all subjects. Coregistration is performed in both an affine and non-rigid manner, to correct for both global and local differences [169].

When only local regions are thought to be affected, an expert may manually annotate these Regions of Interest (ROIs), or the ROIs can be automatically labeled based on an atlas [181]. Also, automatic segmentation of a specific tissue type such as the white matter can provide an area of interest. Average scalar values are then computed per ROI and compared.

Studies into brain connectivity and tissue integrity along white matter tracts are facilitated by tractography methods. Tracked fibers provide a frame of reference along which profiles are reconstructed [89].

## 1.5 THE GRID - A VIRTUAL LABORATORY

Performing large experiments to determine biomarkers is challenging, because of heavy computation times and the potentially large number of datasets involved. In addition, novel algorithms need to be optimized by studying the effect of parameter variations on the result. Grid computing provides an environment in which these experiments can be performed. The Virtual Laboratory for e-Science (VL-e) Project (<http://www.vl-e.nl>) is such a project that tried to build a grid-enabled infrastructure for research in The Netherlands. A Proof of Concept environment was built, in which collaborative research can be performed and algorithms can be shared.

## 1.6 THESIS OBJECTIVES

The primary goal of the research in this thesis is to build novel methods aiming at better quantification of pathology in DW-MRI.

We wish to simultaneously model and identify multiple white matter biomarkers that jointly describe a complex disease such as schizophrenia. Due to high inter-brain connectivity, we hypothesize that it is likely that multiple brain regions are involved in schizophrenia. Conventional analyses treat measurements in each voxel or region independently, such that correlated regions may not be found. We perform a discriminant analysis by adopting and extending pattern recognition techniques that can encapsulate these correlations. We iteratively isolate those brain regions that jointly separate a group of patients from healthy controls. This technique, called ‘shaving’, is described in chapter 2. In chapter 3 we opt to ‘prune’ the dataset, such that discrimination is only based on representative prototypes. We explore the use of either a linear or non-linear kernel discrimination function for detecting differences (see chapter 4).

Biomarkers residing in crossing fiber regions will not be identified by the single diffusion tensor, which cannot model multiple fiber orientations within one voxel. Approximately 30% of the white matter was accounted for having more than one fiber orientation [41]. More complex models are needed for adequately modeling tissue in such regions, such that possible differences between patients and controls can be detected. We propose a new method for creating experimental phantom data of fiber crossings, by mixing the DWI-signals from high FA-regions in chapter 5. This method can be used for evaluating more complex diffusion models. Such a model can only be estimated in High Angular Resolution Diffusion Imaging (HARDI)-data, requiring longer scanning times than clinically available. We are able to estimate a dual tensor in non-HARDI data, by creating a dual tensor atlas from multiple coregistered non-HARDI datasets (see chapter 6). An

optimization framework for the acquisition parameters of HARDI-data and estimating a proper dual tensor model is described in chapter 7.

For particular pathologies, a specific hypothesis may exist stating that only certain white matter tracts are involved. Rather than employing the voxels as coordinate frame for statistics, the fiber tracts may then be used as a frame of reference. To find inter-subject correspondence between tracts, we introduce a new method for non-rigid matching of white matter fiber tracts (chapter 8). Then we perform tract-based comparison of DTI-indices between groups, based on a representation that is estimated while matching fiber tracts. Spatial consistency is taken to reflect a significant difference between groups (chapter 9).

Two comparative studies are included in this thesis. In chapter 10, we determine if childhood medulloblastoma and acute lymphoblastic leukaemia survivors have decreased chemotherapy induced white matter fractional anisotropy (WMFA) and if WMFA is related to the speed of processing and motor speed. In chapter 11, we study timing and level of onset of involvement of the upper motor neuron in various phenotypes of motor neuron disease by means of a longitudinal fiber tracking study.

High throughput analysis of our data is realized within a grid environment (chapter 13). Chapter 13 explores the potential of grid computing to facilitate the analysis of DW-MRI data.

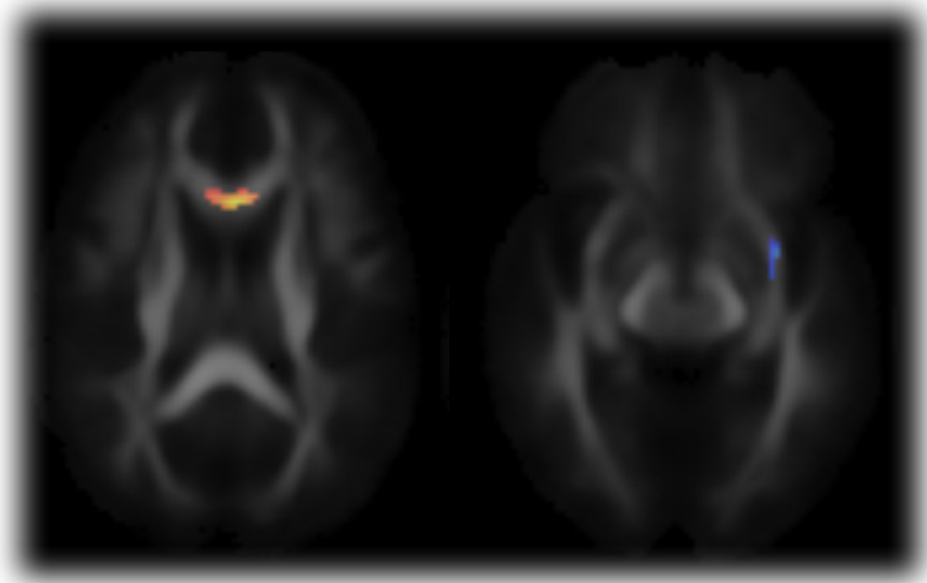






# DISCRIMINANT ANALYSIS

# I





## SHAVING DIFFUSION TENSOR IMAGES IN DISCRIMINANT ANALYSIS: A STUDY INTO SCHIZOPHRENIA

---

A technique called ‘shaving’ is introduced to automatically extract the combination of relevant image regions in a comparative study. No hypothesis is needed, as in conventional pre-defined or expert selected Region Of Interest (ROI)-analysis. In contrast to traditional Voxel Based Analysis (VBA), correlations within the data can be modeled using Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). A study into schizophrenia using Diffusion Tensor Imaging (DTI) serves as an application. Conventional VBA found a decreased Fractional Anisotropy (FA) in a part of the genu of the corpus callosum and an increased FA in larger parts of white matter. The proposed method reproduced the decrease in FA in the corpus callosum and found an increase in the posterior limb of the internal capsule and uncinate fasciculus. A correlation between the decrease in the corpus callosum and the increase in the uncinate fasciculus was demonstrated.

Published as:

M.W.A. Caan, K.A. Vermeer, L.J. van Vliet, C.B.L.M. Majoie, B.D. Peters, G.J. den Heeten and F.M. Vos, *Shaving Diffusion Tensor Images in Discriminant Analysis: a study into Schizophrenia*, *Medical Image Analysis*, 10 (5), 1-9, 2006 [66].

## 2.1 INTRODUCTION

Over the past few years, Diffusion Tensor Imaging (DTI) [33, 34] has provided important insights in the structure of the brain. DTI measures the amount and direction of the diffusion of water. In white matter, the diffusion of water is anisotropic, parallel to the nerve direction, whereas in gray matter the diffusion is more isotropic. The diffusion is often characterized by a rank-2 tensor, from which scalar measures can be derived such as the Fractional Anisotropy (FA) and Apparent Diffusion Coefficient (ADC).

A still growing number of studies analyze DTI to determine changes in brain structure in schizophrenia. Schizophrenia is a cognitive disorder occurring in about 1% of the world population. The first symptoms usually occur at early adolescence, with signs of abnormal social behavior and hallucinations. Studying DTI-data of schizophrenics and controls is challenging, since there is no clear hypothesis specifying where differences are to be expected. What is generally assumed however, is that schizophrenia affects white matter structure in the brain [224].

Typically, schizophrenia is studied by way of a Region Of Interest (ROI)-analysis or Voxel-Based Analysis (VBA) [147]. In the former method, mean values of pre-defined or expert selected regions in a patient and control group are compared. The latter method consists of a comparison per individual voxel. The main findings reported are a decreased FA in the corpus callosum [6, 21, 106, 132] and cingulum [21, 160, 229, 253], but changes in various other white matter regions have been reported as well [147]. Clearly, the regions in the brain are highly interacting with each other, such that a correlation between the reported regions with a decreased FA might be expected. This correlation is modeled neither by ROI-analysis nor VBA, because they consider the regions or voxels independently and thus analyze them separately. Another drawback of ROI-analysis is that expert knowledge, needed in annotating the ROIs, is subject to variation. A voxel-based analysis method is therefore preferred, although it is extremely sensitive to a misregistration of the data, a problem that gets worse with increasing resolution of modern scanners. A voxel-wise comparison of the data is then unjustified.

Despite of the mentioned drawbacks, doing ROI-analysis or VBA is still common practice in clinical studies. The analysis would be greatly strengthened by a new model that combines the advantages of both previous methods. As in VBA, no prior knowledge on correlating regions should be required. The result would be regions, containing multiple voxels related to brain regions, showing differences between the populations, comparable to a ROI-analysis. Such a model may have more discriminative power and better describe the underlying process of the disease.

In this paper, a machine learning framework will be presented that meets the previously mentioned requirements. Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are used to mutually weight the voxels, such that correlation within the data can be modelled. A small study using PCA/LDA has been done earlier [266]. We propose a technique called 'shaving' [125] to automatically extract the set of voxels

separating patients from controls. By seeking a set of voxels, spatial correlation is incorporated in shaving. No expert knowledge is introduced in the whole process. A study into schizophrenia serves as an application; the proposed framework can potentially be applied to many other comparative studies.

## 2.2 MATERIALS AND METHODS

### 2.2.1 *Subjects*

Male patients admitted to the Adolescent-clinic of the Academic Medical Center in Amsterdam, The Netherlands and between 18 and 28 years of age were consecutively included. Patients had a clinical diagnosis of recent-onset schizophrenia or a related disorder according to DSM-IV<sup>1</sup> criteria [16]. All patients received antipsychotic medication.

Exclusion criteria were: history of a demonstrable neurological or endocrine disease, history of a head trauma with loss of consciousness for more than 15 minutes, mental retardation, gross brain abnormalities on conventional MRI and substance abuse within one week of MRI-acquisition. Additional exclusion criteria for controls were: lifetime diagnosis of substance abuse and a personal or family history of a major psychiatric illness such as schizophrenia. Controls were matched to patients for gender, age, educational level and handedness. This study was approved by the local medical and ethical committee. After complete description of the study to the subjects, written consent was obtained.

After the clinical condition of the patients had stabilized, patients were interviewed by a research staff member who was not involved in the treatment. Date of the first psychotic episode and educational level were assessed. The start of the first psychotic episode was defined as the moment when patients experienced hallucinations, delusions or disorganized behavior/speech during most part of the day, during at least one week. Handedness was determined using the Annett Handedness Questionnaire [19].

### 2.2.2 *Data acquisition*

MR diffusion tensor imaging was performed on a 1.5T Siemens Visions (VB33E, Siemens, Erlangen, Germany). A spin-echo EPI sequence with an extra  $180^\circ$  pulse and balanced diffusion sensitizing gradients was used to minimize artifacts induced by eddy currents [213]. Other imaging parameters were: diffusion weighting  $b = 1000$  s/mm<sup>2</sup>, voxel-size  $2 \times 2 \times 6.5$  mm, TE (echo time) 109 ms, six icosahedric diffusion directions [9].

### 2.2.3 *Data preprocessing*

The diffusion is described by a  $3 \times 3$  symmetric tensor  $\mathbf{D}$ , from which eigenvectors  $e_i$  and eigenvalues  $\lambda_i$  can be calculated. The diffusion tensor can be visualized by an

<sup>1</sup> Diagnostic and Statistical Manual of Mental Disorders, fourth edition.

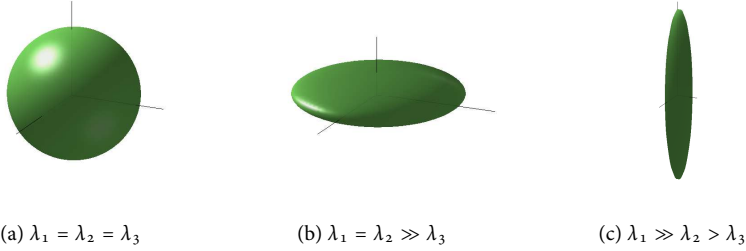


Figure 2.1: Three tensor representations with FA-values 0,  $\sim \frac{1}{\sqrt{2}}$  and  $\sim 1$  respectively.

ellipsoid. The eigenvectors are the axes of the ellipsoid, the eigenvalues determine the shape and size of the ellipsoid. The Fractional Anisotropy FA [32] is a scalar measure derived from the tensor and written as

$$FA = \frac{\sqrt{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}}{\sqrt{2}\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}, \quad (2.1)$$

which ranges from 0 (isotropic) to 1 (anisotropic). Tensor representations for some FA-values are displayed in figure 2.1. The FA is commonly used in brain studies in which changes in brain structure are expected [147]. This is based on the assumption that a reduction of FA in white matter corresponds to a reduction in white matter integrity.

The diffusion weighted images (DWIs) were isotropically resampled to  $2 \times 2 \times 2$  mm, allowing sub-voxel translations in the  $z$ -plane in registration. Moreover, the DWIs were smoothed with a Gaussian kernel of 2 mm, thereby reducing the occurrence of high FA due to noise and limiting the negative side-effects of possible misregistration. When isotropically smoothing DWIs, there is a risk of mixing up intensities related to bundles with different orientations. However, considering the relatively low spatial resolution of the acquisitions, this effect will be negligible.

In a voxel-based analysis, a proper registration is not only essential but also a challenging task, due to large variability in inter-subject anatomy. Solely an affine transformation will not map the volumes onto the same frame of reference. Therefore, a non-rigid registration using the demons algorithm [232] was added to an affine transformation using 12 degrees of freedom. The registration was done directly on the FA-images in 3D-Slicer [4], using the FA computed from an DTI-atlas [252] as target. In order to avoid local minima in the non-rigid registration, an image pyramid was built with subsampled copies of the data. Registration was performed at each level, starting at the top and moving one level down after convergence of the demons algorithm. The deformation field, a 3D-vector field, was regularized by smoothing using a Gaussian kernel with  $\sigma = 1$ .

After registration, a threshold of 0.25 was applied to the FA, restricting the analysis to white matter [11], which is expected to be affected by schizophrenia [224].

#### 2.2.4 Algorithm design

A new algorithm will now be described, aiming to automatically detect regions that discriminate patients from controls. This is done by ‘shaving’ the voxel space, based on the mapping computed by PCA/LDA. Classification is done during shaving to validate if the populations can still be discriminated based on the information stored in the preserved regions.

##### Mapping computation

Let us now briefly review PCA/LDA-analysis. The patient and control group are denoted  $\omega_p$  and  $\omega_c$ , the so-called *classes*. Let  $m_p$  and  $m_c$  be the number of patients and controls, summing to  $m$ , the size of the cohort. Each image can be regarded as a point in an  $n$ -dimensional space  $\mathbb{R}^n$ . The  $n$  voxels to be studied are reshaped to  $n$ -dimensional vectors  $\mathbf{x}_i$  ( $i = 1 \dots m$ ), where the ordering is chosen arbitrarily.

In general, the cohort size is much smaller than the number of voxels ( $m \ll n$ ), making the problem of discriminating data in this space highly ill-posed. This is called the *small sample size problem* [113], for which a unique solution does not exist. PCA/LDA solves this problem in a two-step way of dimensionality reduction,  $\mathbb{R}^n \xrightarrow{PCA} \mathbb{R}^r \xrightarrow{LDA} \mathbb{R}$  [131], such that a one-dimensional  $x_i^*$  is obtained:

$$x_i^* = \mathbf{a}^T (\mathbf{x}_i - \bar{\mathbf{x}}), \quad (2.2)$$

where  $\bar{\mathbf{x}}$  is the mean over all  $\mathbf{x}_i$  and  $\mathbf{a}$  is written as

$$\mathbf{a} = \mathbf{P}\mathbf{q}. \quad (2.3)$$

The matrix  $\mathbf{P}_{n \times r}$  maps a point from  $\mathbb{R}^n$  to  $\mathbb{R}^r$ . It contains only non-singular eigenvectors or principal components of the total scatter matrix  $\mathbf{S}_t = \sum (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T$  [240]. The order of eigenvectors is such that the corresponding eigenvalues are monotonically increasing. If  $r = m - 1$ , all variance in the data is retained; lowering  $r$  yields a loss of information. A proper choice of  $r$  will be discussed later on. Projecting  $(\mathbf{x}_i - \bar{\mathbf{x}})$  onto  $\mathbf{P}$  results in vectors  $\mathbf{x}'_i = \mathbf{P}^T(\mathbf{x}_i - \bar{\mathbf{x}})$  of length  $r$ .

The vector  $\mathbf{q}$  of length  $r$  is found by LDA, aiming to discriminate patients from controls in a one-dimensional space  $\mathbb{R}$ .  $\mathbf{q}$  is computed in the low-dimensional subspace  $\mathbb{R}^r$  and written as [99]

$$\mathbf{q} = \mathbf{S}_w^{-1} \mathbf{S}_b, \quad (2.4)$$

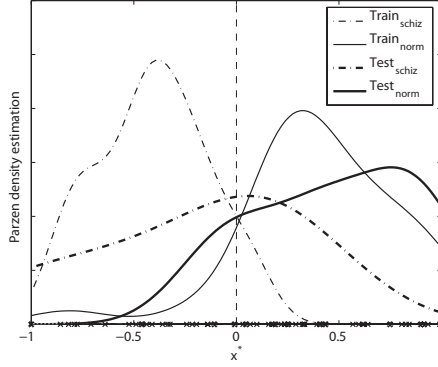


Figure 2.2:  $x^*$  with Parzen density estimation plotted for a training and testing set. The vertical dashed line indicates the decision boundary.

with the class-within scatter matrix  $\mathbf{S}_w = \frac{1}{m} \sum_{j=1}^2 \sum_{i=1}^{m_j} (\mathbf{x}_i' - \bar{\mathbf{x}}_j')(\mathbf{x}_i' - \bar{\mathbf{x}}_j')^T$  and the class-between scatter matrix  $\mathbf{S}_b = \frac{1}{m} \sum_{j=1}^2 (\bar{\mathbf{x}}_j' - \bar{\mathbf{x}}')(\bar{\mathbf{x}}_j' - \bar{\mathbf{x}}')^T$  for the two-class case and the mean of class  $j$  notated as  $\bar{\mathbf{x}}_j'$ .

### Classification

Now the data can be projected onto a one-dimensional  $x^*$  using equation 2.2. In the remainder of this paper, it is assumed that  $\bar{x}_{\omega_p}^* < \bar{x}_{\omega_c}^*$ , i.e. the average value of the patients is smaller than the controls ( $x^*$  is mirrored if this is not the case). Classification with a simple classifier, the nearest-mean classifier, is done on  $x^*$ . The classification error is computed by means of cross-validation [99].

Five-fold cross-validation randomly divides the data into five groups of approximately equal size and with the same proportion of the classes. Now four groups are used as training set, for computing  $a$ , and one group is used as testing set, which is mapped onto  $a$  for computing the classification error. This is done five times, each time rotating the data in the training and testing sets, resulting in five errors computed on the individual groups, which are averaged. The cross-validation is repeated ten times, with different composition of the cross-validation groups, after which a mean error is computed. The 95%-confidence region of the error depends on the expected error and on the size of the cohort [99]. If the upper boundary is below 50%, a significant difference between patients and controls is concluded. An estimate of the stability of the classification error is given by the standard deviation of the error over the ten repeated cross-validations. As an illustration,  $x^*$  is plotted for one training and testing set in the cross-validation, with a corresponding Parzen density estimation, in figure 2.2.



Cross-validation is used to choose  $r$ , the dimensionality of the subspace where LDA is performed. The first principal components are expected to describe most variance in the data, whereas the last only describe noise. The error is therefore expected to lower until a certain minimum, which defines the value of  $r$ . The classification error is computed ‘forward’, starting with the first and using an increasing number of principal components. A second experiment, in which the classification error is computed ‘backward’, verifies if the last components indeed describe noise. Starting with the trailing principal component, more and more leading principal components are added until at the end all are used. The error should in this case remain close to 50% until the leading principal components are added.

### *Visualization*

It is clinically important to know which regions of the brain manifest a significant difference between the populations (as indicated by the FA). The mapping  $a$  computed by the PCA/LDA-algorithm will therefore be analyzed to identify those regions.  $a$  describes how the voxels are mutually weighted for discriminating patients from controls; in other words, it reflects how voxels are combined to separate the populations. The regions where  $|a|$  is high contribute most to the separation in  $x^*$  (notice that  $\bar{x}^* = 0$  due to the subtraction of the mean from the data).

Differences between patients and controls are expected to emerge in regions related to known anatomical brain regions. ‘Shaving’ is proposed as a way to automatically extract the set of voxels separating patients from controls. By seeking a set of discriminating voxels, spatial correlation is incorporated in shaving. This technique was originally applied to genes, to identify subsets of genes with similar expression patterns [125]. It iteratively removes voxels with small  $|a|$  and then re-trains the remaining voxels, such that the mutual weighting of the voxels is adapted. A constraint is added to this procedure to remove in each iteration small objects from  $a$ .

The shaving procedure is as follows:

1. Consider all  $n$  voxels,  $V = \{1, 2, \dots, n\}$
2. Compute  $\mathbf{a}$  for all voxels in  $V$  in  $m$  images  $\{\mathbf{x}_i(j) \mid i = 1 \dots m, j \in V\}$
3. Let  $P = \{j \mid a(j) > 0\}$  and  $N = \{j \mid a(j) \leq 0\}$
4. Remove objects smaller than  $s_{min}$  voxels from  $P$  and  $N$
5. Let  $V = P \cup N$
6. Discard 25% of the voxels with the smallest value of  $|\mathbf{a}(V)|$  from set  $V$
7. Repeat from step 2 until the desired percentage of voxels is left

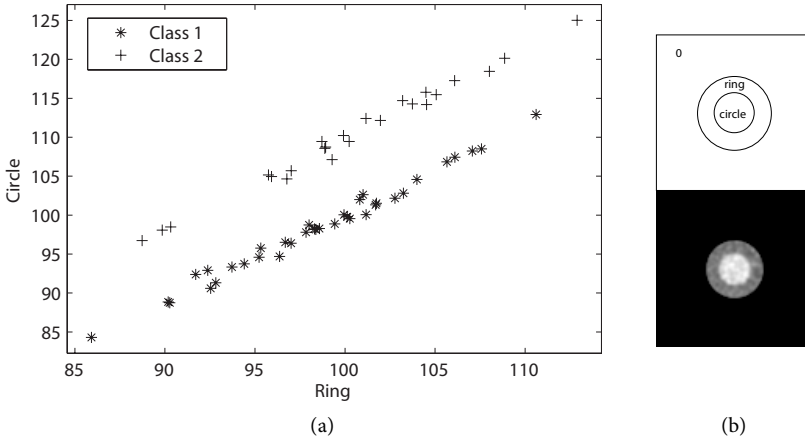


Figure 2.3: (a) Scatterplot of the mean intensity values in the simulated data, generated with class means  $(100\ 100)$  and  $(100\ 110)$ , and equal covariance matrix  $\begin{pmatrix} 40 & 35 \\ 35 & 40 \end{pmatrix}$ . (b) The design used to generate the simulated data (top) and an example image of class 2, at intensity interval  $(95\ 125)$  to enhance contrast (bottom).

The stopping criterion is determined by studying the classification error as a function of the fraction of retained voxels. The error is expected to initially decrease due to removal of noise, after which it increases when the discriminating regions are being shaved off. The expected minimum of this error curve is used to determine at which fraction of voxels shaving is terminated. Note that the number of principal components  $r$  is determined in a separate cross-validation, before shaving. As a result of the shaving process, the combination of regions discriminating the two populations is automatically extracted from the data.

### 2.2.5 Algorithm evaluation

The behavior of the proposed shaving algorithm is studied by means of simulated data, containing regions whose mean values are correlated. It will be investigated if  $r$  can properly be chosen, how the classification error behaves during shaving and if the resulting discriminating regions match with the generated ones. For comparison, VBA of this data will be done.

In order to study the reliability of the algorithm, two tests are performed, regarding overtraining and stability. In the first test, patients and controls are randomly divided into two classes, containing the same proportion of the initial populations. If the classification error on this data is significantly lower than 50%, the used classifier is overtrained. In the second test the stability of the algorithm is evaluated. The regions the algorithm comes up with should not change much while adding or removing a subject to or from the cohort. The stability is tested by a jack-knife procedure [99]. Given  $m$  images, the

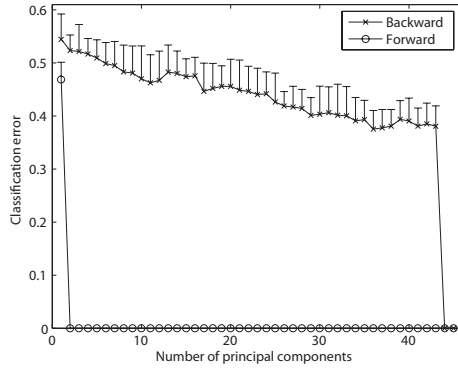


Figure 2.4: Classification error using the simulated data as a function of the number of principal components used, with error bars denoting the standard deviation of the error. See the text for an explanation of the ‘forward’- and ‘backward’-experiment.

classifier is trained  $m$  times, each time removing one of the images from the full set. For each object in  $a$  that is found, the number of occurrences in each result is computed. Objects showing up in more than 90% of the steps are considered to be stable and are displayed as a final result.

The hypothesis of correlation in the data is verified by doing LDA on the mean values of the regions resulting from the shaving algorithm. The extent to which the regions are mutually weighted gives an indication of the existence of correlation.

## 2.3 RESULTS

### 2.3.1 Simulated data

The behavior of the shaving algorithm was studied using simulated data. Two classes containing 36 and 24 images of size  $100 \times 100$  were generated, containing a circle and a ring both of constant intensity. The intensities of the two objects were chosen such that they are correlated, see figure 2.3.

Zero-mean Gaussian white noise with  $\sigma = 10$  was added to the data, such that the SNR was approximately equal to the SNR of DTI [133]. The data were smoothed using a Gaussian kernel with  $\sigma = 1$ .

A parameter sweep was performed for choosing the optimal number of principal components  $r$ . The classification error after cross-validation is plotted in figure 2.4, using up to 45 principal components computed on the training set. Both in the ‘forward’- and ‘backward’-experiment, the error drops to zero when the second principal component is

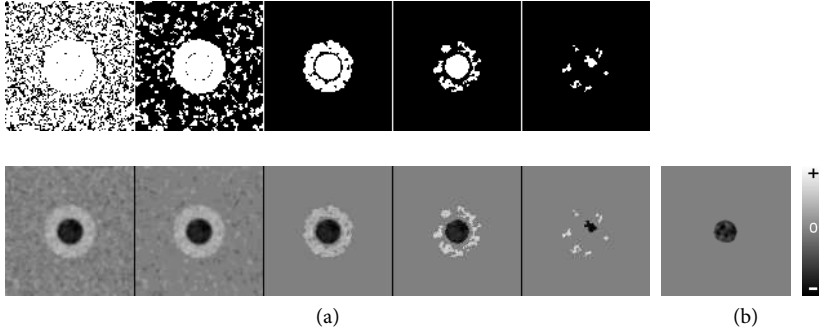


Figure 2.5: (a) Illustration of the ‘shaving’-procedure: the remaining voxels in white (top) and the corresponding mapping  $a$  (bottom) of the simulated data after shaving to 75%, 32%, 10%, 6% and 1.3% respectively. (b) Two-sided conventional VBA ( $P < 0.001$ ) of the simulated data.

added. This component is perpendicular to the direction of most variation; it can be concluded from figure 2.3(a) that best discrimination is achieved in this direction.

Figure 2.5(a) shows the results of the shaving procedure. In the top image, the remaining voxels during shaving are shown, the bottom image displays the corresponding mapping  $a$ . The result of a VBA using Statistical Parametric Mapping (SPM) [111] with  $p < 0.001$  and a minimal object size of 10 voxels is given in figure 2.5(b). What catches the eye, is that VBA only found the mean difference in the circle and not in the ring, because the data was generated with different class means in the former and equal class means in the latter feature. Because the proposed algorithm was able to model correlation in the data, it resulted in a combination of the ring and the circle.

The classification error cannot be used as stopping criterion in this experiment, because the classes were perfectly separable using the subsequent mappings presented in figure 2.5(a). Based on visual inspection, 10% is an optimal fraction of voxels to use, because then background noise has been removed whereas the generated structures remain intact. This example demonstrates the potential power of shaving; its performance on real data will now be discussed.

### 2.3.2 Experimental data

34 male patients were included with an average age of 22.3 years and a standard deviation of 2.6 years. They were compared with 24 healthy controls with an average age of 22.5 years and a standard deviation of 3.2 years. Handedness of patients was (right/left/ambidexter) 85/12/3%, of controls it was 87/13/0%. Educational level of patients was: 29% skilled training, 32% bachelor level and 38% master level; of controls 29%, 29% and 42% respectively. No gross abnormalities could be detected on conventional MR imaging by an expert. Eddy current induced morphing in the phase direction was visually esteemed to be negligible.

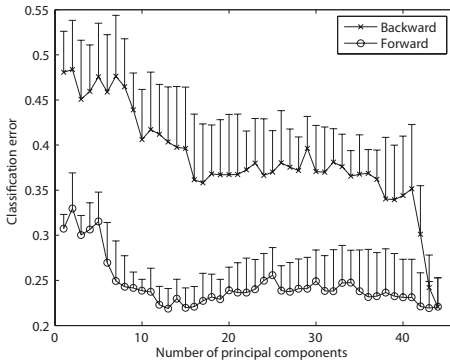


Figure 2.6

Classification error using the FA as a function of the number of principal components used, with error bars denoting the standard deviation of the error.

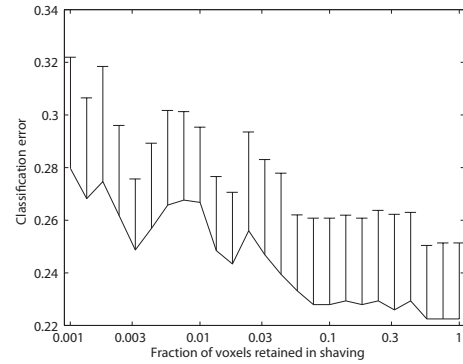


Figure 2.7

Classification error using the FA as a function of the fraction of voxels retained in shaving, with error bars denoting the standard deviation of the error.

In order to apply PCA/LDA and shaving, a threshold value  $s_{min}$  had to be chosen for objects to be removed in the shaving process, as described in section 2.2.4.  $s_{min}$  was set to 50 voxels ( $1.6 \text{ cm}^3$ ), which is a typical size of brain tracts where changes are expected [127]. The number of principal components  $r$  was chosen based on the classification error after cross-validation, as can be seen in figure 2.6. The ‘forward’-experiment makes clear that  $r = 13$  is the optimal number of components, since adding more components does not lower the error and increases the standard deviation in the error. The ‘backward’-experiment, starting with the last and adding leading components, reveals that the trailing principal components are indeed describing noise, because the error is close to 50%. As in the simulated data experiment, the standard deviation in the error decreases when leading principal components are included.

The classification error as a function of the fraction of voxels retained after shaving is shown in figure 2.7. This figure reveals only a slight increase in the error during shaving; shaving is chosen to be terminated at 0.3%. The minimal classification error of 25%, with the upper boundary of the 95%-confidence interval at 35% [99], means a significant difference between patients and controls.

The PCA/LDA mapping **a** before shaving is shown in figure 2.8. The mapping after shaving to 0.3% and with unstable regions discarded can be seen in figure 2.9. These results are interpreted as a decrease in the genu of the corpus callosum and an increase in the right uncinate fasciculus and posterior limb of the internal capsule.

VBA using SPM [111] was applied, keeping voxels with  $p < 0.001$  and a minimal object size of 20 voxels. The result is displayed in figure 2.10. A significantly lower FA can be seen in the genu of the corpus callosum, which disappears when increasing the object size threshold to 50 voxels as used in shaving. A higher FA is observed in the posterior limb of the internal capsule, as well as in several inferior white matter regions.

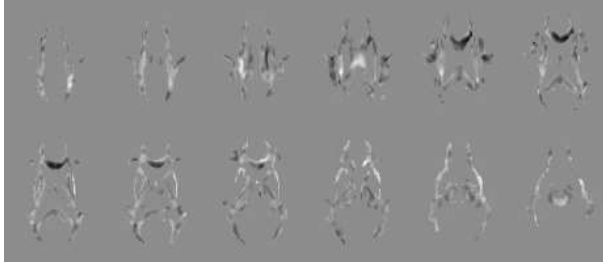


Figure 2.8: The mapping  $\mathbf{a}$  computed using PCA/LDA of FA-images before shaving.  $a$  is zero on the gray background, negative in black and positive in white regions.

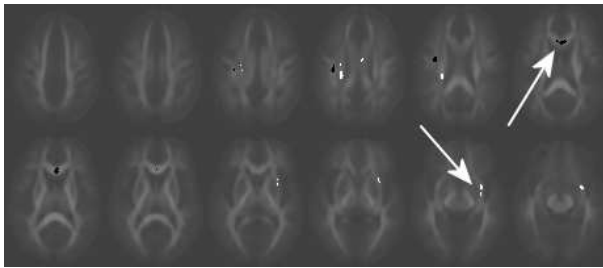


Figure 2.9: Only the stable regions in  $\mathbf{a}$  after shaving to 0.3% of the total number of voxels. White regions indicate an increased FA, black regions a decreased FA. The average FA is shown in gray on the background. The arrows point to the decrease in the corpus callosum and the increase in the uncinate fasciculus.

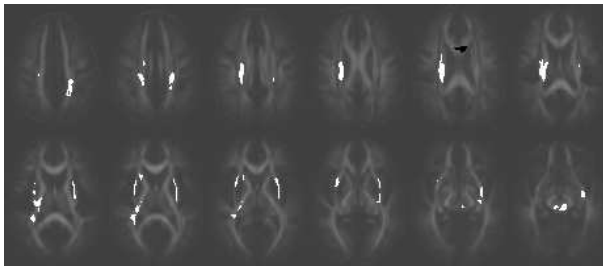


Figure 2.10: Two-sided VBA ( $p < 0.001$ ) of the FA. White regions indicate an increased FA, black regions a decreased FA. The average FA is shown in gray on the background.

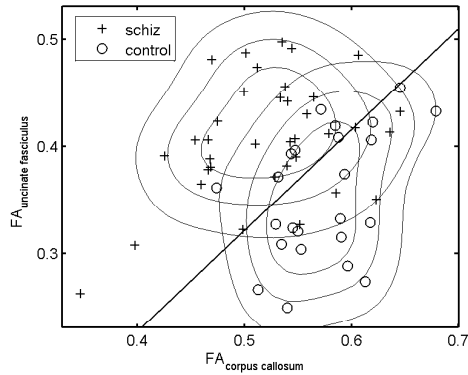


Figure 2.11: Scatter plot of the mean FA-values of the regions in the corpus callosum and uncinate fasciculus found with shaving, with the classification boundary given by LDA.

After that the results were calculated, the algorithm was tested on overtraining. The classification error of randomly composed classes before shaving yielded 62%, lowering to 42% when 0.1% of the voxels are used. The upper boundary of the 95%-confidence regions then equals 55%, such that a significant difference between these classes cannot be concluded.

Finally, LDA was performed on the mean values of two resulting regions of the shaving procedure, situated in the corpus callosum and uncinate fasciculus. These mean values are scatter plotted in figure 2.11, with the classification boundary computed by LDA. The slope of this boundary indicates that the classifier equally weights the regions, indicating a correlation between them. The classification error after cross-validation equals 25%, comparable to the error after shaving.

## 2.4 DISCUSSION AND CONCLUSION

We introduced a new machine learning framework for comparative studies on volumetric data. By 'shaving' the mapping computed by PCA/LDA, a combination of characteristic regions is automatically extracted. In this way, a solution is found to a shortcoming of both ROI-analysis and VBA, namely that they do not take correlation between regions into account.

The method was applied to DTI-data of schizophrenics and controls, which resulted in a classification error of 25% after shaving to 0.3% of the voxels. The discriminating regions correlated with a decrease in FA in the genu of the corpus callosum and an increase in the right uncinate fasciculus and posterior limb of the internal capsule. Conventional VBA found a decrease in FA in the corpus callosum and an increased FA in various white matter regions, including those found using the proposed algorithm, which is

unprecedented [147]. The shaving algorithm localizes the differences more accurately, such that they can be better interpreted. The correlation in the data was demonstrated by LDA of the means of the regions of the corpus callosum and the uncinate fasciculus (see figure 2.11).

The number of principal components is a critical parameter, affecting the behavior of the algorithm. Including too few components will only allow for describing global differences, whereas including too many might result in fitting the model to the noise. Based on a parameter sweep embedded in cross-validation, 13 principal components were used. With only 60 subjects available, LDA was computed in a sparsely sampled space, which can make the algorithm unstable and thus unreliable. The stability was ensured by only keeping the regions showing up in more than 90% of the steps in a jack-knife procedure. The reliability was demonstrated by an insignificant classification error of 42% on two randomly composed classes.

It was not our intention to use the classifier in clinical practice, judging whether a new subject is schizophrenic or not, although it is theoretically possible. Rather, we use the method only to identify significant differences between patients and controls.

PCA is driven by the total variance, such that non-brain regions largely influence the classifier, as they cover approximately 50% of the voxels. The experiment with simulated data has shown that shaving is able to remove these regions before eliminating characteristic regions. Still, a threshold was applied to the experimental data for two reasons. First, only differences in white matter were a priori expected [224]. Second, only the main white matter tracts were observed to be aligned after registration, such that voxel-wise analysis of other brain regions was not allowed.

Not having any knowledge about the type of correlation present in the data, a linear correlation was modeled in the algorithm. A possible extension is using a non-linear (kernel-based) classifier instead of PCA/LDA. This may lead to a lower classification error, if the correlation is indeed better modeled non-linearly. However, there is no standard solution for the problem of how to localize and visualize differences using such a classifier.

The observed decrease in the corpus callosum was found several times before [6, 21, 106, 132], whereas a finding of higher FA is almost unprecedented. The increase in the posterior limb of the internal capsule may be explained by the fact that in this region tracts are crossing [257]. Then the second-order diffusion tensor model does not hold and care must be taken in interpreting changes in FA.

In this paper, the focus was on the methodology of the shaving algorithm. Future work will be on gathering and studying higher-resolution data, in order to validate the findings. Incorporating correlation in analysis is a new step in pathological studies, opening the way to more knowledge of brain diseases in the future.



## PRUNING DATASETS IN DISCRIMINANT ANALYSIS: A DTI STUDY TO SCHIZOPHRENIA

---

A comparative study is commonly performed by means of pre-defined or expert selected Region of Interest (ROI)-analysis or Voxel Based Analysis (VBA). In contrast to these methods, correlations within the data can be modeled by using Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). The mapping computed by PCA/LDA is displayed to identify the discriminative regions.

A technique called 'pruning' is introduced to iteratively discard misclassified subjects from the cohort. These subjects reside in the region in feature space where the classes are overlapping. As the exact cause of this overlapping is unknown, it is preferable to base the mapping merely on representative prototypes, residing in the non-overlapping parts of the feature space.

After pruning the PCA/LDA mapping, a more pronounced decrease in FA in larger parts of the corpus callosum was observed, compared to conventional VBA.

Published as:

M.W.A. Caan, K.A. Vermeer, L.J. van Vliet, C.A. Grimbergen, F.M. Vos, *Pruning datasets in discriminant analysis: a DTI study to Schizophrenia*, in: Proc. ISBI, 2007, 1340-1344 [62].

### 3.1 INTRODUCTION

Over the past few years, Diffusion Tensor Imaging (DTI) [34] has provided important insights in the structure of the brain. DTI measures the amount and direction of the diffusion of water. In white matter, the diffusion of water is anisotropic, parallel to the nerve direction, whereas in gray matter the diffusion is more isotropic. The diffusion is often characterized by a rank-2 tensor, from which scalar measures can be derived such as the Fractional Anisotropy (FA) and Apparent Diffusion Coefficient (ADC). A still growing number of studies analyzes DTI to determine changes in brain structure in schizophrenia. Studying images of schizophrenics is challenging, since there is no clear hypothesis specifying where differences are to be expected. What is generally assumed however, is that schizophrenia affects white matter structure in the brain [224].

Typically, schizophrenia is studied by way of a Region Of Interest (ROI)-analysis or Voxel-Based Analysis (VBA) [147]. The main findings are a decreased FA in the corpus callosum and cingulum, but changes in various other white matter regions have been reported as well [147]. Clearly, the regions in the brain are highly interacting with each other, such that a correlation between the reported decreases is to be expected. This correlation is modeled neither by ROI-analysis nor VBA, because they consider the regions or voxels independently and thus analyze them separately. Another drawback of ROI-analysis is that expert knowledge is needed in annotating the ROIs, which is subject to variation. By using Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), a combination of voxels is sought that is mutually weighted to discriminate patients from controls. Two small studies using PCA/LDA have been done earlier [233, 266].

It is clinically important to know in which regions of the brain a significant difference in structure (as indicated by the FA) manifests itself. The mapping computed by the PCA/LDA-algorithm will be analyzed to identify those regions. The data in the feature space might be partially overlapping, due to a modeling error or an a priori non-perfect separation between the classes. It is therefore chosen to minimize the influence of the subjects in the overlapping region on the mapping and base the mapping merely on representative prototypes, i.e. the non-overlapping parts of the data in feature space. This is achieved by ‘pruning’ the dataset, which iteratively discards the subjects in the overlapping region from the cohort. The algorithm is thus expected to better describe the way the studied disease is affecting the human brain. A study into schizophrenia serves as an application; the proposed framework can potentially be applied to many other comparative studies that have already been done before.

### 3.2 METHOD

#### 3.2.1 Mapping computation

Let us first briefly review PCA/LDA-analysis. The patient and control group are denoted  $\omega_p$  and  $\omega_c$ , the so-called *classes*. Let  $m_p$  and  $m_c$  be the number of patients and controls,

summing to  $m$ , the size of the cohort. Each image can be regarded as a point in an  $n$ -dimensional space  $\mathbb{R}^n$ . The  $n$  voxels to be studied are reshaped to  $n$ -dimensional vectors  $\mathbf{x}_i$  ( $i = 1, \dots, m$ ), where the ordering can be chosen arbitrarily.

In general, the cohort size is much smaller than the number of voxels ( $m \ll n$ ), making the problem of discriminating data in this space highly ill-posed. This is called the *small sample size problem* [113], for which a unique solution does not exist. PCA/LDA solves this problem in a two-step way of dimensionality reduction,  $\mathbb{R}^n \xrightarrow{PCA} \mathbb{R}^r \xrightarrow{LDA} \mathbb{R}$  [131], such that a one-dimensional  $x_i^*$  is obtained:

$$x_i^* = \mathbf{a}^T (\mathbf{x}_i - \bar{\mathbf{x}}), \quad (3.1)$$

where  $\bar{\mathbf{x}}$  is the mean over all  $x_i$  and  $\mathbf{a}$  is written as

$$\mathbf{a} = \mathbf{P}\mathbf{q}. \quad (3.2)$$

The matrix  $\mathbf{P}_{n \times r}$  maps a point from  $\mathbb{R}^n$  to  $\mathbb{R}^r$ . It contains only non-singular eigenvectors or principal components of the total scatter matrix  $\mathbf{S}_t = \sum (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T$  [240]. The order of eigenvectors is such that the corresponding eigenvalues are monotonically increasing. If  $r = m - 1$ , all variance in the data is retained; lowering  $r$  yields a loss of information. A proper choice of  $r$  will be discussed later on. Projecting  $(\mathbf{x}_i - \bar{\mathbf{x}})$  onto  $\mathbf{P}$  results in vectors  $\mathbf{x}'_i = \mathbf{P}^T(\mathbf{x}_i - \bar{\mathbf{x}})$  of length  $r$ .

The vector  $\mathbf{q}$  of size  $r$  is found by LDA, aiming to discriminate patients from controls in a one-dimensional space  $\mathbb{R}$ .  $\mathbf{q}$  is computed on  $\mathbf{x}'$  in the low-dimensional subspace  $\mathbb{R}^r$  and written as [99]

$$\mathbf{q} = \mathbf{S}_w^{-1} \mathbf{S}_b, \quad (3.3)$$

with the class-within and -between scatter matrices  $\mathbf{S}_w$  and  $\mathbf{S}_b$

$$\mathbf{S}_w = \frac{1}{m} \sum_{j=1}^2 \sum_{i=1}^{m_j} (\mathbf{x}'_i - \bar{\mathbf{x}}'_j) (\mathbf{x}'_i - \bar{\mathbf{x}}'_j)^T \quad (3.4)$$

$$\mathbf{S}_b = \frac{1}{m} \sum_{j=1}^2 (\bar{\mathbf{x}}'_j - \bar{\mathbf{x}}') (\bar{\mathbf{x}}'_j - \bar{\mathbf{x}}')^T \quad (3.5)$$

for the two-class case and the mean of class  $j$  notated as  $\bar{\mathbf{x}}'_j$ .

### 3.2.2 Classification

Now the data can be projected to  $x^*$  using equation 3.1. In further investigation, it is assumed that  $\bar{x}_{\omega_p}^* < \bar{x}_{\omega_c}^*$ , i.e. the average value of the patients is smaller than the controls ( $x^*$  is mirrored if this is not the case). Classification with a simple classifier, the nearest-

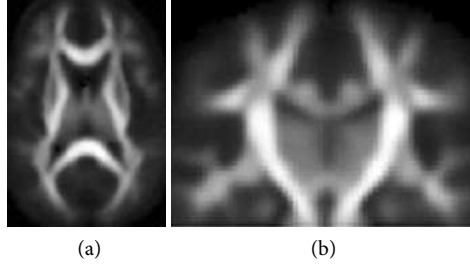


Figure 3.1: (a) Axial and (b) coronal slice of the average FA-volume after registration.

mean classifier, is done on  $x^*$ . The classification error is computed by means of ten-times repeated five-fold cross-validation [99]. The 95%-confidence region of the error depends on the expected error and on the size of the cohort [99]. If the upper boundary is below 50%, a significant difference between patients and controls is found. An estimate of the stability of the classification error is given by the standard deviation of the error over the ten repeated cross-validations.

Cross-validation is also used to choose  $r$ , the dimensionality of the subspace. The classification error is computed ‘forward’, starting with the first and using an increasing number of principal components. It is expected the data resides in a low-dimensional subspace of dimensionality  $r < m - 1$ , and that the last principal components only describe noise. The error is therefore expected to lower until a certain minimum, which defines the value of  $r$ . In a second experiment, in which the classification error is computed ‘backward’, it is verified if the last components indeed describe noise. Starting with the trailing principal component, more and more leading principal components are added until at the end all are used. It is expected that the error in this case will remain close to 50% until the leading principal components are added.

In order to study the reliability of the algorithm, patients and controls are randomly divided into two classes, gathering the same proportion of the initial populations. If the classification error on this data is significantly lower than 50%, the algorithm is overtraining the data.

### 3.2.3 Pruning

It is clinically important to know in which regions of the brain a significant difference in structure (as indicated by the FA) manifests itself. The mapping  $\mathbf{a}$  computed by the PCA/LDA-algorithm will be analyzed to identify those regions.  $\mathbf{a}$  describes how the voxels are mutually weighted for discriminating patients from controls; it reports the combination of regions that separates the populations.

The precision by which  $\mathbf{a}$  can be estimated depends on the amount of available data. After doing PCA/LDA and projecting the data onto  $\mathbb{R}$ , the classes are expected to be

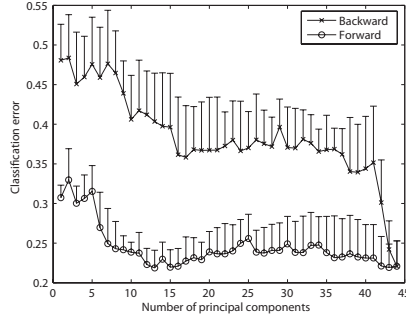


Figure 3.2: Classification error using the FA as a function of the used principal components, with errorbars denoting the standard deviation of the error. The results of the ‘forward’- and ‘backward’-experiment, using a set of leading and trailing components respectively, are only comparable when they end using all components.

compact and partially overlapping. This overlapping originates from a modeling error or from an a priori non-perfect separation between the classes in  $\mathbb{R}^r$ . It is unknown which of these effects is causing the overlap in this study. Still, a report of discriminative regions in the data is needed. It is therefore chosen to minimize the influence of the subjects in the overlapping region on the mapping. Consequently, the mapping is merely based on representative prototypes, i.e. the non-overlapping parts of  $\mathbb{R}$ .

We propose to enhance the mapping by ‘pruning’ the dataset. Pruning iteratively discards the subjects in the overlapping region from the cohort. This is realized by tracking the misclassifications of subjects in cross-validation. Subjects that are misclassified in at least nine out of ten times are removed from the dataset. The remaining cohort is classified again, which is repeated until no more subjects are removed.

#### 3.2.4 Thresholding

The regions where  $|\mathbf{a}|$  is highest contribute most to the separation in  $x^*$ , since  $\overline{x^*} = 0$  due to the extraction of the mean from the data.  $|\mathbf{a}|$  is therefore thresholded, such that only the discriminating regions remain. The threshold value  $e$  is determined on the non-pruned dataset by doing cross-validation using the remaining voxels with an increasing threshold value. The classification error is expected to initially decrease due to removal of noise, after which it increases until the upper boundary of the 95%-confidence region of the error reaches 50%. The corresponding threshold value will be used when visualizing  $\mathbf{a}$ .

#### 3.2.5 Subjects

Male patients admitted to the Adolescent-clinic of the Academic Medical Center in Amsterdam, the Netherlands and between 18 and 28 years of age were consecutively included. Patients had a clinical diagnosis of recent-onset schizophrenia or a related dis-

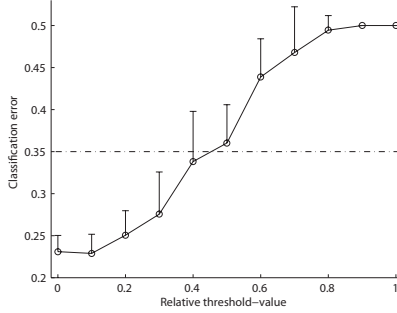


Figure 3.3: Classification error as a function of the threshold on  $\mathbf{a}$ , with the dash-dotted line denoting the boundary of significant difference.

order according to DSM-IV criteria [16]. All patients received antipsychotic medication. Handedness was determined using the Annett Handedness Questionnaire [19].

Exclusion criteria were: history of a demonstrable neurological or endocrine disease, history of a head trauma with loss of consciousness for more than 15 minutes, mental retardation, gross brain abnormalities on conventional MRI and substance abuse within one week of MRI-acquisition. Additional exclusion criteria for controls were: lifetime diagnosis of substance abuse and a personal or family history of a major psychiatric illness such as schizophrenia. Controls were matched to patients for gender, age, educational level and handedness. This study was approved by the local medical and ethical committee. After complete description of the study to the subjects, written consent was obtained.

### 3.2.6 Data acquisition and preprocessing

Magnetic Resonance DTI of the subjects was performed on a 1.5T Siemens Visions (VB33E, Siemens, Erlangen, Germany). A spin-echo EPI sequence with an extra  $180^\circ$  pulse and balanced diffusion sensitizing gradients was used to minimize artifacts induced by eddy currents [213]. Other imaging parameters were: diffusion weighting  $b = 1000$  s/mm<sup>2</sup>, voxel-size  $2 \times 2 \times 6.5$  mm, TE (echo time) 109 ms, six icosahedric diffusion directions [9].

The diffusion is described by a  $3 \times 3$  symmetric tensor  $\mathbf{D}$ , from which eigenvectors  $v_{1...3}$  and eigenvalues  $\lambda_1 \leq \lambda_2 \leq \lambda_3$  can be calculated. The Fractional Anisotropy FA [32] is a scalar measure derived from the tensor and written as

$$FA = \frac{\sqrt{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_1 - \lambda_3)^2}}{\sqrt{2}\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}. \tag{3.6}$$

The diffusion weighted images (DWIs) were isotropically resampled, allowing sub-voxel translations in the  $z$ -plane in registration. Moreover, the DWIs were smoothed with

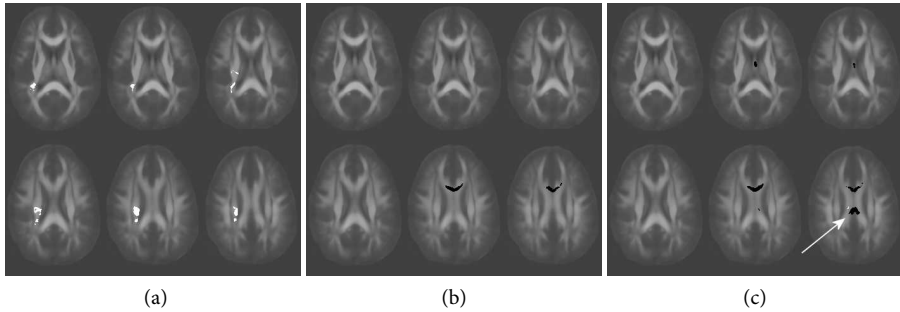


Figure 3.4: (a) Two-sided conventional VBA ( $p < 0.001$ ), (b) PCA/LDA-mapping **a** of original and (c) pruned dataset, thresholded at 0.4. White regions indicate an increased FA for schizophrenics, black regions a decreased FA, the average FA is shown in gray. The arrows point at a decrease in the corpus callosum.

a Gaussian kernel of 2 mm, thereby reducing the occurrence of high FA due to noise and limiting the negative side-effects of possible misregistration. An affine registration and a non-rigid registration using demons with a multi-resolution approach [232] of the FA-images were performed in 3D-Slicer ([www.slicer.org](http://www.slicer.org)), using the FA computed from a DTI-atlas [252] as target. The average FA-volume after registration is displayed in figure 3.1.

### 3.3 RESULTS

In this study, 34 male patients were included with an average age of 22.3 years and a standard deviation of 2.6 years. They were compared with 24 healthy controls with an average age of 22.5 years and a standard deviation of 3.2 years. Handedness of patients was (right/left/ambidexter) 85/12/3%, of controls it was 87/13/0%. Educational level of patients was: 29% skilled training, 32% bachelor level and 38% master level; of controls 29%, 29% and 42% respectively. No gross abnormalities could be detected on conventional MR imaging by an expert. Eddy current induced morphing in the phase direction was visually esteemed to be negligible.

The number of principal components  $r$  is chosen based on the classification error after cross-validation, as can be seen in figure 3.2. The ‘forward’-experiment makes clear that  $r = 13$  is the optimal number of components, since adding more components does not lower the error. The error of 22%, with the upper boundary of the 95%-confidence interval at 37%, means a significant difference between the classes. The ‘backward’-experiment, starting with the last and adding leading components, reveals that the trailing principal components are indeed describing noise, because the error is close to 50%. Also note the higher standard deviation in the error in the latter experiment. Classification of randomly composed classes resulted in an error of 42%, a non-significant difference.

The classification error as function of the threshold on  $|a|$  is displayed in figure 3.3. While thresholding, a minimal object size of 50 voxels ( $1.6 \text{ cm}^3$ ) was used, which is a typical size of brain tracts where changes are expected [127]. Based on the 95%-confidence region, given a cohort size of 60 subjects, a classification error below 35% is considered significant. A threshold of  $e = 0.4$  on  $|a|$  is therefore chosen.

Pruning was applied to the cohort, resulting in a removal of 6 schizophrenics and 2 controls after 3 iterations. All discarded subjects were misclassified using the mapping computed with the pruned dataset. The result of a VBA is displayed in figure 3.4(a). The PCA/LDA mappings computed on the total and pruned dataset are shown in figure 3.4(b) and (c) respectively. The mapping of the pruned dataset shows a larger area of decreased FA in the corpus callosum (annotated with the white arrow), compared to the mapping of the total dataset. If the significance level in VBA is increased, a decreased FA in the corpus callosum as reported by PCA/LDA is appearing.

### 3.4 DISCUSSION

A machine learning framework has been proposed for comparative studies on volumetric data, that is able to correct for artifacts present within the data. First, a PCA/LDA mapping of the cohort is computed, which mutually weights the voxels to discriminate the populations. This mapping is thresholded to automatically extract the combination of differing regions. A technique called 'pruning' is then introduced to iteratively discard misclassified subjects from the cohort. These subjects reside in the region in feature space where the classes are overlapping. As the exact cause of this overlapping is unknown, it is preferable to base the mapping merely on representative prototypes, residing in the non-overlapping parts of  $\mathbb{R}$ .

The method, applied to schizophrenia, found a decreased FA in the corpus callosum, which was less pronounced when pruning was omitted. VBA was only capable of finding regions of increased FA, which is almost unprecedented.

Comparing the results of this study with earlier findings, a decrease in the corpus callosum has been found several times before, whereas a finding of higher FA is almost unprecedented. It can thus be said that the results of the proposed algorithm are in better correspondence with the literature [147] than the findings of VBA.

Incorporating correlation in a comparative study is proven to be a gain in discriminative power. Therefore, we have made a new step in pathological studies, opening the way to more knowledge of brain diseases in the future.



## LINEAR AND KERNEL FISHER DISCRIMINANT ANALYSIS FOR STUDYING DIFFUSION TENSOR IMAGES IN SCHIZOPHRENIA

---

A new method is explored to study schizophrenia using Diffusion Tensor Imaging (DTI). Both Linear Discriminant Analysis (LDA) and Kernel Fisher Discriminant Analysis (KFDA) are combined with Principal Components Analysis (PCA). Thus, a linear and non-linear combination of voxels is sought that separates patients from controls. PCA/KFDA does not show an improvement over PCA/LDA in classification. Because the PCA/LDA-mapping can be visualized, which enables localisation of differences, this method is preferred for analysis.

Published as:

F.M. Vos, M.W.A. Caan, K.A. Vermeer, C.B.L.M. Majoie, G.J. den Heeten and L.J. van Vliet, *Linear and kernel Fisher discriminant analysis for studying diffusion tensor images in Schizophrenia*, in: Proc. ISBI, 2007, 764-767 [249].

## 4.1 INTRODUCTION

Schizophrenia is a psychiatric disorder occurring in about 1% of the world population. The first symptoms usually happen at early adolescence, with signs of abnormal social behaviour and hallucinations. Schizophrenia is expected to affect the white matter structure in the brain [224]. Ten years ago, Magnetic Resonance Diffusion Tensor Imaging (MR-DTI) came up as a promising new way to analyze schizophrenia *in vivo* by investigating white matter structure [33, 34]. DTI measures the diffusion of water. In white matter, the diffusion of water is anisotropic, parallel to the nerve direction, whereas in grey matter the diffusion is more isotropic.

Schizophrenia is in the medical literature often studied by means of Region Of Interest (ROI) or voxel-by-voxel analysis of the Fractional Anisotropy (FA) in DTI. In the former method, mean values of regions in a patient and control group are compared, the latter method consists of a comparison per individual voxel.

It is likely that in schizophrenia multiple brain regions are involved. Clearly, the regions in the brain are highly interacting and influencing each other and one therefore expects correlation within the data. Obviously, neither ROI nor voxel-by-voxel analysis searches for this correlation, raising the need for a new method.

In this paper we explore a different method to analyze differences between schizophrenics and controls, that takes correlation between data values into account. Our approach is based on the use of sophisticated pattern recognition techniques. Both Linear Discriminant Analysis (LDA) and Kernel Fisher Discriminant Analysis (KFDA) will be used, preceded by Principal Component Analysis (PCA). LDA may be considered a special case of KFDA, with the advantageous possibility of visualizing the computed mapping, enabling localization of differences. KFDA on the other hand has more degrees of freedom in classification, which may lead to a better discrimination of patients and controls. Experiments will make clear which method is to be preferred. Two small studies using PCA/LDA have been done earlier [233, 266].

Thus, the aim of this work is two-fold. First, a classification framework will be build aiming to optimally separate schizophrenics from controls. Second, the localized differences in the brain will be determined.

## 4.2 METHOD

### 4.2.1 Discrimination

We present a new algorithm to study brain disorders, aiming at distinguishing patients from controls based on intensity changes in images. Let  $m_p$  and  $m_c$  be the sizes of the patient and control classes, summing to  $m$ , the size of the cohort. Given 3D-images will

be reshaped to  $n$ -dimensional vectors  $\mathbf{x}$ , with  $n$  denoting the number of voxels in one image.

An image can be seen as a point in a  $n$ -dimensional space. As only  $m$  points in this space are available and  $m \ll n$ , the problem of finding a classification boundary in this space is ill-posed. This is called the *small sample size problem* [113] for which a unique solution does not exist. Many attempts have been done to overcome this problem, see for instance the review in [79].

We proceed by using Principal Component Analysis (PCA) to project the data onto a low-dimensional subspace. In this subspace classification is done, both linear by means of Linear Discriminant Analysis (LDA) and non-linear by means of Kernel Fisher Discriminant Analysis (KFDA).

### PCA

PCA is a way of dimensionality-reduction that retains the directions describing most variation in the data. These directions are eigenvectors of the total scatter or covariance matrix  $\mathbf{S}_t$ , defined as

$$\mathbf{S}_t = \Phi_t \Phi_t^T, \quad (4.1)$$

where  $\Phi_t$ , with components

$$\Phi_{t,i} = \mathbf{x}_i - \bar{\mathbf{x}}, \quad i = 1, \dots, m \quad (4.2)$$

is a  $n \times m$  matrix with  $\bar{\mathbf{x}}$  the mean over all  $\mathbf{x}$ . Eigenvectors are found by eigenanalysis of  $\mathbf{S}_t$ . As  $\mathbf{S}_t$  is singular<sup>1</sup>, this cannot be done directly. The  $m - 1$  eigenvectors  $\mathbf{v}_{1,\dots,m-1}$  with corresponding non-zero eigenvalues  $\lambda_{1,\dots,m-1}$  can however be computed [240]:

$$\mathbf{v}_i = \lambda_i^{-\frac{1}{2}} \Phi_t \mathbf{v}'_i, \quad (4.3)$$

where  $\mathbf{v}'_i$  is an eigenvector with corresponding eigenvalue  $\lambda_i$  of the matrix  $\mathbf{S}_t$ . Now the projection matrix  $\mathbf{P} = (\mathbf{v}_1 \mathbf{v}_2 \dots \mathbf{v}_{m-1})$  is found, which projects an  $n$ -dimensional vector  $\mathbf{x}$  onto an  $(m - 1)$ -dimensional subspace:  $\mathbf{y} = \mathbf{P}^T(\mathbf{x} - \bar{\mathbf{x}})$ .

---

<sup>1</sup> The rank of  $\mathbf{S}_t$  is  $m - 1$ .

## LDA

In the by PCA computed subspace, LDA seeks a mapping  $\mathbf{q} : \mathbb{R}^m \mapsto \mathbb{R}$  in such a way that optimal separation between the two classes is achieved. The vector  $\mathbf{q}$  is found by maximizing the Fisher-criterion  $J(\mathbf{q})$  [99]

$$J(\mathbf{q}) = \frac{\mathbf{q}^T \mathbf{S}_b \mathbf{q}}{\mathbf{q}^T \mathbf{S}_w \mathbf{q}}, \quad (4.4)$$

with  $\mathbf{S}_b$  and  $\mathbf{S}_w$  denoting the class-between and class-within scatter respectively. The class-between scatter  $\mathbf{S}_b$  is defined as

$$\mathbf{S}_b = \frac{1}{m} \sum_{j \in \{p, c\}} m_j (\bar{\mathbf{y}}_j - \bar{\mathbf{y}}) (\bar{\mathbf{y}}_j - \bar{\mathbf{y}})^T \quad (4.5)$$

and the class-within scatter  $\mathbf{S}_w$  as

$$\mathbf{S}_w = \frac{1}{m} \sum_{j \in \{p, c\}} \sum_{i=1}^{m_j} (\mathbf{y}_i - \bar{\mathbf{y}}_j) (\mathbf{y}_i - \bar{\mathbf{y}}_j)^T, \quad (4.6)$$

with  $\bar{\mathbf{y}}_j$  the mean of class  $j$  and  $\bar{\mathbf{y}}$  the overall mean. The solution for  $\mathbf{q}$  is [99]

$$\mathbf{q} = \mathbf{S}_w^{-1} (\bar{\mathbf{y}}_p - \bar{\mathbf{y}}_c), \quad (4.7)$$

resulting in a projection of  $\mathbf{y}$  onto a one-dimensional  $z_L$ :  $z_L = \mathbf{q}^T \mathbf{y}$ .

The combined PCA/LDA projection vector  $\mathbf{a} : \mathbb{R}^n \mapsto \mathbb{R}$ ,

$$\mathbf{a} = \mathbf{P} \mathbf{q}, \quad (4.8)$$

which maps  $\mathbf{x}$  to  $z_L$ , can be analyzed to study where differences between patients and controls are found. The positive and negative part of  $\mathbf{a}$  are thresholded, to show regions where  $\mathbf{x}$  is respectively increased or decreased for patients.

## KFDA

LDA assumes equal covariance structures for patients and controls, because the within-class covariance matrices are averaged. As this assumption may be untrue for clinical data, KFDA [175] can be used instead of LDA. KFDA searches for a non-linear mapping  $\Phi : \mathbb{R}^m \mapsto \mathcal{F}$  from the input space to some high-dimensional feature space  $\mathcal{F}$ , where

the linear Fisher-criterion is maximized. Implicitly this yields a non-linear discriminant in input space. The Fisher-criterion  $J^\Phi(\mathbf{q})$  in  $\mathcal{F}$  can be written as:

$$J^\Phi(q) = \frac{\mathbf{q}^T \mathbf{S}_b^\Phi \mathbf{q}}{\mathbf{q}^T \mathbf{S}_w^\Phi \mathbf{q}}, \quad (4.9)$$

with now  $\mathbf{q} : \mathcal{F} \mapsto \mathcal{F}$  and  $\mathbf{S}_b^\Phi$  and  $\mathbf{S}_w^\Phi$  the between and within scatter matrices in  $\mathcal{F}$ .

The feature space  $\mathcal{F}$  may be very high-dimensional, which implies that directly solving of (4.9) in  $\mathcal{F}$  is impossible. This problem is avoided by only writing dot-products  $\Phi(\mathbf{y}_1) \cdot \Phi(\mathbf{y}_2)$  of the data<sup>2</sup>, such that an explicit mapping to  $\mathcal{F}$  is not needed. Writing in dot-products is achieved by using a kernel  $k(\mathbf{y}_1, \mathbf{y}_2) = \Phi(\mathbf{y}_1) \cdot \Phi(\mathbf{y}_2)$ , for which the Gaussian Radial Basis Function (RBF) is chosen:  $k(\mathbf{y}_1, \mathbf{y}_2) = \exp(-\|\mathbf{y}_1 - \mathbf{y}_2\|^2/c)$ , with  $c$  a positive constant.  $k(\mathbf{y}_1, \mathbf{y}_2)$  can be interpreted as a distance measure: it equals one for similar values of  $\mathbf{y}_1$  and  $\mathbf{y}_2$  and approaches zero for increasing distance with respect to  $c$  of input patterns in the input space  $\mathbb{R}^m$ . LDA may be considered a special case of KFDA, with a linear kernel  $k(\mathbf{y}_1, \mathbf{y}_2) = \mathbf{y}_1 \cdot \mathbf{y}_2$ .

It is derived that (4.9) can be rewritten in  $\mathcal{F}$  in terms of only dot-products [175], yielding

$$J(\alpha) = \frac{\alpha^T \mathbf{M} \alpha}{\alpha^T \mathbf{N} \alpha}. \quad (4.10)$$

Here  $\mathbf{M} = (\mathbf{M}_p - \mathbf{M}_c)(\mathbf{M}_p - \mathbf{M}_c)^T$  with  $(\mathbf{M}_j)_i = \frac{1}{m_j} \sum_{l=1}^{m_j} k(\mathbf{y}_i, y_{j,l})$  and  $\mathbf{N} = \sum_{j \in \{p,c\}} \mathbf{K}_j (\mathbf{I}_{m_j} - \mathbf{1}_{m_j}) \mathbf{K}_j^T$ , with  $(K_j)_{nm} = k(\mathbf{y}_n, y_{j,m})$  of size  $m \times m_j$ , the identity matrix  $\mathbf{I}$  and  $\mathbf{1}_{m_j}$  the matrix with all entries  $1/m_j$ . The vector  $\alpha$  that maximizes  $J$  is the leading eigenvector of  $\mathbf{N}^{-1} \mathbf{M}$ .

A vector  $\mathbf{y}$  can now be mapped, forming  $z_K$ :

$$z_K = (\mathbf{q} \cdot \Phi(\mathbf{y})) = \sum_{i=1}^m \alpha_i k(y_i, \mathbf{y}), \quad (4.11)$$

#### 4.2.2 Classification

Classification with a simple classifier, the nearest-mean classifier, is done on  $z$ . The classification error is computed by 10-times repeated 5-fold crossvalidation [99]. Since equal prior probabilities are assumed, an error around 50% means that no separation between patients and controls is found. A significant difference between patients and controls is concluded, if the upper boundary of the 95%-confidence region of the error is below 50% [99].

<sup>2</sup> This requires normalization of the data to achieve zero mean and equal variance.

To study the reliability of the classifiers, patients and controls are randomly divided into two classes, after which classification is done. If the error is significantly lower than 50%, the classifier is finding a difference where none is expected. The classifier is then concluded to be classifying on noise.

#### 4.2.3 *Subjects*

Male patients admitted to the Adolescent-clinic of the Academic Medical Center in Amsterdam and between 18 and 28 years of age were consecutively included. The patients had a clinical diagnosis of recent-onset schizophrenia or a related disorder according to DSM-IV criteria [16]. All patients received antipsychotic medication. Healthy control subjects were recruited through mouth-to-mouth advertisement and advertisement posters at the Academic Medical Center Only male controls were included between 18 and 28 years of age.

After the clinical condition of the patients had stabilized, they were interviewed by a research staff member who was not involved in the treatment. Date of the first psychotic episode and educational level were assessed, handedness was determined using the Annett Handedness Questionnaire [19].

#### 4.2.4 *Data acquisition and preprocessing*

MR-DTI was performed on a 1.5T Siemens Visions (VB33E, Siemens, Erlangen, Germany). A spin-echo EPI sequence with an extra  $180^\circ$  pulse and balanced diffusion sensitizing gradients was used to minimize artifacts induced by eddy currents [213]. Other imaging parameters were diffusion weighting  $b = 1000$  s/mm<sup>2</sup>, voxel-size  $2 \times 2 \times 6.5$  mm, TE (echo time) 109 ms, two acquisitions. Six icosahedric diffusion directions were used [9].

Using the diffusion weighted images, the  $3 \times 3$  symmetric diffusion tensor was calculated [256]. The diffusion weighted images (DWIs) were isotropically resampled, allowing sub-voxel translations in the  $z$ -plane in registration. Moreover, the DWIs were smoothed with a Gaussian kernel of 2 mm, thereby reducing the occurrence of high FA due to noise and limiting the negative side-effects of possible misregistration. An affine registration and a non-rigid registration using demons with a multi-resolution approach [232] of the FA-images were performed in 3D-Slicer ([www.slicer.org](http://www.slicer.org)), using the FA computed from a DTI-atlas [252] as target.

From the tensor, the three eigenvalues  $\lambda_1 \geq \lambda_2 \geq \lambda_3$  were calculated, describing the shape and size of the tensor. Based on the eigenvalues, scalar measures describing the anisotropy were computed. The Fractional Anisotropy (FA) [32] is commonly used in brain studies in which changes in brain structure are expected. This is based on

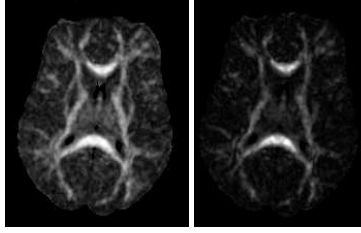


Figure 4.1: Axial slice denoting the FA (left) and  $c_l$  (right).

the assumption that high anisotropy corresponds with high white matter tracts. The Fractional Anisotropy (FA) [32] is written as

$$\text{FA} = \frac{\sqrt{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_1 - \lambda_3)^2}}{\sqrt{2}\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}. \quad (4.12)$$

Clearly,  $\lambda_3$  is the most noisy eigenvalue and suffers from sorting bias. The linear component of the tensor  $c_l$  [256] is an approximation to the FA with  $\lambda_3$  discarded from the numerator:

$$c_l = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2 + \lambda_3}. \quad (4.13)$$

Figure 4.1 shows that  $c_l$  has most signal in regions of white matter and that the FA also describes gray matter suitably. Both measures will be studied in analysis.

### 4.3 RESULTS

In this investigation 34 male patients were included with an average age of  $22.3 \pm 2.6$  years. They were compared with 24 healthy controls with an average age of  $22.5 \pm 3.2$  years. Handedness of patients was (right/left/ambidexter) 85/12/3%, of controls it was 87/13/0%. Educational level of patients was: 29% skilled training, 32% bachelor level and 38% master level; of controls 29%, 29% and 42% respectively. No gross abnormalities could be detected on conventional MR imaging by a radiologist.

After registration, the FA and  $c_l$  were computed and values below 0.1 were discarded. The classification error for PCA/LDA on FA is  $15.7\% \pm 2.1\%$  and on  $c_l$   $12.5\% \pm 1.4\%$ . The upper boundary of the 95%-confidence-region of the minimal error found using the FA is 29% and using  $c_l$  25%. As the upper boundary is below 50% for both measures, a significant difference between patients and controls is found.

The classification error as function of the kernel parameter  $c$  is given in figure 4.2. For high values of  $c$ , the errors computed by PCA/LDA and PCA/KFDA are equal. For lowering  $c$ , the error increases monotonically for PCA/KFDA.

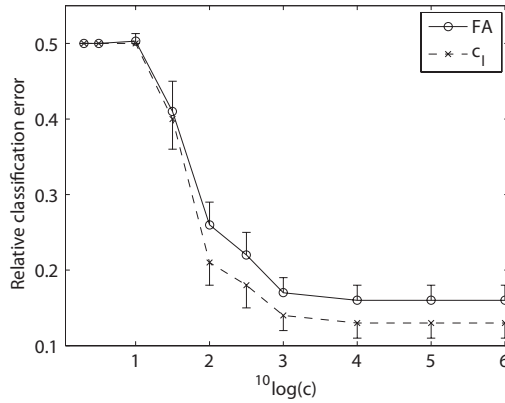


Figure 4.2: Relative classification error of PCA/KFDA on FA and  $c_1$ , as function of the kernel parameter  $c$ , with errorbars denoting the standard deviation of the error.

Classification of randomly divided data using PCA/LDA resulted in a minimal error of  $56\% \pm 3\%$  on FA-images and  $48\% \pm 4\%$  on  $c_1$ -images. These results indicate that PCA/LDA is not finding a difference based on noise.

The mappings **a** found by PCA/LDA using both the FA and  $c_1$  are displayed in Fig. 4.3. They show a common decrease in parts of the corpus callosum. The mapping of  $c_1$  also reports a slight increase. The PCA/KFDA-mapping cannot be displayed likewise, because of its non-linearity.

#### 4.4 DISCUSSION

A new framework has been developed for comparative studies on volumetric data. Using the voxels as features, PCA is applied to reduce the dimensionality, after which both LDA and KFDA are performed to classify the data. In this way, correlation in the data is taken into account, thereby overcoming one of the main deficits of voxel-by-voxel analysis.

This method was applied to schizophrenia, using the FA and  $c_1$  as diffusion tensor derived scalar measures for anisotropy. Using PCA/LDA, a significant difference was found using both measures, which is not based on noise. Although  $c_1$  contains less information about brain structure than the FA, it yields a slightly smaller classification error.  $c_1$  is therefore concluded to be a useful measure to be used in addition to the FA in comparative studies.

Visualization of the mapping computed by PCA/LDA using the FA and  $c_1$ , showed a common decrease in parts of the corpus callosum. This is in agreement with the



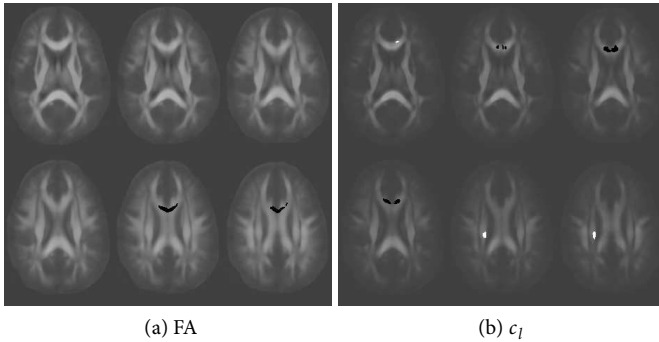


Figure 4.3: Visualization of the mapping  $\mathbf{a}$  computed by PCA/LDA using the FA and  $c_l$ . The average FA and  $c_l$  is shown in gray, with a threshold of the positive and negative part of  $\mathbf{a}$  in white and black respectively.

discussed literature findings. Based on additional found increases in  $c_l$ , analysis of merely the FA might be preferable, but more experiments are needed to support this statement.

For high values of  $c$ , the Gaussian RBF kernel used in PCA/KFDA can be written as a linear kernel, by means of a Taylor expansion. This explains the equal performance with respect to PCA/LDA. As  $c$  decreases, PCA/KFDA is becoming more and more non-linear and tends more and more to overtrain the data. A linear classifier is therefore preferred in investigation, because of less complexity and the ability of interpretation of the mapping.

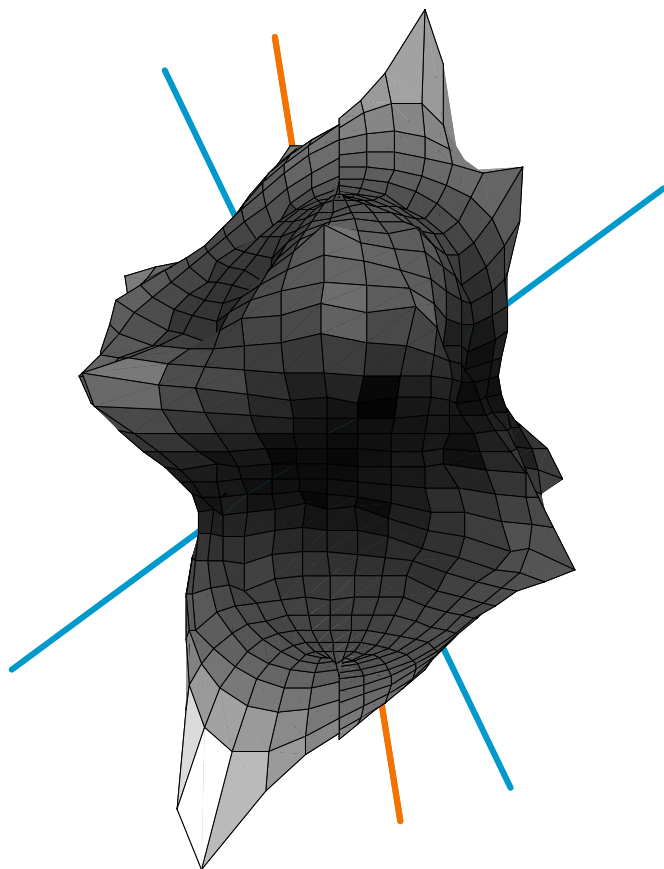
Another shortcoming of voxel-by-voxel analysis, namely that it does not take spatial correlation into account, is, besides smoothing, not solved yet. A solution could be a more restricted extraction of regions out of the mapping  $\mathbf{a}$ , for instance on the minimal size of the remaining regions.

Schizophrenia is a complex disease, yet only partially understood. The introduction of classification based on a combination of voxels opens the way to more knowledge of this disease in the future.



# RESOLVING CROSSING FIBERS

# III





## GENERATING FIBER CROSSING PHANTOMS OUT OF EXPERIMENTAL DWIS

---

In Diffusion Tensor Imaging (DTI), differently oriented fiber bundles inside one voxel are incorrectly modeled by a single tensor. High Angular Resolution Diffusion Imaging (HARDI) aims at using more complex models, such as a two-tensor model, for estimating two fiber bundles.

We propose a new method for creating experimental phantom data of fiber crossings, by mixing the DWI-signals from high FA-regions with different orientation. The properties of these experimental phantoms approach the conditions of real data. These phantoms can thus serve as a 'ground truth' in validating crossing reconstruction algorithms. The angular resolution of a dual tensor model is determined using series of crossings, generated under different angles. An angular resolution of  $0.6\pi$  was found in data scanned with a diffusion weighting parameter  $b=1000$  s/mm<sup>2</sup>. This resolution did not change significantly in experiments with  $b=3000$  and  $5000$  s/mm<sup>2</sup>, keeping the scanning time constant.

Published as:

M.W.A. Caan, A.W. de Vries, H.G. Khedoe, E.M. Akkerman, L.J. van Vliet, C.A. Grimbergen and F.M. Vos, *Generating fiber crossing phantoms out of experimental DWIs*, in: Proc. MICCAI, LNCS 4791, 2007, 169-176 [61].

## 5.1 INTRODUCTION

In the last decade, specifications of MR-scanners have increasingly improved. Consequently, Diffusion Weighted Images (DWIs) can be acquired at higher resolution and with better image quality. A more precise analysis of the diffusion profile of water in the human brain has thus become possible. This diffusion process is modeled by a rank-two symmetric positive-definite tensor [37], describing both the average distance and the orientation of the displacement. Anisotropic diffusion is associated with white matter structures, whereas isotropic diffusion is generally related to grey matter and cerebral spinal fluid. The Fractional Anisotropy (FA) has been introduced as a distance measure to an isotropic tensor [35], which is of use in analyzing changes in white matter integrity. The first eigenvector of the diffusion tensor represents the orientation of fibrous white matter structures, which can be tracked throughout the image volume [46]. Such tractography is of practical interest for pre-operative planning. Additionally, studies to brain diseases are increasingly using fiber tracts as features in analysis [89].

Still, the minimal size of a voxel in a DWI is orders of magnitude higher than the diameter of the individual axons. Only macroscopic brain properties can therefore be described. What is more, it is likely that multiple fiber bundles are crossing inside a single voxel, due to the partial volume effect. Multiple anisotropic diffusion processes in different directions are then modeled as a single process, introducing a severe bias in the tensor estimation. Consequently, regions of crossing fibers have to be excluded from comparative studies to brain diseases using tensor properties. Additionally, the measured principal diffusion direction is ill-defined and not related to the underlying anatomical structure. Fiber tracking based on a single tensor fit will erroneously connect distinct tracts or be terminated in a crossing. A dual tensor fit in regions of crossing fibers is thus indispensable for correctly tracking fibers through the brain.

Several solutions were proposed for overcoming the crossing-fiber problem. Non-parametric methods aim to find the probability density function of the spin diffusion displacement. For instance,  $q$ -ball imaging was proposed [239] for calculating diffusion orientations from High Angular Resolution Diffusion Images (HARDI). Parametric methods include a two-tensor model for analyzing the diffusion signal [239, 109, 159]. This work was continued by using spherical harmonics to characterize anisotropy [110], which was elsewhere shown to be related to the coefficients of higher order tensors [194]. Recently, high angle fiber crossings were reconstructed in clinical data with a diffusion weighting parameter  $b = 1000\text{s/mm}^2$  [201].

In order to validate the reconstruction of crossings, various experiments with simulated data have been performed. For instance, data with additive uncorrelated Gaussian noise was simulated [13]. Scanning artifacts and patient movement are thereby ignored. Additionally, using a spinal chord, an experimental crossing phantom was created [27]. Still, a ground truth is missing for properly facilitating a validation of crossings found in the human brain.

We propose a new method for generating fiber crossing phantoms out of experimental data, containing realistic imaging artifacts. These phantoms can serve as a ‘ground truth’ in validating algorithms aiming to resolve crossing fibers. Series of crossings, generated under different angles, give insight into the precision that can be achieved in estimating the fiber bundle orientations. Therefore, a restricted two-tensor model is fitted on the projected data onto the plane of the crossing. Included in our analysis is the influence of the diffusion weighting parameter  $b$  on the angular resolution.

Finally, the relevance of a proper modeling of crossing fiber regions is demonstrated. Streamlines generated through a crossing are shown to erroneously depict the fiber paths. Using a dual tensor model, it is shown that the FA of two crossing fiber bundles can be distinctively estimated.

## 5.2 METHOD

### 5.2.1 *Experimental phantom data*

#### *Signal*

The MR-signal measured in a certain voxel contains the contributions of the individual underlying volume fractions of different tissue. Dependent on the heterogeneity of the local tissue, more or less distinct fractions are present within a voxel. A summation of the fractions is assumed to correctly model the acquired MR-signal. This assumption is justified by the fact that there is no reason to expect inherent phase signal differences within a voxel. Previously, this summability was for instance silently used in [17], where an MR-signal is Monte Carlo simulated by aggregating a number of spins per voxel. And as an illustration, the figures 5.1(b) and 5.2(b) can be compared, showing that adjacent perpendicular fiber bundles contribute equally to the signal of the intermediate voxel, inducing partial voluming.

We now propose to mix voxels out of experimental Diffusion Weighted Images (DWIs) whose FA is high enough to assume a single fiber bundle inside. Thus we are able to generate fiber crossings with different known properties, derived from single tensor fits of the individual parts. Here, only two voxels are mixed; dependent on the application even a higher number of orientations could be combined.

By using this approach, most known and unknown imaging artifacts that are present in real data are included in the phantoms. A realistic validation of a crossing reconstruction algorithm of interest has thus become possible.

### Noise

Diffusion is measured as an attenuated signal, such that noise is expected to significantly affect the acquired images. This noise follows a Gaussian distribution, added independently to both the real and imaginary part of the signal. DWIs, being magnitude images with relatively low SNR, are thereby containing Rician noise [123]. The noise in the generated phantom data now yields the sum of two Rician noise processes. Following the central limit theorem, this noise distribution can be expected to slightly tend back towards Gaussian. For simplicity, a Gaussian noise model is now used in reconstructing the generated crossing fibers.

#### 5.2.2 Two-tensor model

In the case of two fiber orientations in a single voxel, distinction between these orientations is needed. We model the diffusion process of two fibers as the sum of two Gaussian distributions with different diffusion tensors,  $\mathbf{D}^1$  and  $\mathbf{D}^2$ , closely following [199]. Note that this model intuitively closely relates to the way the experimental phantoms are generated. Without loss of generality, the two fibers are assumed to reside in the  $(x, y)$ -plane. The equation that models the MR-signal then becomes:

$$S(\tilde{\mathbf{g}}) = S_o \left( f \exp(-b\tilde{\mathbf{g}}^T \mathbf{D}^1 \tilde{\mathbf{g}}) + (1 - f) \exp(-b\tilde{\mathbf{g}}^T \mathbf{D}^2 \tilde{\mathbf{g}}) \right), \quad (5.1)$$

with

$$\mathbf{D}^{1,2} = \begin{pmatrix} D_{xx}^{1,2} & D_{xy}^{1,2} & 0 \\ D_{xy}^{1,2} & D_{yy}^{1,2} & 0 \\ 0 & 0 & D_z \end{pmatrix}, \quad (5.2)$$

where  $f$  is the volume fraction of the first tensor  $D_z$  is the same for both tensors. The gradient directions  $\mathbf{g}$  are rotated to  $\tilde{\mathbf{g}}$ , such that the fiber bundles are situated in the  $(x, y)$ -plane. The tensor components (after rotation)  $D_{xz}^{1,2}$  and  $D_{yz}^{1,2}$  are assumed to be negligible and are thus set to zero. Moreover, the diffusion in the direction perpendicular to the plane of interest is assumed to be equal for both fibers, i.e. the smallest eigenvalues of both tensors are equal. The expected error made due to this model simplification is small, due to the small contribution of this eigenvalue to the signal. The data is projected onto this direction. As a reference, the normal to the plane spanned by the principal diffusion directions in the voxels, used to generate the crossing, is used.

In measuring the diffusion profile of a crossing fiber, proper contrast is needed between the signal contributions of both fibers. Theoretically, altering the diffusion weighting parameter  $b$  highly influences this contrast. In figure 5.2(a), a logarithmic signal profile ( $\ln S_o - \ln S(\mathbf{g})$ ) of a crossing fiber is computed for different  $b$ -values. The iso-lines further away from the origin (corresponding to high  $b$ ) reveal more information about the



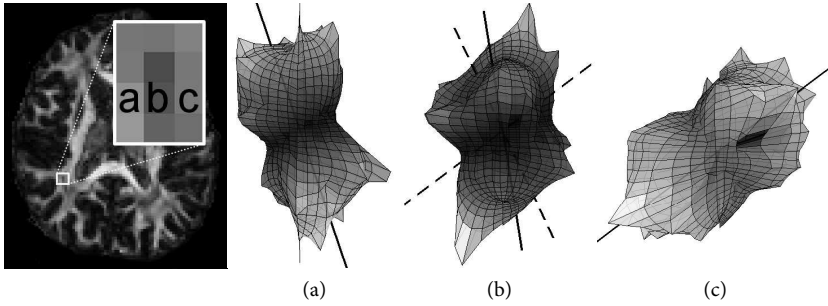


Figure 5.1: Axial slice of the FA with logarithmic signal profiles of three adjacent voxels (a-c), measured at  $b=3000 \text{ mm/s}^2$ . Solid lines represent single tensor fits, dashed lines dual tensor fits to the data. Note the correspondence between the dashed lines in (b) and the solid lines in (a) and (c).

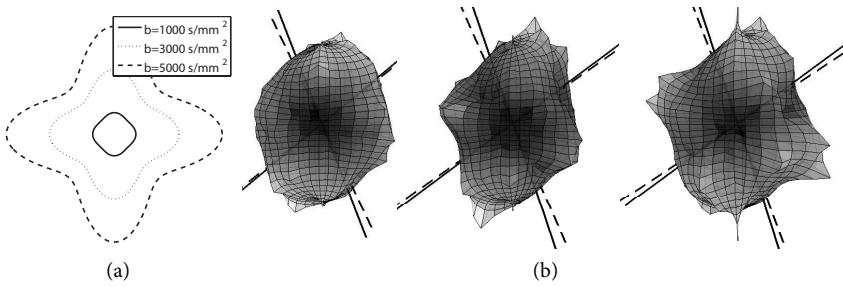


Figure 5.2: (a) Computed logarithmic signal profiles for two perpendicular tensors with typical white matter properties [204] and (b) measured profiles after mixing voxels a and c in figure 5.1, for  $b=1000$ ,  $3000$  and  $5000 \text{ s/mm}^2$  respectively. Solid lines represent single tensor fits before mixing, dashed lines dual tensor fits to the data after mixing. The profiles are scaled for proper visualization.

crossing than the iso-lines near the origin. However, in measured data, noise increasingly distorts measurements at higher  $b$ -values.

The angular resolution for different  $b$ -values will be determined in generated experimental phantoms, using the dual tensor method. The minimal angle between fiber bundles that can be estimated is defined at the point where both the bias and variance clearly start increasing.

### 5.3 RESULTS

DTI-data of four healthy volunteers, between 20 and 30 years of age, were acquired on a Philips Intera 3.0 Tesla MRI scanner (Philips Intera, Philips Medical Systems, Best, The

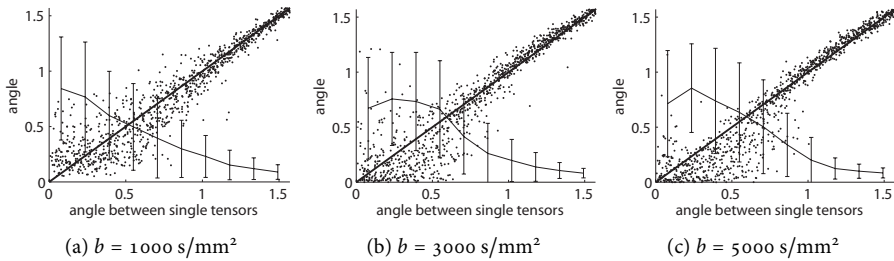


Figure 5.3: Scatter plot of the estimated angle between the principal directions in the generated crossing fiber, as function of the angle between the single tensors. The diagonal solid bold line refers to ground truth. The solid line with errorbars denotes the error and standard deviation in the projection direction. Crossings were generated by mixing DWIs of two voxels with  $FA > 0.6$ , for  $b=1000$ ,  $3000$  and  $5000 \text{ s/mm}^2$  (a,b,c).

Netherlands) by means of a spin-echo EPI sequence. The diffusion weighting was along 92 three-fold tessellated icosahedric gradient directions. Other parameters were: TE 84 msec, TR 7912 msec,  $b=1000$ ,  $3000$  and  $5000 \text{ s/mm}^2$ , FOV 220 mm, scan matrix  $112 \times 110$ , image matrix  $128 \times 128$ , slice thickness 3mm. Eddy current induced distortion was visually inspected to be negligible.

Only voxels for which  $FA > 0.6$  were included. 1000 pairs of voxels were chosen to generate a uniform distribution of angles between the bundles in these voxels in the range  $[0, \pi/2]$ . Next, the measured signals  $S$  of these pairs were mixed, forming the crossing fiber phantoms.

For illustration purposes, logarithmic signal profiles of three adjacent voxels measured at  $b=3000 \text{ s/mm}^2$  are displayed in figure 5.1. Here, the superior longitudinal fasciculus (a) and posterior corona radiata (c) are crossing (b). Logarithmic signal profiles after mixing DWIs of voxels a and c in figure 5.1 for  $b=1000$ ,  $3000$  and  $5000 \text{ s/mm}^2$  are given in figure 5.2(b).

Optimization was performed based on the interior-reflective Newton method (Matlab, Mathworks) [88]. A quadratic regularization term for the volume fraction  $f$  was added to the cost function, with a scaling factor of 0.2 around the mean value 0.5. The results of fitting the generated crossing fibers for  $b=1000$ ,  $3000$  and  $5000 \text{ s/mm}^2$  are displayed in figure 5.3. The reproducibility of the method was analyzed by computing the angular deviation with respect to the input angle for data with  $b=3000 \text{ s/mm}^2$  of four healthy individuals, depicted in figure 5.4. An angular resolution of  $0.6\pi$  was found, for both  $b=1000$ ,  $3000$  and  $5000 \text{ s/mm}^2$ . Reproducible results in four subjects were achieved.

Finally, a single streamline was generated in the body of the corpus callosum, passing through the crossing with the corticospinal tract [46]. This streamline incorrectly follows

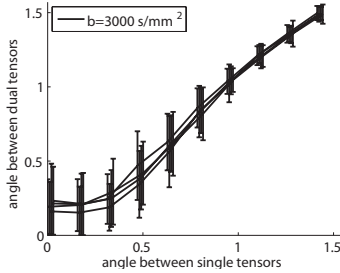


Figure 5.4: Estimated mean angle with errorbars denoting the standard deviation, as function of the input angle for  $b=3000$  s/mm<sup>2</sup> for four datasets of healthy individuals.

the intermediate orientation of the two fiber bundles. A dual tensor fit was performed along this tract, and estimated fiber bundle orientations were plotted. The decreased FA and increased planar tensor component  $c_p$  of the single tensor indicate the presence of a crossing. This is confirmed by the higher FAs of the dual tensors along the tract. The results of this experiment are given in figure 5.5.

#### 5.4 DISCUSSION

We propose a new method for creating experimental phantom data of fiber crossings, by mixing the DWI-signals from high FA-regions with different orientation. The properties of these experimental phantoms approach the conditions of real data. These phantoms can thus serve as a ‘ground truth’ in validating crossing reconstruction algorithms. In this paper, the angular resolution of a dual tensor model is determined. Series of crossings, generated under different angles, give insight into the precision that can be achieved in estimating the fiber bundle orientations. Not only are the results based on realistic imaging artifacts, but also can the reconstructed fiber orientations be compared to a ground truth, being the single tensor fits on the individual voxels.

An equal angular resolution of  $0.6\pi$  was found, for both  $b=1000$ ,  $3000$  and  $5000$  s/mm<sup>2</sup>. A higher resolution for higher  $b$ -values was to be expected, due to more non-linear behaviour of the diffusion profile. Apparently, the decreased SNR in high  $b$ -value data (with constant scanning time), due to a stronger signal attenuation, strongly affected the algorithm’s performance. Using a Rician noise model in a Maximum Likelihood estimator might reduce the bias in the reconstructed angle. The error in the estimated projection direction clearly increases for lower input angles, resulting in an ill-defined crossing fiber plane. Nevertheless, the deviation in estimated output angles did not decrease significantly when the correct projection direction was used instead.

Future work consists of acquiring DTI-data at  $b$ -values around  $1000$  s/mm<sup>2</sup>, to determine the optimal  $b$ -value in maximizing the angular resolution in the reconstruction of crossing fibers. Building a theoretical model will aid in choosing proper values for  $b$

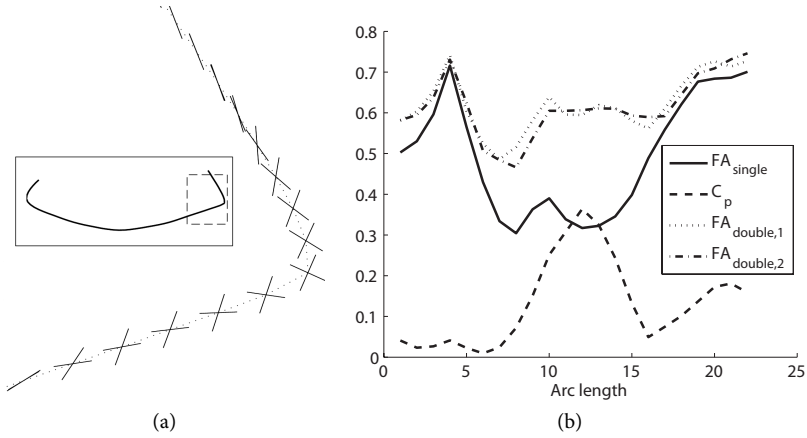


Figure 5.5: (a) Selected part of a single streamline in the body of the corpus callosum (inset), with solid lines representing the principal directions after a dual tensor fit, and (b) anisotropy measures based on single ( $FA_{single}$ ,  $C_p$ ) and dual tensor fit ( $FA_{dual,1}$ ,  $FA_{dual,2}$ ) along this tract.

and other imaging parameters, the number of repetitive measurements and gradient directions. Eventually, tractography with increased precision and an unbiased anisotropy estimation will lead to an improved analysis of both healthy and pathological brain data.

Instead of generating phantoms out of magnitude data, the real and imaginary part of the DWIs could be individually mixed. A Gaussian noise model could then be obtained in both parts, making even more realistic experimental phantoms. However, phase distortions, between and within acquisitions, hamper usage of the complex data. Until now, no correction method for these distortions is available.

The proposed validation method is expected to be useful in various crossing reconstruction algorithms. As a result, crossing fiber regions can be reliably included in clinical studies to brain diseases.

## DUAL TENSOR ATLAS GENERATION BASED ON A COHORT OF COREGISTERED NON-HARDI DATASETS

---

We propose a method to create a dual tensor atlas from multiple coregistered non-HARDI datasets. Increased angular resolution is ensured by random variations of subject positioning in the scanner and different local rotations applied during coregistration resulting in dispersed gradient directions. Simulations incorporating residual coregistration misalignments show that using 10 subjects should already double the angular resolution, even at a relatively low  $b$ -value of  $b = 1000 \text{ smm}^{-2}$ . Commisural corpus callosum fibers reconstructed by our method closely approximated those found in a HARDI dataset.

Published as:

M.W.A. Caan, C.A. Sage, M.M. van der Graaf, C.A. Grimbergen, S. Sunaert, L.J. van Vliet, F.M. Vos, *Dual Tensor Atlas Generation based on a Cohort of Coregistered non-HARDI Datasets*, in: Proc MICCAI'09, LNCS 5761, 2009, 869-876 [64].

## 6.1 INTRODUCTION

The white matter (WM) structure of the human brain can be studied by means of Diffusion Weighted Magnetic Resonance Imaging (DW-MRI). The diffusion is measured in multiple directions, from which the principal diffusion orientation field is reconstructed [33]. Fiber tracking throughout this field allows reconstruction of WM bundles, such as the corpus callosum [252].

The validity of local comparison of WM properties between different cohorts depends on good inter-subject anatomical correspondence. Variability in anatomical correspondence can be minimized by coregistration of the data on voxel basis. Both an affine and a non-rigid transformation can be computed to compensate for global and local variability, respectively [169].

Coregistration of data requires an appropriate reference space to be chosen. The reference space to which the subjects are transformed, can either be represented by one subject out of the cohort, or by an atlas that is built by averaging coregistered datasets. The ICBM FA atlas was generated by averaging 81 affinely transformed DTI datasets [181]. Recently, several DTI atlas-building approaches have been proposed with some yielding single subject-based atlases, e.g. [22], and others yielding population-based atlases [244].

By coregistering the data to such an atlas, fiber tracking can be performed in atlas space, which allows comparing tract statistics between patient and control groups [118]. However, tracking is known to be biased in regions where fibers are crossing. Fiber tracts may either deviate or truncate in such regions, due to inadequate modeling of the diffusion process by a single tensor. Fiber tracking through complex tissue regions is facilitated by higher order diffusion models, such as a dual tensor model or q-ball [98, 159]. In order to apply such complex models, High Angular Resolution Diffusion Imaging (HARDI) data needs to be acquired [238]. One of the properties of a HARDI-acquisition is a high number of gradient directions, being more than 100. Clinical data is generally sampled at 6-45 gradient directions.

The design of a comparative study to find structural changes along fibers may be regarded as sampling a high-dimensional space, spanned by respectively three spatial coordinates, gradient directions and included subjects. Let us assume that the amount of available scanning time is constant, such as in a large population study in which only a few minutes of acquisition time is reserved for DW-MRI. In order to reduce bias in fiber tracking, one might prefer to acquire HARDI-data, yielding a denser sampling along the gradient axes. A fixed-time constraint implicitly imposes a lower sampling density along either the spatial or subject axes. Both are generally unacceptable, since not only small anatomical structures but also statistical power in the comparison need to be retained. Therefore, acquiring HARDI-data for clinical studies is not always realistic, as scanning time needs to be increased or scanner hardware needs to be upgraded to facilitate a higher SNR. Also, ongoing studies with conventional DW acquisition protocols, may not comply with the previously mentioned criteria.

In this study, we propose a method to fit a higher order diffusion model to a cohort of coregistered non-HARDI data. We consider the diffusion attenuation profiles of multiple subjects as realizations of one underlying fiber distribution. After non-rigid coregistration of the datasets, the resulting signal profiles are gathered to generate an artificial HARDI-dataset (disregarding the low  $b$ -value), with a high sampling density along the gradient axes. First, simulations incorporating residual coregistration misalignments are used to demonstrate the potential increase in angular resolution. Second, a dual tensor model is fitted to the data to generate a *dual tensor atlas*. We expect that in this atlas, tracking will be as accurate as in HARDI-data.

## 6.2 METHOD

### 6.2.1 Dual Tensor Model

We propose to estimate a dual tensor model from non-HARDI data of multiple subjects, with the purpose of precisely estimating two independent diffusion orientations per voxel. We assert a model for the diffusion weighted signal  $S_j(\mathbf{q}_j)$  of a subject  $j$ , with two tensors combined with an isotropic compartment:

$$\frac{S_j(\mathbf{q}_j)}{S_{j,o}} = f_1 \exp(-\mathbf{q}_j^T \mathbf{D}_1 \mathbf{q}_j) + (1 - f_1 - f_{iso}) \exp(\mathbf{q}_j^T \mathbf{D}_2 \mathbf{q}_j) + f_{iso} \exp(-\mathbf{q}_j^T \mathbf{q}_j D_{iso}). \quad (6.1)$$

Here  $S_{j,o}$  is the signal for subject  $j$  measured without diffusion weighting and  $f_{\dots}$  are the normalized volume fractions, while  $D_{iso}$  is the isotropic diffusion constant. The vector  $\mathbf{q}_j = \sqrt{b} \mathbf{g}_j$  reflects the diffusion weighting  $b$  in gradient direction  $\mathbf{g}_j$ . The diffusion tensors are spectrally decomposed,  $\mathbf{D}_{1,2} = \mathbf{R}_{1,2} \mathbf{E} \mathbf{R}_{1,2}^T$ , with  $\mathbf{E} = \text{diag}(\lambda_{\parallel} \quad \lambda_{\perp} \quad \lambda_{\perp})$  being the eigenvalue matrix with axial and planar diffusion values (assuming axially symmetric tensors). The two rotation matrices  $\mathbf{R}_{1,2}$  are parametrized using Euler angles  $\alpha_{1-4}$ . The vector of eight parameter values  $\boldsymbol{\theta} = (\lambda_{\parallel} \quad \lambda_{\perp} \quad \alpha_{1-4} \quad f_1 \quad f_{iso})$  is estimated in a least-squares sense.

### 6.2.2 Simulating Variation in the Cohort

A cohort of subjects that is studied for epidemiological processes or pathology is carefully defined by setting precise inclusion criteria. Thus, undesirable variations unrelated to the process of interest are minimized and may include handedness, sex, age and education [66]. Remaining variations are to be seen as normally distributed axial and planar diffusion values. Anatomical variations between subjects are apparent as differences in size and position of WM bundles. Non-rigid coregistration of the data accounts for the majority of these differences, residual misalignment of the principal diffusion direction was found to be  $20^\circ$  [243, 73].

In an experiment on synthetic data, the expected angular resolution (i.e. the absolute error in the angle between the major axes of the tensors) will be computed, as a function of the number of included datasets. The abovementioned variations are incorporated

by generating a signal profile per subject with normally perturbed eigenvalues and angles (using equation 6.1). Also, Rician noise is included to distort the simulated data. Gradient directions are randomly rotated per subject (see section 6.2.3). The parameter vector  $\theta$  is subsequently estimated on the combined data of multiple subjects. In this experiment, a diffusion weighting of  $b = 1000 \text{ smm}^{-2}$  is used to reflect our conventional clinical protocol.

### 6.2.3 Building a Dual Tensor Atlas

#### *Coregistration*

The DWI datasets of multiple subjects need to be transformed to a common reference space. In this study, a population-based atlas is build. To generate this atlas, all datasets are first affinely coregistered to the ICBM FA template [181]. As an intermediate step, we construct a (single tensor) DTI atlas that contains the full diffusion information [244]. The original DWIs were coregistered to our DTI atlas using an algorithm based on a viscous fluid model and optimization of mutual information as the coregistration criterion [243]. This coregistration consists of a combination of an affine and a non-rigid transformation to correct for global and local morphological differences, respectively. The effect of deformation discontinuities is removed by smoothing of the final deformation field with a Gaussian kernel of FWHM = 3mm.

#### *Transformation*

A diffusion-weighted image describes the MR signal attenuation due to diffusion in a specific gradient direction. If it is rotated, in our case during registration, signal correspondence with the gradient direction needs to be retained. This was done previously by fitting a single tensor to the data, after which preservation of principal direction (PPD) was applied to realign tensors with the underlying fiber structure after transforming the tensor field [243]. The DWIs were then recomputed from the rotated tensor field. In the latter step, the number of degrees of freedom (DOF) is reduced to six (i.e. the DOF of the single tensor model), which prohibits a higher order model fit to the data. We therefore transform the original DWIs without intensity correction. Instead, the gradient directions are rotated, such that the vectors  $\mathbf{q}_j$  become:

$$\mathbf{q}_j = \sqrt{b} \cdot \mathbf{R}_{\text{nonrigid},j}(x, y, z) \cdot \mathbf{R}_{\text{affine},j} \cdot \mathbf{g}_j. \quad (6.2)$$

The rotational component of the affine transformation  $R_{\text{affine},j}$  is applied globally, while the rotation due to non-rigid transformation  $\mathbf{R}_{\text{nonrigid},j}(x, y, z)$  is applied per voxel  $(x, y, z)$ .

An advantage of our reorientation method is that both  $R_{\text{affine},j}$  and  $\mathbf{R}_{\text{nonrigid},j}$  will be dispersed among subjects (as demonstrated in figure 6.2(c)). Patient positioning in the



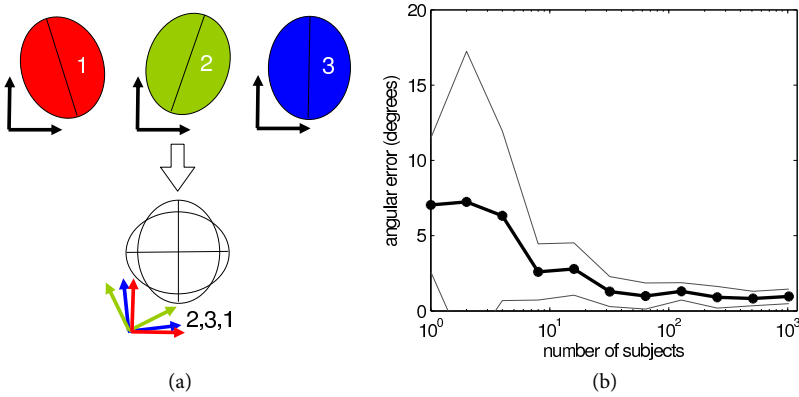


Figure 6.1: (a) Illustration of the dual tensor atlas generation, the gradient directions are randomly dispersed after coregistration, such that a dual tensor fit becomes possible. (b) Mean angular error (bold) with standard deviation (gray) in degrees as a function of the number of included datasets. The two tensors per subject were slightly perturbed, such that the angles were normally distributed with  $\sigma = 5^\circ$ . The average angle between both tensors was  $45^\circ$ , 32 random gradient directions were used per subject.

scanner and inter-subject variations in bundle trajectories contribute to this distribution. The combined set of multiple subjects is sampled at a higher angular resolution than the original acquisition protocol. Hence, the merged dataset not only has an increased SNR, but indeed approaches a HARDI-dataset due to random dispersion of gradient directions per subject after coregistration.

### *Tensor Estimation and Fiber Tracking*

The dual tensor model can now be estimated (equation 6.1) on the merged dataset, using the dispersed sampling of gathered gradient directions. The least squares difference of the model and data is minimized using Levenberg-Marquardt optimization. Error-function parametrization ensures positiveness of the parameters. Next, WM fibers are reconstructed using High Angular Fiber Tracking (HAFT) [237]. As an extension to common streamline tracking algorithms, this method generates branches of fibers if within a voxel the angle between both tensors exceeds a certain threshold. One seed Region of Interest (ROI) and one AND-ROI through which fibers should pass have to be defined.

### 6.3 RESULTS

#### 6.3.1 *Synthetic Data*

Synthetic data was generated according to the model in equation 6.1. The eigenvalues were  $\lambda_{\parallel} = 1.5 \cdot 10^{-3} \text{ mm}^2\text{s}^{-1}$  and  $\lambda_{\perp} = 0.4 \cdot 10^{-3} \text{ mm}^2\text{s}^{-1}$  [204] and the angle between both tensors was  $45^\circ$ . For second and further realizations, the eigenvalues and the angle were normally perturbed with a standard deviation of 10%.  $S_0$  was set to 230, whereas a volume fraction  $f_1 = 0.5$  was used. Sixteen icosahedric and dodecahedric gradient directions on a half sphere [9] and along the 16 corresponding, antipodal directions were chosen, conform to our clinical scanning protocol. Rician noise (SNR = 20:1) distorted the data.

Next, data of different numbers of datasets with dispersed gradient directions were gathered and the dual tensor model was fit to the data. The isotropic volume fraction  $f_{iso}$  was included to account for inter-subject diffusion variations, whereas the isotropic diffusion was set constant:  $D_{iso} = 0.5 \cdot 10^{-3} \text{ mm}^2\text{s}^{-1}$ . The mean and standard deviation of the angular error (defined in section 6.2.2) as a function of the number of subjects are depicted in figure 6.1. Initially, both the mean error and the standard deviation remained high when data of up to four subjects were used. For small sample sizes, modeled residual misalignments induce a slight fitting instability. Adding more subjects lowered the mean error and the standard deviation to  $2^\circ \pm 1^\circ$  for 64 subjects. Further increase of data did not result in a lower error, again due to the applied perturbations.

This experiment shows that the angular resolution may be increased by estimating tensors based on a cohort.

#### 6.3.2 *Clinical Data*

The proposed method was applied to a small cohort of 11 healthy subjects, of which DTI data were acquired on a 3.0T scanner (Intera, Philips, Best, The Netherlands). The spatial resolution was  $2.0 \times 2.0 \times 2.2 \text{ mm}$ , such that 64 axial slices of matrix size  $128 \times 128$  were acquired, while 32 gradient directions were used (see section 6.3.1) with a diffusion weighting of  $b = 1000 \text{ smm}^{-2}$ . Additionally, one set of images was acquired without any diffusion-weighting. Eddy current distortions were corrected for by an affine registration in the phase direction [173]. For comparison, a HARDI-dataset was acquired, with 92 gradient directions and  $b = 1800 \text{ smm}^{-2}$ . The acquired datasets were coregistered as described in section 6.2.3. The DWIs were resampled without intensity correction, and diffusion weighting vectors  $\mathbf{q}_j$  were rotated (equation 6.2) and stored per voxel per deformation field of a subject. The merged dataset in total contained  $32 \cdot 11 = 352$  DWIs.

The dual tensor fit was initialized by positioning the principal eigenvectors in the plane of the first two eigenvalues of a single tensor fit. Dual tensor parameters were then estimated in each intercranial voxel. The results of the dual tensor atlas estimation are

shown in figure 6.2. In figure 6.2(a), two distinct orientations within one voxel are mainly seen in the region where the corpus callosum is known to cross the corticospinal tract. Aggregated signal values within one voxel with the fitted model are depicted in figure 6.2(b). Due to the low  $b$ -value, low contrast is seen in the modeled signal. Moreover, significant noise and/or inter-subject variation may be observed (e.g. in the extension of the blue vertical vector in figure 6.2(b)). This confirmed the need for a sufficient amount of data in order to precisely estimate the orientation. In fact, high inter-voxel consistency in orientation in figure 6.2(a) is observed, indicating high fitting stability. Unfortunately, the gradient directions were not distributed uniformly, as displayed in figure 6.2(c). Still, a gain in angular resolution was achieved using our approach.

Fibers were reconstructed through the crossing and are displayed in figure 6.2(a). The fiber branching threshold was set to  $12.5^\circ$ , the angular threshold was  $25^\circ$  and the single tensor FA threshold 0.1. The obtained fibers in the HARDI-dataset and the dual tensor atlas are shown in figures 6.2(d) and (e) respectively. Both corpus callosum fibers (in green) as well as commissural fibers (in magenta) were successfully reconstructed in both datasets. A higher number of 38 commissural fibers was measured in the HARDI-dataset, compared to 11 in the dual tensor atlas. This is explained by the fact that in the HARDI-dataset a high number of fibers was running over the same trajectory.

#### 6.4 DISCUSSION

We proposed a method to create a dual tensor atlas from multiple coregistered non-HARDI datasets. Increased angular resolution was ensured by random variations of subject positioning in the scanner and different local rotations applied during coregistration resulting in dispersed gradient directions. Simulations showed that using 10 subjects should already double the angular resolution, even at a relatively low  $b$ -value of  $b = 1000 \text{ smm}^{-2}$ . Experiments on clinical data revealed distinct orientations within a fiber crossing as well as inter-voxel consistency. Commissural corpus callosum fibers reconstructed by our method closely approximated those found in a HARDI dataset.

Although the adopted coregistration method uses DWIs generated with a single tensor model, as argued in section 6.2.3, we have apparently obtained a good correspondence in regions with crossing fibers. In a single tensor model, fiber crossings manifest themselves by an oblate spheroid. Although the oblate spheroid does not contain orientation information in the plane of the crossing, it permits accurate location of the crossing. The required orientational information for correct steering of the coregistration comes from its surrounding regions in which a single tensor model suffices.

We aim to apply our method to a large cohort in order to create an atlas that can be made available to the neuro-imaging community. WM bundles traveling through fiber crossings can then be included in comparative studies.

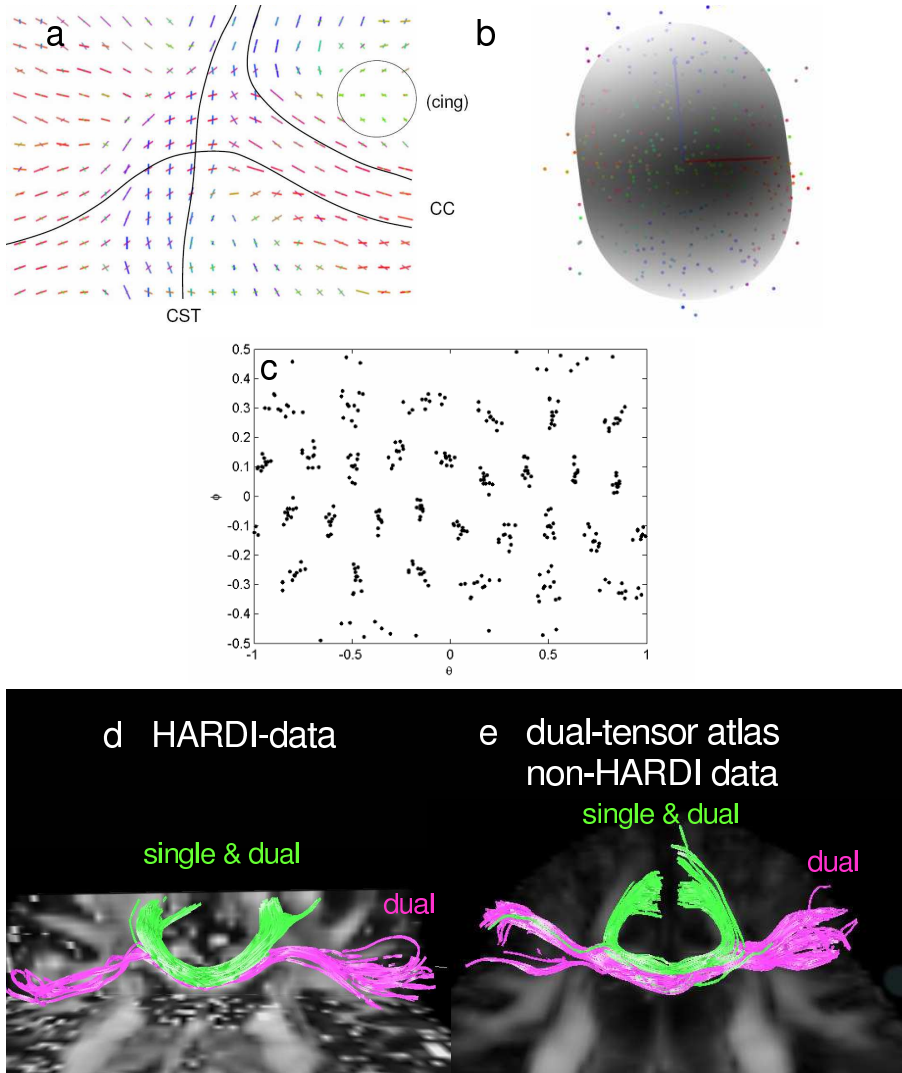


Figure 6.2: Dual tensor atlas results. (a) Coronal cross-section through the estimated dual tensor atlas field, with principal diffusion orientations per voxel scaled by the volume fraction. Anatomical references are given (CST = corticospinal tract, CC = corpus callosum, cing = cingulum) together with an impression of bundle trajectories (drawn by the authors). (b) Aggregated ADC-profile plotted as points and fitted model in gray with principal orientations. Conventional color coding is used. (c) Aggregated gradient directions plotted in spherical coordinates. (d-e) HARDI-data and dual tensor atlas tracking result: part of the corpus callosum is tracked (green fibers), as well as commissural fibers (magenta fibers).

ESTIMATION OF DIFFUSION PROPERTIES IN CROSSING  
FIBER BUNDLES

---

There is an ongoing debate on how to model diffusivity in fiber crossings. We propose an optimization framework for the selection of a dual tensor model and the set of diffusion weighting parameters  $\mathbf{b}$ , such that both the diffusion shape and orientation parameters can be precisely as well as accurately estimated. For that, we have adopted the Cramér-Rao lower bound (CRLB) on the variance of the model parameters, and performed Monte Carlo simulations. We have found that the axial diffusion  $\lambda_{\parallel}$  needs to be constrained, while an isotropic fraction can be modeled by a single parameter  $f_{iso}$ . Under these circumstances, the Fractional Anisotropy (FA) of both tensors can theoretically be independently estimated with a precision of 9% (at SNR=25). Levenberg-Marquardt optimization of the Maximum Likelihood function with a Rician noise model approached this precision while the bias was insignificant. A two-element  $\mathbf{b}$ -vector  $\mathbf{b} = [ 1.0 \quad 3.5 ] \cdot 10^3 \text{ mm}^{-2}\text{s}$  was found to be sufficient for estimating parameters of heterogeneous tissue with low error. This has allowed us to estimate consistent FA-profiles along crossing tracts. This work defines fundamental limits for comparative studies to correctly analyze crossing white matter structures.

Submitted as:

M.W.A. Caan, H.G. Khedoe, D.H.J. Poot, A.J. den Dekker, S.D. Olabbarriaga, L.J. van Vliet and F.M. Vos, *Estimating diffusion properties in crossing fiber bundles*, IEEE Transactions on Medical Imaging

## 7.1 INTRODUCTION

The integrity of white matter structures in the brain is frequently studied by means of Diffusion Weighted Magnetic Resonance Imaging (DW-MRI). Classically, the diffusion in the brain is described by a rank-two diffusion tensor which is estimated from the Diffusion Weighted Images (DWIs). The local geometry of axons is quantified by measures such as the Fractional Anisotropy (FA), which measures the anisotropy of the diffusion process within a voxel. The orientation information, extracted from the DWIs, allows the reconstruction of white matter bundles by means of fiber tractography [39].

There is an ongoing debate on how to model diffusivity in fiber crossings. Here the diffusion profile cannot be adequately described by a single tensor [238]. We propose an optimization framework for the estimation of diffusion parameters at such locations. Our focus is on unbiased and highly reproducible estimation of both the diffusion shapes and the orientations of the fiber bundles. The diffusion shapes reveal fiber integrity that can be used as biomarkers: both in animal [227] and human brain tissue [54], axial and radial diffusivities were related to underlying biology. Concurrently, the optimal acquisition parameters  $\mathbf{b}$  are determined to facilitate the diffusion parameter estimation. It has already been noted that when diffusivity information in multiple compartments is required, data need to be acquired at multiple  $\mathbf{b}$ -values [128]. We assume constant scanning time and uniformly distributed gradient directions. We will use the framework to determine the optimal dual tensor model. It will be shown that the axial diffusion of the two tensors involved must be constrained (taken identical), while an isotropic fraction can be simultaneously modeled by a single parameter. Moreover, it will be demonstrated that under these circumstances the FA of the both tensors can be independently estimated with high precision.

We adopt the Cramér-Rao lower bound (CRLB) on the variance of the dual tensor model parameters to determine in theory the optimal precision of the parameters. Monte Carlo simulations are performed to study the degree to which this precision is achieved in practice and to address a potential bias in the estimation. We will verify the findings by reconstructing FA-profiles along crossing fiber bundles.

The relevance of our work is in comparative studies where the goal is to assess subtle differences in diffusion shape between patients and matched controls. As an example, an increase in FA in the posterior limb of the internal capsule, as reported in a study of schizophrenia [66], might need re-interpretation due to possibly incorrect single tensor modeling in the presence of the crossing superior longitudinal fasciculus. A study of Amyotrophic Lateral Sclerosis reported decreased FA-values in the corticospinal tract, except at the crossing with the corpus callosum [221]. Preliminary work on estimation of diffusion shape parameters of a dual tensor model resulted in unstable results [13].

## 7.2 BACKGROUND

### 7.2.1 Diffusion modeling

We aim to quantify the independent diffusion profiles of two axonal bundles that appear as ‘crossing fibers’ within the typical resolution of DW-MRI (2 mm at a field strength of 3.0 T). A number of solutions has been proposed for these situations, mainly focusing on reconstructing fiber orientations. Two- or multi-tensor approaches [109, 159, 201], higher order tensors [194] or constrained spherical deconvolution [235] parametrically extend conventional diffusion tensor imaging. Non-parametric approaches have also emerged, including q-ball imaging [239]. Tractography results are becoming more accurate when performed in a dual tensor field [206, 171, 170] or in a fiber orientation distribution field [98].

Assuming the diffusion to be mono-exponential and Gaussian, the measured diffusion weighted signal  $S_j$  is initially modeled to contain a contribution of up to two fiber bundles and an isotropic part:

$$S_j = S_0 \left\{ \sum_{i=1,2} f_i \exp(-b_j \mathbf{g}_j^T \mathbf{D}_i \mathbf{g}_j) + f_{iso} \exp(-b_j D_{iso}) \right\}, \quad (7.1)$$

with  $S_0$  being the signal without diffusion weighting,  $b_j$  the amount of diffusion weighting corresponding to a gradient direction  $\mathbf{g}_j$ , selected from a vector  $\mathbf{b}$  of length  $n_b$  with unique components,  $\mathbf{D}_1$  and  $\mathbf{D}_2$  the  $(3 \times 3)$  diffusion tensors of the two fibers,  $D_{iso}$  the scalar amount of isotropic diffusion and  $f_i$  the volume fractions of the diffusion processes summing up to one, i.e.  $f_1 + f_2 + f_{iso} = 1$ .

An incorrect representation of the noise properties, particularly assuming a Gaussian instead of a Rician noise distribution in the DWIs, may also render an inappropriate (i.e. biased) signal model [96]. A single tensor model was extended by estimating the Rician noise level in a maximum likelihood framework [18, 161]. Rician noise reduction by spatial regularization [105, 40] can be used to limit the bias in estimates of a single tensor model. Recently, the Linear Minimum Mean Square Error (LMMSE) estimator [8] was proposed to decrease the noise in DWIs in an efficient way.

Previously, the Fisher Information matrix was used to assess the expected (co)variance of parameters of a single tensor model [18], which was also explored by means of an error propagation framework [157].

### 7.2.2 Optimization of diffusion imaging

The diffusion weighting parameter  $b$  selects the diffusivity scale of the measurements. The amount of diffusion weighting is determined by the diffusion weighting parameter  $b$ , which is calculated from the selected gradient strength and diffusion time  $\Delta - \frac{1}{3}\delta$ , where

$\Delta$  is the time between both gradients and  $\delta$  the diffusion weighting gradient duration [163].

The expected diffusivity values determine the range for which  $b$  is optimal. Measurements of the Apparent Diffusion Coefficients (ADC) in white matter show values in the range of  $0.25 \cdot 10^{-3}$  to  $1.5 \cdot 10^{-3} \text{ mm}^2\text{s}^{-1}$  [204]. At first glance the inverse<sup>1</sup> diffusion values indicate that  $b$ -values in the range of  $b = 0.7 \cdot 10^3 \text{ mm}^{-2}\text{s}$  to  $2.5 \cdot 10^3 \text{ mm}^{-2}\text{s}$  should be selected. For lower  $b$ -values, only isotropic diffusion between axonal boundaries is measured. Measuring at higher  $b$ -values ( $b > 3.0 \cdot 10^3 \text{ mm}^{-2}\text{s}$ ) will introduce a sensitivity to different compartments such as the myelin sheet [198], giving rise to restricted and hindered diffusion [25]. In the latter case, the Gaussian diffusion assumption is no longer valid, which is beyond the scope of this work.

For a broad range of applications (such as stroke), these diffusivity values are best determined with a  $b$ -value of  $1.0 \cdot 10^3 \text{ mm}^{-2}\text{s}$  [155], by means of a spin-echo EPI sequence with typically 6 to 45 gradient directions [221]. Multiple fiber orientations are best modeled at a higher diffusion weighting contrast using  $b = 2.0 \cdot 10^3 \text{ mm}^{-2}\text{s}$  [13, 41]. Multi-exponential signal decay emanating from multi-compartments in a study to Multiple Sclerosis was only visible at an even higher  $b$ -value of  $3.0 \cdot 10^3 \text{ mm}^{-2}\text{s}$  [26]. In order to independently estimate two diffusion profiles, a vector  $\mathbf{b}$  containing two distinct values ( $n_{\mathbf{b}} = 2$ ) must be used [128].

The DW-signal is increasingly attenuated with higher  $b$ -values, as is shown in Eq. 7.1, such that its signal-to-noise ratio ( $\text{SNR}_{\text{DW}}$ ) is optimal at low  $b$ -values ( $b < 0.5 \cdot 10^3 \text{ mm}^{-2}\text{s}$  [141]). For high values ( $b > 3.0 \cdot 10^3 \text{ mm}^{-2}\text{s}$ ) noise dominates the signal for all expected diffusion values. Concurrently, a low diffusion weighting implies poor contrast between low and high diffusivities. The contrast-to-noise ratio is thus expected to be optimal for intermediate  $b$ -values.

By minimizing the CRLB,  $\mathbf{b}$  was optimized for higher ratios of the maximum and minimum diffusion coefficients in the sample, modeled by a single tensor [187]. Maximal gradient strengths for varying axon diameters were found in a composite hindered and restricted model of diffusion, allowing for measuring direct tissue-microstructure features [12].

## 7.3 THEORY

### 7.3.1 Dual tensor model parameterization

The dual tensor model (Eq. 7.1) needs to be parameterized such that its estimated values reside in a well-defined range. We will parameterize  $\mathbf{D}_i$  by its eigenvalues  $\mathbf{E}_i$  and rotations  $\mathbf{R}_i$ ,  $\mathbf{D}_i = \mathbf{R}_i^T \mathbf{E}_i \mathbf{R}_i$ . The matrices  $\mathbf{R}_{i=1,2}$  describe rotations around the  $x$ -,  $y$ - and  $z$ -axes:  $\mathbf{R}_i(\alpha_{1-4}) = \mathbf{R}_x(\alpha_1) \mathbf{R}_y(\alpha_2) \mathbf{R}_z(\alpha_3 \pm \alpha_4)$ . The first two rotations  $\mathbf{R}_x(\alpha_1) \mathbf{R}_y(\alpha_2)$

<sup>1</sup> The naive inverse balances reported optimal values for  $b$  of  $0.85/D$  [143] and  $1.1/D$  [45].



determine the orientation of the plane in which the principal eigenvectors of both tensors reside;  $\mathbf{R}_z(\alpha_3 \pm \alpha_4)$  represents the mean orientation  $\alpha_3$  with deviations  $\pm\alpha_4$  of the principal eigenvectors of the two tensors  $i = 1, 2$ . The eigenvalues are contained in the diagonal matrices  $\mathbf{E}_i$ . The diffusion perpendicular to the fiber orientation is assumed to be isotropic such that its eigenvalues are identical [128], which reduces the number of free parameters. Let  $\lambda_{i\parallel}$  and  $\lambda_{i\perp}$  denote respectively the axial and perpendicular diffusion. Now  $\mathbf{E}_i$  yields:

$$\mathbf{E}_i = \text{diag}(\lambda_{i\parallel} \quad \lambda_{i\perp} \quad \lambda_{i\perp}), \quad i \in \{1, 2\}.$$

Non-negativity of diffusivity values is imposed by adopting the exponential mappings  $\exp(\cdot)$ . The model needs to be further constrained such that the eigenvalues are sorted, i.e.  $\lambda_{i\parallel} > \lambda_{i\perp}$ . For that purpose, the axial diffusivity  $\lambda_{i\parallel}$  is computed by adding the average perpendicular diffusivity  $\frac{1}{2}(\lambda_{1\perp} + \lambda_{2\perp})$  to a positive difference between axial and perpendicular diffusivity  $\exp \Delta \tilde{\lambda}_{i\parallel}$ . This formulation also allows us to impose an equality constraint on  $\lambda_{1\parallel}$  and  $\lambda_{2\parallel}$  (see Table 7.1 below). Volume fractions are defined in the range  $0 < f_i < 1$  by error functions  $\text{erf}(\cdot)$ , in which  $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt$ . All these constraints are implemented in the following parametrization:

$$\begin{aligned} \lambda_{i\perp} &= \exp \tilde{\lambda}_{i\perp} \\ \lambda_{i\parallel} &= \exp \Delta \tilde{\lambda}_{i\parallel} + \frac{1}{2}(\lambda_{1\perp} + \lambda_{2\perp}) \\ f_{\{1, iso\}} &= \frac{1}{2}(\text{erf}(\tilde{f}_{\{1, iso\}}) + 1) \\ f_2 &= 1 - f_1 - f_{iso} \\ D_{iso} &= \exp(\tilde{D}_{iso}). \end{aligned} \quad (7.2)$$

Employing this parameterization, the DW-signal (Eq. 7.1) is described by the 11-dimensional parameter vector  $\boldsymbol{\theta}$ :

$$\boldsymbol{\theta} = \{\Delta \tilde{\lambda}_{1\parallel}, \tilde{\lambda}_{1\perp}, \Delta \tilde{\lambda}_{2\parallel}, \tilde{\lambda}_{2\perp}, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \tilde{f}_1, \tilde{f}_{iso}, \tilde{D}_{iso}\}. \quad (7.3)$$

The Fractional Anisotropy  $\text{FA}_i$  [32] of an axially symmetric tensor is given by

$$\text{FA}_i = \sqrt{\frac{(\lambda_{i\parallel} - \lambda_{i\perp})^2}{\lambda_{i\parallel}^2 + 2\lambda_{i\perp}^2}} = \frac{\lambda_{i\parallel} - \lambda_{i\perp}}{\sqrt{\lambda_{i\parallel}^2 + 2\lambda_{i\perp}^2}}. \quad (7.4)$$

### 7.3.2 Maximum likelihood estimation

The DWIs contain the magnitude of the complex MR signal. It is assumed that the constituting components are independently affected by Gaussian noise, such that the result is Rician distributed [123]. Consequently, the probability density function of a

measured signal  $S_{m,j}$  in a certain gradient direction  $\mathbf{g}_j$  with a chosen diffusion weighting  $b_j$  is given by

$$p(S_{m,j}|S_{\theta,j}, \sigma) = \frac{S_{m,j}}{\sigma^2} \exp\left(-\frac{S_{m,j}^2 + S_{\theta,j}^2}{2\sigma^2}\right) I_0\left(\frac{S_{m,j}S_{\theta,j}}{\sigma^2}\right), \quad (7.5)$$

with  $S_{\theta,j}$  the true underlying value given the parameter vector  $\boldsymbol{\theta}$  (and  $\{\mathbf{g}_j, b_j\}$ ),  $\sigma$  the standard deviation of the noise and  $I_0$  the zeroth order modified Bessel function of the first kind. Since the DWIs are independent, the joint probability density function  $p(\mathbf{S}_m|\boldsymbol{\theta}, \sigma)$  of the signal profile  $\mathbf{S}_m$  in a voxel is given by the product of the marginal distributions for the measured signal  $S_m$  in each of the  $N_g$  diffusion weighted directions  $\mathbf{g}_j$

$$p(\mathbf{S}_m|\boldsymbol{\theta}, \sigma) = \prod_{j=1}^{N_g} p(S_{m,j}|\boldsymbol{\theta}, \sigma). \quad (7.6)$$

However, given measurements  $\mathbf{S}_m$  and unknown  $[\boldsymbol{\theta}, \sigma]$ , the function  $p(\mathbf{S}_m|\boldsymbol{\theta}, \sigma)$  is considered a likelihood function, indicated by  $L(\boldsymbol{\theta}, \sigma|\mathbf{S}_m)$ . When the DWIs are assumed to be Rician distributed, the likelihood function is given by

$$f(\boldsymbol{\theta}, \sigma) = \ln L(\boldsymbol{\theta}, \sigma|\mathbf{S}_m) = -2 \ln(\sigma) + \sum_{j=1}^{N_g} \left( \ln(S_{m,j}) - \frac{S_{m,j}^2 + S_{\theta,j}^2}{2\sigma^2} + \ln I_0\left(\frac{S_{m,j}S_{\theta,j}}{\sigma^2}\right) \right). \quad (7.7)$$

The Maximum Likelihood (ML) estimate of the parameters  $\boldsymbol{\theta}$  is obtained by maximizing the log likelihood function:  $\hat{\boldsymbol{\theta}}_{ML} = \arg \{ \max_{\boldsymbol{\theta}} (\ln L) \}$ .

For high signal to noise ratio (SNR =  $S_o/\sigma > 3$  [96]) the Rician distribution can be approximated by a Gaussian distribution. In that case the ML estimator reduces to a least-squares estimator.

The ML estimator has a number of favorable statistical properties. First, it can be shown that, under very general conditions, the ML estimator asymptotically achieves the Cramér-Rao Lower Bound (CRLB, introduced in the next section) i.e. for an increasing value of  $N_g$ . Second, it can be shown that the ML estimator is consistent, which means that it asymptotically converges to the true value of the parameter in a statistically well-defined way [242]. Whether these properties also apply to a small number of observations can be assessed by experiments on synthetic (noise corrupted) data.

### 7.3.3 Cramér-Rao analysis for model selection and optimizing $\mathbf{b}$

Any chosen model and its parameterization  $\boldsymbol{\theta}$  should be such that the parameters can be estimated precisely. Simultaneously, the acquired data should support the model to be fitted: the model, as described in section 7.2.1, is to be selected with respect to

the precision in the estimation of  $\theta$ . The dimensionality of  $\theta$  should be constrained in case the precision in specific parameters is low. In addition, the diffusion weighting  $b$  influences the model fit (as argued in section 7.2.1) and needs to be carefully selected. For these purposes, we will adopt a Cramér-Rao analysis. The Cramér-Rao Lower Bound (CRLB) is the theoretical lower bound on the variance of any unbiased estimator of the parameters.

The Fisher information matrix is a measure of the amount of information about  $\theta$  that is present in the data. It is derived from the probability density function of the data [97, 225],

$$\mathbf{I} = -\mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{S}_m | \theta)}{\partial \theta \partial \theta^T} \right], \quad (7.8)$$

where  $\mathbb{E}$  denotes the expectation value. Diagonal elements of the inverse of the Fisher Information matrix  $\mathbf{I}^{-1}$  reflect the variance of parameters, whereas off-diagonal elements relate to covariance between parameters. Effectively, the CRLB is a function for the curvature of the cost function, given by  $\mathbf{I}^{-1}$ . The CRLB-inequality yields [96]:

$$\text{Cov}(\tau, \tau) \geq \text{CRLB} = \left( \frac{\partial \tau}{\partial \theta} \right) \mathbf{I}^{-1} \left( \frac{\partial \tau}{\partial \theta} \right)^T. \quad (7.9)$$

Here  $\left( \frac{\partial \tau}{\partial \theta} \right)$  is the Jacobian of a potential transformation  $\tau(\theta)$  of the parameter vector to a measure of interest, as explained in the appendix.

The CRLB implies that no unbiased estimator can be found for which the variance of the estimated parameters is lower than the diagonal elements of the CRLB matrix [242]. This inequality does not directly bound the off-diagonal elements, which specify the covariance between two parameters. It may be shown however, that the covariance of two ML-estimated parameters  $\text{cov}(\hat{\tau}_{m\text{ML}}, \hat{\tau}_{n\text{ML}})$  will converge to the corresponding off-diagonal element  $\text{CRLB}_{mn}$  when the number of samples goes to infinity [242].

The CRLB is quadratic in the parameters and is thus not directly comparable to the parameter values. We examine, therefore, the dimensionless relative CRLB (rCRLB), defined by,

$$\text{rCRLB}_{mn} = \sqrt{\frac{|\text{CRLB}_{mn}|}{\bar{\tau}_m \bar{\tau}_n}}. \quad (7.10)$$

where the indices  $m, n$  denote the respective matrix element of CRLB and  $\bar{\tau}_{m,n}$  the true parameter value.

The type of noise that affects the measurements needs to be properly addressed in the Fisher Information Matrix. For the Rician noise model the Fisher information matrix is given by [148],

$$\mathbf{I}_R = \frac{1}{\sigma^2} \sum_{j=1}^{N_g} \left( -\frac{S_{\theta,j}^2}{\sigma^2} + \mathbb{E} \left[ \left( \frac{\mathbf{S}_{\theta,j} I_1(z_j)}{\sigma I_0(z_j)} \right)^2 \right] \right) \cdot \frac{\partial S_{\theta,j}}{\partial \boldsymbol{\theta}} \frac{\partial S_{\theta,j}}{\partial \boldsymbol{\theta}^T}, \quad (7.11)$$

with  $z_j = \frac{S_{m,j} S_{\theta,j}}{\sigma^2}$  and  $I_1$  a modified Bessel function of the first kind. The expected value can be calculated numerically [148]. For high SNR, the noise distribution can be approximated by a Gaussian. Eq. 7.11 then simplifies to

$$\mathbf{I}_G = \frac{1}{\sigma^2} \sum_{j=1}^{N_g} \left( \frac{\partial S_{\theta,j}}{\partial \boldsymbol{\theta}} \frac{\partial S_{\theta,j}}{\partial \boldsymbol{\theta}^T} \right). \quad (7.12)$$

## 7.4 EXPERIMENTS

First, we build the dual tensor model by analyzing which model parameters can be sufficiently precisely estimated and which need to be constrained (section 7.4.1). Second, the optimal  $\mathbf{b}$ -values are determined for the chosen diffusion model in which we consider both a ‘homogeneous’ and a ‘heterogeneous’ sample, the latter containing a distribution of parameter values (section 7.4.2). Cramér-Rao and Monte Carlo analyses allow us to do so. Third, in brain data FA-profiles are derived along fibers tracked through a crossing of white matter tracts in data of varying quality using different model parameterizations (section 7.4.3).

### 7.4.1 Model selection

We aim to build a dual tensor model such that its parameters can be estimated both precisely and accurately given a specific SNR and the number of diffusion encoding directions (see below). Different models of increasing complexity are proposed in Table 7.1. For these models, we study how the axial and perpendicular diffusion parameters of the tensors need to be constrained to be estimated with low error. Also, the parameterization of the isotropic compartment is studied. Including such a compartment, for a single tensor model, yields a degenerated problem [203].

#### *Maximal precision: Cramér-Rao analysis*

The relative Cramér-Rao Lower Bound (rCRLB) is defined in Eq. 7.10. The upper limit for the rCRLB to be maintained throughout this paper is 15%. This value is  $\sim \sqrt{2}$  times

Table 7.1: Dual tensor models with increasing degrees of freedom. Parameters may be estimated ('+'), constrained to other parameters (e.g.  $\lambda_{2\parallel} = \lambda_{1\parallel}$ ), set constant, or omitted ('o'). In each model the noise level  $\sigma$  is set to  $\sigma_{\#2}$ , the value estimated with model #2. The units of the diffusion values are  $\{\lambda, D\} \cdot 10^{-3} \text{ mm}^2 \text{ s}^{-1}$ . Mean  $\bar{\theta}$  and standard deviation of the parameter values  $\text{std}(\theta)$  used in the experiments, in which the angles  $\alpha_{1-3}$  are randomly chosen (rnd).

model	$\lambda_{1\parallel}$	$\lambda_{1\perp}$	$\lambda_{2\parallel}$	$\lambda_{2\perp}$	$\alpha_{1-3}$	$\alpha_4$	$f_1$	$f_{iso}$	$D_{iso}$	$\sigma$
#1	+	+	+	+	+	+	+	o	o	$\sigma_{\#2}$
#2	+	+	$\lambda_{1\parallel}$	+	+	+	+	o	o	+
#3	+	+	$\lambda_{1\parallel}$	+	+	+	+	+	3.0	$\sigma_{\#2}$
#4	+	+	$\lambda_{1\parallel}$	+	+	+	+	+	+	$\sigma_{\#2}$
$\bar{\theta}$	1.4	0.4	1.4	0.3	rnd	$\frac{\pi}{5}$	0.4	0.15	3.0	
$\text{std}(\theta)$	0.2	0.07	0.2	0.07	rnd	$0.1 \cdot \frac{\pi}{5}$	0.05	0.03	o	

larger than the 10% limit by which *single* tensor model parameters are estimated [18], since the length of our parameter vector is approximately twice that of [18].

The CRLB depends on the hypothetical true values of the parameters and as such representative parameters need to be defined. To determine representative diffusion values, we selected the anisotropic voxels having  $\text{FA} > 0.5$  and  $\text{MD} > 1 \cdot 10^{-3} \text{ mm}^2 \text{ s}^{-1}$  after a single tensor fit in a randomly selected dataset. The mean (diffusion) value and corresponding standard deviation were determined for the largest and smallest eigenvalues  $\lambda_1$  and  $\lambda_3$  ( $\lambda_2$  may be biased in crossing fiber voxels), and assigned to  $\lambda_{1,2\parallel}$  and  $\lambda_{1\perp}$ . Notice that marginally perturbed values were chosen for  $\lambda_{1,1}$  and  $\lambda_{1,2}$ . The volume fraction of the first tensor is  $f_1 = 0.4$ . The isotropic compartment is taken small ( $f_{iso} = 0.15$ ), while the isotropic diffusion approximates the diffusivity of free water at 37 °C:  $D_{iso} = 3 \cdot 10^{-3} \text{ mm}^2 \text{ s}^{-1}$ . The representative parameters  $\bar{\theta}$  are given in Table 7.1.

A **b**-vector of length two,  $\mathbf{b} = [ 1.0 \quad 3.0 ] \cdot 10^3 \text{ mm}^{-2} \text{ s}$ , is chosen, the selection of these values is addressed in the next section. The number of gradient directions  $N_q$  is assumed constant: for each unique element of the **b**-vector an identical series of 92 three-fold tessellated icosahedric gradient directions is used. We adopt a modest  $\text{SNR}=25$ .

The calculated rCRLB-matrices for the four models (see Table 7.1) are depicted in figure 7.1(a). The Cramér-Rao analysis reveals that even if  $f_{iso}$  and  $D_{iso}$  are constant  $\lambda_{1,2\parallel}$  cannot be estimated independently for the two tensors (model #1): the values on the diagonal indicate  $\text{rCRLB} > 15\%$  for all parameters except  $\alpha_4$ . We hypothesize that the imprecision in  $\lambda_{1,2\parallel}$  emanates from the high diffusion and hence the low signal measured along the axial orientation. In model #2  $\lambda_{1\parallel}$  and  $\lambda_{2\parallel}$  are set to be equal, while  $\lambda_{1\perp}$  and  $\lambda_{2\perp}$  are still included in the model. The second image in figure 7.1 (a) shows that the constraint enables precise estimation of the remaining parameters. What is

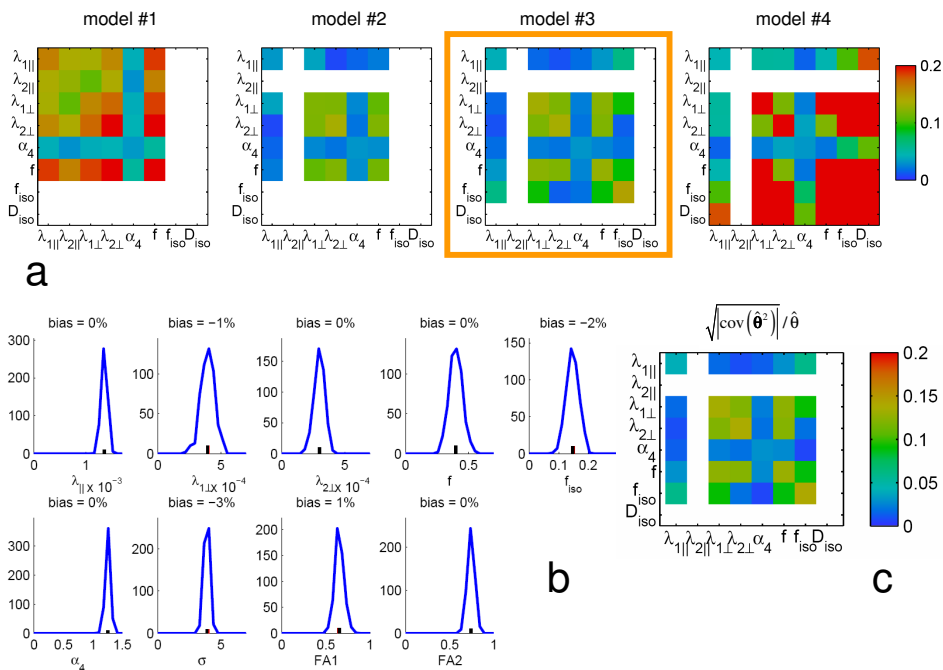


Figure 7.1: (a) rCRLB-matrices for models involving different constraints (c.f. Table 7.1). The diagonals of these matrices contain the relative Cramer-Rao lowest bounds. Model #3 is consistently used in our further analyses. Values above 0.2 are trimmed. (b) Distributions of estimated parameters determined in a Monte-Carlo experiment involving model #3. The small black bars (at the bottom of the graphs) correspond to the true parameter values; relative biases are annotated. (c) Square root of the absolute covariance matrix emanating from the Monte Carlo experiment. The diffusion weighting was  $\mathbf{b} = [1.0 \ 3.0] \cdot 10^3 \text{ mm}^{-2}\text{s}$ .

more, we find that constraining  $\lambda_{1,2\parallel}$  even allows for estimating  $f_{iso}$  with sufficient precision (rCRLB < 15%) as shown by the third image (model #3): adding  $f_{iso}$  does not significantly increase the CRLB of other parameters. The fourth image (model #4) confirms that  $f_{iso}$  and  $D_{iso}$  cannot be estimated simultaneously.

#### *Estimated accuracy and precision: Monte Carlo simulation*

Additionally, a Monte Carlo simulation is performed to study whether the theoretical lower bounds on the precision can be practically reached by the ML-estimator. Moreover, a potential bias and the uniqueness of the obtained ML-estimation are assessed. Parameter values are directly substituted into the diffusion model equation to generate signal values (Eq. 7.1). Subsequently, 500 of such simulated measurements are distorted by Rician noise. The ensembles of estimated parameter values for these measurements are expected to be Gaussian distributed [74].

The two-tensor model was implemented in Matlab (The MathWorks, Natick, MA). Levenberg-Marquardt optimization was used in all parameter estimations. The two-tensor model was initialized based on a single tensor estimation with three (sorted) eigenvalues  $\lambda_{1\dots 3,s}$  and a rotation matrix  $\mathbf{R}_{\text{single}}$  (without  $\alpha_4$ -term). We initialized anisotropic tensors by  $\lambda_{1,2\parallel\text{init}} = \lambda_{1s} + \lambda_{2s}$ ,  $\lambda_{1,2\perp\text{init}} = \lambda_{3s}$ , and with opening angle  $\alpha_{4,\text{init}} = \arctan(\lambda_{2s}/\lambda_{1s})$ . The initial volume fractions were  $f_{1,\text{init}} = 0.4$  and  $f_{iso,\text{init}} = 0.2$ .

A Monte-Carlo experiment involving model #3 rendered the histograms of the estimated parameter values depicted in figure 7.1(b). It shows that the distributions of the estimated parameters in each experiment are unimodal, while the bias is negligible (the maximal bias is -3% in  $\sigma$ ). This proves both a good initialization and the uniqueness of the solution.

The estimated relative standard deviations in the Monte-Carlo experiment approximate the rCRLBs: see figure 7.1(c). All estimated precisions were within 10% above the corresponding rCRLBs.

#### *7.4.2 Optimizing $\mathbf{b}$ for different tissue configurations*

We aim to find the optimal  $\mathbf{b}$  for estimating the diffusion shape of two crossing fiber tracts using the previously selected model.  $S_0$  images (no diffusion weighting involved) are assumed to be acquired in all cases. Since the diffusion parameters are to be estimated with sufficient precision for two compartments, a vector with at least two supplementary non-zero values  $\mathbf{b}$  is needed [128], in order to prevent the Fisher Information matrix from being rank deficient. Further, a larger vector  $\mathbf{b}$  with a high range of values may be needed to adequately capture the heterogeneous diffusion in the data, as explained in section 7.2.2. The experiments for optimal  $\mathbf{b}$  selection focus on model #3 (independent perpendicular diffusion  $\lambda_{1,2\perp}$  and  $f_{iso}$  included).

### *Single configuration*

We will first determine the optimal  $\mathbf{b}$  for a single configuration of crossing fibers  $\bar{\theta}$ , see Table 7.1. The selected model (see section 7.4.1) is adopted, and rCRLB-values will be computed for a two-element  $\mathbf{b}$ -vector. To minimize the rCRLB we also explore the high  $b$ -value range, although in section 7.2.2 we argued that Gaussian diffusion cannot be assumed beyond  $b > 3.0 \cdot 10^3 \text{ mm}^{-2}\text{s}$ . The optimum will be determined for SNR=25.

Figure 7.2(a) presents the rCRLB for the mean parameter vector  $\bar{\theta}$  (Table 7.1) for varying supplementary  $\mathbf{b}$ -values (along the axes). Note that values on the diagonals effectively involve merely one  $b$ -value. A detailed inspection of the results presented in figure 7.2(a) shows that the  $\text{rCRLB}_{\text{FA}_{1,2}}$  are minimal at  $\mathbf{b} = [ 0.75 \dots 1.0 \quad 3.0 \dots 4.25 ] \cdot 10^3 \text{ mm}^{-2}\text{s}$ , with  $\text{rCRLB}_{\text{FA}_1} = 6 \dots 7\%$ . The rCRLB of all parameters remain close to their optimal rCRLB around these  $\mathbf{b}$ -values.

### *Varying perpendicular diffusion*

Then, the optimal  $\mathbf{b} = [ b_1 \quad b_2 ]$  is determined for varying perpendicular diffusion values  $\lambda_{1,2\perp}$  involved in the model configuration  $\bar{\theta}$ . rCRLB-values are computed for  $\lambda_{1,2\perp} = [ 0.3 \quad \dots \quad 0.8 ] \cdot 10^{-3} \text{ mm}^2\text{s}^{-1}$ . Simultaneously, the  $b_2$ -value is varied, taking the  $b_1$ -value fixed at  $b_1 = 1.0 \cdot 10^3 \text{ mm}^{-2}\text{s}$ . If we observe significant variation in rCRLB at the location of the minimum, this can be considered to be an indication that multiple  $b_{1,2}$ -values are needed to model the entire diffusion range with low error.

The rCRLBs under these circumstances are depicted in Figure 7.2(b). The rCRLB at optimal  $b_2$  are not significantly different from the rCRLB at constant  $b_2 = [ 2.5 \quad \dots \quad 3.0 ] \cdot 10^3 \text{ mm}^{-2}\text{s}$ , which is about the optimum for the entire parameter vector. Only  $\text{rCRLB}_{\lambda_{1,2\perp}}$  slightly improves from 14% ( $b_2 = 3.0 \cdot 10^3 \text{ mm}^{-2}\text{s}$ ) to 10% ( $b_2 = 5.0 \cdot 10^3 \text{ mm}^{-2}\text{s}$ ) for a low diffusion value  $\lambda_{1,2\perp} = 0.3 \cdot 10^{-3} \text{ mm}^2\text{s}^{-1}$ .

### *'Heterogeneous' sample*

Third, we looked into the number of values of  $\mathbf{b}$  needed to measure a 'heterogeneous' sample precisely and accurately, by considering a series of different crossing fiber configurations. The mean parameter values  $\bar{\theta}$  and corresponding standard deviations in Table 7.1 are at the basis to generate 100 Gaussian distributed parameter values.

In all these experiments, the total number of DWIs stays constant. This is realized by randomly assigning an equal number of gradient directions to each supplementary



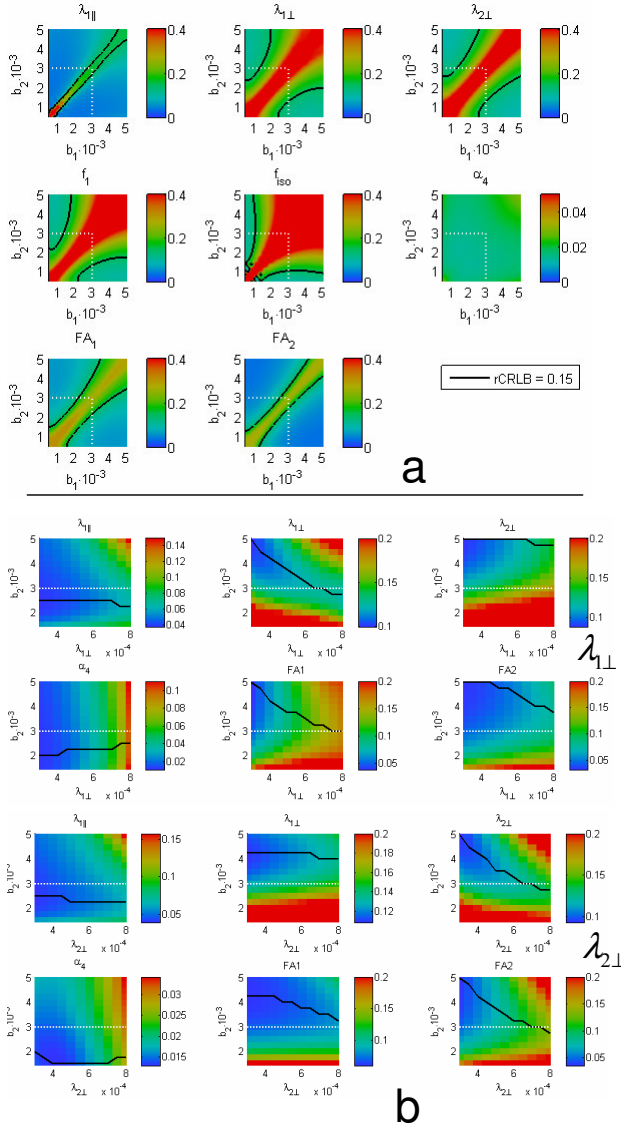


Figure 7.2: (a) rCRLB-values for different parameters as a function of  $b_{1,2}$ , SNR=25. Values above 0.4 are trimmed. (b) The configurations in which the rCRLB is determined are based on  $\hat{\theta}$ . (Table 7.1) while adjusting  $\lambda_{1\perp}$  and  $\lambda_{2\perp}$  respectively. Along the vertical axis the  $b_2$ -value is given, maintaining another non-zero  $b$ -value fixed at  $b_1 = 1.0 \cdot 10^3 \text{ mm}^{-2}\text{s}$ . (top) Horizontally  $\lambda_{1\perp}$  is varied while  $\lambda_{2\perp}$  is constant. (bottom)  $\lambda_{2\perp}$  is varied while  $\lambda_{1\perp}$  is constant. Values above 0.2 are trimmed. The solid line in the figures give the  $b_2$ -value at which the rCRLB is minimal. The white dashed line denotes the upper-boundary of the region where the Gaussian diffusion assumption holds (see text).

Table 7.2:  $\mathbf{b}$ -vector yielding a minimum mean rCRLB on the FA and a minimum mean rmse on the FA, on configurations drawn from Table 7.1. Different  $b$ -value configurations, as described in section 7.4.2, were used. Note that in the Cramér-Rao analysis a sub-optimal result within the Gaussian diffusion range (\*) is included.

Cramér-Rao analysis				Monte-Carlo simulation			
rCRLB <sub>FA</sub>	$\mathbf{b} \cdot 10^3 \text{ mm}^{-2}\text{s}$			RMSE(FA)	$\mathbf{b} \cdot 10^3 \text{ mm}^{-2}\text{s}$		
0.26	5.0			0.23	3.0		
0.065	0.75	4.5		0.074	1.0	3.5	
0.073*	0.75	3.0*					
0.064	0.75	4.25	4.5	0.074	0.5	1.0	3.5
0.065	0.75	0.75	4.0 4.75	0.074	0.5	0.5	1.0 3.5

$b$ -value. Optimizing this distribution over  $\mathbf{b}$  is beyond the scope of this paper.  $b$ -values in the range of  $[ 1.0 \dots 5.0 ] \cdot 10^3 \text{ mm}^{-2}\text{s}$  are considered. Note that Gaussian diffusion can be safely assumed only if  $b < 3.0 \cdot 10^3 \text{ mm}^{-2}\text{s}$  [25].

The mean rCRLB<sub>FA</sub> (averaged for both tensors) is determined using all 100 configurations for an increasing length of  $\mathbf{b}$ . The selected  $\mathbf{b}$  and the corresponding mean rCRLB are given in Table 7.2. The table shows that the mean rCRLB remained constant when employing a  $\mathbf{b}$  with more than two values. This experiment shows that a  $\mathbf{b}$  of length two is also sufficient to reach the smallest rCRLB over a range of expected parameter values.

Additionally, a Monte-Carlo experiment is performed by simulating measurements of these configurations that are distorted by Rician noise to yield 100 noise realizations per parameter configuration. The bias, standard deviation and root mean squared error (RMSE) in FA and MD (averaged for both tensors) are computed for each combination of  $\mathbf{b}$ .

The outcome of the Monte-Carlo experiment is given in figure 7.3. This figure shows that the optimal value of  $\mathbf{b}$  to be chosen for a heterogeneous sample is  $\mathbf{b} = [ 1.0 \ 3.5 ] \cdot 10^3 \text{ mm}^{-2}\text{s}$ , because for these values of  $\mathbf{b}$  the RMSE in both FA and MD is minimal. The RMSE balances between a minimum in the bias at a lower  $\mathbf{b} = [ 1.0 \ 2.5 ] \cdot 10^3 \text{ mm}^{-2}\text{s}$  and the standard deviation at a slightly higher  $\mathbf{b} = [ 1.0 \ 4.0 ] \cdot 10^3 \text{ mm}^{-2}\text{s}$  respectively.  $n_{\mathbf{b}} = 2$  is the optimal number of elements in  $\mathbf{b}$ , since RMSE(FA) remains constant for  $n_{\mathbf{b}} > 2$ . The relative standard deviation closely approaches the rCRLB determined in Table 7.2. Also, notice the high standard deviation along the diagonal (reflecting the usage of merely one  $b$ -value), which is in accordance with the Cramér-Rao analysis.

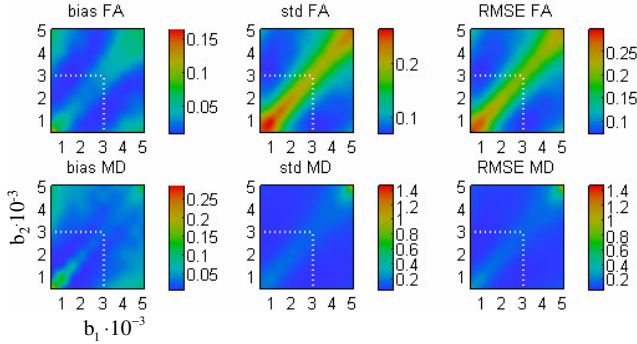


Figure 7.3: Bias in the estimated parameter value (left), the corresponding standard deviations (middle) and the root mean squared error, all averaged over both tensors, on simulated measurements (drawn from the distributions indicated in Table 7.1). The axes correspond to the  $b$ -values at which this Monte-Carlo experiment was performed. The white dashed line denotes the upper-boundary of the region where the Gaussian diffusion assumption holds.

### 7.4.3 Validation in brain data

A quantitative validation of estimated diffusion properties within fiber crossings in the human brain is complicated by the lack of a ground truth. To the best of our knowledge, a phantom with known diffusion properties (apart from fiber orientations [236]) is unavailable. We have chosen to track fibers through a crossing fiber region, where gyral projections of the corpus callosum (commissural fibers) are crossing the corticospinal tract. FA-profiles along the tracts will be compared for data acquired with different  $\mathbf{b}$ .

#### Data acquisition and preprocessing

Diffusion Weighted Image acquisition of 20 axial slices (mid-brain) of a healthy volunteer was performed on a Philips Intera 3.0 Tesla MRI scanner (Philips Intera, Philips Healthcare, Best, The Netherlands) by means of a spin echo EPI-sequence. An imaging matrix of  $112 \times 110$  was used, from which a  $128 \times 128$  sized slices were reconstructed, with a voxel size of  $1.7 \times 1.7 \times 2.2 \text{ mm}^3$ . An echo time of  $TE = 84 \text{ ms}$  and repetition time of  $TR = 3800 \text{ ms}$  were used. The diffusion weighting was along 92 three-fold tessellated icosahedric gradient directions, with five  $b$ -values:  $\mathbf{b} = [ 1.0 \ 1.4 \ 1.8 \ 2.2 \ 3.0 ] \cdot 10^3 \text{ mm}^{-2}\text{s}$ . Per  $b$ -value, one non-diffusion weighted image  $S_0$  was acquired. 20 axial slices were acquired, resulting in a total scanning time of 30 minutes. Deformation induced by eddy currents was corrected for using an affine registration in the phase encoding direction [173]. In addition, a rigid registration of the  $S_0$ -images and coregistration of the DWIs corrected for patient head motion (up to 2 voxels). One average  $S_0$ -image was computed and used in all experiments with subsets of  $b$ -values. Subse-

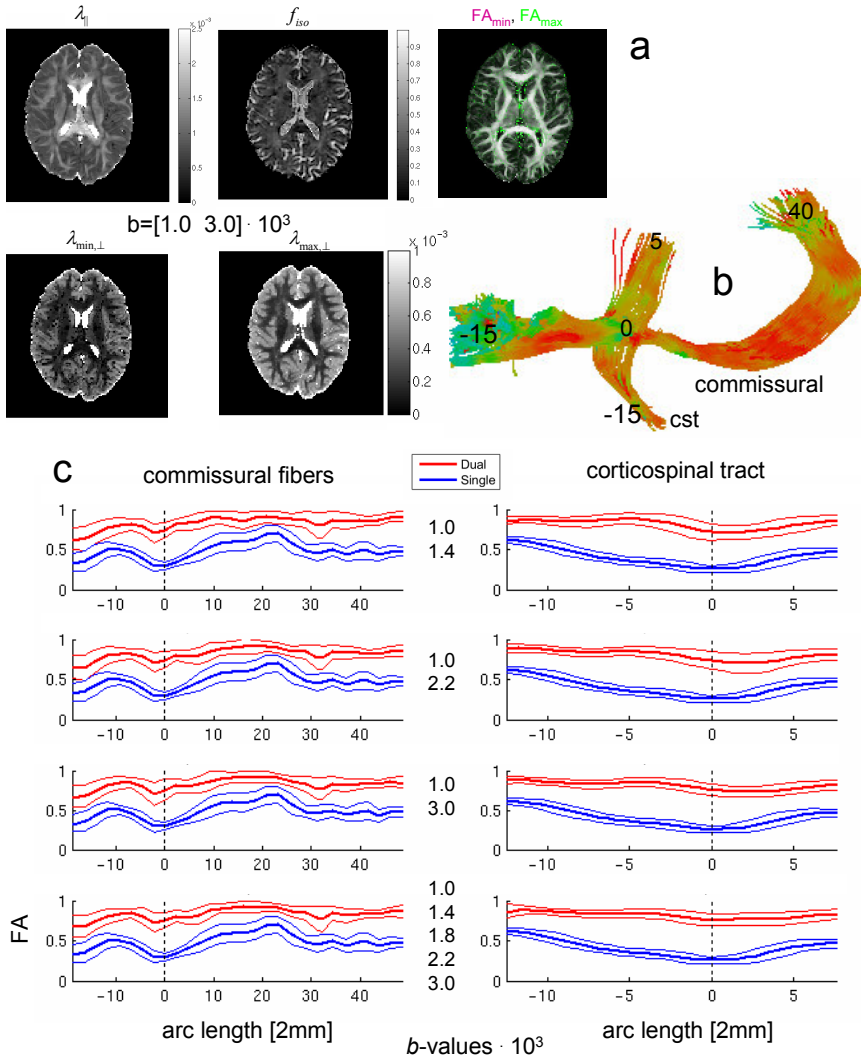


Figure 7.4: Results in brain data. (a)  $\lambda_{\parallel}$ ,  $\lambda_{\min,\max,\perp}$ ,  $f_{iso}$  and  $FA_{\{\min,\max\}}$  estimated in one axial slice,  $\mathbf{b} = [1.0 \ 3.0] \cdot 10^3 \text{ mm}^{-2}\text{s}$ .  $FA_{\{\min,\max\}}$  are colored magenta and green respectively, such that white colored voxels reflect  $FA_{\min} = FA_{\max}$  and green voxels  $FA_{\min} < FA_{\max}$ . (b) Tracked fibers: the corticospinal tract (CST) and commissural fibers. Color denotes the FA. (c) Mean FA-profiles (in bold) with standard deviations calculated through dual (red) and single (blue) tensor fits as a function of the arc length (in units of 2 mm) of the commissural fibers (left) and the corticospinal tract (right). The data was acquired at different  $b$ -value configurations as annotated. The crossing is indicated with a dashed line and is considered the origin.

quently, Rician noise in the data was reduced using an adaptive Linear Minimum Mean Square Estimator (LMMSE) [8, 68]. This filtering approach uses an estimate of the noise level  $\sigma$ .  $\sigma$  is estimated per voxel by fitting a dual tensor model which does not include an isotropic compartment: model #2 from Table 7.1.

### *Consistent parameter estimation*

From the acquired DWIs, different data sets with a two-element  $\mathbf{b}$ -vector were generated. From the antipodally symmetric set of gradient-directions, one gradient direction was randomly chosen and assigned to the first  $b_1$ -value. Simultaneously, the antipodal gradient direction was assigned to the second  $b_2$ -value, resulting in 46 gradients per  $b$ -value. In this way, the angular resolution over all permutations of  $\mathbf{b}$  was equal. In addition, the entire scanned dataset - involving a five-element  $\mathbf{b}$ -vector and all DWIs - was used to assess an improved FA-estimation when sampling a broader range of diffusion values. The dual-tensor model was fit to all created datasets. The estimated perpendicular diffusivity and  $FA_{1,2}$ -values are sorted into  $\lambda_{\min,\perp}$  and  $\lambda_{\max,\perp}$  and  $FA_{\{\min,\max\}}$  respectively.

Model #3 was estimated taking constant  $b_1 = 1.0 \cdot 10^3 \text{ mm}^{-2}\text{s}$  and varying  $b_2$ -values. Figure 7.4(a) shows axial slices depicting the parameter values estimated on data involving  $\mathbf{b} = [ 1.0 \quad 3.0 ] \cdot 10^3 \text{ mm}^{-2}\text{s}$ . Notice that there is high inter-voxel ‘consistency’. The  $\lambda_{\parallel}$  image displays good contrast between white and gray matter, whereas the  $f_{iso}$ -image mainly highlights the CSF. Isolated black voxels may be observed in the  $\lambda_{\min,\perp}$  image. These correspond to samples containing only a single anisotropic compartment. Here  $\lambda_{\min,\perp}$  and  $FA_{\{\min,\max\}}$  are undefined since the volume fraction  $f$  of the corresponding tensor approximates zero.

We adopted a grid approach to perform the experiments using the infrastructure of the Dutch Virtual Laboratory for Medical Image Analysis [192]. Thus, computing times were reduced from one week (estimated) to six hours, a 28:1 improvement in speed.

### *FA-profiles along crossing fiber tracts*

Fiber tracking was performed using High Angular Fiber Tracking [237]. As an extension to common streamline tracking algorithms, this method generates branches of fibers if, within a voxel, the angle between both tensors exceeds  $12.5^\circ$ . The curvature threshold was set to  $20^\circ$  per voxel, and the single tensor FA threshold yielded  $FA_{\text{single}} > 0.1$ . One seeding Region of Interest (ROI) and one additional ROI through which fibers should pass were defined. Fibers were tracked through the crossing of the corticospinal tract with gyral projections of the corpus callosum (commissural fibers). The corticospinal tract was only partially tracked, due to the limited scanning region; it was confirmed

that the crossing region was completely present in the data, though. Two ROIs per tract sufficed to successfully track the fibers.

During tracking, the most closely aligned tensor per voxel is selected. Along the tracked fibers, FA-profiles are computed using the aligned tensors. The mean and standard deviation of the FA in voxels through which fibers traverse are estimated in the plane perpendicular to the centerline of the tracts. At the location of the fiber crossing, oblate single tensors are expected whose FA is biased, i.e. lower than the FA-values of the prolate dual tensors. Dual-tensor FA-profiles are higher (i.e. less biased) than their single-tensor counterparts at the location of the fiber crossing. An additional global offset in the FA may be apparent due to the included isotropic fraction in the dual tensor model. In addition, there should be spatial consistency in FA in adjacent single- and dual-tensor regions.

Figure 7.4(b) shows the corticospinal and commissural fibers tracked in this data. The FA-profiles along those tracts are displayed in figure 7.4(c) in which we estimate parameters for two models: a single tensor model and dual tensor model #3. At the position of the crossing, the bias of the single tensor model is most distinct. It can also be seen that dual tensor FA-profiles are consistent at the transition of a single to a dual orientation region, whereas single tensor profiles show a reduction in FA. Additionally, the standard deviation of the estimated FA slightly decreases for higher  $b_2$ -values, particularly in the corticospinal tract, and is minimal for  $\mathbf{b} = [ 1.0 \quad 3.0 ] \cdot 10^3 \text{ mm}^{-2}\text{s}$ . Notice that the dual tensor FA-profiles are systematically higher than the single tensor profiles (also remote from the crossing). This effect emanates from the inclusion of an isotropic component in the model. Finally, estimating the model with a five-element  $\mathbf{b}$ -vector renders only marginally different profiles.

## 7.5 DISCUSSION

A framework was presented for model selection and optimization of acquisition parameters for estimating diffusion properties in fiber crossings. A dual tensor model was studied for which the optimal diffusion weighting  $\mathbf{b}$  was determined. The framework is directly applicable to other parameterized diffusion models and different experimental conditions.

### 7.5.1 Model selection

The Cramér-Rao analysis showed that at a modest SNR the perpendicular diffusion  $\lambda_{1,2\perp}$  can be precisely estimated in model #3, whereas the axial diffusions  $\lambda_{1,2\parallel}$  cannot be independently modeled, since they result in a large rCRLB in model #1. This may be explained by higher DW-signal attenuation along the axial orientation, which has a higher diffusion and thus a lower  $\text{SNR}_{\text{DW}}$ . In other words, our inability to independently estimate  $\lambda_{1,2\parallel}$  is a limitation of the DW-MRI modality under the described circumstances. The strength of our work is in independently capturing myelin degradation

in crossings. Histology in ischemic mice confirms myelin degradation and because of that increased mobility of the water molecules perpendicular to the axons, measured by DW-MRI as an increased  $\lambda_{\perp}$  [227].

An isotropic compartment was modeled by a single parameter in model #3. In fact, it was shown that adding the isotropic compartment to model #2 does not reduce the precision of the other parameters. Simultaneously estimating  $f_{iso}$  and  $D_{iso}$  led to a degeneration in model #4 (figure 7.1(a)).

In previous work, particularly the angular resolution attracted attention. An angle of  $\pi/4$  between the tensors could be accurately determined by a model that only estimated orientations, (the experiment involved a single  $b = 2000 \text{ mm}^{-2}\text{s}$ ) [41]. Q-ball imaging and the Diffusion Orientation Transform also achieved an angular resolution of  $\pi/4$  [205]. The experiments that we have performed involve a lower angular resolution ( $2 \cdot \alpha_4 = \pi/2.5$ ) since our main interest is to ensure a low error in the estimated diffusivity parameters. Lower values of  $\alpha_4$  will e.g. result in  $\text{rCRLB}_{\text{FA}} > 0.15$  under equal measurement conditions.

The Monte Carlo simulation showed that the estimator converged to a unique solution. Moreover, the observed precisions were close to the CRLB, which demonstrates a good estimator performance. Simultaneously estimating the diffusion ‘size’ and ‘shape’ parameters was previously reported to be a difficult problem [13]. The eigenvalue parameterization (Eq. 7.2) prevents degeneration of the parameter estimation.

The dual tensor model (#3) was estimated independently per voxel. In other work, the spatial continuity of the parameters was included as a constraint [197]. The maximal precision to be achieved in this approach might also be quantified by means of our Cramér-Rao framework. The framework is directly applicable to multi-tensor approaches [109, 159, 201], higher order tensors [194], constrained spherical deconvolution [235] or a series expansion of the PDF made of Gaussian-Laguerre and Spherical Harmonics functions [28], which all are parametric models. Additionally, the ML-approach could be extended to a Bayesian estimation scheme by incorporating prior distributions on parameters [41].

### 7.5.2 Optimizing $\mathbf{b}$ for different tissue configurations

The Cramér-Rao analysis also showed that a two-element  $\mathbf{b}$ -vector is needed for independently estimating axial and perpendicular diffusion as well as an isotropic compartment (model #3, figure 7.2(a)). This result confirms previous statements [128]. A minimal  $\text{rCRLB}_{\text{FA}_1} = 6.5\%$  was found for  $\mathbf{b} = [ 0.75 \quad 4.5 ] \cdot 10^3 \text{ mm}^{-2}\text{s}$ , which only marginally raised  $\text{rCRLB}_{\text{FA}_1} = 7.5\%$  for  $\mathbf{b} = [ 0.75 \quad 3.0 ] \cdot 10^3 \text{ mm}^{-2}\text{s}$  within the Gaussian diffusion range. In the Monte Carlo analysis, the RMSE was minimal at a lower  $\mathbf{b} = [ 1.0 \quad 3.5 ] \cdot 10^3 \text{ mm}^{-2}\text{s}$ , due to an increasing bias at higher  $\mathbf{b}$ -values. Notice that the precisions were well below the maximally allowed value, e.g.  $\text{rCRLB}_{\text{FA}} = 6.5\% < 15\%$ .

We observed that the minimum in rCRLB and RMSE as a function of  $\mathbf{b}$  is rather ‘shallow’ (see e.g. figure 7.2). Accordingly, the results on brain data were fairly robust with respect to modification of  $b_2$ . A significantly higher bias in the FA for low  $b$ -values ( $b = 1.0 \cdot 10^3 \text{ mm}^{-2}\text{s}$ , see figure 7.3) is in agreement with earlier findings [141].

The  $\mathbf{b}$  associated with a minimum rCRLB did not vary as  $\lambda_{\perp}$  was changed (Figure 7.2(b)). Moreover, it was shown that two elements in  $\mathbf{b}$  suffice for estimating the parameters in a heterogeneous sample:  $\text{rCRLB}_{\text{FA}}$  and  $\text{RMSE}(\text{FA})$  did not decrease when a  $\mathbf{b}$  up to length four was used. In previous work, the optimal  $b$ -value was shown to be mainly dependent on the mean diffusivity [13], which is in contrast to our findings. We attribute this to a difference in model selection: Alexander and coworkers constrained the diffusivity values to be equal for both tensors. A single non-zero  $b$ -value then suffices for estimating the diffusion. This value appears to be more sensitive to the expected diffusivity values. Our independent parameterization of the perpendicular diffusivity minimizes bias and variance in the parameters when the distance between the  $b$ -values is large. This allows us to independently estimate the FA in both compartments.

It may be expected that the optimal  $b$ -values found in the Cramér-Rao analysis correspond to the underlying diffusivity scales. The inverse diffusion weighting values  $\frac{1}{b_1} = 1 \cdot 10^{-3} \text{ mm}^{-2}\text{s}$  and  $\frac{1}{b_2} = 0.3 \cdot 10^{-3} \text{ mm}^{-2}\text{s}$  indicate that lower diffusion values are emphasized (we used diffusion values of around  $1.5 \cdot 10^{-3} \text{ mm}^2\text{s}^{-1}$  and  $0.4 \cdot 10^{-3} \text{ mm}^2\text{s}^{-1}$  in the model of the experiments). Optimization of acquisition parameters previously resulted in  $b \sim 2.2 - 2.8 \cdot 10^3 \text{ mm}^{-2}\text{s}$  [13], which is close to our optimum for  $b_2$ .

### 7.5.3 Validation in brain data

Model #3 was used to estimate the diffusion shape in the crossing of the corticospinal tract and the corpus callosum. It yielded a ‘consistent’ FA-profile at the location of the crossing, whereas single tensor profiles showed a reduction in FA. The standard deviation of the estimated FA was minimal for  $\mathbf{b} = [ 1.0 \quad 3.0 ] \cdot 10^3 \text{ mm}^{-2}\text{s}$ . It was demonstrated that the FA profiles for  $n_{\mathbf{b}} = 2$  did not differ significantly from the FA-profiles emanating from  $\mathbf{b}$  with more  $n_{\mathbf{b}} = 5$  elements.

FA-values estimated in brain data were systematically lower for the single tensor model than for a two tensor model even in non-crossing tissue regions (figure 7.4). We attribute this to the isotropic compartment that is lacking in the single tensor model. Previous work considered the relative proportion of anisotropic signal rather than the FA per tensor [129]. Note that exploring data-driven model selection, such as automatic relevance detection [41] is beyond the scope of this work.

This work defines fundamental limits for comparative studies to correctly analyze complex white matter structures.



## 7.A APPENDIX

## 7.A.1 Derivatives to estimated parameters

The performance of the Levenberg-Marquardt optimization is significantly increased by including derivatives of the signal to the parameters  $\frac{\partial S_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ . To obtain the derivatives to our reparameterization, we first write:

$$\frac{\partial S_j}{\partial \lambda_{i\parallel}} = -b_j \mathbf{g}_j^T \mathbf{R}_i \text{diag} \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \mathbf{R}_i^T \mathbf{g}_j f_i S_{ij} \quad (7.13)$$

$$\frac{\partial S_j}{\partial \lambda_{i\perp}} = -b_j \mathbf{g}_j^T \mathbf{R}_i \text{diag} \begin{pmatrix} 0 & 1 & 1 \end{pmatrix} \mathbf{R}_i^T \mathbf{g}_j f_i S_{ij} \quad (7.14)$$

$$\frac{\partial S_j}{\partial \alpha_j} = -b_j \sum_{i=1,2} \mathbf{g}_j^T \left( \frac{\partial \mathbf{R}_i}{\partial \alpha_j} \mathbf{E}_i \mathbf{R}^T + \mathbf{R}_i \mathbf{E}_i \frac{\partial \mathbf{R}_i^T}{\partial \alpha_j} \right) \mathbf{g}_j f_i S_{ij}, \quad j \in \{1 \dots 4\} \quad (7.15)$$

$$\frac{\partial S_j}{\partial f_1} = S_{1j} - S_{2j} \quad (7.16)$$

$$\frac{\partial S_j}{\partial f_{iso}} = S_{iso} - S_{2j} \quad (7.17)$$

$$\frac{\partial S_j}{\partial D_{iso}} = -b_j f_{iso} S_{iso}. \quad (7.18)$$

Note that  $f_2 = 1 - f_1 - f_{iso}$ . The rotation matrices for both tensors are written as:

$$\mathbf{R}_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha_1 & -\sin \alpha_1 \\ 0 & \sin \alpha_1 & \cos \alpha_1 \end{bmatrix} \cdot \begin{bmatrix} \cos \alpha_2 & 0 & \sin \alpha_2 \\ 0 & 1 & 0 \\ -\sin \alpha_2 & 0 & \cos \alpha_2 \end{bmatrix} \cdot \begin{bmatrix} \cos(\alpha_3 + (-1)^i \alpha_4) & -\sin(\alpha_3 + (-1)^i \alpha_4) & 0 \\ \sin(\alpha_3 + (-1)^i \alpha_4) & \cos(\alpha_3 + (-1)^i \alpha_4) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (7.19)$$

from which derivatives  $\frac{\partial \mathbf{R}_i}{\partial \alpha_j}$  are obtained and substituted in equation 7.15. The computed signal per tensor  $i$  equals

$$S_{ij} = S_0 \exp(-b_j \mathbf{g}_j^T \mathbf{D}_i \mathbf{g}_j). \quad (7.20)$$

Derivatives to our parameterization then are:

$$\frac{\partial S_j}{\partial \Delta \tilde{\lambda}_{i\parallel}} = \frac{\partial S_j}{\partial \lambda_{i\parallel}} s \exp \Delta \tilde{\lambda}_{i\parallel} \quad (7.21)$$

$$\frac{\partial S_j}{\partial \tilde{\lambda}_{i\perp}} = \left\{ \frac{\partial S_j}{\partial \lambda_{i\perp}} + \frac{1}{2} \left( \frac{\partial S_j}{\partial \lambda_{1\parallel}} + \frac{\partial S_j}{\partial \lambda_{2\parallel}} \right) \right\} s \exp \lambda_{i\perp} \quad (7.22)$$

$$\frac{\partial S_j}{\partial \tilde{f}_{\{1, iso\}}} = \frac{1}{\sqrt{\pi}} \exp \left( -(\tilde{f}_{\{1, iso\}})^2 \right) \quad (7.23)$$

$$\frac{\partial S_j}{\partial \tilde{D}_{iso}} = \frac{\partial S_j}{\partial D_{iso}} \exp \tilde{D}_{iso} \quad (7.24)$$

with  $s = 10^3$  a constant scaling factor to achieve approximately equal order of magnitude for all parameter values in numerical optimization.

#### 7.A.2 Cramér-Rao Lower Bounds on eigenvalues and FA

We will now derive the CRLB on eigenvalues and FA, as mentioned in Section 7.3.3. Exponential and error function mappings are ignored in calculating the CRLB, since they merely aim to constrain the optimization. Equation 7.2 then reduces to a parametrization of the axial diffusivity

$$\lambda_{i\parallel} = \Delta \tilde{\lambda}_{i\parallel} + \frac{1}{2} (\lambda_{1\perp} + \lambda_{2\perp}) \quad (7.25)$$

and an adapted parameter vector  $\theta'$

$$\theta' = \{ \Delta \tilde{\lambda}_{1\parallel}, \lambda_{1\perp}, \Delta \tilde{\lambda}_{2\parallel}, \lambda_{2\perp}, \alpha_1, \alpha_2, \alpha_3, \alpha_4, f_1, f_{iso}, D_{iso} \}. \quad (7.26)$$

Equation 7.9 describes the CRLB-inequality including  $\left( \frac{\partial \tau}{\partial \theta'} \right)$ , which is the Jacobian of a potential transformation  $\tau(\theta')$  of the parameter vector to a measure of interest. For the purpose of calculating the CRLBs of the eigenvalues, such as  $\text{CRLB}_{\lambda_{1\parallel}} = \left( \frac{\partial \lambda_{1\parallel}}{\partial \theta'} \right) \mathbf{I}^{-1} \left( \frac{\partial \lambda_{1\parallel}}{\partial \theta'} \right)^T$ , we write the following transformations,

$$\begin{bmatrix} \frac{\partial \lambda_{1\parallel}}{\partial \theta'} \\ \frac{\partial \lambda_{1\perp}}{\partial \theta'} \\ \frac{\partial \lambda_{2\parallel}}{\partial \theta'} \\ \frac{\partial \lambda_{2\perp}}{\partial \theta'} \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \cdots 0 \\ 0 & 1 & 0 & 0 & 0 \cdots 0 \\ 0 & \frac{1}{2} & 1 & \frac{1}{2} & 0 \cdots 0 \\ 0 & 0 & 0 & 1 & 0 \cdots 0 \end{bmatrix}, \quad (7.27)$$

based on the adopted parametrization described above in Equation 7.25.

To compute  $\text{CRLB}_{\text{FA}_{1,2}} = \left( \frac{\partial \text{FA}_{1,2}}{\partial \theta'} \right) \mathbf{I}^{-1} \left( \frac{\partial \text{FA}_{1,2}}{\partial \theta'} \right)^T$ , the derivatives of  $\text{FA}_{1,2}$  to  $\theta'$  are derived to be,

$$\begin{bmatrix} \frac{\partial \text{FA}_1}{\partial \theta'} \\ \frac{\partial \text{FA}_2}{\partial \theta'} \end{bmatrix} = \begin{bmatrix} \frac{\partial \text{FA}_1}{\partial \Delta \lambda_{1\parallel}} & \frac{\partial \text{FA}_1}{\partial \lambda_{1\perp}} & 0 & \frac{\partial \text{FA}_1}{\partial \lambda_{2\perp}} & 0 \cdots 0 \\ 0 & \frac{\partial \text{FA}_2}{\partial \lambda_{1\perp}} & \frac{\partial \text{FA}_2}{\partial \Delta \lambda_{2\parallel}} & \frac{\partial \text{FA}_2}{\partial \lambda_{2\perp}} & 0 \cdots 0 \end{bmatrix}, \quad (7.28)$$

The derivatives of  $\text{FA}_1$  to the eigenvalues are

$$\frac{\partial \text{FA}_1}{\partial \Delta \lambda_{1\parallel}} = 4\lambda_{1\perp}(5\lambda_{1\perp} + 2\Delta\lambda_{1\parallel} + \lambda_{2\perp})c_1$$

$$\frac{\partial \text{FA}_1}{\partial \lambda_{1\perp}} = -2(\lambda_{2\perp} + 2\Delta\lambda_{1\parallel})(5\lambda_{1\perp} + 2\Delta\lambda_{1\parallel} + \lambda_{2\perp})c_1$$

$$\frac{\partial \text{FA}_1}{\partial \lambda_{2\perp}} = 2\lambda_{1\perp}(5\lambda_{1\perp} + 2\Delta\lambda_{1\parallel} + \lambda_{2\perp})c_1$$

with

$$c_1 = (4\Delta\lambda_{1\parallel}^2 + 4\Delta\lambda_{1\parallel}\lambda_{1\perp} + 4\Delta\lambda_{1\parallel}\lambda_{2\perp} + 9\lambda_{1\perp}^2 + 2\lambda_{1\perp}\lambda_{2\perp} + \lambda_{2\perp}^2)^{-3/2}$$

while similarly for  $\text{FA}_2$ :

$$\frac{\partial \text{FA}_2}{\partial \Delta \lambda_{2\parallel}} = 4\lambda_{2\perp}(5\lambda_{2\perp} + 2\Delta\lambda_{2\parallel} + \lambda_{1\perp})c_1$$

$$\frac{\partial \text{FA}_2}{\partial \lambda_{1\perp}} = 2\lambda_{2\perp}(5\lambda_{2\perp} + 2\Delta\lambda_{2\parallel} + \lambda_{1\perp})c_1$$

$$\frac{\partial \text{FA}_2}{\partial \lambda_{2\perp}} = -2(\lambda_{1\perp} + 2\Delta\lambda_{2\parallel})(5\lambda_{2\perp} + 2\Delta\lambda_{2\parallel} + \lambda_{1\perp})c_1$$

with

$$c_2 = (4\Delta\lambda_{2\parallel}^2 + 4\Delta\lambda_{2\parallel}\lambda_{1\perp} + 4\Delta\lambda_{2\parallel}\lambda_{2\perp} + \lambda_{1\perp}^2 + 2\lambda_{1\perp}\lambda_{2\perp} + 9\lambda_{2\perp}^2)^{-3/2}.$$

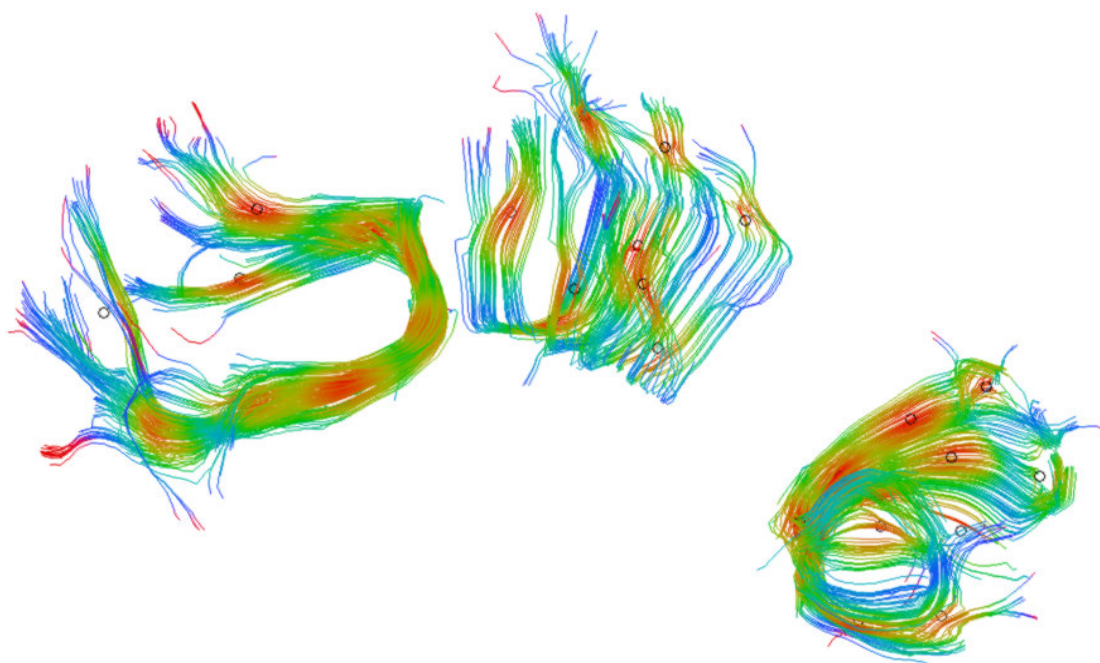
#### ACKNOWLEDGEMENTS

We are grateful to P. de Boer, M. Stam, T. Glatard, J.J. Keizer, J. Engelberts and grid.support@sara.nl for their contributions to this work. We thank Prof. I.T. Young for revising the manuscript. This work was sponsored by the BiG Grid project for the use of the computing and storage facilities, with financial support from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Organisation for Scientific Research, NWO).



FIBER TRACT STATISTICS

# III





## NON-RIGID POINT SET MATCHING OF FIBER TRACTS IN DIFFUSION TENSOR IMAGING

---

Patient studies based on Diffusion Tensor Images (DTI) require spatial correspondence between subjects. We propose to obtain the correspondence from white matter tracts, by introducing a new method for non-rigid matching of white matter fiber tracts in DTI. The method boils down to point set registration which involves simultaneously clustering and matching of the data points. The tracts are implicitly warped to a common frame of reference to avoid the potential bias towards one of the datasets. The algorithm gradually refines from global to local registration, which is implemented through deterministic annealing. Special care was taken to incorporate the spatial relation between fiber points and the uncertainty in principal diffusion orientation. As a result, the computed clusters are oriented along the fiber tracts and discriminate between adjacent but distinct fiber tracts. This is validated on synthetic and clinical data. The root mean squared distance with respect to expert-annotated landmarks is low (3 mm). In contrast to a state-of-the-art non-rigid registration technique, the proposed method is more robust to residual misalignments in terms of measured FA-values.

Submitted as:

M.W.A. Caan, L.J. van Vliet, C.B.L.M. Majoie, M.M. van der Graaff, C.A. Grimbergen and F.M. Vos, *Non-rigid Point Set Matching of Fiber Tracts in Diffusion Tensor Imaging*, IEEE Transactions on Biomedical Engineering

## 8.1 INTRODUCTION

Magnetic Resonance Diffusion Tensor Imaging (DTI) measures the diffusion of water. Complementary to anatomic MR imaging, DTI can visualize structural information about the integrity of white matter tracts [10]. It was first described around a decade ago by Basser [33, 34]. A still growing number of studies now specifically uses DTI to analyze changes in brain structure due pathological processes.

Prior to a group-wise comparison of feature values, for instance by voxel- or tract-based analysis of the Fractional Anisotropy (FA), spatial normalization of the data is needed. Conventionally, registration is based on a predefined set of features derived from the data, which act as fiducials. Clearly, one must take care that these features do not interfere with those used in the analysis afterwards. For instance, if changes in white matter structure are to be expected, features describing this structure should not be used. Doing so may cause differences to become "registered away".

In this paper we propose a new registration algorithm for white matter fiber tracts, based on white matter tract geometries, leaving the FA for subsequent patient studies. Registration and data analysis then use different image features.

Over the last few years, several non-rigid registration algorithms have been proposed for DTI data analysis. Ruiz-Alzola [216] identifies landmarks in the tensor field by means of structure analysis, which drive the registration through local template matching. An elastic registration scheme of Alexander et al. [10] optimizes a tensor similarity measure per voxel. The tensors are reoriented only after registration. The problem of applying image warps to DTI was addressed in [14], in which the principal orientation as well as the plane spanned by the two largest eigenvectors of the tensor are preserved. Both Guimond [124] and Jones [144] used rotationally invariant metrics derived from the tensor to achieve spatial normalization. Rohde [214] and Park [196] treated tensor image registration as a multi-channel problem. They showed that the inclusion of orientational information can improve the accuracy of the registration result. Cao proposed a diffeomorphic mapping of both principal eigenvector [71] and tensor fields [72]. The tensor reorientation was incorporated in the optimization. Zhang [264] also explicitly optimized the orientation using a similarity measure, based on the tensor as a whole. Local changes are corrected for by using a piecewise affine transform. Alternatively, a pairwise comparison of tensor components based on mutual information was proposed, with a viscous fluid deformation model [126]. Yeo [261] employed an exact finite-strain differential for accurate diffusion tensor registration. Leemans [164] presented a registration method that focuses on fiber tracts. The method particularly implements a rigid registration, making it suitable for intra-subject registration. Corouge [89] used Procrustes analysis for affinely matching fibers and Maddah achieved arc-length tract representations by joint probabilistic clustering and point-by-point mapping [168]. Recently, Eckstein chose to estimate a deformable fiberbundle model, simultaneously with the tensor field, skipping the fiber tracking step [100]. This work relies on a accurate initialization by a non-rigid registration algorithm.



We consider the registration to be just one step in the whole process of analyzing DTI-data. In our opinion, the objective of the analysis is a crucial aspect in designing the registration framework. Previous work on DTI image registration (described above) concentrates on the alignment, without considering the effect on subsequent analysis. We will demonstrate that doing so discards small, but important, changes in white matter structure that may be crucial for early diagnosis of neuro-diseases.

We propose a registration method, based on a tractography representation derived from the DTI data. Thereby, not the diffusion shape, but its orientation is used to realize a spatial normalization. It is assumed that there may be local differences in the shape of the tensors, but that the global course of white matter tracts is not affected. Consequently, comparing tensor shape information after the data have been registered is allowed.

Registration aims at maximizing correspondence between the scans of interest. Fortunately, the inter-subject anatomical correspondence is high in the main white matter tracts. We restrict the analysis to these tracts, thereby avoiding confusion, induced by low-correspondence regions such as the gyri. A limited number of degrees of freedom is assumed to suffice for adequate modeling inter-subject variability in these tracts.

In the proposed method, the fiber tracts are represented as point sets. It will be shown that the concept of Chui's point set matching framework [84] applies. A clustering of the data results in a series of landmarks, driving the registration. Previously, we presented preliminary work using the classic Chui approach on a limited amount of data [67]. Chui's algorithm is extended by simultaneously incorporating information about the orientation of the fiber tracts as well as the uncertainty in principal orientation. As a result, the computed clusters are predominantly oriented along single fiber tracts even when fiber tracts are in close proximity. By following a deterministic annealing scheme, the algorithm gradually refines from global to local matching. A thin plate spline models the non-rigid deformation. An atlas space is formed to which all the input data sets are warped, thus preventing a possible bias towards one of the datasets.

## 8.2 BACKGROUND

### 8.2.1 *Data acquisition and reconstruction*

The diffusion can be described by a rank two  $3 \times 3$  symmetrical tensor  $\mathbf{D}$ . Diffusion manifests itself as an attenuated signal if a direction encoded diffusion weighting is applied:

$$S_i = S_o \exp(-b \mathbf{g}_i^T \mathbf{D} \mathbf{g}_i). \quad (8.1)$$

In this equation,  $\mathbf{g}_i$  is a normalized vector, corresponding to the  $i^{\text{th}}$  direction of diffusion encoding and  $b$  is a constant reflecting the degree of diffusion encoding [256]. The attenuated signal  $S_i$  relates to a reference measurement  $S_o$  without diffusion weighting.

At least six Diffusion Weighted Images (DWIs) in different gradient directions  $\mathbf{g}_i$  are needed. Then, a linear system of equations (8.1) is build, from which  $\mathbf{D}$  may be computed [256, 154].

The Fractional Anisotropy (FA), [32] is a commonly used scalar measure derived from the eigenvalues  $\lambda_{1\dots 3}$  of  $\mathbf{D}$ , written as:

$$\text{FA} = \frac{(\lambda_1 - \lambda_2)^2 + (\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_3)^2}{\sqrt{2}\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}. \quad (8.2)$$

The FA describes the anisotropy of the diffusion and is sensitive to subtle changes in white matter.

Conventionally, an antipodal diffusion profile is assumed, i.e. measuring in positive and negative gradient direction will result in an equal signal attenuation. By diagonalizing  $\mathbf{D}$ , eigenvectors  $\mathbf{v}_i$  and -values  $\lambda_i$  are obtained, indicating the principal orientations and the diffusion component along these orientations.

### 8.2.2 *Uncertainty in principal diffusion orientation*

All MR images are hampered by noise, which will induce an uncertainty in the estimation of the diffusion parameters. The precision of the principal direction can be determined from repetitive image acquisitions. However, this is time-expensive and will be susceptible to motion artifacts. Instead, a bootstrapping approach is used [142], to obtain a robust measure for the precision in the principal direction, based on a single dataset.

An imaging sequence containing pairs of positive and negative gradient directions is assumed. Subsets of the DWIs with either the positive or negative gradient direction of each pair are randomly chosen. Subsequently, the principal diffusion direction field is computed, based on this subset of DWIs. Based on these results, a cone whose opening angle  $\theta_{95\%}$  is defined in such a way that 95% of the found eigenvectors fall inside the cone. For anisotropic tensors,  $\theta_{95\%}$  will be small, whereas for isotropic tensors  $\theta_{95\%} \approx \pi$ .

Here, the uncertainty in principal direction is calculated in 50 iterations. Then,  $\theta_{95\%}$  is defined by the angle between eigenvector 47 ( $\sim 95\%$  of 50) and the mean eigenvector. Figure 8.1 displays  $\theta_{95\%}$  for one axial slice, as well as the corresponding FA. It illustrates that the uncertainty is smallest for the large tracts and larger in the areas of grey matter.

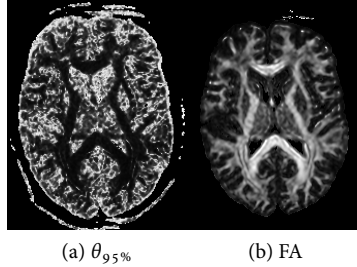


Figure 8.1: (a) One axial slice showing the uncertainty in the principal diffusion orientation,  $\theta_{95\%}$ , (b) corresponding FA-image.

### 8.2.3 Dyadic coherence

The dyadic tensor  $\langle \mathbf{v}_i \mathbf{v}_i^T \rangle = \sum_{j=1}^P \mathbf{v}_i^j \mathbf{v}_i^{jT}$  [36] is used to describe the coherence between  $P$  principal eigenvectors. The dyadic coherence  $\kappa$  yields

$$\kappa = 1 - \sqrt{\frac{\beta_2 + \beta_3}{2\beta_1}}, \quad (8.3)$$

where  $\beta_{1,2,3}$  are the sorted eigenvalues of the dyadic tensor.  $\kappa$  equals one for aligned and zero for uniformly distributed eigenvectors.

## 8.3 FIBER TRACT MATCHING

### 8.3.1 Feature selection

The fiber tract matching is initiated by first identifying the structures of interest. The focus in this paper will be on the corticospinal tract, forceps major, cingulum and inferior longitudinal fasciculus. The selection of the tracts is supported by the continuous tracking or FACT method, as implemented in DTIStudio software [140]. After initial tracking from all brain voxels (i.e. "full brain tractography"), subsets of fibers (corresponding to structures of interest) are selected in a multi-ROI approach. Each individual fiber is represented as a point set via the spatial coordinates at which the fiber is sampled. The connection between subsequent points is not used in the matching. However, it will be shown that by using the principal eigenvector field, orientation information is included in the algorithm.

### 8.3.2 Classic point set registration

We will now go into the fundamentals of classic point set registration [84] and then extend the method to the registration of diffusion tensor data. The whole process of

point set registration contains three aspects that act simultaneously: 1) clustering the points in each point set, 2) mapping estimation to warp the cluster center points to one another and 3) atlas (mean shape) estimation. In fact, the method aims to derive a “common frame of reference”, the atlas, which represents the mean shape of the point sets.

Let  $P$  denote the number of datasets to be aligned. Each dataset  $p \in \{1, \dots, P\}$ , is represented by a point set (in our case obtained by sampling the fiber tracts). A point-set  $X^p$  consists of  $N^p$  points  $X^p = \{\mathbf{x}_i^p, i = 1 \dots N^p\}$ . A Gaussian mixture model is used to model each point-set. Accordingly, a set of cluster center points is defined, one for each point set, denoted by  $C^p = \{\mathbf{c}_a^p | a = 1 \dots K\}$ . Notice that there are  $K$  cluster center points for each point set. An atlas cluster point set  $\mathcal{Z} = \{\mathbf{z}_a | a = 1 \dots K\}$  is defined that is to explain the cluster centers over all point sets.  $f^p$  denotes the transformation that performs a warping of points.

A mixture of Gaussians models the density of the point set:

$$p(X^p | C^p, \Sigma_a^p) = \prod_{i=1}^{N^p} \sum_{a=1}^K p(\mathbf{x}_i^p | \mathbf{c}_a^p, \Sigma_a^p), \quad (8.4)$$

in which

$$p(\mathbf{x}_i^p | \mathbf{c}_a^p, \Sigma_a^p) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma_a^p|^{\frac{1}{2}}} e^{-\frac{1}{2} (\mathbf{x}_i^p - \mathbf{c}_a^p)^T (\Sigma_a^p)^{-1} (\mathbf{x}_i^p - \mathbf{c}_a^p)} \quad (8.5)$$

are the probability distributions associated with each cluster, with  $\Sigma_a^p$  the covariance matrix of the cluster. For the sake of simplicity Chui defined the covariance matrices to be isotropic, diagonal, and identical. We will use the structure of our data and extend the algorithm to non-identical anisotropic covariance matrices in the next section. For that reason we continue a notation that incorporates  $\Sigma_a^p$ .

The objective function to be minimized was derived to be [84]:

$$E_{TPS}(f) = \sum_{i=1}^{N^p} \sum_{a=1}^K m_{ai}^p |\mathbf{x}_i^p - \mathbf{c}_a^p|^2 + \sum_{a=1}^K |z_a - f^p(\mathbf{c}_a^p)|^2 + \gamma_s \|L f^p\|^2 + T \sum_{i=1}^{N^p} \sum_{a=1}^K m_{ai}^p \log m_{ai}^p \forall a \in \{1, \dots, K\}, \forall p \in \{1, \dots, P\}. \quad (8.6)$$

in which the new membership variable  $m_{ai}^p$  indicates the degree to which a point “feature”  $\mathbf{x}_i^p$  belongs to cluster center point  $\mathbf{c}_a^p$ . It is assumed that  $m_{ai}^p > 0$  and  $\sum_{a=1}^K m_{ai}^p = 1$ .  $L$  is a regularization operator, the Laplacian, which results in a thin-plate spline [250, 48].  $T$  is a temperature parameter as in simulated annealing (see below), preceding an entropy term. The first term in the equation,  $\sum_{i=1}^{N^p} \sum_{a=1}^K m_{ai}^p |\mathbf{x}_i^p - \mathbf{c}_a^p|^2$ , reflects the extent to which the cluster center points represent the point set. The cluster center points are warped

towards the atlas point set, the deviation from which is accounted for by the second term  $\sum_{a=1}^K |\mathbf{z}_a - f^P(\mathbf{c}_a^P)|^2$ . The amount of bending is represented by the third term  $\gamma_s \|L f^P\|^2$ . The regularization parameter  $\gamma_s$  defines the weight of the bending term in relation to terms 1 and 2.

A deterministic annealing approach is used to assist the registration process to gradually refine from global to local matching. In effect, the temperature  $T$  in the fourth term  $T \sum_{i=1}^{N^p} \sum_{a=1}^K m_{ai}^P \log m_{ai}^P$  corresponds to the scale at which matching is taking place.  $T$  is gradually decreased such that the clusters first globally and then locally describe the data, i.e. the influence of the fourth term is gradually reduced. The initial temperature  $T_{init}$  and final temperature  $T_{final}$  are a priori defined.  $T$  is lowered after each iteration:  $T_{ind+1} = \alpha T_{ind}$ , with  $0 < \alpha < 1$ .

The membership  $m_{ai}^P$  is defined by Chui as

$$m_{ai}^P = \frac{q_{ai}^P}{\sum_{a=1}^K q_{ai}^P}, \forall a, I, \quad (8.7)$$

with

$$q_{ai}^P = \exp\left(-\frac{1}{2\sigma^2} |\mathbf{x}_i^P - \mathbf{c}_a^P|^2\right). \quad (8.8)$$

assuming isotropic covariance matrices ( $\Sigma_a^P = \sigma^2 \mathbf{I}$ ). The temperature is defined to be the variance in equation 8.8,  $T = \sigma^2$ . The algorithm proceeds in an iterative manner. The cluster center points are initiated at random offsets near the center of gravity. They are recomputed after each iteration, using

$$\mathbf{c}_a^P = \frac{\sum_{i=1}^{N_p} m_{ai}^P \mathbf{x}_i^P}{\sum_{i=1}^{N_p} m_{ai}^P}, \forall a \in \{1, \dots, K\}, \forall p \in \{1, \dots, P\}. \quad (8.9)$$

The atlas points are formed by averaging the cluster center points,

$$\mathbf{z}_a = \frac{1}{P} \sum_{p=1}^P \mathbf{c}_a^P \quad (8.10)$$

The transformation  $f^P$  is a combination of an affine and non-rigid thin plate spline transform. Using homogeneous coordinates, it is expressed as

$$f^P(\mathbf{c}_a, \mathbf{D}, \mathbf{W}) = \mathbf{c}_a^P \cdot \mathbf{D} + \Phi(\mathbf{c}_a^P) \cdot \mathbf{W}, \quad (8.11)$$

with an  $4 \times 4$  sized affine transformation matrix  $\mathbf{D}$  and an  $K \times 4$  sized matrix  $\mathbf{W}$  containing the spline transform coefficients. The  $1 \times K$  kernel function  $\Phi(\mathbf{c}_a^P)$  yields

$$\Phi(\mathbf{c}_a^P) = -\left(\|\mathbf{c}_1^P - \mathbf{c}_a^P\| \quad \|\mathbf{c}_2^P - \mathbf{c}_a^P\| \quad \dots \quad \|\mathbf{c}_K^P - \mathbf{c}_a^P\|\right). \quad (8.12)$$

The optimization of the transformation can be written explicitly [83]. Let  $\mathbf{Z}$  and  $\mathbf{C}$  be matrices of concatenated atlas points  $z_a^p$  and cluster center points  $\mathbf{c}_a^p$ . A QR-decomposition results in

$$\mathbf{Z} = [\mathbf{Q}_1 \mathbf{Q}_2] \begin{pmatrix} \mathbf{R} \\ \mathbf{o} \end{pmatrix} \quad (8.13)$$

The transformation matrices ( $\hat{\mathbf{D}}, \hat{\mathbf{W}}$ ) are regularized using a variable  $\gamma_s$  and have been derived to be [83]

$$\hat{\mathbf{W}} = \mathbf{Q}_2 (\mathbf{Q}_2^T \Phi \mathbf{Q}_2 + \gamma_s \mathbf{I})^{-1} \mathbf{Q}_2^T \mathbf{C} \quad (8.14)$$

$$\hat{\mathbf{D}} = (\mathbf{R}^T \mathbf{R} + \gamma_s \mathbf{I})^{-1} (\mathbf{R}^T \mathbf{Q}_1^T (\mathbf{C} - \mathbf{K} \hat{\mathbf{W}}) - \mathbf{R}^T \mathbf{R}) \quad (8.15)$$

### 8.3.3 Clustering fiber tracts

The algorithm above was successfully applied to describe object boundaries in 2D and 3D images [84]. Isotropic clusters (represented in the covariance matrices) were assumed while doing so. We now wish to extend the method to the registration of fiber tracts, which not only improves the method, but also makes it suitable for use in comparative studies of diffusion shape. Let us shortly go into two fundamental differences associated with our application.

Firstly, each fiber tract is described as a volumetric, 3D point cloud in contrast to the 1D respectively 2D boundary representations in the original application. Note that the mixture of Gaussians model implicitly allows for a volumetric representation. Actually, it will be demonstrated that such a 3D cluster representation is a prerequisite needed to avoid mis-registration (see below).

Secondly, we will explicitly benefit from the diffusion orientation information that is available in the diffusion tensor per voxel. The cluster covariance matrix  $\Sigma_a^p$  will be adopted to locally reflect the mean as well as the uncertainty of the principal diffusion orientation field. What is more, the clustering is to be further restricted to single tracts (such as the cingulum), if multiple tracts are in close proximity. This restriction aids in determining correspondence between datasets in registration. Imposing such a constraint will be discussed in the next subsection.

Let us first incorporate the uncertainty of the fiber tract orientation into the cluster covariance matrix  $\Sigma_a^p$  of equation 8.5. To that end an uncertainty tensor  $\mathbf{U}_i^p$  is defined in each point  $\mathbf{x}_i^p$ , oriented along the local diffusion tensor  $\mathbf{D}_i^p$ , whose eccentricity is determined merely by the uncertainty in the principal diffusion orientation.

The eigenvalues  $\lambda_{a,1\dots 3}$  of  $\mathbf{U}_i^P$  are defined as an angle deviation of  $\theta_{95\%}$  with respect to  $\pi$ :

$$\lambda_{a,2} = \lambda_{a,3} = \frac{\theta_{95\%}}{\pi} \lambda_{a,1}. \quad (8.16)$$

Additionally, the determinant  $|\mathbf{U}_i^P|$  is determined by the temperature,

$$|\mathbf{U}_i^P| = \lambda_{a,1} \lambda_{a,2} \lambda_{a,3} = T^{2/3}, \quad (8.17)$$

which implies

$$\lambda_{a,1} = \left( \frac{\pi}{\theta_{95\%}} \right)^2 T^{2/3}. \quad (8.18)$$

The opening angle of the 95%-confidence cone in these equations  $\theta_{95\%}$  is calculated as proposed in section 8.2.2. Equation 8.17 implies that the determinants of the covariance matrices are the same for all clusters.

Given the matrix  $\mathbf{V}$  of normalized eigenvectors  $\mathbf{v}_{1\dots 3}$  from a tensor  $\mathbf{D}_i^P$ ,  $\mathbf{U}_i^P$  is written as

$$\mathbf{U}_i^P = \mathbf{V}^T \begin{pmatrix} \lambda_{a,1} & 0 & 0 \\ 0 & \lambda_{a,2} & 0 \\ 0 & 0 & \lambda_{a,3} \end{pmatrix} \mathbf{V}, \quad (8.19)$$

so that both the uncertainty and the orientation field are represented. Notice that  $\mathbf{V}$  may be considered to be a rotation matrix that is used to properly align the uncertainty tensor.

Let  $S^P$  be the set of voxels included in the tractography. Now  $\mathbf{U}_i^P$  can be locally averaged, to obtain  $\Sigma_a^P$ ,

$$\Sigma_a^P = \sum_{\mathbf{x}_i^P \in S^P} p(\mathbf{x}_i^P | \mathbf{c}_a^P, \Sigma_a^P) \cdot \mathbf{U}_i^P, \quad (8.20)$$

in which the weight  $p(\mathbf{x}_i^P | \mathbf{c}_a^P, \Sigma_a^P)$  is calculated via equation 8.5 which involves  $\Sigma_a^P$  taken from the previous iteration (we start with an isotropic tensor).

The covariance matrix thus calculated is needed for the calculation of the cluster center points (equation 8.9). It may be observed that this computation (which uses the underlying diffusion tensor) is performed in the frame of each individual dataset. Consequently, no tensor reorientation is needed. Interpolation is done directly on the components of the logarithmic tensors, such that positive definiteness is retained [23], using a second order b-spline fit [231].

### 8.3.4 Membership in fiber tracts

It is preferred that the clustering is restricted to single fiber tracts (see above). Such a spatial relation may be included in the membership variables in a straightforward manner. We propose to construct  $q_{ai}^p$  (equation 8.7) from two terms:

$$q_{ai}^p = q_{ai,mah}^p q_{ai,orient}^p \quad (8.21)$$

The first term serves for normalization, so that (typically longer) distances along the fiber tract are weighted equally to distances perpendicular to the tract:

$$q_{ai,mah}^p = \exp\left(-(\mathbf{x}_i^p - \mathbf{c}_a^p)^T (\Sigma_a^p)^{-1} (\mathbf{x}_i^p - \mathbf{c}_a^p)\right) \quad (8.22)$$

in which  $\Sigma_a^p$  is calculated as defined previously. Effectively,  $q_{ai,mah}$  represents the Mahalanobis distance of points  $\mathbf{x}_i^p$  to the cluster center point  $\mathbf{c}_a^p$ . The orientation uncertainty is incorporated by means of the covariance matrix (see above).

What is more, points along the same tract should be considered nearer than points on different tracts. Such a spatial relationship is enforced by incorporating the extend to which points are on parallel tracts

$$q_{ai,orient}^p = \exp\left(-\rho(T) \left(1 - \mathbf{v}_1(x_i^p) \cdot \mathbf{v}_1(c_a^p)\right)\right), \quad (8.23)$$

with  $\mathbf{v}_1(\cdot)$  the interpolated principal eigenvector and  $\rho(T)$  a weighting term as a function of the temperature:

$$\rho(T) = \frac{T_{init} - T}{T_{init} - T_{final}} \rho_0. \quad (8.24)$$

The term  $q_{ai,orient}^p$  favours points that have similarly oriented principal eigenvectors to the principal eigenvector of the cluster center point.

Notice that both  $q_{ai,mah}^p$  (via  $\Sigma_a^p$ ) and  $q_{ai,orient}^p$  are affected by the temperature  $T$ . Consequently,  $q_{ai}^p$  is influenced by the temperature which, in turn, controls the scale of the registration. We opted to do so to force the membership to gradually become more precise as the registration refines from global to local matching.

Formulated in pseudo-code, the algorithm can be summarized as follows:

- Compute the tensor field, principal eigenvectors, FA and  $\theta_{95\%}$ .
- Track corresponding fibers of interest in all datasets.
- Initialize  $T = T_{init}$ ,  $\gamma$ .
- Initialize the atlas points  $z_a$  as a random subset of the point sets  $x_i^p$ .



- Update until either  $T < T_{final}$  or the set  $\{m_{ai}^p < \epsilon | \epsilon = 1e^{-5} \forall a, i, p\}$  includes more than 5% of the points:
  - cluster center points  $c_a^p$  (eq. 8.9)
  - atlas point set  $z_a$  (eq. 8.10)
  - transformations  $f^p$  (eq. 8.11)
  - membership  $m_{ai}^p$  (eq. 8.7)
  - $T = \alpha T$

### 8.3.5 Parameters

A number of parameters is involved in the fiber tract matching. Below, we indicate some global effects; the exact setting will be discussed in the next section.

- $T_{init}$  represents the maximum scale of registration, which equals the number of voxels to the power 2/3 (equation 8.17).
- $T_{final}$  reflects the local scale of registration. It is bounded by the resolution of the data.
- The annealing rate  $0 < \alpha < 1$  must be set to allow the algorithm sufficient iterations for convergence. A little conservative choice is preferred, although it extends computation times. The total number of iterations (by definition) equals  $n_{iter} \approx \log(T_{final}/T_{init})/\log \alpha$ .
- The regularization parameter  $\gamma_s$  needs to be chosen in such a way, that the matrix  $\mathbf{Q}_2^T \Phi \mathbf{Q}_2$  (size  $K \times K$ ) in equation 8.14 is invertable. A way to do so, is to let  $\gamma_s$  anneal together with  $T$ .
- The number of clusters must be balanced between a proper modeling anatomical variations and robustness to noise. The number of voxels per cluster  $n_c = \text{Vol}_{\text{tracts}}/n_c$  is invariant to differences in tract volume (i.e. the total volume of the voxels intersected by tracts). The latter parameter will be used hereafter.
- A certain sampling distance has to be chosen for tractography. A fraction of all points  $N_p$  emanating from the tractography is used in the registration, that is inversely proportional to the temperature. At each scale, sufficient points per cluster should be available for accurate estimation of the covariance matrices  $\Sigma_a^p$ .
- $\rho_o$ , the weighting parameter in equation 8.24, is related to the data. If the diameter of fiber tracts is known and  $T_{final}$  is given,  $\rho_o$  should be such that the membership value  $m_{ai}^p$  is low for points belonging to adjacent tracts, i.e. having a different orientation.

We consider  $n_c$  to be the critical parameter in the proposed algorithm. For that reason, a number of experiments will be performed in which particularly this parameter is varied (in the next section).

#### 8.4 EXPERIMENTS

The proposed algorithm will be evaluated in a series of experiments using both synthetic and brain data: The *added value* of the algorithm is demonstrated with respect to three aspects. Initially, the focus is on how the algorithm copes with volumetric data. Subsequently, the inclusion of orientation information into the algorithm for clustering adjacent, perpendicular fibers will be assessed. Lastly, we believe that the application should drive the design of the registration algorithm. The proposed algorithm is designed for studying FA-profiles over fiber tracts. Hence, the consistency in FA-values along tracts over subjects is studied after matching using the proposed method. The results are compared to a non-rigid registration method available in DTI-TK [264]. The value for  $n_c$  (and thereby the degrees of freedom) will be varied during the experiments as it directly determines the registration flexibility.

##### 8.4.1 Synthetic data

A synthetic dataset was generated such that there were two adjacent perpendicular tracts. The principal eigenvector field was defined accordingly. The resolution of the dataset was  $30 \times 30 \times 10$  voxels. The eigenvalues of the tensors were chosen to be  $\mathbf{E} = \text{diag} \left( \begin{bmatrix} 1.5 & 0.5 & 0.5 \end{bmatrix} \right) \cdot 10^{-3} \text{ mm}^2\text{s}^{-1}$ , such that  $\text{FA}=0.60$ . From the eigenvectors and -values, tensors were computed, following  $\mathbf{D} = \mathbf{V}^T \mathbf{E} \mathbf{V}$ , in which  $\mathbf{V}$  is the matrix of eigenvectors with  $v_{2,3}$  chosen in the plane perpendicular to the defined  $v_1$ . No noise was added to the data. Merely performing a clustering on fiber tracts generated in this data (omitting atlas generation) yields the outcome in figure 8.2. In all experiments,  $T_{init}$  and  $T_{final}$  were set to  $10^{-1}$  and  $10^{-4}$  respectively. The temperature was step-wise attenuated by  $\alpha = 0.9$ . In the membership computation,  $\rho_0 = 1$ , to have an attenuation factor  $e^{-1} \approx 0.4$  for perpendicular eigenvectors (equation 8.24). The clustering was done twice, by the classic and proposed method respectively. The dyadic coherence  $\kappa$  was determined for different numbers of voxels per cluster  $n_c$ . The two plots in figure 8.2(a) and (b) relate to  $n_c = 50$ , where the difference in  $\kappa$  between both methods is largest. Fiber tracts are colored corresponding to the orientation of the principal eigenvectors of the cluster center points with the highest membership value. Apparently, the conventional method results in cluster center points holding fibers from both tracts. This may be noticed by one tract asserting the color label of the adjacent tract.

In the next experiment, two synthetic tensor fields were generated which are non-rigidly warped versions of each other. The tensor fields consisted of  $80 \times 80 \times 10$  voxels, with an isotropic voxel size of 2 mm. The datasets were registered by means of the classic (original) and the proposed point set clustering algorithm. The point sets were scaled

to the unit circle by a factor  $0.01^1$  and the center of gravity was placed at the origin. The spline-transformation regularization parameter  $\gamma_s$  was empirically chosen to be  $\gamma = 10^{-2} \cdot T/T_{final}$ . The membership  $m_{ai}^p$  for three selected cluster center points at three iterations in one data set is depicted in figure 8.3.

Going from iteration 1 to 2, the cluster center points can be observed to move to the center of the points with the highest  $m_{ai}^p$ . At lower temperature, in iteration 50, the clusters are aligned along the fibers. The registration outcome for varying  $n_c$ , i.e. the number of voxels per cluster, is shown in figure 8.4. It may be observed that as the cluster center points remain co-linear, mis-registration results (top left). When using the proposed membership method, the ellipsoidal shaped clusters earlier start to distribute over the volume, resulting in a higher overlap. In other words, a volumetric representation of the fibers is a prerequisite to accurate registration. Concurrently, one may observe that at a high  $n_c$  the proposed algorithm already yields a fairly good outcome (top right). This behaviour is confirmed in figure 8.5, in which the relative overlap is given as a function of  $n_c$ . Notice that a high  $n_c$  is typically preferred to keep the number of parameters low.

#### 8.4.2 Brain data

DTI-data were acquired on a Philips Intera 3.0 Tesla MRI scanner (Philips Intera, Philips Medical Systems, Best, The Netherlands) by means of a spin-echo EPI sequence. Balanced diffusion sensitizing gradients were used to minimize artifacts induced by eddy currents [213]. The diffusion weighting was along 16 icosahedric and dodecahedric directions on a half sphere [9] and along the 16 corresponding, antipodal directions. Other parameters were: TE 94 msec, TR 4831-6248 msec,  $b=1000$  s/mm<sup>2</sup>, FOV 250 mm, scan matrix 70 x 112, image matrix 128 x 128, slice thickness 2.2 mm, such that the voxel size was 2.0 x 2.0 x 2.2 mm. Eddy current induced morphing was corrected for by an affine registration in the phase direction [173].

In one dataset, fibers were tracked at the junction of the corticospinal tract, cingulum and corpus callosum. A clustering of this data was performed, again using the classic and proposed method. The results for  $n_c = 30$  are displayed in figure 8.6, together with the dyadic coherence  $\kappa$  as a function of  $n_c$ . The proposed method obtained a higher  $\kappa$ , independent of  $n_c$ . Here, part of the corpus callosum is erroneously labeled as if it were part of the cingulum by the classic method, whereas the proposed method is able to distinct between these tracts. Correct clustering of fiber tracts stimulates a unique correspondence between datasets.

From an ongoing clinical study ten subjects were selected in which an expert did not identify visible pathology. In these datasets, the right corticospinal tract, right inferior longitudinal fasciculus, right cingulum and forceps major were tracked.

<sup>1</sup> Doing so typically scales the translation parameters down by an order of magnitude, such that the affine transformation can be estimated with higher stability.

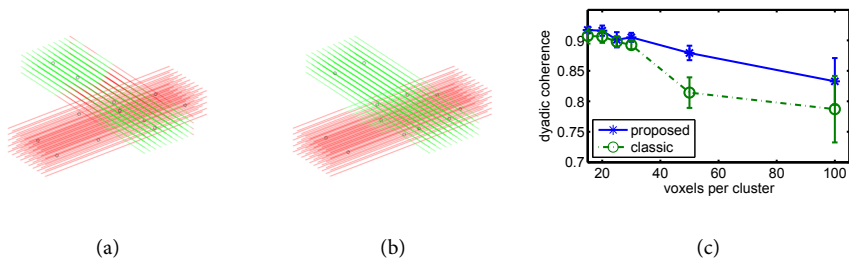


Figure 8.2: Synthetic fiber tracts, with labeling based on the orientation of the cluster center points, after clustering using the classic and proposed membership method. Here,  $n_c = 50$  is used, determined from the plot of  $n_c$  as a function of  $\kappa$  for both clustering methods.



Figure 8.3: Membership  $m_{ai}^p$  for three cluster center points colored by red, blue and green, for iterations 1, 2 and 50 (left to right).

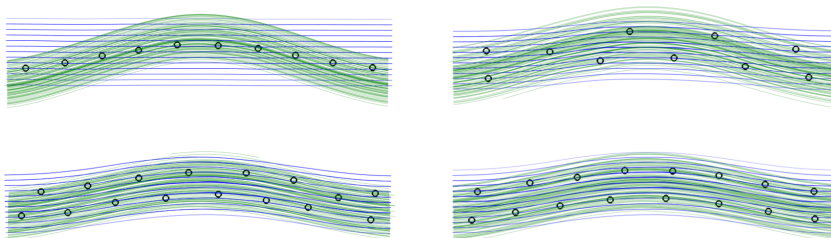


Figure 8.4: 2D-projection of the results of the synthetic fiber tract matching experiment, with  $n_c = 500$  (top) and  $n_c = 300$  (bottom), by means of the classic (left) and the proposed method (right). Note that, for equal  $n_c$ , the proposed method has a higher overlap.

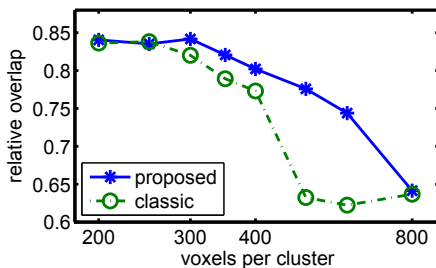


Figure 8.5: Relative overlap of the two synthetic fiber tracts after registration, as function  $n_c$ , using the classic and proposed membership method.

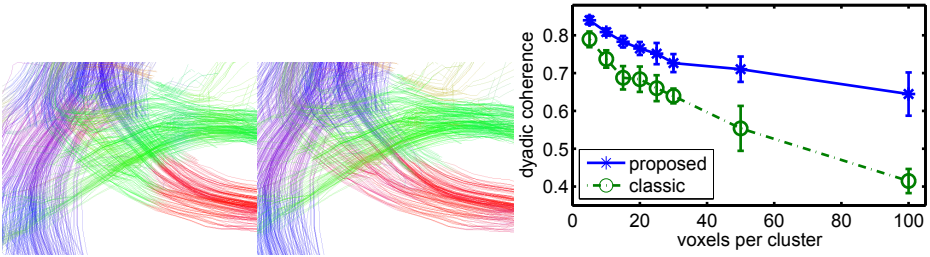


Figure 8.6: Fiber tracts in the brain, with colored labeling based on the orientation of the cluster center points, after clustering using the classic and proposed membership method. Here,  $n_c = 30$  is used, determined from the plot of  $n_c$  as a function of  $\kappa$  for both clustering methods in the plotted region.

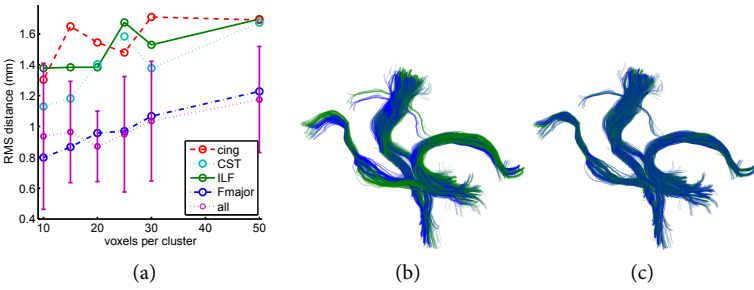


Figure 8.7: (a) RMS distance of the four tracts and the combination of them all (with error bars denoting the standard deviation over the 10 experiments) after reconstructing the applied non-rigid transform. (b) Transformation, applied to all tracts in one dataset and (c) reconstructed transformation using the proposed algorithm.

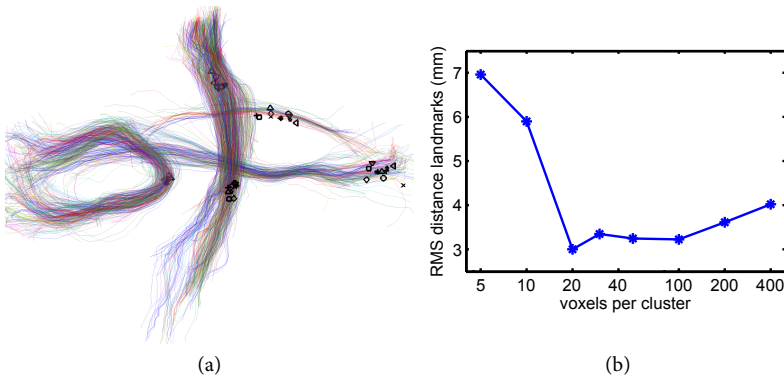


Figure 8.8: (a) An impression of the tracts and landmarks after registration, datasets are colored individually and (b) the RMS distance (in mm) of the landmarks as a function of the number of voxels per cluster, after registration using the classic and proposed membership method.

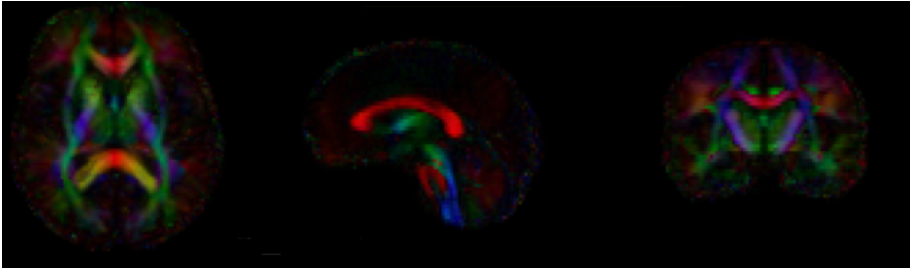


Figure 8.9: Three cross-sections of the FA color-coded by orientation of the DTI atlas generated with DTI-TK.

The FA-volumes of the ten subjects were registered to an available FA-atlas [252], using the Statistical Parametric Mapping (SPM)-method. In SPM, volumes are non-rigidly registered to a template using a Discrete Cosine Transform (DCT) [24]. The DCT consisted of  $12 \times 12 \times 6$  basis functions, enabling detailed registration of the full brain (visually checked). The resulting (considered representative) deformation fields were applied to the tracked fibers of the ten subjects. The algorithm's ability to reproduce these deformations was then determined. The original and deformed tracts were individually as well as jointly matched by means of the proposed method. The RMS distance after registration, for varying  $n_c$  is displayed in figure 8.7. A careful observation reveals a higher error along elongated tracts, due to a translational invariance along these tracts. The lowest error is achieved if the combination of tracts is used. The standard deviation of the estimated distance yields approximately 0.3 mm, indicating a sufficiently low precision in repetitive measurements.

A neuro-radiologist annotated 5 landmarks in the FA-volumes of all 10 subjects on locations through which the tracts pass. These landmarks were randomly picked by the expert provided that they corresponded to unique locations with respect to the surrounding cerebral structures. The landmarks were co-registered with the fiber tracts. An illustration of the result is shown in figure 8.8 (a). It may be noticed that those points on elongated tracts are less precisely registered (as may be expected). The graph in the figure 8.8 (b) gives the root mean square distance to the center of gravity of the points as a function of  $n_c$ . It may be observed that there is a broad optimum for  $n_c$ . The optimal choice boils down to the bias-variance dilemma: sufficient clusters are needed to code the degrees of freedom of the deformation. Too few voxels per cluster will increase the variance per cluster. Fortunately, the method is not highly sensitive to this parameter setting. A lower value for  $n_c$  is to be preferred though,  $20 < n_c < 40$ , in order for the clustering to successfully distinct between tracts.

All tensor volumes were registered to a template and an average tensor field (atlas) was computed [264, 265]. The atlas is depicted in figure 8.9. Subsequently, we studied the FA-values along several fiber tracts after fiber tract matching and DTI-TK registration respectively. These fiber tracts were reconstructed in the DTI-TK atlas, see figure 8.10,



Figure 8.10: Fiber tracts reconstructed in the DTI-TK atlas, clustered subsequently by the proposed method.

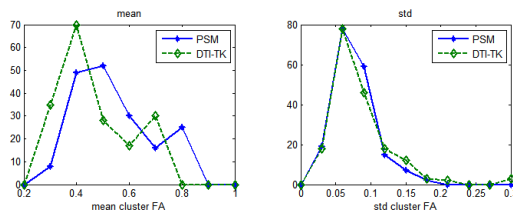


Figure 8.11: Histograms of the mean and standard deviation of the mean FA per cluster over subjects, for the DTI-TK and proposed method. The vertical axis denotes the number of clusters per histogram bin.

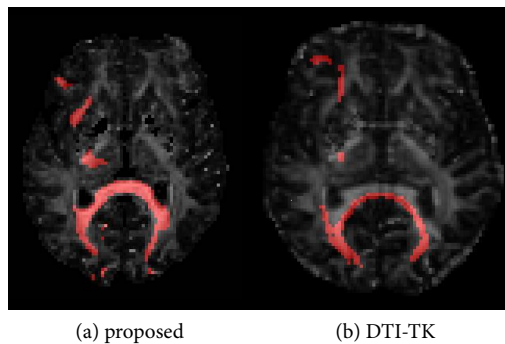


Figure 8.12: An axial slice of one subject, with voxels through which fibers pass labeled red. (a) The proposed method tracks fibers in subject space. (b) Atlas-based tracking after DTI-TK registration.

after which the fibers were clustered with 20 voxels per cluster  $b$ . The resulting fiber points with corresponding membership values and cluster centers were then superimposed on the FA-volumes of the warped tensor datasets. A weighted mean FA per cluster per subject was then computed,  $FA_{DTI-TK_b}^p = \sum_{j=1}^{N^{atlas}} m_{b_j}^p FA_{b_j}^p$ . Similarly, FA-values along the fibers tracked in the original datasets were derived. After point set matching the fibers, mean FA values of the resulting clusters (different from the DTI-TK clustering)  $FA_{PSM_a}^p = \sum_{i=1}^{N^p} m_{a_i}^p FA_{a_i}^p$  were computed. Figure 8.11 depicts histograms of the mean and standard deviation of the FA over subjects for both methods. A clear trend towards a lower mean FA for the DTI-TK method is seen while the standard deviations do not differ significantly. Visual inspection revealed that atlas fibers occasionally do not overlap subjects' FA volumes due to small residual misalignments (figure 8.12).

The total computation time of registering the four tracked bundles in ten subjects was 23 minutes using the proposed method and 11 minutes using classic point set matching, on an AMD Opteron 2GHz processor with 8 Gbyte of RAM memory. Profiling revealed that interpolation of the eigenvectors (equation 8.23) and computation of the transformation parameters (equations 8.14 and 8.15) was most costly.

## 8.5 CONCLUSION

We introduced a new method for non-rigid registration of Diffusion Tensor Images by means of white matter fiber tracts. The method boils down to point set registration which involves simultaneously clustering and matching of the data points. The tracts are implicitly warped to a common frame of reference to avoid the potential bias towards one of the datasets. The algorithm gradually refines from global to local registration, which is implemented through deterministic annealing. Special care was taken to incorporate the spatial relation between fiber points and the uncertainty in principal diffusion orientation. We implemented this by properly weighting both the cluster center point covariance matrices and the membership assignments of points to specific clusters.

The concept of point set matching was applied to volumetric data. The number of cluster center points (reflecting the flexibility of the registration) was varied in each experiment. Looking back at the inter-subject registration, it may be observed that there is a broad optimum (figure 8.8). Apparently, the method is not very sensitive to this parameter setting. Furthermore, it was shown that both a high curvature and low inter-tract variability induce a lower error. If a registration of an elongated tract, such as the corticospinal tract, is desired, a higher precision in registration is obtained by including a perpendicular and adjacent tract, such as the inferior longitudinal fasciculus.

The method has been build with the aim of comparing FA-profiles along fiber tracts. In contrast to a state-of-the-art non-rigid registration technique, the proposed method is more robust to residual misalignments. Fiber tracts and FA-values are sampled in subject space prior to transformation, while in atlas-based tracking sampling occurs after correspondence is found. Meanwhile, the root mean squared distance with respect



to expert-annotated landmarks is low (3 mm). Population studies will benefit from this robustness, as potentially fewer subjects are needed to obtain a significant difference between cohorts. Our future work will focus on such an assessment of registration algorithms in a machine learning framework.

A thin plate spline transformation was used to model the non-rigid transformation. In a thin plate spline, folding may occur, such that the inverse transformation is ill-defined. Concurrently, the inverse operation is needed to go back and forth from image to atlas space. Instead, a diffeomorphic mapping such as the geodesic interpolating spline may be used [262]. For now, the problem is circumvented by properly regularizing the transformation.

Consistent fiber tracking in all subjects is vital for a successful registration. Instead of a naive streamline technique, stochastic fiber tracking may yield in a more robust tracking and thus a better registration.

To the best of our knowledge, there is no other work on non-rigidly matching DTI data based on the white matter tracts. Incorporating information about the course of tracts into the registration may add to pathological studies, and open the way to more knowledge of brain diseases in the future.



## SPATIAL CONSISTENCY IN 3D TRACT-BASED CLUSTERING STATISTICS

---

We propose a novel technique for tract-based comparison of DTI-indices between groups, based on a representation that is estimated while matching fiber tracts. The method involves a non-rigid registration based on a joint clustering and matching approach, after which a 3D-atlas of cluster center points is used as a frame of reference for statistics. Patient and control FA-distributions are compared per cluster. Spatial consistency is taken to reflect a significant difference between groups. Accordingly, a non-parametric classification is performed to assess the continuity of pathology over larger tract regions. In a study to infant survivors treated for medulloblastoma with intravenous methotrexate and cranial radiotherapy, significant decreases in FA in major parts of the corpus callosum were found.

Published as:

M.W.A Caan, L.J. van Vliet, C.B.L.M. Majoie, E.J. Aukema, C.A. Grimbergen and F.M. Vos, *Spatial Consistency in 3D Tract-Based Clustering Statistics*, in: Proc. MICCAI, Part I, LNCS 5241, 2008, 535–542 [63].

## 9.1 INTRODUCTION

Over the past decade Magnetic Resonance Diffusion Tensor Imaging (MR-DTI) has been used to characterize the local white matter structure in the human brain. Typically, properties derived from the tensor, such as the Fractional Anisotropy (FA) [32], are involved to study differences between diseased subjects and normals.

It may be observed that not all the tensor information (represented by orientation and shape) is used in the latter analysis. For instance, connectivity information is neglected as tensor shape properties are studied on a voxel- or region-of-interest-basis. Alternatively, fiber tracts are used as a subspace in which properties are averaged. However, high variation in white matter anatomy exists along white matter tracts [115]. Not surprisingly, analysis is currently focusing on profiles along tracts. For instance, in a study to Amyotrophic Lateral Sclerosis (ALS), it was recently found that only parts of the corticospinal tract were affected by the disease [221]. Here, a Mann-Whitney U-test was performed to assess local spatial consistency.

Prior to such inter-subject comparison of tract-based features, correspondence between the tracts is to be achieved. Anatomical variability across subjects needs to be accounted for in a non-rigid manner. Preferably, correspondence is not only obtained along, but also in the plane normal to the local tract orientation.

In case of small patient and control groups, diffusion properties need to be spatially combined to ensure sufficient sensitivity for discrimination. This may be achieved by a statistical description, e.g. some mean property such as FA over the tract and the corresponding standard deviation in case of a unimodal distribution. Clearly, such averaging goes at the cost of losing spatial information. A more specific approach is to employ the correlation in the data, after which it is found which voxels do or do not contribute to an observable difference between patients and controls [66].

Relevant related work includes affine matching of fibers [89], such that diffusion properties are represented along tracts. Inter-subject comparison is not facilitated by either approach, though. Joint probabilistic clustering and point-by-point mapping achieves arc-length tract representations [168]. Tract-based morphometry also uses such an arc-length representation, based on matched tract regions to a chosen prototype. Subsequently, it is used for detecting white matter differences between groups of subjects [188]. Neither do these methods achieve non-rigid tract correspondence, nor is an unsupervised partitioning of tracts facilitated. Tract based spatial statistics projects non-rigidly registered volumes onto a skeletonization of the FA [226], but tensor orientation information is not used in the alignment, nor in the analysis.

We propose a framework for tract-based analysis that partitions tracts in an unsupervised manner. The method involves a non-rigid registration based on a joint clustering and matching approach, after which a 3D-atlas of cluster center points is used as a frame of reference for statistics. Underlying shape coherence of trajectories is captured

in anisotropic cluster shapes. Patient and control FA-distributions are compared per cluster. Spatial consistency is taken to reflect a significant difference between groups. Accordingly, a non-parametric classification is performed to assess the continuity of pathology over larger tract regions.

## 9.2 METHOD

### 9.2.1 Feature selection

The analysis starts with selecting fiber tracts of interest both in patients and controls. We perform a full brain tractography using the FACT algorithm in DTIStudio software [140], after which ROIs are placed that, combined by logical operators, define a sub-selection of fibers, here the corpus callosum.

### 9.2.2 Registration

Initially, the fibers need to be brought into spatial correspondence. The concept of point set matching is applied to points along the fiber tracts in order to do so [84].

Suppose that in total,  $P$  datasets are to be aligned. Each dataset  $p \in \{1, \dots, P\}$ , is represented by sampling the selected fiber tracts, yielding point-sets consisting of  $N^p$  points  $X^p = \{x_i^p, i = 1 \dots N^p\}$ . For each point set, a set of  $K$  cluster center points is defined, denoted by  $C^p = \{c_a^p | a = 1 \dots K\}$ . An atlas cluster point set  $\mathcal{Z} = \{z_a | a = 1 \dots K\}$  is defined that is to explain the cluster centers over all point sets. By means of a deterministic annealing approach, the registration process gradually refines from global to local matching. The initial temperature  $T_{init}$  and final temperature  $T_{final}$  are a priori defined.  $T$  is lowered after each iteration:  $T_{ind+1} = \alpha T_{ind}$ , with  $0 < \alpha < 1$ .

A mixture of Gaussians models the density of the point set, in which the membership variable  $m_{ai}^p$  indicates the degree to which a point "feature"  $x_i^p$  belongs to cluster center point  $c_a^p$ . The membership  $m_{ai}^p$  is now defined as

$$m_{ai}^p = \frac{q_{ai}^p}{\sum_{a=1}^K q_{ai}^p}, \forall a, i, \quad (9.1)$$

with

$$q_{ai}^p = \exp\left(-\frac{1}{2\sigma^2} |x_i^p - c_a^p|^2\right). \quad (9.2)$$

assuming isotropic covariance matrices ( $\Sigma_a^p = \sigma^2 I$ ), and, after normalization,  $\sum_{a=1}^K m_{ai}^p = 1$ . The temperature is defined to be the variance in equation 9.2,  $T = \sigma^2$ .

The algorithm proceeds in an iterative manner. The cluster center points are initiated at randomly chosen points in the dataset. They are recomputed after the membership has been updated (eq. 9.1), using

$$c_a^p = \frac{\sum_{i=1}^{N_p} m_{ai}^p x_i^p}{\sum_{i=1}^{N_p} m_{ai}^p}, \forall a \in \{1, \dots, K\}, \forall p \in \{1, \dots, P\}. \quad (9.3)$$

Then, the atlas points are formed by averaging the cluster center points,  $z_a = 1/P \sum_{p=1}^P c_a^p$ . The atlas points can be seen as a “common frame of reference”, representing the mean shape of the point sets. An affine and non-rigid thin plate spline transform is now computed to warp the atlas points to the cluster center points and vice versa. The next iteration is started by lowering the temperature and transforming the atlas points to the subsequent datasets. In the latter step, information from other datasets is implicitly included in clustering a single dataset.

### 9.2.3 Statistics

DTI is sensitive to detecting pathology that manifests itself along white matter pathways. The challenge is to be precise in describing which local regions specifically characterize a disease. We assume for our analysis that the voxels through which a fiber tract passes are somehow connected. By jointly analyzing such samples, statistical power is to be gained compared to per-voxel comparison. What is more, a multi-modal analysis of the data may actually be necessary: simply averaging data over the connected voxels delivers a physiologically feasible mean only if there is a unimodal distribution.

We propose to use the clustering outcome of the registration as the statistical frame of reference; i.e. correspondence is implicitly defined via the atlas cluster points  $z_a$ . The membership function (equation 9.1) defines the extend to which points  $x_i^p$  are assigned to cluster centers point  $c_a^p$ . It may be observed that in such a way fiber tract statistics are modeled in 3D since the data points emanate from several fibers that constitute a tract. To the best of our knowledge, previously proposed arc-length representations have not yet been generalized to higher dimensions. A cumulative distribution of FA-values per patient per cluster is build, in which the membership values act as relative weights. Consequently, distant voxels will get a low weight.

The clinical study described in this paper (see below) involves paired data of patients and controls. For each such pair, the distributions of FA data in the neighbourhood of a cluster center point is compared. In other words, a pair-wise comparison of FA-distributions is needed. As no prior information on the distribution is available, a non-parametric test is required. We use the Kolmogorov-Smirnov (KS)-test to do so in a two-tailed fashion: for each patient/control pair and for each cluster center point, it is determined if either the FA-distribution of the one is significantly above or below the other or if the difference is insignificant. The latter comparison is performed over all patient/control

pairs from which the majority vote is retained. Effectively, this score boils down to either a higher or lower FA, ties are discarded.

Subsequently, the consistency of the outcome along the tract is checked. Here, the underlying hypothesis is that if the voting score is similar over a larger number of nearby clusters, it is an indication of a significant difference between the groups. In order to check the 'consistency' two classes are asserted, corresponding to dominance of either the patient or control distributions. Class assignment is obtained by unsupervised classification by means of a k-nearest neighbour classifier. A leave-one(-cluster)-out cross validation scheme is involved to train and test the classifier, resulting in the relative classification error over all clusters. It is taken that a small classification error indicates that the dominating distribution can be predicted from neighbouring data. This is considered to reflect a systematic difference between the groups.

For comparison purposes a paired t-test is performed on averaged FA-values over the clusters, using the membership values as relative weights. Notice that such a comparison does not include any spatial connectivity.

To assess the validity of both approaches, statistics are computed on ten randomly permuted data sets (i.e. by randomly assigning the class labels 'patient' or 'control' to the data).

### 9.3 RESULTS

Six infant survivors treated for medulloblastoma with intravenous methotrexate and cranial radiotherapy were included in the study as well as age-, education- and sex-matched healthy controls. The aim of this study was to confirm the hypothesis that the treatment affects a major white matter tract, the corpus callosum. The number of subjects emanated from a power analysis performed a priori, expecting a major effect. All subjects were scanned on a 3.0T Philips Intera (Best, The Netherlands) MRI-scanner. After data inspection, it was found that unfortunately one patient had incomplete data and had to be excluded (and simultaneously the matched control).

Fiber tracking was performed in the genu, center and splenium of the corpus callosum. Because of white matter degeneration, other corpus callosum parts could not be successfully tracked in all subjects. Fiber coordinates (in mm) were scaled by a factor of 0.01, such that increased numerical stability between translation and rotation parameters was obtained. The resulting tracts from the 10 datasets were jointly matched. The optimal cluster size is determined experimentally below (see figure 9.4). The temperature was step-wise lowered by a factor of 0.9, starting at  $T_{init} = 10^{-2}$ , until the final temperature of  $T_{final} = 10^{-3}$  was reached. As an illustration, registered tracts of one randomly chosen patient/control pair are displayed in figure 9.1. To account for the inherent stochastic nature of the clustering, ten realizations of atlas points were combined.

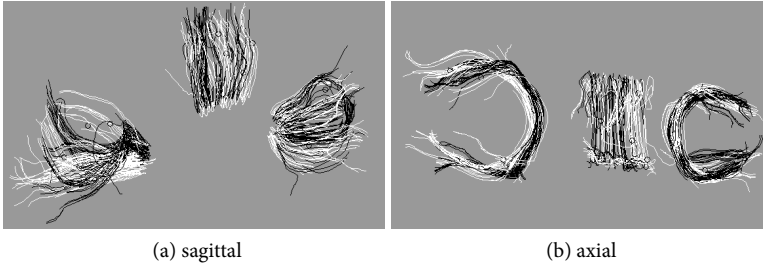


Figure 9.1: Transformed fibers of the corpus callosum (that could be reproducibly tracked in all subjects) after matching. Fibers of one patient and one control are shown in black and white.

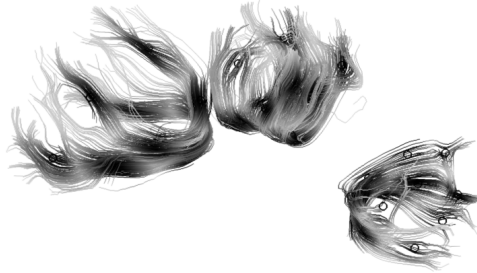


Figure 9.2: The maximum membership per fiber-point  $\max_i m_{ai}^p$  and cluster center points after one matching instance, in one dataset.

Next, the membership  $m_{ai}^p$  was computed for all voxels included in the tracking. Here,  $T = 5 \cdot 10^{-3}$  was chosen, yielding a uniform coverage of the clusters over the data. The result is depicted in figure 9.2. Subsequently, the cumulative FA-distributions were calculated using the membership function as weighting. In the Kolmogorov-Smirnov test to assess the difference per patient/control pair, the number degrees of freedom equaled the cluster size in voxels. The distributions in one cluster for all patients and corresponding KS-significances are given in figure 9.3.

In a leave-one-out cross-validation, a  $(3 \cdot 10)$ -nearest neighbour classifier was trained on the voting outcome over all pairs of the 10 matching instances, and the classification error was determined. Registration and classification were executed ten times with a range of voxels per cluster. The results are shown in figure 9.4. The error for randomly composed classes is proportional to the cluster size. Due to the local continuity of the FA, neighbouring values can be predicted, regardless of the class composition. The difference with randomly assigned class labels was significant ( $p < 0.05$ ) for cluster sizes larger and 50 voxels [15]. The outcome of the voting and classification process for a cluster size of 200 voxels is displayed in figure 9.5. A globally decreased FA can be observed.



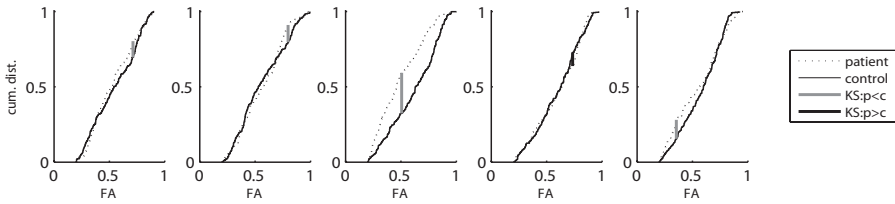


Figure 9.3: Cumulative FA-distributions for five patient/control pairs in one arbitrarily chosen cluster, with Kolmogorov-Smirnov test results. Voting yields a lower FA for patients for this cluster.

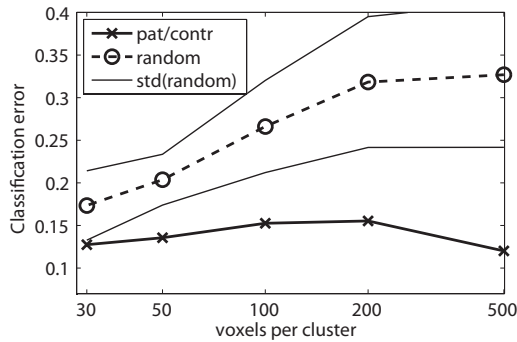


Figure 9.4: Classification error as function of the cluster size for both patients/controls and randomly composed classes. The latter is a mean over ten different class compositions, for which the corresponding standard deviation is plotted in gray.

For comparison, a paired t-test per cluster center point was done (we did not include a correction for the number of comparisons at this stage). 22% of the total number of cluster center points yielded a significant difference ( $p < 0.05$ ). These are annotated in figure 9.5. In comparison to the results of the 2-nn classifier, less consistent and more evenly distributed differing points can be observed, which we consider not trustworthy. Randomly permuting the class labels 500 times yielded that only in 4% of the permutations a significant difference was found in 22% or more cluster center points. From this we conclude that the differences might be conceived as mathematically significant (since the percentage is lower than 5%), despite clinical reserves.

#### 9.4 DISCUSSION

We presented a novel approach to 3D, tract-based statistical analysis of FA data. Inter-subject correspondence was achieved by non-rigid registration based on a joint clustering and matching. The clustering delivered atlas points that served as a frame of reference for performing the analysis. A distribution of FA-values per cluster and per subject was cal-

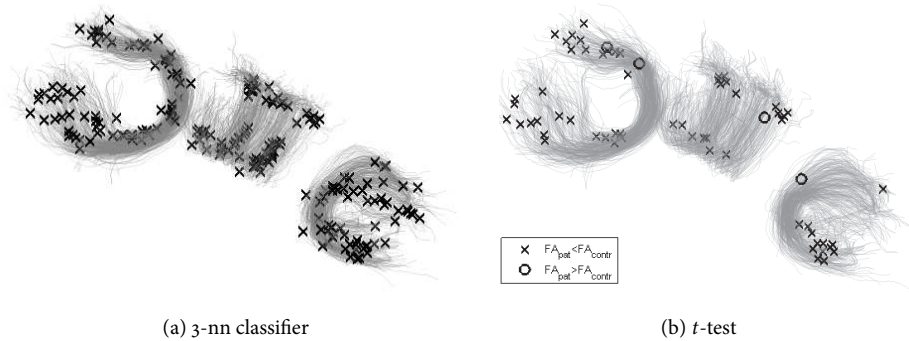


Figure 9.5: Result of the voting and classification procedure, depicting correctly classified clusters with decreased ( $\times$ ) and increased FA ( $\circ$ ) for patients. Also, the results of  $t$ -test statistics (without multiple comparison correction) are shown.

culated. Subsequently, the Kolmogorov-Smirnov test for inequality of distributions was used to pair-wise compare patients and controls. The latter comparison was performed over available patient/control pairs from which the majority vote was retained.

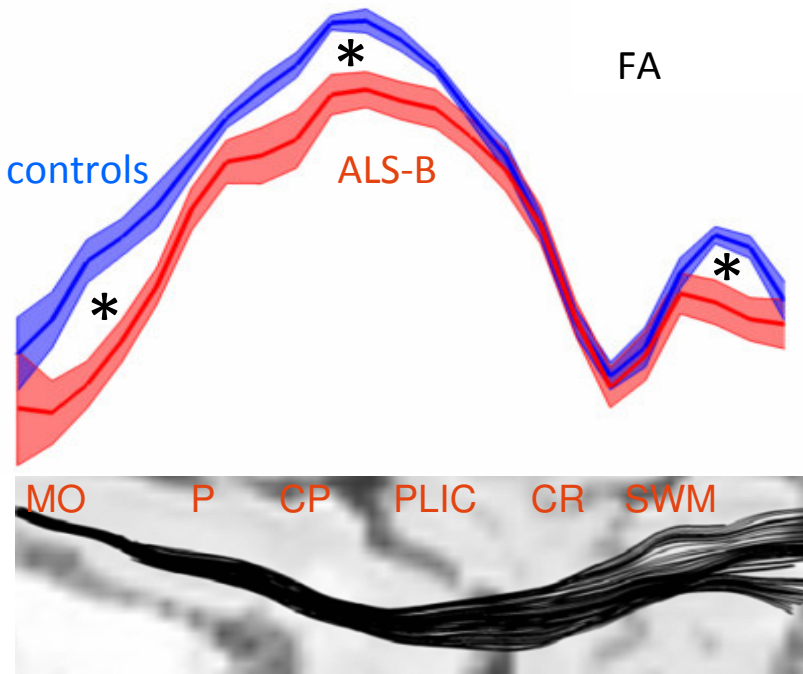
In a study into infant medulloblastoma survivors, corpus callosum tracts from patients and controls were pair-wise studied. A decreased FA in larger parts of the corpus callosum was observed. The decrease could be predicted by a 3-NN classifier with an error of 14% when 200 voxels per cluster were used, which indicated a significant difference. In comparison to  $t$ -test statistics per cluster, the proposed method was able to detect larger regions in which patients differed from controls.

Voxel-based analysis is well-known to be relatively sensitive to small mis-registration. In contrast, the proposed method is merely affected by the mis-registration component that is parallel to the tract. Due to the continuity of the FA, this effect could be limited. Actually, in figure 9.1, we illustrated a good overlap between tracts of an arbitrarily chosen patient/control pair.

Although manual tracking was performed in this study, we believe that this work can easily be adapted to (un)supervised full brain tractography clustering methods. Additionally, analyzing larger cohorts that are partitioned in several subgroups, could be facilitated by a multi-sample Kolmogorov-Smirnov test.

# CLINICAL STUDIES

# IV





## WHITE MATTER FRACTIONAL ANISOTROPY CORRELATES WITH SPEED OF PROCESSING AND MOTOR SPEED IN YOUNG CHILDHOOD CANCER SURVIVORS

---

### **Purpose**

To determine if childhood medulloblastoma and acute lymphoblastic leukaemia (ALL) survivors have decreased white matter fractional anisotropy (WMFA) and if WMFA is related to the speed of processing and motor speed.

### **Methods and Materials**

For this study, 17 patients (6 medulloblastoma, 5 ALL treated with high dose methotrexate (MTX) (4 x 5 g/m<sup>2</sup>) and 6 with low dose MTX (3 x 2 g/m<sup>2</sup>)) and 17 age-matched controls participated. On a 3.0-T magnetic resonance imaging (MRI) scanner, diffusion tensor imaging (DTI) was performed, and WMFA values were calculated, including specific regions of interest (ROIs), and correlated with the speed of processing and motor speed.

### **Results**

Mean WMFA in the patient group, mean age 14 years old (range 8.9 – 16.9), was decreased compared to the control group ( $P = 0.01$ ), as well as WMFA in the right inferior fronto-occipital fasciculus (IFO) ( $P = 0.03$ ) and in the genu of the corpus callosum (gCC) ( $P = 0.01$ ). Based on neurocognitive results, significant positive correlations were present between processing speed and WMFA in the splenium (sCC) ( $r = 0.53$ ,  $P = 0.03$ ) and the body of the corpus callosum (bCC) ( $r = 0.52$ ,  $P = 0.03$ ), while the right IFO WMFA was related to motor speed ( $r = 0.49$ ,  $P < 0.05$ ).

### **Conclusions**

White matter tracts, using a 3.0-T MRI scanner, show impairment in childhood cancer survivors, medulloblastoma survivors, and also those treated with high doses of MTX. In particular, white matter tracts in the sCC, bCC and right IFO are positively correlated with speed of processing and motor speed.

Published as:

Eline J. Aukema, Matthan W.A. Caan et al., *White Matter Fractional Anisotropy Correlates With Speed of Processing and Motor Speed in Young Childhood Cancer Survivors*, Int J Rad Onc Biol Phys, 74(3), 837–843, 2009 [29].

## 10.1 INTRODUCTION

Today, in developed countries such as the U.S., about one in every 450 adolescents reaching the age of 20 will be a long-term cancer survivor [184]. Due to multimodal treatment strategies, the overall survival rate of children with a brain tumor has increased dramatically [58]. Unfortunately, progress in treatment strategies has not been able to prevent treatment-related side effects, such as neurotoxicity. For instance, children treated with craniospinal radiotherapy (CSRT) and chemotherapy for a medulloblastoma often experience serious neurocognitive impairments. Deficits in attention, memory and speed of processing are commonly found in survivors [58, 180, 186, 59]. Children treated for acute lymphoblastic leukaemia (ALL) have also shown treatment-induced neurotoxicity; this is probably due to intrathecal and/or high dose intravenous methotrexate (MTX), which is a replacement of cranial radiation therapy as central nervous system (CNS) prophylaxis, although published data are inconsistent [212, 70, 177].

The neurocognitive impairments have been associated with white matter changes due to cranial (spinal) radiation therapy (C(S)RT), some chemotherapeutic agents including MTX, and other factors, such as tumor infiltration and hydrocephalus [183, 210, 211]. The treatment-induced neurotoxicity may be caused by either a failure of “normal” maturation and myelination of the brain at an age-appropriate rate or by damage to already-existing white matter tracts. In children treated for cancer, negative effects based on both assumptions seem likely.

Diffusion tensor imaging (DTI), an advanced brain imaging technology, enables the study of the integrity of white matter structures, which are most vulnerable for toxic treatment. The diffusion of water molecules is high along and low perpendicular to coherent white matter tracts, resulting in an anisotropic diffusion profile. The white matter fractional anisotropy (WMFA) value quantifies this anisotropy, with zero for isotropic and one for fully anisotropic diffusion profiles. WMFA reflects the myelination and axonal integrity [32]. Increase of WMFA during childhood and adolescence parallels the development of important basic cognitive functions [31, 185].

Diminished WMFA seems potentially useful for detecting and monitoring white matter damage. Thereby, significant positive correlations were found for WMFA in different brain regions with intelligence and neuropsychological functions in medulloblastoma survivors [150, 223, 152]. Likewise, these positive correlations have been found in children with traumatic brain injury [260] and in patients suffering from a wide range of psychiatric disorders [116]. White matter plays an important role in the speed of processing, which is crucial for learning and coping in daily life, one of the main problems for childhood cancer survivors, especially brain tumor survivors. It is unknown whether diminished WMFA is related to these problems in speed of processing; increased insight into this relationship might help to predict neurocognitive decline in the future.

The aims of this study were 1) to estimate the functioning of cancer survivors (survivors from a medulloblastoma and ALL) with respect to intelligence, speed of processing

and motor speed, 2) to measure WMFA by MRI in childhood cancer survivors after treatment with CRST and/or high dose intravenous MTX compared to peers, and 3) to relate neurocognitive function (intelligence, speed of processing and motor speed) with WMFA in different regions of interest (ROIs). We prospectively studied WMFA using a 3.0-T MRI and neurocognitive functioning in a group of childhood cancer survivors compared to healthy peers.

## 10.2 METHODS AND MATERIALS

### 10.2.1 *Patients*

Survivors treated in the Emma Children's Hospital for a medulloblastoma or ALL, between 8 and 16 years old and at least 3 years after the end of treatment in medulloblastoma survivors or 3 years after intravenous MTX as CNS prophylaxis in ALL survivors, were eligible for this study.

To compare different treatment modalities, our study design consisted of 3 subgroups of childhood cancer survivors ("patient group"); 6 medulloblastoma survivors treated with surgery, radiation (whole brain and spine: total dose ranging from 25.2 to 34.5 Gy and posterior cranial fossa boost: ranging from 53.3 to 55.4 Gy) and chemotherapy, including lomustin, vincristin and cisplatin; 6 ALL survivors treated with 4 x 5 gr/m<sup>2</sup> intravenous MTX ("high dose ALL") according to the DCLSG protocol 1993-1997; and 6 survivors treated with 3 x 2 gr/m<sup>2</sup> intravenous MTX ("low dose ALL") according to the DCLSG ALL-9 protocol 1997-2004 as CNS prophylaxis [55, 56].

Parents or adolescents were first contacted by phone and then received written information about the study.

We approached 8 medulloblastoma survivors and received 6 positive reactions. We then searched for age- and sex-matched survivors treated for ALL with comparable time after completion of treatment. Twenty ALL survivors were selected, of whom 7 did not want to be confronted with their cancer history again. We received 13 approvals for participation, and they were divided into the two different treatment groups.

After inclusion of the survivors, we selected a "control group" of classmates of the participating survivors. A suitable classmate, matching the survivor as closely as possible according to age, sex and the level of education, was chosen by the school, unless survivors picked their own classmate due to privacy reasons, or a suitable classmate of a different patient was approached by us.

The survivor and control were scheduled together for the MRI and neurocognitive evaluation to reduce possible anxiety through peer support and to increase participation. Age-specific illustrated information about the MRI procedure was provided, while professional child-directed support and the possibility of using a distracting audiotape

or videotape were available during the MRI. Specific requirements as described by Dutch law and the behavioral research code of the Dutch Association for Pediatrics were met in this study design, which was approved by the local medical ethics committee.

Finally, 17 survivors (6 medulloblastoma; 5 high dose ALL and 6 low dose ALL) participated in this study, since one survivor was unable to finish his MRI session due to anxiety and one survivor had a steel splinter in his eye.

#### 10.2.2 *Measurements*

Diffusion tensor imaging (DTI) was performed on a 3.0T MRI scanner (Philips Intera, Philips Medical Systems, Best, the Netherlands). DTI acquisition was along 16 nonlinear and 16 antipodal directions. The other parameters were echo-time: 94 msec, repetition time: 4831-6248 msec, diffusion weighting parameter  $b$ : 1000 s/mm<sup>2</sup>, FOV: 240 mm, scan matrix: 70 x 112, and slice thickness: 3 mm. Eddy current-induced morphing was corrected by a 2D affine registration of the Diffusion Weighted Images to the Bo-image [173].

The participants were only informed about the MRI results if health-related data resulted, which was not the case in any of the children.

The neurocognitive tests were individually administered by two psychologists (EA, NO) in approximately 2.5 hours. All subjects completed a test battery to assess general intelligence, speed of processing and motor speed. This battery included the Dutch version of the Wechsler Intelligence Scale for Children, 3rd Edition [158], and the Purdue peg-board [114]. Participants were informed about their cognitive functioning and possible consequences for school and daily life.

The interval between neurocognitive testing and MRI ranged between 0.0 and 3.5 months.

#### 10.2.3 *MRI Data Analyses*

Structural images were judged by an experienced neuroradiologist for macroscopic white matter lesions, atrophy, status of the primary tumor, if appropriate (i.e., in medulloblastoma survivors), and possible other new lesions. FA-images were computed using Teem-software (<http://teem.sf.net>). Further analysis was performed in Matlab using Statistical Parametric Mapping (SPM5)-software (Wellcome Department of Cognitive Neurology, London, England) and Matlab software (Mathworks). As an initialization, all data were co-registered to the Echo Planar Imaging-template available within the SPM-toolbox.

The data were segmented into white matter, gray matter and cerebral spinal fluid (CSF) based on the Bo-image. A two-step procedure was performed. In the first iteration, the a priori white matter, gray matter and CSF-maps available in SPM were used to segment



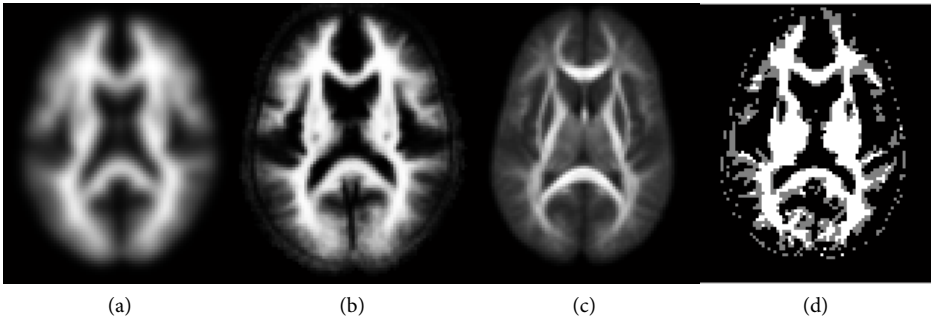


Figure 10.1: (a) A priori white matter segmentation map provided by SPM (see text) and (b) based on the data in this study. (c) FA averaged over all subjects after spatial normalization. (d) Voxels segmented into white matter of one patient-control pair, gray denoting either one of the two and white denoting both subjects, the latter being used in analysis.

the data. After this segmentation, average maps of the cohort were computed, and the second iteration was initiated using these maps. The resulting white matter mask was obtained using the following operation:  $i_2 > i_1 \ \& \ i_2 > i_3 \ \& \ i_2 > 1 - i_1 - i_2 - i_3$ , where  $i_1$ ,  $i_2$  and  $i_3$  are the grey, white and CSF-maps, respectively, resulting from the segmentation as shown in Figure 10.1. The WMFA-volumes were smoothed using a Gaussian kernel, with a size of 6 mm (FWHM). Next, all images were spatially normalized, using both an affine and non-rigid transformation. In this way, tiny segmentation errors could be corrected through pair-wise intersection of the masks.

Based on the literature [263, 167], white matter regions of interest (ROI) with a presumed relation with neurocognitive function were selected: the genu (gCC), the splenium (sCC) and the body of the corpus callosum (bCC) and the bilateral inferior fronto-occipital fasciculus (IFO). ROIs were outlined on color-coded WMFA maps. We first used a control subject and manually drew the ROIs. The ROIs identified in this subject were then used as a guide to manually define ROIs for other subjects as reproducibly as possible; subsequently, the ROIs were outlined manually by 1 operator (L.R.). Mean WMFA was measured in each ROI.

#### 10.2.4 Statistical Analyses

Overall WMFA (mean WMFA and mean ROI WMFA) was calculated, and correlations with cognitive functioning were analyzed using SPSS version 14.2. First, demographics and cognitive functioning of the participants were described. One-sample t-tests were performed to test whether mean scores on cognitive tests in the “patient group” (medulloblastoma, low dose ALL and high dose ALL) and the “control group” differed from the norm population [158, 114]. Differences in cognitive functioning between the patient subgroups were analyzed using multiple univariate analyses of variance (ANOVA). Second,

Table 10.1: Characteristics of the participants.

	MED	High-dose ALL	Low-dose ALL	Patient group	Control group
<i>n</i>	16	5	6	17	17
Gender					
Female	5	0	6	11	11
Male	1	5	0	6	6
Age at study (y)					
Mean (SD)	13.6 (3.0)	15.4 (1.4)	13.2 (2.7)	14.0 (2.5)	13.9 (2.9)
Min, max	8.9, 16.8	13.4, 16.9	10.1, 16.7	8.9, 16.9	8.8, 17.0
Age at diagnosis (y)					
Mean (SD)	4.7 (1.3)	3.7 (1.5)	7.1 (4.5)	5.2 (3.1)	–
Min, max	2.9, 6.7	2.2, 5.7	2.0, 13.2	2.0, 13.2	
Interval (y)					
Mean (SD)	8.8 (4.0)	11.5 (1.2)	5.9 (2.4)	8.4 (3.5)	–
Min, max	2.7, 13.3	10.6, 13.6	3.4, 9.9	2.7, 13.6	

Abbreviations: MED = medulloblastoma group; High-dose ALL = leukemia group, high-dose MTX ( $4 \times 5 \text{ g/m}^2$ ); low-dose ALL = leukemia group low-dose MTX ( $3 \times 2 \text{ g/m}^2$ ); Interval = time since end of treatment.

we studied group differences in overall WMFA between the patient group and the control group using t-tests. If overall tests for differences in WMFA were significant, we further analyzed the differences between the 3 patient subgroups with ANOVA. For assurance of the results, we also performed nonparametric Mann Whitney U tests because of the non-normal distribution of WMFA in these small subgroups. Finally, we calculated correlations of overall WMFA with intelligence, speed of processing and motor speed in the patient group. We followed Cohen in considering correlation coefficients of 0.1 as small, 0.3 as medium and 0.5 as large [87].

### 10.3 RESULTS

#### 10.3.1 Participants

The characteristics of the participants are listed in Table 10.1. The control group was well-matched for age and gender. Within the patient group, the subgroups did not differ in ‘age at testing’ ( $F(2, 14) = 2.04, P = 0.17$ ) or in ‘age at diagnosis’ ( $F(2, 14) = 1.13, P = 0.35$ ). However, ‘interval since treatment’ differed significantly ( $F(2, 14) = 5.36, P = 0.02$ ) between the patient subgroups. As expected, further analysis showed a longer interval for the high dose ALL group compared to the low dose ALL group ( $t = 4.76, df = 9, P = 0.00$ ). These results were confirmed by non-parametric Mann-Whitney U-tests.

#### 10.3.2 Neurocognitive functioning

Neurocognitive functioning and differences from the norm population are presented in Table 10.2. The medulloblastoma group scored the worst on almost all cognitive measures, especially on processing speed and motor speed compared to the norm population, followed by the high dose ALL group and the low dose ALL group. Lower scores were also found in the control group; moreover, cognitive functioning differed significantly from the norm scores.

Table 10.2: Cognitive functioning of the participants compared to the norm group.

	MED	High-dose ALL	Low-dose ALL	Patient group	Control group
<i>n</i>	6	5	6	17	17
Intelligence *					
FSIQ					
Mean (SD)	82.8 (14.8)	97.4 (12.4)	102.2 (13.7)	93.9 (15.5)	88.4 (13.8)
Min, max	61, 104	81, 111	87, 123	61, 123	68, 120
<i>p</i> Value	<b>0.04</b>	0.66	0.71	0.13	<b>0.00</b>
VCF					
Mean (SD)	88.8 (13.1)	96.6 (12.9)	98.0 (11.0)	94.4 (12.3)	90.1 (13.0)
Min, max	72, 109	80, 109	86, 114	72, 114	69, 120
<i>p</i> Value	0.09	0.59	0.67	0.08	<b>0.01</b>
POI					
Mean (SD)	88.0 (13.4)	102.8 (16.4)	106.2 (15.8)	98.8 (17.4)	89.9 (11.4)
Min, max	58, 108	80, 123	83, 123	58, 123	67, 110
<i>p</i> Value	0.13	0.73	0.38	0.77	<b>0.00</b>
Processing speed <sup>†</sup>					
PSF					
Mean (SD)	73.7 (11.6)	92.8(12.7)	106.5 (7.3)	90.9 (17.4)	96.9 (14.9)
Min, max	55, 91	72, 105	99, 114	55, 114	72, 124
<i>p</i> Value	<b>0.00</b>	0.27	0.08	0.05	0.41
Motor speed <sup>‡</sup>					
MS (Z)					
Mean (SD)	-1.79 (1.43)	-1.07 (0.95)	-1.10 (0.53)	-1.34 (1.03)	-0.56 (0.98)
Min, max	-4.0, -0.10	-1.97, 0.46	-1.88, -0.48	-4.0, 0.46	-2.62, 1.24
<i>p</i> Value	<b>0.03</b>	0.06	<b>0.00</b>	<b>0.00</b>	<b>0.03</b>

Abbreviations: FSIQ = full scale IQ; VCF = verbal comprehension factor; POF = perceptual organization factor; PSF= processing speed factor; MS= motor speed score in Z-scores; MED = medulloblastoma group; high-dose ALL = leukemia group, high-dose MTX (4 × 5 gr/m<sup>2</sup>); low-dose ALL = leukemia group, low-dose MTX (3 × 2 gr/m<sup>2</sup>).

\* WISC-III-NL; mean = 100, SD = 15; results of 1-sample *t* tests with test value = 100.

<sup>†</sup> Processing speed; mean = 100, SD = 15; results of 1-sample *t* tests with test value = 100.

<sup>‡</sup> Motor speed: total Z score; mean = 0, SD = 1 (-1 SD indicates worse; +1 SD indicates better); results of 1-sample *t* tests with test value = 0.

Analysis of variance showed a trend towards a difference in full scale IQ ( $F(2, 14) = 3.19$ ,  $P = 0.07$ ) between the patient subgroups. Further analysis revealed significantly lower full scale IQ in the medulloblastoma group ( $P = 0.04$ ) compared to the low dose ALL group. The verbal comprehension index score and perceptual reasoning score did not differ between the patient subgroups.

In addition, ANOVA showed significant differences ( $F(2, 14) = 14.50$ ,  $P = 0.00$ ) between the patient subgroups in the processing speed index score. Further analysis revealed significantly lower processing speed scores in the medulloblastoma group compared to the high dose ALL group ( $P = 0.03$ ) and the low dose ALL group ( $P = 0.00$ ). A trend towards a difference between the high dose ALL group compared to the low dose ALL group ( $P = 0.05$ ) was found in favor of the latter group. No group differences in motor speed were found ( $F(2, 14) = 0.90$ ,  $P = 0.43$ ). All results were confirmed by non-parametric Mann-Whitney U-tests.

### 10.3.3 Structural MR Images

All controls and 12 patients (low dose ALL and high dose ALL) had normal findings for the T<sub>2</sub>-weighted and 3D-T<sub>1</sub> weighted scans. Six patients, (medulloblastoma group) showed structural abnormalities on the T<sub>2</sub>-weighted and 3D-T<sub>1</sub>-weighted scans, including tissue loss of the cerebellar hemispheres ( $n=4$ ) or vermis ( $n=3$ ), hemosiderin

Table 10.3: Mean White Matter Fractional Anisotropy (WMFA) values in different regions of interest.

	High-dose		Low-dose	Patient	Control
	MED	ALL	ALL	group	Group
<i>n</i>	6	5	6	17	17
Mean WMFA	0.399**	0.418	0.408	0.407*	0.418
SD	0.072	0.010	0.011	0.012	0.010
Right IFO WMFA	0.217	0.225	0.212	0.217*	0.243
SD	0.029	0.019	0.031	0.026	0.038
Left IFO WMFA	0.221	0.230	0.212	0.220	0.235
SD	0.019	0.019	0.028	0.023	0.035
gCC WMFA	0.472	0.455	0.499	0.474*	0.517
SD	0.064	0.029	0.032	0.047	0.004
sCC WMFA	0.494**	0.666	0.623	0.590	0.615
SD	0.014	0.021	0.097	0.012	0.053
bCC WMFA	0.494**	0.666	0.623	0.590	0.615
SD	0.014	0.021	0.097	0.012	0.053

Abbreviations: MED = medulloblastoma group, high dose ALL = leukemia group high-dose MTX (4 x 5 gr/m<sup>2</sup>), low-dose ALL = leukemia group dose MTX (3 x 2 gr/m<sup>2</sup>), WMFA = white matter fractional anisotropy values, IFO = inferior fronto-occipital fasciculus, gCC, sCC, bCC = genu, splenium and body of the corpus callosum.  
 \* significant (< 0.05) differences between the patient group and the control group.  
 \*\* significant (<0.05) differences between MED and high-dose ALL.

deposits related to small previous occipital (n=2) or temporal (n=2) hemorrhages and subtle signal increases in the bilateral parietal white matter (n=1).

10.3.4 WMFA findings

WMFA is presented in Table 10.3. As anticipated, mean WMFA was lower in the patient group compared to the control group (P = 0.01). WMFA was lower in the right IFO (P = 0.03) and the gCC (P = 0.01), and a trend was found in the bCC (P = 0.07). These results were confirmed by non-parametric Mann-Whitney U-tests.

Analysis of variance showed significant differences in mean WMFA (F (2, 14) = 5.61, P = 0.02), sCC WMFA (F (2, 14) = 4.21, P = 0.04) and bCC WMFA (F (2,14) = 4.79, P = 0.03) between the patient subgroups. Further analyses revealed significantly lower mean WMFA (P = 0.01), lower sCC WMFA (P = 0.03) and lower bCC WMFA (P = 0.03) in the medulloblastoma group compared to the high dose ALL group. Trends towards lower sCC (P = 0.10) and bCC WMFA (P = 0.07) compared to the low dose ALL group were found. No differences between the two ALL groups were found.

With respect to the different patient groups, the medulloblastoma group had lower mean WMFA (mean difference = 0.02, P = 0.00) and lower bilateral IFO WMFA (mean difference right IFO = 0.05, P = 0.01 and mean difference left IFO = 0.04, P = 0.00) compared to their age and sex matched controls. A trend was found towards a difference in the sCC WMFA (mean difference = 0.13, P = 0.09). The high dose ALL group had significantly lower gCC WMFA (mean difference = 0.07, P = 0.01) compared to the

Table 10.4: Correlations of mean WMFA values with cognitive functioning in the patient group (N = 17).

	Full scale intelligence	Processing speed factor	Motor Speed
Mean WMFA	0.22	0.35	0.07
<i>p</i> Value	0.40	0.17	0.78
Right IFO WMFA	0.23	0.21	0.49
<i>p</i> Value	0.38	0.41	<b>0.045</b>
Left IFO WMFA	0.10	0.06	0.16
<i>p</i> Value	0.70	0.81	0.54
gCC WMFA	-0.16	0.22	-0.33
<i>p</i> Value	0.55	0.40	0.20
sCC WMFA	0.16	0.53	0.46
<i>p</i> Value	0.54	<b>0.03</b>	<b>0.06</b>
bCC WMFA	0.10	0.52	0.33
<i>p</i> Value	0.71	<b>0.03</b>	0.19

Abbreviations: WMFA = white matter fractional anisotropy; IFO = inferior fronto-occipital fasciculus; gCC, sCC, bCC = genu, splenium and body of the corpus callosum.

controls. No significant differences between the low dose ALL group and their controls were found. These results were confirmed by non-parametric Mann-Whitney U-tests.

### 10.3.5 Correlations of WMFA with cognitive functioning

Correlations between WMFA and cognitive functioning in the patient group are presented in Table 10.4.

Mean WMFA was not significantly correlated with total intelligence in the patient group. The sCC WMFA ( $r = 0.53$ ,  $P = 0.03$ ) and the bCC WMFA ( $r = 0.52$ ,  $P = 0.03$ ) showed a significantly positive correlation with the processing speed index score. The right IFO WMFA showed significantly positive ( $r = 0.49$ ,  $P < 0.045$ ) correlation and a trend towards a positive correlation between sCC ( $r = 0.46$ ,  $P = 0.06$ ) and motor speed score. No correlations between other ROI WMFA values and cognitive scores were found. The correlations were strong according to Cohen ( $r = \sim 0.5$ ). In the control group the right IFO as well as the left IFO were also correlated to speed of processing ( $r = 0.55$ ,  $P = 0.02$  and  $r = 0.58$ ,  $P = 0.02$ ).

## 10.4 DISCUSSION

To our knowledge, this is the first study reporting WMFA changes after childhood cancer using a 3.0-T MRI scanner. The results of this study showed that WMFA is decreased in childhood cancer survivors and is associated with neurocognitive skills, including speed of processing and motor speed. Mean WMFA and WMFA in ROIs, especially right IFO and gCC, were reduced. Since WMFA reflects the myelination and axonal integrity [32], these results indicate that the integrity of the white matter tracts are affected in childhood cancer survivors of medulloblastoma and survivors of ALL treated with high doses MTX. These findings are in agreement with a number of other studies showing

white matter impairment [211, 136, 151] and, in particular, vulnerability of the frontal lobes and the corpus callosum [207, 63] after treatment for these types of cancer during childhood.

In addition, the current study showed evidence for positive correlations between detailed WMFA impairment, especially in the right IFO and in the sCC and bCC, and neurocognitive impairment, especially speed of processing and motor speed. Lower WMFA in these regions correlated with slower speed of processing and motor speed. This supports the idea that more anisotropic white matter tracts facilitate more processing and faster processing of information. The splenium and body of the corpus callosum play important roles in the communication between the different brain areas, in particular the occipital and motor regions, and this could explain the relation with visual speed of processing and motor speed [246]. Mabbot et al. found that the right frontal–parietal region contributes to the speed of visual–spatial searching [167]. In a wide range of childhood neuropsychiatric illnesses, including attention-deficit/hyperactivity disorder (ADHD), size differences in the corpus callosum have been reported [116]. In patients with traumatic brain injury (TBI), which also usually induces diffuse axonal injury, damage in the corpus callosum is related to a poorer neurocognitive outcome [135].

Several limitations to our study can be identified. First, the compilation of our control group, derived from the same school level with a low average intelligence, hindered the generalization of the patient data. Second, since many childhood brain tumors are located infratentorially, the impact of cerebellar damage and hydrocephalus in the past requires specific attention for its influence on motor speed, attention and executive functions [215]. Thus, cerebellar damage may also have contributed to the cognitive decline in our medulloblastoma subgroup. Third, we did not correct for age since we did not find any age-related increase of overall WMFA. WMFA increases more rapidly during the first few years; in the corpus callosum, the increase occurs up to the age of 6, and in the center semiovale, the increase occurs up to the age of 11 [42]. The reason that we did not find age-related increases of WMFA is likely because the majority of our patients were over 12 years of age. However the application of DTI at a 3.0-T MRI still enabled us to detect differences at a more specific detailed level, despite the older age of our participants.

## 10.5 CONCLUSION

We conclude that DTI on a 3.0-T MRI is sensitive for the detection of region specific changes in white matter integrity in pediatric cancer survivors, and WMFA correlates with speed of processing and motor speed, both serious problems for childhood brain tumor survivors.

The impact of different doses of MTX on neurocognitive functioning should be studied more thoroughly, not only in childhood cancer survivors but also in patient groups where MTX is a common treatment modality. Future longitudinal studies with DTI

collected from the start of the treatment for different types of malignancies and treatment, integrated with neurocognitive measures, could lead to a better understanding of causative neurotoxic factors and its relations with adverse cognitive functions. This could ultimately result in a predictive model of neurotoxicity on neurocognitive outcome — leading to changes in treatment modalities or possibly white matter protection, thereby preventing these adverse late effects.





## SPATIAL PROFILING OF THE CORTICOSPINAL TRACT USING DTI IN RECENT ONSET MND

---

*Objectives.* Upper motor neuron (UMN) degeneration varies in different phenotypes of motor neuron disease (MND). We performed a longitudinal diffusion tensor tractography (DTT) study measuring fractional anisotropy (FA) along the corticospinal tract (CST) in the brain of patients with various phenotypes of MND at an early stage of disease. We aimed at visualizing profiles of CST degeneration, to contribute to further delineation of the phenotypes.

*Methods.* We performed DTT of the CST in cohorts (n=12) of patients with limb-onset amyotrophic lateral sclerosis (ALS-L), bulbar-onset ALS (ALS-B), progressive muscular atrophy (PMA), all at an early stage of disease (onset < 1 year), primary lateral sclerosis (PLS) and healthy controls. Disease duration in PLS patients was longer per definition. We compared 1) patient groups with controls and 2) data at baseline with data at follow up after 6 months. We performed correlation analysis between FA and clinical parameters.

*Results.* Spatial profiling along the CST showed significantly decreased FA along the entire course in PLS and in the brainstem in ALS-B. In ALS-L, a decrease of FA in the posterior limb of the internal capsule and in the subcortical white matter was borderline significant. In PMA, we found changes in the corona radiata compatible with CST degeneration at that level. FA values correlated significantly with clinical measures. Over time we found no significant FA changes in any of the groups.

*Conclusion.* The CST is involved in all MND phenotypes, with the extent and anatomical location of involvement being variable between the phenotypes.

Submitted as:

M.M. van der Graaff, M.W.A. Caan, C.A. Sage, E.M. Akkerman, C. Lavini, C.B. Majoie, A.J. Nederveen, A.H. Zwinderman, F.M. Vos, S. Sunaert, F. Brugman, L.H. van den Berg, M.C. de Rijk, P.A. van Doorn, M. de Visser, *Spatial profiling of the corticospinal tract using diffusion tensor imaging in recent onset MND*, Neurology.

## 11.1 INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is a fatal disease, characterized by progressive upper motor neuron (UMN) and lower motor neuron (LMN) signs and symptoms with onset in the limbs or the bulbar muscles [91, 267]. Substantial differences between these two phenotypes exist concerning epidemiology, prognosis and response to treatment [267, 172]. The primary target of disease remains unknown [81, 101]. Progressive muscular atrophy (PMA) is defined by progressive LMN signs and diagnosed after exclusion of other LMN syndromes [247]. Its prognosis resembles that of ALS. UMN signs may occur in the course of the disease [134, 153, 248]. Primary lateral sclerosis (PLS) is a pure UMN syndrome of slowly progressive spinobulbar spasticity [119]. The exact relationship between PMA, ALS and PLS remains unresolved.

Diffusion Tensor Imaging (DTI) is a MRI modality measuring orientation and integrity of fiber tracts [202]. Diffusion tensor tractography (DTT) allows for the reconstruction of white matter (WM) tracts in vivo [251]. Fractional anisotropy (FA) describes the degree of directionality of diffusion quantitatively [182].

Previous studies found decreased FA to be related to clinical and electrophysiological measures of UMN degeneration in ALS, indicating that DTI can be used to evaluate UMN degeneration in ALS [137, 219].

We aimed to further delineate the various phenotypes within the spectrum of motor neuron disease (MND) with respect to the profile of degeneration along the corticospinal tract (CST) at an early stage of disease. Results of a voxel-based analysis DTI study in the same cohorts are described elsewhere [218].

## 11.2 MATERIALS AND METHODS

### 11.2.1 *Subjects*

We studied cohorts of patients with bulbar-onset ALS (ALS-B), limb-onset ALS (ALS-L), PMA and PLS, and healthy controls. ALS and PMA patients were recruited from neuromuscular clinics in the Netherlands (University Medical Centers of Amsterdam, Utrecht, Rotterdam, and the Catharina Hospital in Eindhoven). All ALS patients met the revised El Escorial criteria for probable (laboratory-supported) or definite ALS [50]. PLS patients were recruited from a cohort studied at the University Medical Center in Utrecht, for which alternative diagnoses had been ruled out [52, 51, 53]. PMA was diagnosed when clinical and electrophysiological evidence of progressive pure LMN involvement was present in two or more regions, excluding patients with focal or segmental muscular atrophy, and after ruling out other LMN diseases [50, 248]. All patients had weakness for less than one year, except for the PLS patients who by definition had a longer disease duration. Healthy age-matched controls were recruited from hospital personnel and

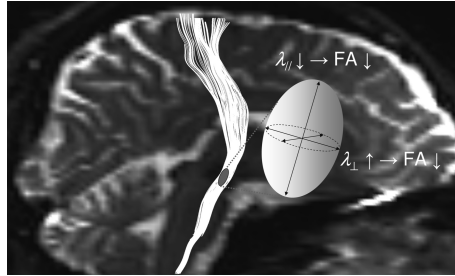


Figure 11.1: Sagittal view of a reconstructed corticospinal tract, overlaid on the non-diffusion-weighted image.

partners of patients. Follow-up visit for all cohorts, except PLS, was 6 months after baseline. PLS patients only provided baseline data, meant to obtain reference values for longstanding and prominent UMN involvement. The protocol was approved by the local ethical committee, all subjects gave written informed consent.

#### 11.2.2 Clinical parameters

We measured finger tapping speed (taps/10 seconds, subsequently expressed as taps/second after averaging values of the left and right side). Deceleration of tapping speed suggests a UMN lesion [149]. Vital capacity (VC) was measured with a handheld spirometer and expressed as a percentage of expected normal value. The ALSFRS-R score (range 0-48), was administered to all subjects [78]. Furthermore, we calculated the disease progression rate (disease progression rate =  $(48 - \text{ALSFRS-R score})/\text{disease duration}$ ) [102].

#### 11.2.3 Data acquisition

DTI data were acquired on a 3 Tesla MRI system (Philips Intera, Philips Medical Systems, Best, The Netherlands) using a spin-echo echo-planar imaging (EPI) sequence. The diffusion-weighting was performed along 32 directions, with a b-value of  $1000 \text{ s/mm}^2$  [9]. Additionally, one set of images was acquired without diffusion weighting ( $b=0 \text{ s/mm}^2$ ). Other sequence parameters were: 64 contiguous axial slices, TE=94 ms, TR=8115 ms, FOV=250 mm, scan matrix =  $109 \times 112$ , interpolated image matrix =  $256 \times 256$ , slice thickness = 2.2 mm, noninterpolated voxel size =  $2.2 \times 2.2 \times 2.2 \text{ mm}^3$ . DTI scan time was approximately 8 minutes.

#### 11.2.4 DTI processing

Eddy current induced deformation of the DTI data was corrected for by an affine registration in the phase direction [173] and the images were filtered for noise using the Linear Minimum Mean Square Estimator (LMMSE) [8]. A single tensor model was

fit to the DTI data, from which the principal diffusion direction was derived for DTT purposes [202].

#### 11.2.5 DTI postprocessing: Fibretracking

The CST in the brain was reconstructed bilaterally in each subject using DTI Studio software, which uses the Fiber Assignment by Continuous Tracking (FACT) method [139, 251]. We set an FA-threshold of 0.2 and an angular threshold of  $50^\circ$  for all reconstructions. Regions-of-interest (ROIs) and algorithm settings were defined in accordance with previous reports on tracking of the CST [220, 251].

DTT is known to be a user-dependent process [254]. Therefore, fiber tracking reproducibility was addressed by two blinded observers (MvdG, CS) reconstructing the CST bilaterally in a subset of 10 subjects, randomly chosen from the entire study population. The overlapping tract volume was computed relative to the total volume tracked by both observers. Subsequently, one blinded observer (MvdG) performed the ROI definition for bilateral CST reconstruction in all subjects. For individuals who completed the study, i.e. with DTI data at two time points, scans of both time points were matched using a rigid transformation of the FA maps as implemented in SPM5 (Statistical Parametric Mapping, Wellcome Department of Imaging Neuroscience, University College London) so that ROIs had to be defined only once for each subject. We computed mean FA values per tract, thus yielding two values (left and right side) per patient per time point. Second, we generated FA profiles along the caudocranial course of the CST as follows. We co-registered the FA maps to the ICBM-81 FA-template [181] using joint affine and non-rigid registration using smooth basis functions [24]. By applying these transformations to the voxel label maps (i.e. binary images of the CST reconstructions, with voxels having a value of 1 being part of the DTT reconstruction), the tract data of all subjects resided in one common reference frame. These datasets were then sampled at a resolution of 4 mm, yielding FA-profiles sampled at 23 positions in a caudocranial direction for both left and right CST of all individual subjects. In case DTT terminated in the pons, we extended tracking in these datasets using the average volume of all transformed label maps.

### 11.3 STATISTICAL ANALYSIS

For all following tests, we set a statistical threshold for significance of  $p < 0.01$ .

#### 11.3.1 Clinical data

We compared the clinical data of patients with those of controls using a Mann-Whitney U test for non-parametric data. Longitudinal data were analyzed with the Wilcoxon signed ranks test for paired data.

### 11.3.2 DTI data

Before pooling values of both hemispheres for further analyses, we compared mean FA of the CST corresponding with the clinical site of disease onset (if applicable) with contralateral FA using the Wilcoxon signed ranks test for paired data. We compared mean FA at baseline between the patient groups and controls, and between ALS-L and ALS-B, using a Mann-Whitney U test. We compared the FA profiles along the CST of the different patient groups with those of healthy controls, as well as those of ALS-B patients with those of ALS-L patients, using a Mann-Whitney U test. We compared mean FA and FA profiles along the CST obtained at baseline and at follow-up in all groups (except for PLS) using the Wilcoxon signed rank test for paired data. Finally, Spearman rank correlation coefficients ( $r$ ) were computed between mean FA and clinical parameters, irrespective of diagnosis: finger tapping speed, disease progression rate, ALSFRS-R and VC.

### 11.3.3 Power analysis

Based on the study of Wong et al., to detect a 5 % decrease of FA in patients when compared to controls with a power of 80% and an alpha of 1%, 12 individuals per cohort were needed [258].

## 11.4 RESULTS

At follow-up, 7 out of 12 ALS-L, 9 out of 12 ALS-B, 10 out of 12 PMA patients and 12 controls were able to undergo a repeat scan session. The baseline and follow-up data of one control subject had to be discarded due to imaging artifacts. The pilot study in which DTT reproducibility was investigated yielded an inter-observer agreement of  $88 \pm 8\%$ . In less than 5% of the cases DTT terminated in the pons.

### 11.4.1 Clinical characteristics

Baseline characteristics of all groups and significant differences between patient groups and controls are summarized in Table 11.1. Disease progression, VC and ALSFRS-R score at baseline did not differ significantly in a mutual comparison of PMA, ALS-L and ALS-B. Finger tapping speed was significantly lower in ALS-L when compared to PMA ( $p=0.006$ ). Findings at follow-up are listed in Table 11.2. None of the clinical parameters changed significantly over time, except for the ALSFRS-R score in PMA. None of the PMA patients developed UMN signs or symptoms during the time of follow-up.

### 11.4.2 DTI: mean FA value of the CST

An example of a CST reconstruction is shown in Figure 11.1. We found no significant differences of mean FA in the CST between the hemisphere corresponding to the clinical site of onset and the contralateral hemisphere. For all 23 samples along the CST,

Table 11.1: Clinical characteristics and mean FA at baseline. Data are means  $\pm$  standard deviation, unless indicated otherwise. Clinical characteristics of the patients were compared to those of controls using the Mann Whitney U test for non-parametric data. Threshold for statistical significance was set at  $p < 0.01$  ( $p$ -values indicating significant differences in superscript). NA: not applicable. <sup>1</sup> disease progression rate =  $(48\text{-ALSFRS-R score at baseline})/(\text{disease duration (days)})$ . <sup>2</sup> finger tapping speed:  $(L+R)/2$ . <sup>3</sup> mean FA CST:  $(\text{left and right})/2$ .

	PLS n=12	ALS-B n=12	ALS-L n=12	PMA n=12	Controls n=12
M:F	7:5	5:7	10:2	11:1	7:4
age, Y (median/range)	59.0 (48-74)	56.5 (42-72)	58.0 (41-78)	60.5 (44-72)	56.5 (46-80)
duration symptoms, months	50.2 $\pm$ 24.8	9.8 $\pm$ 3.2	10.5 $\pm$ 2.5	10.8 $\pm$ 2.0	NA
delay until diagnosis, months	24.4 $\pm$ 16.0	6.7 $\pm$ 2.7	6.6 $\pm$ 2.5	6.6 $\pm$ 2.9	NA
disease progression rate <sup>1</sup>	0.010 $\pm$ 0.005	0.026 $\pm$ 0.018	0.027 $\pm$ 0.016	0.020 $\pm$ 0.010	NA
vital capacity	91.5 $\pm$ 23.3	98.5 $\pm$ 17.9	101.3 $\pm$ 23.8	102.5 $\pm$ 14.9	114.6 $\pm$ 18.5
ALSFRS-R	35.0 $\pm$ 4.6 <sup>&lt;0.001</sup>	40.1 $\pm$ 4.7 <sup>&lt;0.001</sup>	40.1 $\pm$ 3.9 <sup>&lt;0.001</sup>	41.8 $\pm$ 3.7 <sup>&lt;0.001</sup>	48.0 $\pm$ 0
finger tapping/s <sup>2</sup>	2.3 $\pm$ 1.2 <sup>&lt;0.001</sup>	4.0 $\pm$ 1.3	3.3 $\pm$ 1.3 <sup>0.002</sup>	4.8 $\pm$ 0.7	5.1 $\pm$ 0.6
Mean FA CST <sup>3</sup>	0.42 $\pm$ 0.02 <sup>0.001</sup>	0.42 $\pm$ 0.03 <sup>0.007</sup>	0.43 $\pm$ 0.03	0.46 $\pm$ 0.02	0.45 $\pm$ 0.02

Table 11.2: Clinical characteristics and mean FA: changes over time (6 months). Data are means  $\pm$  standard deviation. P-values indicating significant differences are shown in superscript and result from Wilcoxon signed ranks tests for non-parametric data. Threshold for statistical significance was set at  $p < 0.01$ . <sup>1</sup> finger tapping speed: (L+R)/2. <sup>2</sup> mean FA CST: (left and right)/2

	ALS-B (n=9)	ALS-L (n=7)	PMA (n=10)	Controls (n=11)
	baseline	baseline	baseline	baseline
	<i>follow-up</i>	<i>follow-up</i>	<i>follow-up</i>	<i>follow-up</i>
vital capacity %	97.7 $\pm$ 15.7	106.9 $\pm$ 19.3	101.0 $\pm$ 16.0	115.6 $\pm$ 19.0
	80.4 $\pm$ 27.6	101.9 $\pm$ 24.2	91.3 $\pm$ 21.3	118.5 $\pm$ 21.3
ALSFRS-R	41.7 $\pm$ 3.4	42.6 $\pm$ 2.0	41.8 $\pm$ 3.5	48.0 $\pm$ 0.0
	36.3 $\pm$ 6.1	39.3 $\pm$ 2.7	36.8 $\pm$ 5.4 <sup>0.008</sup>	47.7 $\pm$ 0.9
finger tapping	4.2 $\pm$ 1.2	3.3 $\pm$ 1.0	4.8 $\pm$ 0.8	4.8 $\pm$ 0.9
speed <sup>1</sup> (/second)	3.7 $\pm$ 1.6	2.5 $\pm$ 1.7	4.3 $\pm$ 1.9	4.6 $\pm$ 0.9
Mean FA <sup>2</sup>	0.43 $\pm$ 0.03	0.44 $\pm$ 0.03	0.46 $\pm$ 0.02	0.45 $\pm$ 0.02
	0.42 $\pm$ 0.04	0.42 $\pm$ 0.04	0.46 $\pm$ 0.02	0.45 $\pm$ 0.01

corresponding FA values never differed more than 1.5% between hemispheres. Therefore, in each patient we pooled FA values of both hemispheres for further analyses.

At baseline, mean FA of the CST in PLS and in ALS-B was significantly lower than in controls (Table 11.1). Mean FA of the CST in patients with ALS-L or PMA was not statistically different from mean FA in controls. Mean FA of the CST in ALS-L did not differ significantly from values in ALS-B.

At follow-up, compared to baseline data, we found no significant changes of mean FA of the CST in any of the groups (Table 11.2).

#### 11.4.3 DTT: profiles of FA along the CST

At baseline, in PLS, FA was significantly decreased (Figure 11.2A) in the medulla oblongata ( $p=0.001$ ) and cerebral peduncle ( $p=0.002$ ) and in the rostral part of the posterior limb of the internal capsule ( $p=0.007$ ) as well as in the subcortical white matter (WM) ( $p=0.007$ ). In ALS-B, FA was significantly decreased (Figure 11.2B) in the medulla oblongata ( $p=0.003$ ), cerebral peduncle ( $p=0.005$ ) and in the posterior limb of the internal capsule ( $p=0.001$ ). In ALS-L, a decrease of FA values in the subcortical white matter was borderline significant ( $p=0.01$ ). No other significant decreases of FA were found (Figure 11.2C). In PMA, FA was significantly *increased* in the rostral part of the posterior limb of the internal capsule and the corona radiata ( $p=0.006$ ) (Figure 11.2D).

We found no significant FA differences in a direct comparison between ALS-L and ALS-B. Over time, none of the observed changes were statistically significant. For ALS-L,

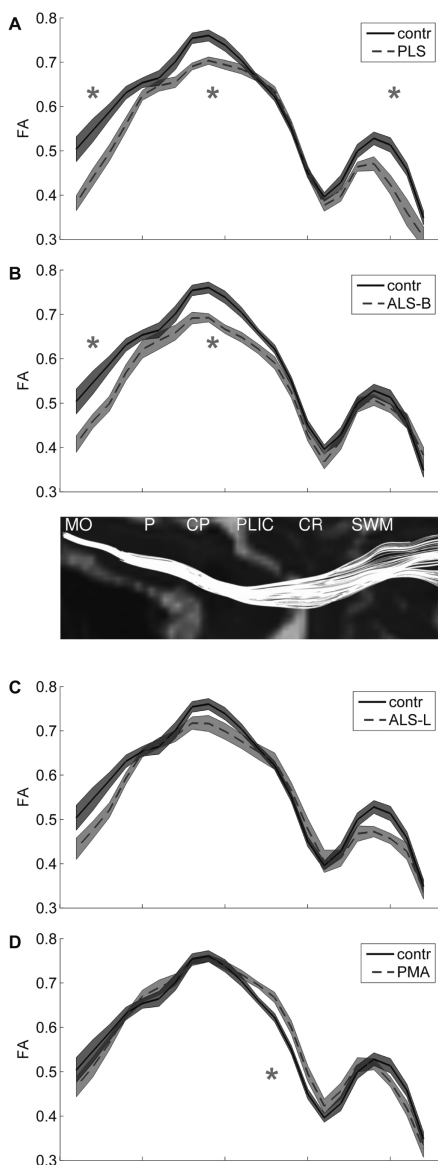


Figure 11.2

FA-profiles along the CST in the various cohorts at baseline FA values along the caudocranial course of the corticospinal tract of patients (dotted line) were compared to FA values in controls (bold line). FA values are shown  $\pm$  the standard error of the mean in patients and controls in light and dark grey, respectively. Asterisks (\*) indicate statistically significant differences ( $p < 0.01$ ) between patient group and controls. The middle image provides an anatomic reference. MO: medulla oblongata; P: pons; CP: cerebral peduncle; PLIC: posterior limb of the internal capsule; CR: corona radiata; SWM: subcortical white matter.

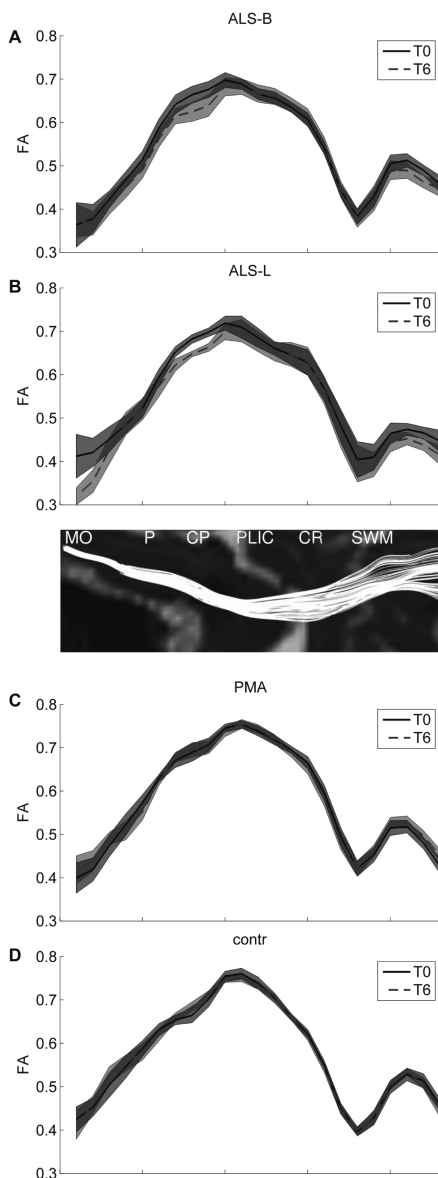


Figure 11.3

FA-profiles along the CST in the various cohorts at baseline and at follow-up FA values along the caudocranial course of the corticospinal tract at baseline (T0, bold line) were compared to those at follow-up (T6, dotted line) were compared to FA values in controls (bold line). FA values are shown  $\pm$  the standard error of the mean at baseline and at follow-up in dark and light grey, respectively. The middle image provides an anatomic reference. Abbreviations as in Figure 11.2.



ALS-B, PMA and controls, FA profiles at baseline and follow-up after 6 months are shown in Figure 11.3.

#### 11.4.4 *Correlation of clinical parameters with FA values*

We found a significant and strong positive correlation of mean FA with finger tapping speed ( $r=0.554$ ,  $p < 0.001$ ) and with ALSFRS-R score ( $r=0.491$ ,  $p < 0.001$ ). Correlation between mean FA and VC or disease progression rate were not significant.

## 11.5 DISCUSSION

Spatial profiling along the CST showed FA changes in all MND phenotypes. The nature and distribution of these changes varied between the phenotypes. In the following paragraphs, we will compare our findings to previous studies and discuss the strengths and limitations of our study.

### 11.5.1 *Strengths and limitations of the study*

To our knowledge this is the first DTI/DTT study to include various phenotypes of MND at an early stage of disease, enabling acquisition of data for homogeneous groups that are likely to reflect changes in the initial stages of disease, although the actual onset of biological disease remains unknown. We performed a systematic follow-up after 6 months. Delay until diagnosis in our study was 6.6 months on average, while delay in epidemiological surveys is 12-15 months [86, 92]. A shorter delay until diagnosis is associated with a more aggressive disease course [91, 107]. We lost almost a third of our patients to follow-up because of rapid disease progression or death.

Because of the relatively small cohorts, we did not fully correct for multiple comparisons, although we did apply a stricter threshold for significance at  $p < 0.01$  instead of  $< 0.05$  throughout the study.

### 11.5.2 *PLS*

FA was decreased along the entire course of the CST in PLS patients. There are two other DTI studies including a separate group of PLS patients [85, 241]. Ciccarelli et al. found lower FA values in the rostral section of the CST of 6 patients with PLS (mean disease duration 7.4 years) when compared to patients with ALS, based on the statistical results of a tract-based spatial statistics analysis [85]. Ulug et al. found an FA reduction within the posterior limb of the internal capsule of PLS patients (mean disease duration 5.1 years) when compared to controls in a ROI analysis [241]. Since our study showed more extensive involvement of the CST in PLS patients despite the fact that mean disease duration in our study was shorter (4.2 years), the use of DTT reconstructions may provide a more accurate approach for obtaining FA values.

### 11.5.3 ALS-L and ALS-B

#### *Findings at baseline*

There were no statistically significant differences in disease progression rate, vital capacity, ALSFRS-R and finger tapping speed at baseline between ALS-L and ALS-B.

Although direct comparison of ALS-B with ALS-L yielded no significant differences in mean FA nor in FA profile along the CST, we found two distinct patterns of CST degeneration in ALS-L and ALS-B. Observed FA changes were borderline significant in the subcortical WM in ALS-L, and prominent in the medulla oblongata, cerebral peduncle and posterior limb of the internal capsule in ALS-B. One explanation may be that the primary target of degeneration is the rostral CST in ALS-L and the caudal CST in ALS-B. A second explanation is that corticobulbar fibers, which are likely to be more affected in ALS-B than in ALS-L, were also included in the cerebral peduncle and posterior limb of the internal capsule of the CST reconstruction.

Only a limited number of DTI studies showed spatial profiling along the CST. In none of these studies ALS-L and ALS-B patients were analyzed separately. Sage et al. found significantly lower FA values in the internal capsule and in the subcortical white matter section of the CST in ALS patients with a disease duration of 4-34 months [220]. Ciccarelli et al. found lower FA values along the entire course of the CST in a tract-based spatial statistics analysis of ALS patients with a disease duration of  $21 \pm 16$  months, thus representing patients at a more advanced stage of disease [85]. Wong et al. showed decreased FA values in the cerebral peduncles in a spatial profiling DTI study in 14 ALS patients with a disease duration of  $22 \pm 12$  months [258]. However, in their study only the section of the CST between cerebral peduncles and the corona radiata was included.

Thus, our findings cannot be easily compared with previous spatial profiling studies as various phenotypes of MND were lumped, disease duration was heterogeneous or only part of the CST was analyzed.

#### *Findings at follow-up*

We found no statistically significant decrease of mean FA at follow-up. Values of FA as a function of position along the CST showed no significant changes over time in any of the groups. Previous studies also failed to demonstrate a decrease of mean FA of the CST over time [220, 47, 176]. There are no other longitudinal spatial profiling study providing data on FA changes as a function of position along the CST.

#### 11.5.4 PMA

Mean FA values of the CST in PMA patients were neither significantly different from those of healthy controls, nor from those of patients with ALS-L or ALS-B.

However, spatial profiling in our study revealed an *increase* instead of a decrease of FA in the rostral internal capsule, a phenomenon described previously [258]. At the level of the corona radiata, the CST intersects with fiber tracts of the corpus callosum and the superior longitudinal fasciculus, leading to low FA values. It may well be that increased FA in a region with intersecting fibers is indicative of CST degeneration.

Comparison with previous studies is difficult as they are scarce, and included heterogeneous patient groups. Spatial profiling studies in PMA are lacking. Previous DTI studies in PMA were based on ROI analysis and yielded contradicting findings. Two studies found no differences between PMA and controls [176, 90]. One study demonstrated decreased FA in the rostral section of the CST in patients with an LMN phenotype [217]. However, these patients developed UMN signs at follow up, whereas our patients did not. Another study showed FA values similar to ALS patients in a small subset of PMA patients at an advanced stage of disease [121].

Our findings in PMA may thus be compatible with commencing degeneration of the CST in the corona radiata.

#### 11.5.5 Overall conclusion

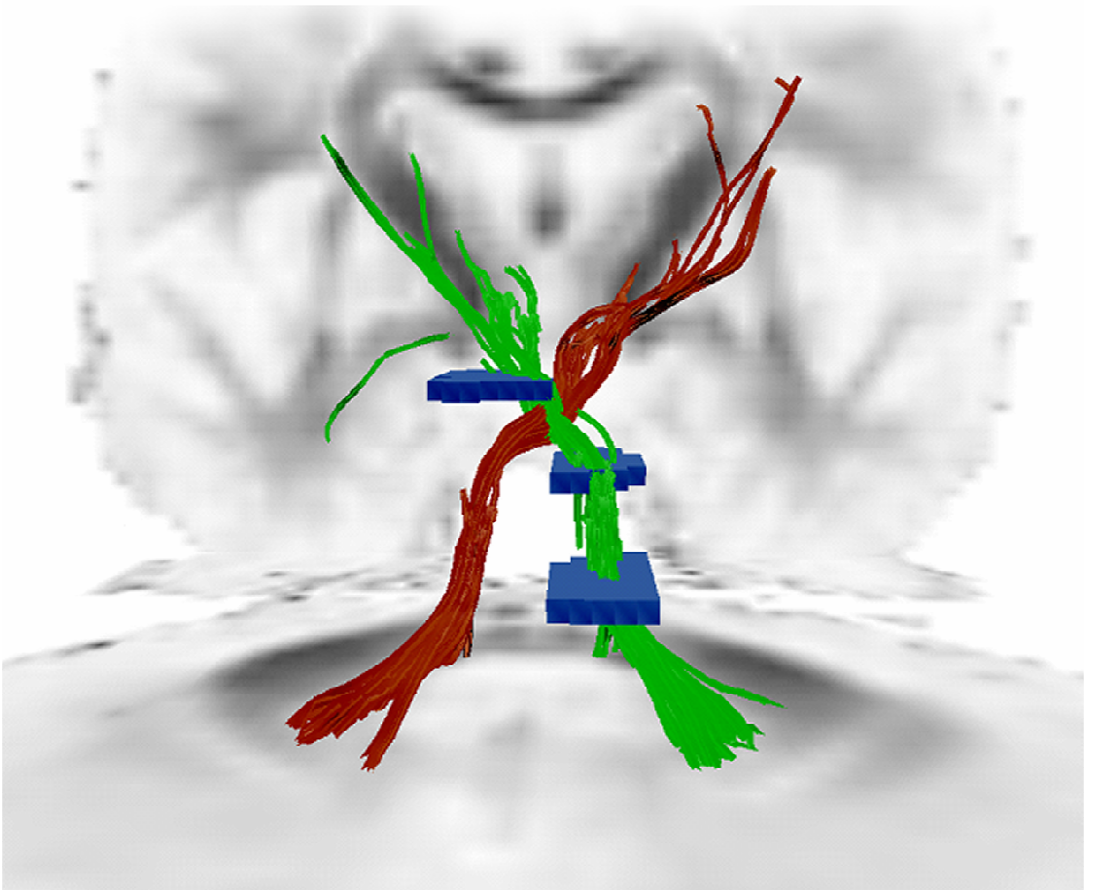
Our study is the first to provide spatial profiling data of the CST in recent onset ALS and PMA. Degeneration of the CST in ALS-B was more prominent than in ALS-L and PMA, and located at a more caudal level of the CST. Changes in ALS-L and in PMA were subtle but suggestive of rostral CST degeneration. Similarities between ALS-L and PMA with respect to the profile of UMN degeneration were larger than similarities between ALS-L and ALS-B. Therefore, PMA patients may benefit from a similar therapeutic approach and inclusion in future therapeutic trials.

#### ACKNOWLEDGEMENTS

We wish to thank all patients and controls for their participation in the study. We thank Professor Vianney de Jong for his conceptual contributions, Hanneke Börger for the logistical support, and Suzan Beyer for her assistance with preparing the figures. This study was funded by the Prinses Beatrix Fund (The Hague, The Netherlands).



# GRID IMPLEMENTATION





## USING DYNAMIC CONDOR-BASED SERVICES FOR CLASSIFYING SCHIZOPHRENIA IN DIFFUSION TENSOR IMAGES

---

Diffusion Tensor Imaging (DTI) provides insight into the white matter of the human brain, which is affected by Schizophrenia. By comparing a patient group to a control group, the DTI-images are on average expected to be different for white matter regions. Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are used to classify the groups. In this work, the number of principal components is optimised for obtaining the minimal classification error. A robust estimate of this error is computed in a cross-validation framework, using different compositions of the data into a training and a testing set. Previously, sequential runs were performed in MATLAB, resulting in long execution times. In this paper we describe an experiment where this application was run on a grid with minimal modifications and user effort. We have adopted a service-based approach that autonomously launches Image Analysis Services onto a campus-wide Condor pool comprising of volunteer resources. This allows high throughput analysis of our data in a dynamic resource pool. The challenge in adopting such an approach comes from the nature of the resources, which change randomly with time and thus require fault tolerance. Through this approach we have reduced the computation time of each dataset from 90 minutes to less than 10. A minimal classification error of 22% was obtained, using 15 principal components.

Published as:

Simon Caton, Matthan Caan, Silvia Olabarriaga, Omer Rana and Bruce Batchelor, Using Dynamic Condor-based Services for Classifying Schizophrenia in Diffusion Tensor Images, in: Proc. CCgrid, 2008, 234-241 [75].

## 12.1 INTRODUCTION AND RELATED WORK

Over the past few years, diffusion tensor imaging (DTI) [38] has provided important insights into the structure of the brain. A still growing number of studies adopt DTI to determine changes in brain structure in schizophrenia, a cognitive disorder affecting approximately 1% of the world's population. Typically, schizophrenia is studied through region of interest (ROI)-analysis or voxel-based analysis (VBA) [147]. To date many regions of the brain have been reported to be linked with schizophrenia, however, exactly which ones are most relevant is still unknown. Computerised image processing of the brain imagery has led to improvements in the understanding of the disease, however disseminating the visual information into usable data is complex because the brain structure is highly inter-connected.

In [66] an algorithm is presented that models correlations between different brain regions using a pattern recognition approach. Within the pattern recognition framework, a classification error is computed, indicating the ability to distinguish between the groups. A robust estimate of the classification error is computed in a series of independent jobs that could run in parallel. However, parallel execution is hindered by relatively high costs of MATLAB licenses. Therefore, computations in [66] were performed sequentially, resulting in excessively long running times.

A number of medical applications have been proposed earlier, based on a parallel implementation of MATLAB scripts. Statistical analysis of Alzheimer's disease was performed using Statistical Parametric Mapping, within MATLAB [30]. Here, MATLAB licenses are required for all nodes. In [145] a problem solving environment automatically generated MATLAB code and carried out simplified Grid computing, thus enabling Finite Element Modeling. A generalisable architecture was presented in [69] for grid-enabled processing of pathological images for computer-aided prognosis, by deploying parallel MATLAB on a cluster.

The closest approach to our work includes the use of compiled MATLAB code, run on a Condor cluster to speed up the computation speed of Electrical Impedance Tomography reconstruction algorithms [112]. However, in this instance, the code was pre-defined for homogeneous tasks and the standard Condor system was utilised. In [162] Condor-G and Pegasus are used to speed up MATLAB-based evaluation of tomographical data. Here, inter-site network latencies and queue wait times significantly impact upon performance. In [130] compiled MATLAB scripts were deployed over a 180 node Condor system to process vibro-acoustic data. Here, dramatic performance improvements were observed as resources were not limited by available MATLAB licenses. However, this does not address the issue of queue time.

In [209, 208] MATLAB-based Condor-G jobs are created on the fly in response to the detection of heavy load on the user's workstation. The user's machine is relieved of work transparently, but the complexity of job generation means that the approach is slower than regular Condor. MATLAB\*G [80] is a parallel MATLAB for the ALiCE Grid [2].



Here, the user submits a job interactively from a MATLAB environment. Each job is then decomposed into a number of tasks and executed in fine-grained parallel manner on (licensed) MATLAB slaves deployed as ALiCE jobs. JavaPorts [174] is a distributed component framework that remaps MATLAB tasks for heterogeneous nodes. In this work, tasks are compiled and distributed to a heterogeneous cluster at runtime, good speedup can be achieved for a small number of nodes, but saturates as more nodes are added due to messaging overheads. In [82] a review of 27 other parallel MATLAB approaches is presented. Most of these approaches either require a license or the use of parallel programming libraries (e.g. MPI and PVM) [209]. The logistics of distributed or multi-institutional software licensing remains an unresolved issue [162, 57] and thus we cannot use such an approach.

We describe a method for porting the application in [66] to an architecture of dynamic Condor-based services for generic image analysis [76]. Each service relates to a continuously running, license free and remotely controllable compiled instance of a MATLAB-based Image Processing Engine (IPE). By using compiled MATLAB sessions we side-step most queue-based latencies, as can be found in the above work. A campus-wide Condor pool made up of volunteer resources provides a source of raw computational power. As Condor was not designed for a reliable infrastructure, we layer our own infrastructure on top of Condor to improve performance and reliability. The image analysis was carried out using the 2,500 node Cardiff University Condor Pool, of which a total of 400 machines were used.

The paper is organised as follows: section 12.2 presents the application and section 12.3 the grid infrastructure. The implementation is described in section 12.4. Results are presented in section 12.5, discussions and conclusions in section 12.6.

## 12.2 APPLICATION

The application consists of an image analysis pipeline for DTI-scans. The goal is to locate regions in which the brain differs in schizophrenic patients and healthy controls. In case of affected brain tissue, signal changes in DTI scans are to be expected. As these changes are subtle, patient and control *groups* are compared to gain statistical power. Still, imaging artifacts, noise and heterogeneity within the patient and control groups hamper the detection of pathological processes in the data. The method proposed in [66] adopts a pattern classification approach using supervised techniques. The system is designed to identify *if* and *where* a significant difference exists in the scans of the patient and the control groups. Note that the relevance of the results obtained with the proposed method resides in analysing the given population, for clinical research purposes, and not (yet) for classifying new incoming patients.

The data adopted in this study consists of DTI scans of 34 schizophrenic patients and 24 controls. Scans were acquired on a dedicated MR-scanner at the Amsterdam Medical Center (AMC). After isotropic resampling, 3D-volumes were obtained (128x128x48 vox-

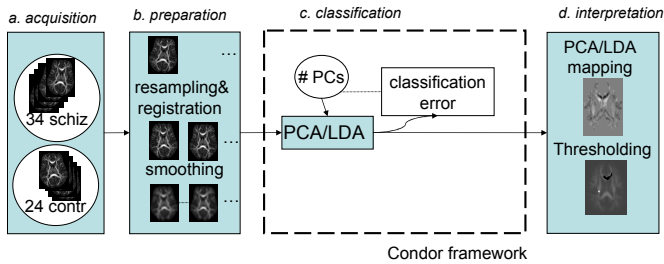


Figure 12.1: Workflow of the application indicating the parallelised steps.

els, resolution of  $2 \times 2 \times 2$  mm). Pre-processing consists of spatially aligning (registration) and smoothing of the data. This step is currently performed on a single workstation.

Classification is done by means of Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). PCA determines representative features, or Principal Components (PCs), out of the data. LDA then decides, using these PCs, if a significant difference can be found between patients and controls.

A training set of a subset of DTI-volumes with known class-labeling is used to train the LDA-classifier. This classifier is subsequently applied to samples in an unseen test set. The classification error is defined as the relative part of the test set that is incorrectly classified. Randomly classifying the subjects would induce an error of 50%. Based on the group sizes, an upper boundary on the error of 35% is concluded to yield a significant difference between patients and healthy controls [66].

The estimated classification error should be independent of the composition of training and test sets, therefore the *cross-validation* strategy is adopted. It consists of splitting the total dataset (all DTI-volumes) in different groups that are partially used as test set, and the remainder of the volumes as the training set. Different compositions of groups lead to repeated cross-validation, and the computed classification errors are averaged for all. Typically, a five-fold cross-validation is performed, and repeated ten times.

An essential parameter in the classification is  $nPCs$ , the number of principal components used by the PCA/LDA algorithm. A low  $nPCs$  yields limited information for the classifier, resulting in a biased estimate. For a high  $nPCs$ , noise will distort the estimation, such that incorrect brain regions will be reported. The optimal  $nPCs$  depends on the application, and needs to be found for each new study.

If the classification error obtained with the optimal parameter configuration is low, such that a significant difference between patients and controls can be concluded, it is of clinical interest to know which brain regions contribute most to this difference. That is, we want to *interpret* the outputs of the pattern recognition system. For that purpose, a

*map* is calculated by the PCA/LDA-algorithm. This map indicates, per voxel, the amount of variation in DTI-values that contributed to the optimal classification into patients and controls. Negative values in the map relate to decreased DTI-values for patients, and positive values indicate increased values. Clinically, it is known that schizophrenia is related to merely a few brain regions, therefore the map is thresholded to extract those regions where the mapping has the highest absolute value. An example of such a map is given in figure 12.1d.

## 12.3 INFRASTRUCTURE

### 12.3.1 *Challenges in distributing the application*

Conventionally, the application presented in the previous section was analysed using MATLAB with two external toolboxes: PRTools [1] and DipLib [3]. In this application, one parameter needs to be tuned: *nPCs*. This means that the cross-validation step has to be performed repetitively for a large range of parameter values, resulting in large running costs. Such long computations require much effort for management and logistics, and additionally prevent users from running other tasks on their machines at the same time. As the tasks involved are independent from each other, they can be executed in parallel.

Porting this application onto distributed resources is a challenging task. Normally a high-performance infrastructure such as provided by the VL-e project<sup>1</sup> would be useful to run the computations. Unfortunately running (licensed) MATLAB programs is not possible in this infrastructure, which holds for public grids in general. Moreover, mechanisms are needed for efficient and meaningful data transportation and control.

It would be unreasonable and unrealistic to expect the adoption of complex low-level parallel approaches, such as MPI, to speed up this application [93]. Firstly, as it is possible to adopt a coarse-grained parallel approach quickly and easily, with minimal changes to application code. Secondly, we cannot expect users to be experts in distributed computing. Hence, preference is given to a system that is easy to use and has reasonable performance (over a system with peak performance but difficulty of use) [82].

As the complexity of an application grows, so too does the necessity to use a significant number of resources to support its execution [95]. It is seldom the case that a single research group possess such resources and must draw on resources from other sources. Such scenarios are common, and greatly benefit from opportunistic approaches. Systems like Condor, Nimrod [5], NetSolve [7], Legion [165] etc. may be used, but Pancake [195] measured that users can spend approximately 10% of their time performing job setup. For large applications this is likely to be a significant period of time and for inexperienced users far greater than 10%. It is often difficult to predict the needs of an application in advance, particularly if those needs change during runtime [146]. This can occur for several reasons: (1) resource requirements change due to the availability of new and

<sup>1</sup> Virtual Laboratory for enhanced Science, <http://www.vl-e.nl>

better resources [43], (2) resources are shared among various users and no single user has control over their allocation [60] or (3) one or more resources fail.

Consequently, the availability of resources varies with time and it is difficult to create a stable environment [60]. The final result is that it is nearly impossible for users to map and then manage the execution of their application by hand [95].

### 12.3.2 *Running Compiled MATLAB Applications on a Condor grid*

We used 400 nodes of the 2,500 node Windows Condor Pool at Cardiff, which are made compiled MATLAB-ready by installing the MATLAB Component Runtime (MCR) libraries. Of these 400 nodes there are on average 120 available at any given time. A consequence of a campus-wide Condor pool with no explicit centralised control is that workstations are both regularly and haphazardly upgraded and modified from multiple sources. We use an extension of Condor's ClassAds to identify machines which have MCR installed. However, over time, we cannot guarantee that this remains the case, as workstation updates can induce anomalous behaviour. This is manifested in one of four problems associated with using the Condor system [230], but it is left up to the user to determine why their job has failed. Condor provides no distinction between failures due to application error and errors which stem from the operating environment.

In addition to the above factors, the MCR provides further challenges: (1) the machine owner has priority over the Condor user. However, a compiled MATLAB job cannot be successfully suspended and resumed (compiler version 4.7), and as in the case above, Condor will not reschedule a job when this occurs. We can also not recover from such errors, without resubmission, as check pointing is not supported for Windows-based Condor pools. (2) The MCR can and often does fail during initialisation. Even without these administrative challenges compiled MATLAB applications can generate other errors. Such errors can be application specific – such as running out of memory, or can simply be irregular events such as a class loader error or corrupted files. However, whatever the cause, handling such errors in a generic way is a significant challenge.

## 12.4 IMPLEMENTATION

### 12.4.1 *Compiling the Image Processing Engine (IPE)*

Our IPE is not a regular compiled MATLAB application. Regular applications (e.g. [130]) cater for a limited number of scenarios, can only be used in batch mode and consequently must be initialised for every distinct job. Our engine differs in the following ways: (1) It caters for many disparate scenarios, as it is essentially a scaled down version of MATLAB. (2) One instance can serve multiple distinct tasks with only one initialisation phrase and thus reduce the impact on performance of unpredictable queueing times. This results in a higher job throughput. (3) It is primarily used interactively but can also be used in batch mode.

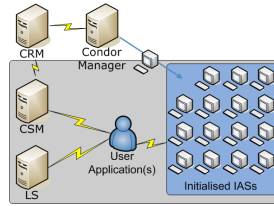


Figure 12.2: System overview. The shaded area relates to the system components known to the user's application – for further detail see: [76]

Compiling entire MATLAB toolkits requires a different approach than compiling regular stand-alone applications. Normally the compiler only needs to be told which function to compile and in the majority of cases all dependencies are handled for the user. In our case, we must explicitly tell the compiler which functions are to be included in the standalone application. When merging toolkits this quickly becomes a list of substantial size, which in the case of PRTools and DipLib is around 1,500 functions. This relates to a compiler call of over 60,000 characters in length. Such a process cannot be performed by hand, and therefore must be automated. For this task we have created a Java-based interface to the MATLAB Compiler, allowing a user to create simple definitions of what to include and, when necessary, exclude from the compiled application.

#### 12.4.2 System Overview

There are four core components that constitute this system (see Fig 12.2). (1) The Central Service Manager (CSM), is responsible for allocating resources to user applications and autonomously and dynamically determines the resource requirements of the system as a whole. The CSM may also run as a Condor job. (2) A number of Multiple Image Analysis Services (IASs), determined by the number of available system resources. Each IAS is an instance of the IPE controlled through a Java Server and runs as a continuous Condor job; terminating only on failure or eviction. (3) The Condor Resource Manager (CRM) is a job submission and monitoring daemon running on a Condor submit node. It attempts to realise the requirements specified by the CSM. As previously mentioned, there are many difficulties to be overcome in performing this realization. Hence, the CRM monitors each IAS job and the Condor pool as a whole in order to determine: (i) to what extent it can satisfy the requirements made by the CSM, (ii) which workstations and IAS jobs are demonstrating anomalous behaviour and (iii) whether resources should be released due to a high demand from other Condor users. (4) One or more Lookup Services (LS), which are used by user applications and system components to locate the CSM. In order to avoid confusion we do not use Task and Job synonymously. A Job corresponds to an instance of an IAS running on a Condor node. A Task corresponds to a MATLAB command sequence, executed by an IAS.

Due to the nature of this approach it is inevitable that IASs join the system haphazardly. In order to reduce system overhead tasks are buffered in remote IAS queues, which enables

the IAS to download any data whilst a previous task executes. However, as more resources become available task reallocation may be required in order to take advantage of new resources. This means that our scheduler must act opportunistically, with likely future task reallocations. Hence, it is important not to plan too far ahead. To facilitate this we autonomously vary the size of IAS queues in relation to the number of unallocated tasks. Therefore, new resources do not result in multiple task reallocations, which would in turn increase system overhead. To determine from which IAS a task is taken for reallocation, the scheduler uses the following heuristic:  $\min(\text{avg}(\text{ExecutionTime}(T_{sim})) - \text{IAS}_{Te})$  where  $Te$  is the currently executing Task for a given IAS, and  $T_{sim}$  is a *similar* Task, which has been monitored in the past. We then take the Task that was most recently submitted to the IAS for reallocation.

In order to control MATLAB sessions we define meta-interpreters (MI), that: (1) execute a command or command sequence, (2) construct results, (3) perform any error handling, and (4) return any result or error events to user applications. Results are returned by the MI in MATLAB's `.mat` format so that they can be easily aggregated into the complete analysis workflow – a Task Specification Object (TSO) was defined relating to this MI.

#### 12.4.3 *Porting the application*

In order to port any application into a parallel context it is inevitable that some changes to code will be required. This was also true for this application, although the changes required were not very significant. The following changes were necessary; all data loads were separated from the execution code. The reasons for encouraging this approach are threefold: (1) our IPE is a modular image processing package in which algorithms are constructed using sequences of interchangeable function calls. (2) It encourages further additions to the package to keep this modular approach, with the assumption that islands of functionality are less likely to occur. (3) The name of files transferred to and from IASs are transparently allocated globally unique names to prevent naming conflicts and to ensure that an IAS does not perform the same data transfer twice. As such, explicit loads made by the programmer would fail. All console and visual output was also suppressed for improved performance. No further modifications were required.

Fig. 12.3 shows an activity diagram illustrating how the application is executed. This process can be summarised as follows: (1) the creation of TSOs by the user application specifying the MATLAB sequence desired, upon what data it is to be performed, and a Task Identifier (TID). The TSO is then passed through the client interface where MATLAB syntax is validated, data is placed into an HTTP server root for transfer and then sent to the scheduler. (2) Locate CSM, register and define the quantity of resources required. (3) The CSM informs the CRM of the new resource requirements. (4) The CSM checks the number of free IASs. (5) IASs are allocated if possible, otherwise the CSM waits for new IASs to become available. (6) The client interface receives information about discovered IAS(s). (7) The client interface schedules all the buffered Tasks. (8) For each subsequent IAS, one Task is reallocated using the above heuristic.

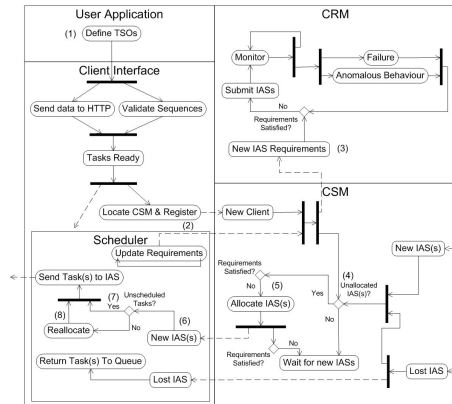


Figure 12.3: Activity Diagram showing interactions between components and the execution cycle of a user application.

As Tasks complete, the resource requirements of the application reduce, and the Client Interface keeps the CSM informed accordingly. As results are received, they are downloaded from the IAS's HTTP daemon and saved under the name specified in the TSO's TID.

#### 12.4.4 Fault Tolerance

The necessity for strong fault tolerance in an approach such as this has been motivated in Section 12.3.2. In order to handle the range of errors that can occur, two distinct mechanisms are required: (1) Infrastructural and (2) Task oriented.

Infrastructural fault tolerance is provided by the CRM. In order to detect that an MCR installation is portraying anomalous behaviour, every instance of an IAS job is closely monitored. This monitoring data is used to determine how long an IAS *should* require for initialisation, once the Condor job has started. The decision making process employed by the CRM is based on the data from over 5,000 successful IAS initialisation. A Job will be terminated if:  $JobRuntime > Max(10\ minutes, avg(initTime) \times 5)$ . There are two possible reasons for this to occur. (1) The workstation has poor performance or is heavily loaded, in which case it is not a suitable choice of workstation or (2) An MCR related error has occurred. In [76] we determined that the average initialisation time for an IAS was 3 minutes and a minimum of 24 seconds.

The IAS job's initialisation sequence also contains an MCR integrity check. Should this fail, the job will terminate with a specific error code. The CRM monitors the Condor logs in order to detect if an IAS Condor Job has been suspended, or if a specific error has occurred with a job during startup. When a job is suspended the CSM is informed and the job terminated. If, however, the CRM terminates a job for a reason other than suspension, or the Job log reports an unexpected error code, the specific workstation

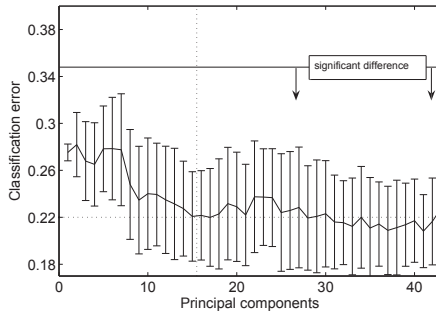


Figure 12.4: Mean classification error as function of *nPCs*, with error bars denoting the standard deviation respective to the mean.

receives 1 fail point. If a workstation receives more than a specific value of fail points it is black listed for 24 hours, for every subsequent point. This period of time was chosen with the assumption that the workstation will be restarted within this time allowing for the network management software to remedy any software abnormalities.

By employing these infrastructural fault measures we know that by the time an IAS becomes available that it will function correctly. However, specific user Tasks can still result in error and hence we also require some fault tolerance within the scheduler. Our error model adopts two tiers. (1) IAS Error (due to an IAS failure) and (2) MATLAB Error. In such cases the task is rescheduled. The second is application specific and relates directly to a fault within a specific Task. These are simply presented to the user at the end of execution and currently not resubmitted. Application specific logic is needed to handle these errors, as they are specific to the particular type of application being executed in MATLAB.

## 12.5 RESULTS

### 12.5.1 Application output

In figure 12.4, the mean classification error is plotted against the number of principal components for the ten-times repeated five-fold cross-validation (yielding 50 runs in total). For this particular dataset with 58 subjects, an error below 35% can be considered significant.

Mean errors decrease from 28% for 1 PC to 22% 15 PC's, after which it stays approximately constant; for larger *nPCs*, error reduction is not significant, so we choose 15 PCs as optimal for further classification. The corresponding thresholded map is overlaid to the FA volume and displayed in figure 12.5. This figure is presented to the radiologist for interpretation.



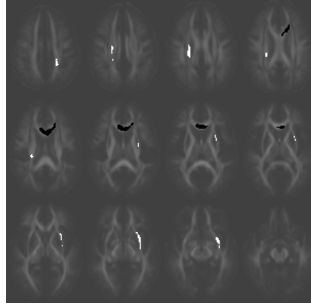


Figure 12.5: The thresholded map for  $nPCs=15$ , displayed as white and black regions for increased and decreased DTI-values for patients respectively, with anatomy on the background. Slices are displayed from top to bottom.

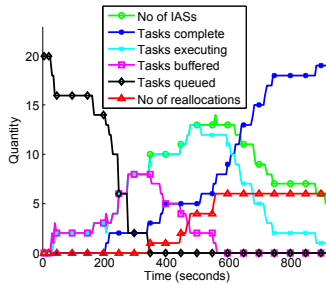


Figure 12.6: Plot of application execution when no resources were available when the application began.

### 12.5.2 Application Execution

When executed sequentially on a single workstation the analysis of a single dataset requires 90 minutes. This consists of 20 independent iterations, the results of which are later aggregated. We have examined three scenarios for the analysis of a given dataset: (1) No IAS jobs have been submitted by the CRM before the application begins. (2) The system has some resources, and (3) enough IASs are available for use by the application. These three scenarios have been chosen to illustrate the opportunistic qualities of the system and the potential speed-up achievable using this approach. Whilst the second and third scenarios may seem a waste of resources, the strategy of spare capacity is inseparable from the need to accommodate high demand on short notice, which in turn is the *raison d'être* for grid and cluster computing [156].

Figure 12.6 shows the analysis of one dataset when no system resources were available when the application started. In this scenario the system must wait for Condor to schedule some IAS jobs and for these IASs to successfully initialise. In this instance some IASs were initialised within 30 seconds, with further resources becoming available

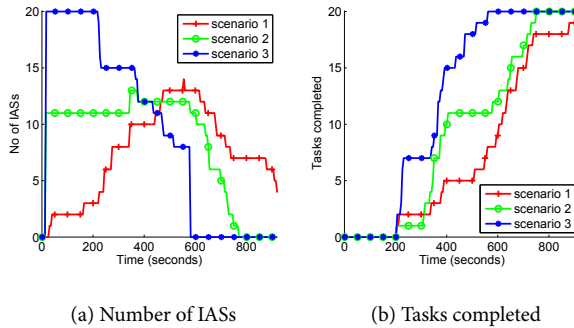


Figure 12.7: Plot of application execution when a handful of resources were available.

in the following 100 seconds. It is clear from the graph that IASs are acquired sporadically, which emphasizes the change in availability of these nodes. We can also see that as the number of IASs reaches 10, task reallocation begins to take advantage of new resources, as no tasks remain in the queue, but are either complete, executing or buffered in remote IAS queues.

The user waited 15 minutes for their application to complete. Considering that a maximum of 13 nodes were used (the 14th was not utilised) this does not initially show good performance. However, consider that the problem size is not a perfect fit for this number of nodes and that with time new resources were gained and released; on average 8 IASs were used. Also, consider the heterogeneity of the workstations being used. The average task completion time was in the order of 4 minutes. However, minimum and maximum execution times are: 3 and 7 minutes respectively. With these considerations a completion time of 15 minutes is respectable.

Fig 12.7 compares the execution of the application for the three above scenarios. For scenarios 2 and 3 there were 11 and > 20 IASs initialised, respectively. The scheduler initially buffers tasks and as such reallocation occurs 4 and 10 times, respectively. It is clear that the system’s performance in either of these two scenarios greatly impacts the overall wait time for the user, scenario 3 in particular. In scenario 3 task reallocation occurred within 30 seconds (note that the dataset was also transferred in this time). The effects of task and resource heterogeneity are also most apparent in scenario 3. However, using this scenario we have reduced the user’s waiting time from 90 minutes to less than 10, despite using unreliable volunteer resources.

### 12.5.3 Infrastructure Performance

The evaluation focused on (1) IAS initialisation times and (2) Effectiveness of the fault tolerance mechanisms. On average an IAS requires 220 seconds. During the development of this work the CRM has submitted > 25,000 IAS jobs. Of these 25,000 the CRM

logged 5,000 successful initialisations; 3,500 of which were later terminated in response to suspension. The remaining 1,500 can be accounted for during the testing of this system.

Of interest is the number of IAS job failures detected by the CRM. In total 9,000 failures occurred and can be attributed to: (1) MCR error, (2) no MCR installation despite the workstation's ClassAd reporting otherwise, (3) the initialisation time exceeded the threshold value mentioned in section 12.4.4. Of these (3) is by far the most frequent, consisting of approximately 7,500 jobs. Without the use of the CRM and other fault monitoring measures it would not be possible to achieve the above performance in such an environment. These 9,000 jobs were distributed over the 400 MCR-ready workstations.

The CRM database is configured using workstation name and IP address which are extracted from the relevant ClassAds. In order to log an event, the CRM must know both the IP address of the machine and its domain name, as not all IP addresses relate to only one Condor node. It is also possible for workstations to be upgraded and IP addresses to be reassigned. The dynamic nature of the Condor pool means that the CRM cannot acquire ClassAds for all machines at any given time, although over 2,500 records exist. Hence, all events cannot be recorded, and therefore there is a difference between the number of jobs recorded and the total number submitted.

## 12.6 DISCUSSION AND CONCLUSIONS

A grid-enabled MATLAB application for performing studies to discriminate among population groups, in this case schizophrenic patients and controls, is described. The ultimate goal of this application is to help the clinical researcher identify brain regions that differ the most between the groups, enabling evidence-based decisions. We describe how the MATLAB-based application was ported to a campus Grid infrastructure, with minor modifications. The application workload, consisting basically of independent tasks, was transparently distributed among a 400 machine Condor Pool. As pre-compiled MATLAB is used, licenses are not required to run the application in parallel. From the user point of view this is very important since it enables running any MATLAB application on grids without the high cost of cluster licenses. The grid-enabled implementation provided a latency reduction from 90 minutes to 10 minutes per dataset, enabling the realization of complex and time consuming experiments in reasonable time. The reduction in latency was obtained through (1) compiling MATLAB code, (2) parallel execution, (3) an improved job submission and management scheme.

Our use of interactive services permits us to side-step the Condor queuing strategy for most tasks. Instead, our management component (CRM) can submit IAS jobs in response to a global demand. The scheduler can then interactively and dynamically allocate tasks to the available resources during the execution of the application as a whole, performing reallocation as required. This means that we can perform tasks near

to their *normal* execution times should IASs be preloaded, and thus greatly improve performance.

With this Condor-based system, we are able to compare results with an earlier implementation [66], demonstrating correctness, but with a reduction in time for data analysis. The success of this experiment opens up the possibility of reducing latency in other parts of the image analysis pipeline, or performing more complex performance assessments within large parameter ranges. In our current workflow, as depicted in figure 12.1, the interpretation and visualisation of the PCA/LDA-mapping are not included in the Condor framework. Instead, a naive thresholding is applied, to find regions where brain tissue of schizophrenics is affected. Future work will be on *shaving* the PCA/LDA mapping [66], to more precisely delineate the regions of interest. Non-discriminating voxels are iteratively discarded, each time updating the PCA/LDA-mapping. We would like to optimise the shaving step size and/or stopping criterion. The interpretation step of the workflow will thereby be migrated into the Condor workflow as well. Even more interesting is the joint optimisation of the number of principal components and the shaving parameters, as it yields an even higher-dimensional parameter space.

There are new challenges in preparing the Condor pool for this migration. Hardware requirements will be increased, as more memory is needed to perform the computations. Besides, scheduling the jobs on the Condor grid becomes more challenging, as the job length will be longer and dependent on the chosen shaving parameters.

The proposed framework can be generalised to other pathologies, among which are brain diseases such as Multiple Sclerosis (MS) and Amyotrophic Lateral Sclerosis (ALS). Also functional MRI clinical studies for neurosurgery planning of oncology patients could benefit from the proposed approach. For each new case, the type of analysis described here must be repeated, so there is interest in porting the current solution to the grid infrastructure available in The Netherlands.

Diffusion Tensor MRI (DTI) is a rather recent image acquisition modality that can help identify disease processes in nerve bundles in the brain. Due to the large and complex nature of such data, its analysis requires new and sophisticated pipelines that are more efficiently executed within a grid environment. We present our progress over the past four years in the development and porting of the DTI analysis pipeline to grids. Starting with simple jobs submitted from the command-line, we moved towards a workflow-based implementation and finally into a web service that can be accessed via web browsers by end-users. The analysis algorithms evolved from basic to state-of-the-art, currently enabling the automatic calculation of a population-specific “atlas” where even complex brain regions are described in an anatomically correct way. Performance statistics show a clear improvement over the years, representing a mutual benefit from both a technology push and application pull.

Published as:

M.W.A. Caan, F.M. Vos, L.J. van Vliet, A.H.C. van Kampen and Silvia D. Olabarriaga, *Gridifying a Diffusion Tensor Imaging Pipeline*, Workshop on Challenges for the Application of Grids in Healthcare, In conjunction with CCGrid'2010[65].

### 13.1 INTRODUCTION

Over the past decades, increasing insight has been gained in the structure and function of the human brain, as well as in brain diseases. Non-invasive imaging techniques such as Magnetic Resonance Imaging (MRI) contribute to this progress. An MR imaging modality called Diffusion Tensor MRI (DTI) is particularly sensitive to subtle changes in the structure of nerve bundles. By adopting this modality markers for diseases such as Multiple Sclerosis, schizophrenia, and Alzheimer's disease can be more accurately characterized.

Grid computing emerged around a decade ago as a promising approach to perform research with virtual organizations spread over multiple institutions on shared resources [108]. Several projects around the globe have attempted since that time to implement this vision in existing and novel research environments. The Virtual Laboratory for e-Science (VL-e) Project (<http://www.vl-e.nl>) is such a project that tried to build a grid-enabled infrastructure for research in The Netherlands. A Proof of Concept environment was built, in which collaborating researchers from various disciplines, among them medical image analysis, could perform their experiments.

In this paper we present our progress in analyzing DTI datasets gathered in clinical studies on the grid over the past four years. Both the application and technology sides show an increase in complexity, resulting today in a workflow-based service that can be accessed by end-users from the web for estimating a 'study-specific atlas' for large datasets.

The paper is organized as follows. We introduce the DTI pipeline (Section 13.2) and discuss its implementation on the grid (Section 13.3). Then we present performance statistics, user experiences and clinical results (Section 13.4). Conclusions are finally drawn in Section 13.5.

### 13.2 THE DIFFUSION TENSOR MRI ANALYSIS PIPELINE

Magnetic Resonance Imaging (MRI) is a non-invasive imaging method measuring the human body at a high magnetic field strength of 3 Tesla. Diffusion Tensor MRI (DTI) is a specific MR-modality enabling the identification of the orientation of human tissue. This is particularly beneficial in the brain, which is highly organized by nerve bundles, connecting different brain regions to one another and to other body parts such as the muscular system. Nerve bundles are built up by axons surrounded by water. The diffusion of water in a nerve bundle is anisotropic, i.e., it is higher along the direction parallel to the axons, and lower along the perpendicular direction. In brain fluid, the diffusion is isotropic, i.e. equal in all directions. DTI measures this diffusion profile in 3D by acquiring commonly 30 to 60 3D volumes of data, each measuring water diffusion in a specific orientation. A model is then fit to each voxel (volumetric image element) to reconstruct the so called "diffusion tensor" (a  $3 \times 3$  symmetric matrix). This may be

interpreted as an ellipsoid that is aligned with the tissue and shaped according to the diffusion profile [256]. In voxels where bundles are crossing, two tensors need to be fit to the data [200].

To facilitate the analysis, tensor-derived measures are computed from the DTI data and visualized using color maps. Examples are the Fractional Anisotropy (FA), between 0 and 1, which describes the degree of anisotropy [32], and the Mean Diffusivity (MD), which quantifies the average size of the tensor and describes the amount of diffusivity. Fiber tracking is the discipline of reconstructing the global nerve bundle trajectories from the measured tensors per voxel. Starting from a seed point, a streamline is generated by integrating along the principal directions of the tensors. Tracking is terminated if the curvature exceeds, or if the FA drops below, a threshold. Multiple seeds and filtering operators (AND/OR/NOT) can also be used to enable the researcher to isolate specific nerve bundles.

DTI is frequently used in comparative studies of brain diseases that are thought to cause local damage to brain tissue, possibly only in specific nerve bundles related to a given brain function. For example, Amyotrophic Lateral Sclerosis (ALS) particularly affects the nerve bundle related to muscular functionality, resulting in a reduced motion ability and eventually heart failure. At the micro scale axonal damage occurs, which in part affects the shape of the diffusion profile, such that the FA decreases and the MD may increase. Although DTI can capture subtle changes in tissue composition, differences may not be apparent in an individual brain. Hence, studies of such disease markers are based on groups of patients and “healthy” controls matched on age and education level.

The analysis of DTI data in such a clinical study includes various steps described below (see Figure 13.1). The volume of computation is proportional to the number of subjects (or scans  $m$ ) and image size ( $n$ ). Common values are  $m = 50$  to 100 subjects, and images of  $128 \times 128$  voxels with 64 slices, of which  $n \sim 10^6$  voxels reside inside the brain. Since many computations are independent and can potentially be performed in parallel for each subject, voxel or slice, below we indicate roughly the order ( $O$ ) of parallelism that could be exploited in each step.

**Data acquisition.** DTI-data of the subjects is acquired on a 3T MRI scanner, involving the measurement of 30 to 60 Diffusion Weighted volumes. Acquisition may be distributed over multiple scanners or hospitals in large studies.

**Preprocessing.** Preprocessing steps include correction of motion artifacts, noise filtering, and computation of tensor fields, FA and MD measures. These steps can be performed independently for each scan ( $O(m)$ ).

**Coregistration.** Spatial correspondence between individuals is obtained by aligning them to each other (coregistration). This corrects for different subject positioning in the scanner, varying head sizes and subtle changes in nerve bundle morphometry. In a naive implementation, all volumes are directly coregistered to an existing average brain,

such as the ICBM-81 FA atlas [181] ( $O(m)$ ). In a more accurate approach, coregistration is performed by estimating the deformations between all pairs of subjects in the study. Each pair of scans can be registered independently ( $O(m^2)$ ).

**Atlas estimation.** When an “atlas” is constructed specifically for the given study, the deformation fields computed by the registration step need to be averaged for all subjects to define a new common space [244]. Using the average deformations, all subjects are transformed (warped) to the common space. The warping operation can be performed independently for each subject ( $O(m)$ ). A single average tensor field is computed, which can then be used for further analysis. This so called atlas is specific for this group of subjects, as it is estimated from the given data. By using all subjects’ data, instead of data available for each subject individually, better SNR (signal-to-noise ratio) is obtained, allowing for fitting a two tensor model (describing fiber crossings), even if the data of individual subjects has poor quality [64]. This operation can be performed independently per voxel - or per slice, for a more reasonable granularity ( $O(n)$ ).

**Statistics and pattern recognition.** Statistics calculated for the group may then indicate if on average the FA or MD is significantly different for the subjects with brain disease respectively to the controls. This analysis is classically done individually per voxel. Since the brain is a complex, highly inter-connected organ, we previously proposed a pattern recognition framework to jointly analyze all brain voxels [66]. The algorithm involves a sweep of  $p$  parameter values and a cross-validation of  $i$  iterations to optimize between variance in the data and noise. The calculations can be done in parallel for each parameter combination and iteration ( $O(p \times i)$ ).

**Reported differences.** The findings of the statistical analysis are discussed with clinicians and other medical experts for interpretation in light of the disease. For example, to investigate whether the brain locations highlighted by the statistical analysis relate to known brain function. This requires presentation of results visually on the anatomical scan.

### 13.3 GRID IMPLEMENTATION

DTI analysis is an interesting use case to be deployed within a grid environment because significant amounts of CPU time and storage may be needed to process an entire clinical study. The algorithms implementing the pipeline steps were developed in Matlab (Mathworks), which allows for rapid prototyping and efficient computation on large matrices. Matlab is a commercial product, and costs about \$1,000 per installation, which represented a challenge for us with respect to parallelization in the early days of the project. Open source Matlab-like environments do not provide the researchers with required functionality. Moreover, although the computing steps are depicted sequentially in Figure 13.1, in practice they are not trivially connected. This is mainly due to the different dimensions along which the data is split into parallel operations, as described in Section II. Also, independently developed algorithms follow their own conventions (e.g.,



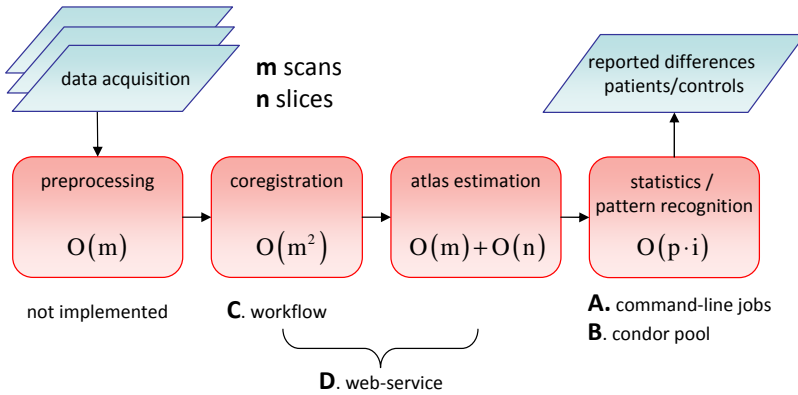


Figure 13.1: Overview of the pipeline for a clinical study, indicating the potential order of parallelism to be achieved ( $O$ , see text) and the corresponding grid implementations (see section 13.3).

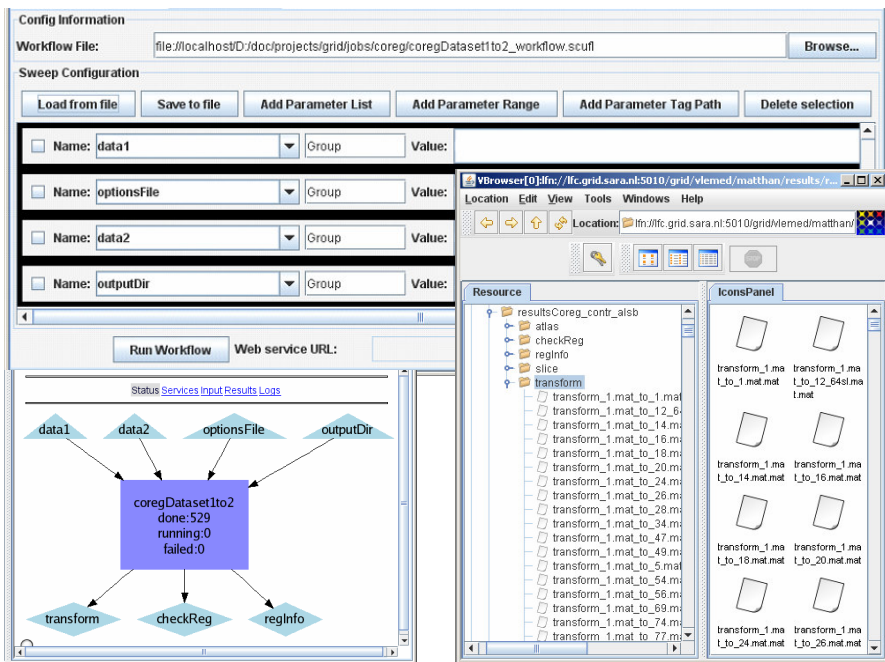


Figure 13.2: Coregistration workflow and VBrowser user interface for workflow submission and data management.

all inputs into one tar archive), such that data conversion steps need to be introduced when connecting them.

Below we describe our various attempts to port the DTI analysis to the grid in chronological order. The initial experiments focused on the pattern recognition component, because this was the most time-consuming pipeline step at the time. The same experiment was implemented on local resources, on the Dutch grid and on a campus grid. Later we ported the coregistration component to the grid, and more recently the complete atlas construction workflow. At this point, the preprocessing has not been gridified yet.

### 13.3.1 *Command-line jobs on a cluster*

The first experiment (2006) aimed at performing the parameter sweep for pattern recognition. The goal was to acquire expertise in using the Dutch grid using the command-line interface. We were granted access to a single 72-node cluster, in which Matlab was installed with sufficient licenses.

We adopted conventional gLite command line utilities to run the parameter sweep. An in-house toolkit of shell scripts took care of data and software staging on the worker node, wrapping the Matlab programs into gLite jobs, and allowing the user to submit the entire range of the parameter sweep at once. General cluster monitoring tools provided statistics (see Figure 13.4). After all jobs completed, the results could be retrieved using the toolkit. Failed jobs needed to be manually identified and resubmitted by the user. Such failures occurred in early days due to unstable cluster installations or unavailability of resources. When all computations were finished, the results were locally merged and further interpreted.

### 13.3.2 *Condor pool of desktop machines*

The same experiment was repeated in 2007 on a campus-grid of volunteer resources at the Cardiff University - see more details in [75]. A 2,500 node Windows Condor pool was potentially available, but without Matlab licenses. The Matlab code was compiled, and the Matlab Component Runtime (MCR) library was installed on 400 nodes. Heterogeneous sources of errors, as well as the dynamic (un)availability of resources, made the use of this pool challenging. We adopted a service-based approach that autonomously submitted jobs to available resources. We investigated the performance statistics for varying numbers of available resources at the start of the experiment.

### 13.3.3 *MOTEUR workflows on the Dutch Grid*

Later (2008) we ported the coregistration step to the grid using the Virtual Lab for Medical Images [193] developed in the VL-e project. It includes a service to execute workflows on the grid using the MOTEUR engine [117]. The Virtual Resource Browser (VBrowser) serves as front-end, allowing for data management, workflow submission, monitoring and result retrieval [191]. Legacy applications implementing the workflow

components (in this case the Matlab code) are automatically wrapped by MOTEUR to execute as jobs on the grid.

The coregistration experiment was described using the SCUFL language of the Taverna workbench [189]. The cross-product operator was used to generate an all-vs-all combination of the subjects. All data (input images and output transformation matrices) are stored on grid resources and interactively manipulated by the user. The Matlab program implementing coregistration was adapted to accept all inputs and outputs as command-line options and to perform extra error detection and logging. To simplify data management, the images are preconverted to Matlab-friendly format and combined into one archive per subject in a pre-processing step performed on the local workstation.

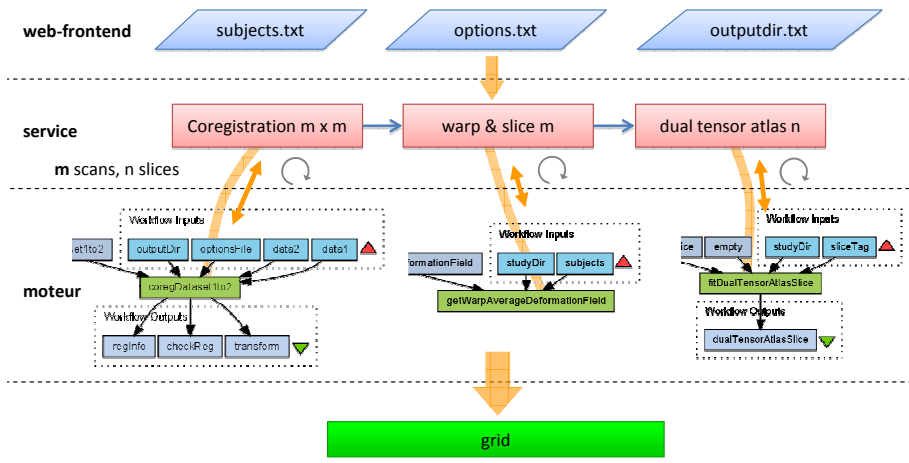
This platform provided a comfortable abstraction layer that enables the end-user to autonomously run large coregistration experiments on the grid (such as in [245]).

#### 13.3.4 *Web-based service with nested fault-tolerant workflows*


Now that the coregistration could be successfully executed, atlas construction became possible for larger groups. We foresee that such population-specific atlases will become tomorrow's routine in clinical studies, and that studies will become larger, requiring more computational effort. Although the interface of the VBrower is intuitive for the medical imaging researcher, it still exposes too many details about the grid, and is therefore unsuitable for the typical clinician. We therefore set-up to build a web service that can be invoked from a web interface (2009).

As seen in Section 13.2, the analysis include steps that differently divide the data in order to exploit parallelism: coregistration (subject), averaging deformation fields and warping the scans into the common space (subject), and tensor estimation (slice). Current limitations in the workflow description languages hamper the manipulation of multi-dimensional data, which is a known problem being addressed by the GWENDIA project [178]. This forced the pipeline to be defined as three separate workflows, imposing user intervention in-between the workflows and hampering its usage by non-experts.

We therefore developed a script that autonomously computes an atlas by encapsulating and connecting all the workflows, and also implements data management and conversion, experiment monitoring and recovery. A schematic overview of the service is shown in Figure 13.3. The MOTEUR service is invoked to execute the sub-workflows on the grid. To provide more external control, the cross-product operators were not used in these workflows. Instead, the dot-product is used, and script generates the list of inputs including all combinations to be computed. The remaining computations are determined based on missing output files. MOTEUR workflows are repeatedly executed until all output is generated. The service then proceeds to the second workflow, which is computationally light and was called once for each subject. Here the data was also sliced as a preparation for the third workflow, where a two tensor model was estimated per voxel.







virtual laboratory for e-science

### Start experiment

Select subjects file:

Select options file:

Select outputdir file:

### createAtlas

Experiment ID: /data/home/mcaan/exp/alsAIAtlas  
 98 subjects. Options: /grid/medim/mathan/bin/regOptions.txt. Output-directory: /grid/medim/mathan/results/alsAI1264Atlas  
 Start time: Mon Dec 14 11:58:29 CET 2009 Mon Dec 14 12:00:50 CET 2009. Now: Mon Dec 14 14:55:43 CET 2009. **Elapsed: 2h57m.**  
 Equipment aborted. Check error log for details.  
 No errors.

Workflows				
No.	Name	Status	Total jobs	Remaining
1	coregDataset1to2	RUNNING	8115	7515
2	getWarpAverageDef	-	-	-
3	fitDualTensorAtlasSlice	-	-	-

Moteur workflows		
No.	moteurID	vWaiting Running Completed Error
1	workflow-8i23eE	0 0 300 16
1	workflow-80aq6B	0 0 300 58
1	workflow-adVBrmj	145 154 0 16

**Log files**

[main.log](#)  
[error.log](#)

Figure 13.3: Schematic overview of the atlas construction service workflow (top). Web front-end for the service: submission and status (bottom).

A web interface was chosen to deploy the full script. The inputs consist of text-files with references (URL) to data on grid storage, configuration options and the output location. In the current implementation, the data must be available on grid storage such as the Logical File Catalogue (LFC), and likewise the results need to be fetched from the grid by the user. The VBrowser provides the functionality to do so, but a pure web-based interface is still missing to facilitate data upload and download by clinical researchers. If the user does not own a valid grid certificate, a robot certificate can be used to run jobs and access data. In such case the user needs to conform with the security policy, in particular taking responsibility for anonymizing the data and implicitly trusting the other Virtual Organization members.

After the service starts a URL to the workflow monitoring page is returned to the user. The monitoring page is periodically updated to report the computed and remaining results, indicating errors if applicable, and linking to the individual monitoring pages of each MOTEUR workflow. Eventually the user is informed that all workflows completed successfully, so the results can be downloaded using the VBrowser. Possible errors are logged and may be reviewed by the user.

## 13.4 RESULTS

Our results in porting the DTI analysis to the grid are discussed below in chronological order. Usage statistics are summarized in Table 13.1. All subject data was anonymized prior to processing. Informed consent was given by all participants.

### 13.4.1 *Parameter sweep for pattern recognition on Dutch grid and on Condor pool*

In this experiment we performed a parameter sweep on the number of “principal components”  $p$ , ranging from 1 to 20. We adopted data from a study into schizophrenia (58 subjects), that was locally preprocessed using conventional methods. Usage statistics after submission to the Dutch grid are depicted in Figure 13.4. The cluster in use was shared with other researchers and heavily loaded. Consequently, after 40 minutes only 10 nodes were running simultaneously. This was the maximum capacity reached at that time and considered a success from the proof-of-concept perspective, but certainly not sufficient to impress the potential users. Additionally, the required Matlab licenses were prohibitive, and the use interface offered too low level access to the grid.

The same experiment was repeated on a Condor pool of voluntary resources as described in [75] – see Figure 13.5. When more resources were available at the beginning, the total computation time could be significantly reduced. In the best actually achieved run, within 15 minutes all jobs were executed, representing a fair speedup compared to the 90 minutes of total CPU time.

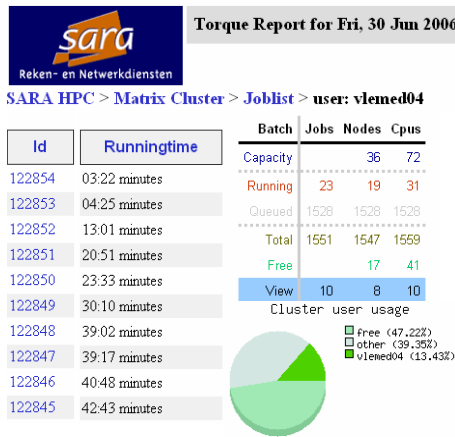


Figure 13.4: Dutch grid cluster usage statistics for the command-line experiment.

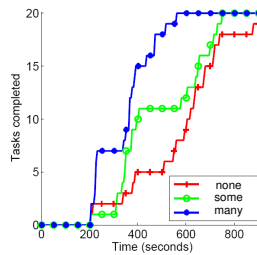


Figure 13.5: Condor pool experiments. Number of completed tasks along time, for different amounts of initial resources (from [75]).

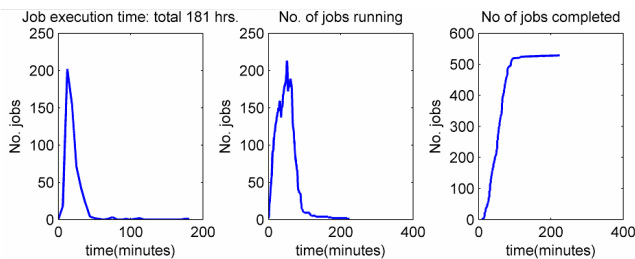


Figure 13.6: Performance of the coregistration workflow along time. Left: Job duration. Center: Number of simultaneous jobs. Right: Number of completed jobs.

#### 13.4.2 *Coregistration as a MOTEUR workflow*

In this experiment we performed a coregistration of 23 subjects (i.e. 529 jobs), adopted from a study into a muscular disease, called Amyotrophic Lateral Sclerosis (ALS).

Figure 13.6 shows that individual job running times were fairly short: 10 min on average. Also, many resources were free at the time, and a maximum of 210 jobs ran simultaneously. The long tails in Figure 13.6 suggest that a few (less than 10) data-related errors had a serious impact on the performance (data uploading only succeeded in a retry after a time-out). Nevertheless, the total computation time reduced from 181 to 3 h wall clock time.

#### 13.4.3 *Atlas building using a web-based service*

In this experiment we built an atlas on data of 46 subjects with 64 slices, a larger sub-population from the same ALS-study as in the previous section. The web-based frontend to the service is displayed in Figure 13.3.

Due to capacity limitations in the current platform, the number of jobs that could be simultaneously submitted was restricted. The service hence performed the coregistration in batches, splitting the workload into sets of a maximum of 300 jobs in these experiments.

The performance of the service is depicted in Figure 13.7. A single coregistration on average took 20 minutes to compute. Due to empirically optimized parameter settings with respect to registration quality, the execution time was doubled compared to the previous experiment. The total amount of CPU time was 656 h, while the entire coregistration completed in 9 h wall clock time. A speed-up of only 73 times was achieved, in spite of the fact that up to 260 jobs were running simultaneously. A significant amount of time is spent in waiting for the completion of a few final coregistrations per batch, as seen in the top right of Figure 13.7. An increase in performance may be achieved by submitting a new batch if more than 90% instead of 100% of the previous workflow is completed. The second and third workflows may be considered optimal already, given the relatively flat curves for running jobs.

#### 13.4.4 *Performance Summary*

A summary of the performance statistics of all experiments is given in Table 13.1. What catches the eye is that the efficiency of the larger scale experiments is still relatively poor.

#### 13.4.5 *Clinical results*

After successful atlas generation, the results were downloaded and locally merged to track fibers. We followed two approaches: the classic ‘single tensor’ method and novel ‘dual tensor’ method [237] – see Figure 13.8. A bundle traversing in both hemispheres can be

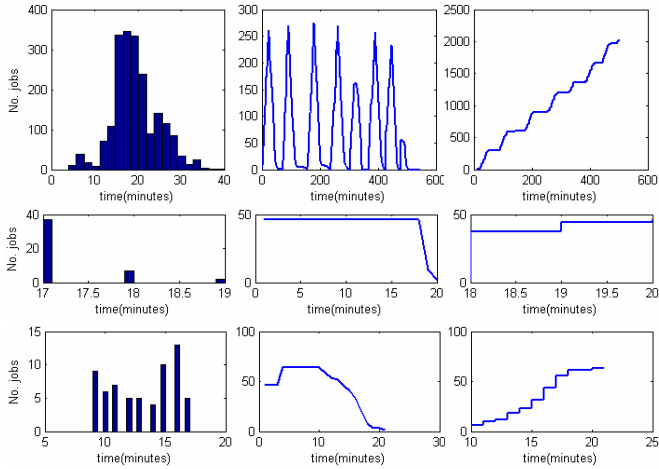


Figure 13.7: Performance of the individual workflows of the atlas construction service: coregistration (top), averaging and warping (middle) and tensor estimation (bottom). Note the varying scales along the axes.

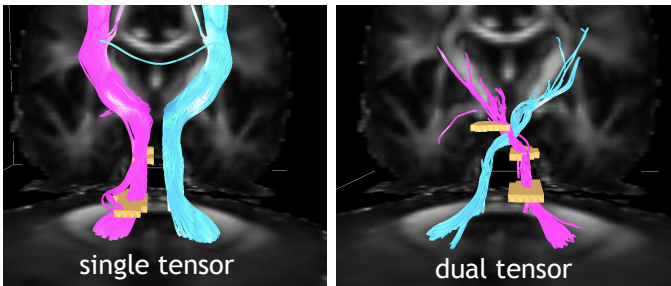


Figure 13.8: Clinical outcomes of the atlas building service: fiber tracking results using a ‘single tensor’ and ‘dual tensor’ fitting approach. The latter provides additional information about crossing bundles, that traverse from one brain hemisphere to the other.

Table 13.1: Summary of performance statistics. Speed-up factor: CPU divided by wall-clock time. Efficiency (eff.): the speed-up factor relative to the maximum number of jobs running at one time. Data for the first experiment are missing.

	# jobs	max jobs	CPU	clock	speed-up
command-line	20	10	-	-	~3
condor pool	20	20	90min	10-15min	7.5
workflow	529	210	181h	3h	60
service 1	2112	269	656h	9h	73
service 2	46	46	13h	20min	39
service 3	64	64	14h	22min	40



reconstructed using the single tensor approach. The dual tensor approach reconstructed a ‘novel’ connection: branches of the main bundles cross and traverse to the other hemispheres. Despite the modest quality of individual datasets, these tracts could be successfully represented in the atlas.

### 13.5 DISCUSSION AND CONCLUSIONS

The early work of [179] pointed out that several medical applications could potentially benefit from grid computing. Since then different approaches have been reported previously to perform medical image analysis on the grid - see an overview in [190]. As an example of directly related work, compiled Matlab code was executed on a Condor cluster to speed up the computation of electrical impedance tomography reconstruction algorithms [112]. Also Condor-G and Pegasus have been used to speed up Matlab-based evaluation of tomographical data [162]. Similar to our earliest findings, queuing times significantly impact performance. Others recently implemented neuroimaging applications using workflow technology [103], similarly to our implementation. The only step that we have not considered so far, grid-based DTI preprocessing, has been implemented using existing toolboxes [166].

We have seen a glimpse of the healthgrid vision becoming partially true: multi-disciplinary research leading to novel results that could not have been achieved otherwise. The project was truly distributed, and in not a single moment were all collaborators gathering at the same location. Medical experts were designing a clinical study and gathering patients within the clinic. They were helped by physicists designing the MR scanning protocol. Image processing experts then took over to devise the algorithms to analyze the data. These were deployed on the grid, together with grid experts. Medical experts assisted in fiber tracking and interpretation of the clinical outcomes. We can state that there was a mutual benefit in development from all sides, to a level that we did not envision at the beginning of the project. Clinical and medical imaging researchers who were previously skeptical about our research approach are now becoming willing to collaborate in the grid environment.

### ACKNOWLEDGMENT

The authors would like to thank M.M. van der Graaff, B. Peters and C.B.L.M. Majoie for providing us with data and interpreting the results. We are grateful to J.J. Keijser, P. de Boer, C. Byrman, A. Luyf, M. Stam, T. Glatard, K. Boulebiar, J. Engelberts, T. Visser, grid.support@sara.nl, C. Sage and W. van Hecke for their contributions. This work was sponsored by the BiG Grid project for the use of the computing and storage facilities, with financial support from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Organisation for Scientific Research, NWO).



## CONCLUSIONS

---

Pathologies affecting the white matter of the human brain can be detected in Diffusion Weighted MRI. The goal of this thesis was to develop methods for improved quantification of such pathological processes. Biomarkers could then be found to detect these processes in a pre-symptomatic stage.

### 14.1 PART I - DISCRIMINANT ANALYSIS

In chapter 2 we introduced a new machine learning framework for comparative studies on volumetric data. By ‘shaving’ the mapping computed by PCA/LDA, a characteristic set of regions is automatically extracted. The method was applied to DTI-data of schizophrenics and controls, which resulted in a classification error of 22% after shaving (versus 31% before shaving) based on regions that comprise only 0.3% of the voxels. The correlation in the data was demonstrated between the corpus callosum and uncinate fasciculus brain regions. Conventional VBA found a decrease in FA in the corpus callosum. Also, an increased FA in various white matter regions including those found using the proposed algorithm was found, which is unprecedented [147].

A technique called ‘pruning’ was then introduced in chapter 3 to iteratively discard misclassified subjects from the cohort. These subjects primarily reside in the feature space where the classes are overlapping. As the exact cause of this overlapping is unknown, it is preferable to base the mapping merely on representative prototypes, residing in the non-overlapping parts of the feature space. The method, applied to schizophrenia, found a more pronounced decrease in FA in the corpus callosum after pruning the dataset.

In addition to the FA, the linear anisotropy  $c_1$  was used in chapter 4 to classify schizophrenia, yielding a slightly smaller classification error (albeit not significantly different from the error when using the FA).  $c_1$  reported a few regions with false positive differences, such that using the FA is preferable. The Fisher Kernel classifier overtrained the data, with a higher error than its linear counterpart. A linear classifier is therefore preferred, because of its lower complexity and the ability to interpret the mapping.

### 14.2 PART II - RESOLVING CROSSING FIBERS

In chapter 5 we proposed a new method for creating experimental phantom data of fiber crossings, by mixing the DWI-signals from high FA-regions with different orientation.

Not only are the results based on realistic imaging artifacts, the reconstructed fiber orientations can be compared to a ground truth, being the single tensor fits on the individual voxels. An angular resolution of  $0.6\pi$  was found for  $b=1000, 3000$  and  $5000 \text{ mm}^{-2}\text{s}$ , while for  $b$ -values above  $1000 \text{ mm}^{-2}\text{s}$  an increase in resolving power was expected. Apparently, the SNR was too low to allow for stronger signal attenuation. A denser sampling of  $b$ -values may further specify the optimum.

We proposed a method to create a dual tensor atlas from multiple coregistered non-HARDI datasets in chapter 6. Random fluctuations in the pose of subjects in the scanner, as well as anatomical heterogeneity resulted in an increased angular resolution: simulations predicted a factor of two per decade of subjects. Commissural corpus callosum fibers reconstructed by our method closely approximated those found in a HARDI dataset.

There is an ongoing debate on how to model diffusivity in fiber crossings. In chapter 7, we propose an optimization framework for the selection of a dual tensor model and the set of diffusion weighting parameters  $b_i$ , such that both the diffusion shape and orientation parameters can be precisely as well as accurately estimated. For that, we have adopted the Cramér-Rao lower bound (CRLB) on the variance of the model parameters, and performed Monte Carlo simulations. We found that the axial diffusion  $\lambda_{\parallel}$  needs to be constrained, while an isotropic fraction can be modeled by a single parameter  $f_{iso}$ . Under these circumstances, the Fractional Anisotropy (FA) of both tensors can theoretically be estimated independently with a precision of 9% (at SNR=25). The Maximum Likelihood estimator based on a Rician noise model approached this precision while the bias was insignificant. Two non-zero  $b$ -values  $b_{1,2} = [ 1.0 \quad 3.0 ] \cdot 10^3 \text{ mm}^{-2}\text{s}$  were found to be sufficient for estimating parameters of heterogeneous tissue with low error. Noise reduction by adaptive filtering was studied as an optional step to increase the SNR in clinical data. It was shown that the data remain Rician distributed after filtering. This has allowed us to estimate consistent FA-profiles along crossing tracts. This work defines fundamental limits for comparative studies allowing them to correctly analyze crossing white matter structures. The framework is also directly applicable to other parameterized diffusion models and different experimental conditions.

### 14.3 PART III - FIBER TRACT STATISTICS

In studies where specific white matter tracts are hypothesized to be affected, fiber tracts may be chosen as a statistical frame of reference. We introduced a new method for non-rigid matching of white matter fiber tracts (chapter 8). Inter-subject correspondence was achieved by non-rigid registration based on a joint clustering and matching.

We presented a novel approach to 3D, tract-based statistical analysis of FA data in chapter 9. The clustering in the fiber tract matching delivered atlas points that served as a frame of reference for performing the analysis. Distribution of FA-values per cluster and per subject were calculated and tested for inequality using the Kolmogorov-Smirnov test. A

study into infant medulloblastoma survivors showed a decreased FA in large parts of the corpus callosum. The difference was more profound than when performing *t*-test statistics per cluster. The decrease could be predicted by a 3-NN classifier with an error of 14%.

#### 14.4 PART IV - CLINICAL STUDIES

The applicability of DW-MRI in clinical studies demonstrate its potential to detect subtle differences in white matter integrity. Chapter 10 shows that DW-MRI on a 3.0-T MRI scanner is sensitive to region specific changes in white matter integrity in pediatric cancer survivors. The white matter fractional anisotropy (WMFA) correlates with speed of processing and motor speed, both serious problems for childhood brain tumor survivors. The corpus callosum and the bilateral inferior fronto-occipital fasciculi were particularly involved. This is the first study reporting WMFA changes after childhood cancer using a 3.0-T MRI scanner.

Chapter 11 is the first study that uses DW-MRI data and tractography to reflect degeneration in recent onset motor neuron diseases (MND). We carefully defined subtypes of MND and included equal numbers per group to avoid bias due to heterogeneity. We performed a systematic follow-up after 6 months. We demonstrated different profiles of corticospinal tract degeneration in different phenotypes of MND. Additionally, DW-MRI and tractography parameters correlated significantly with clinical parameters, indicating that these are (partly) influenced by degeneration of the upper motor neurons. Changes over time were modest. As we found different patterns of corticospinal tract degeneration in the various phenotypes of MND, our study suggests that these phenotypes indeed may have different underlying mechanisms of degeneration and therefore should be studied as separate entities.

#### 14.5 PART V - GRID IMPLEMENTATION

In Part I, we optimized the number of principal components needed to discriminate schizophrenic patients from healthy controls. Here sequential runs were performed, resulting in long execution times. In chapter 12 we describe an experiment where this application was run on a grid with minimal modifications and user effort. We successfully approached a campus-wide Condor pool comprising of volunteer resources, which change randomly with time and thus require fault tolerance. Through this approach we have reduced the computation time of each dataset from 90 minutes to less than 10.

Due to the large and complex nature of DW-MRI data new and sophisticated pipelines are needed for the analysis to be more efficiently executed within a grid environment. Chapter 13 presents our progress over the past four years in doing so. Starting with simple jobs submitted from the command-line, we moved towards a workflow-based implementation and finally into a web service that can be accessed via web browsers by end-users. Performance statistics show a clear improvement over the years, representing

a mutual benefit from both a technology push and application pull. Novel studies can use the developed methodology in a straightforward way via the web portal.

#### 14.6 FUTURE RESEARCH

May it be voxelwise or based on tracts, statistics are commonly analyzed independently, such as in publicly available packages as SPM [111] and FSL [259]. Modeling correlation in the data should become common in comparative studies. We will continue using our shaving algorithm in future studies.

Now that we can also estimate diffusion properties in crossings, patient and control groups can be compared in larger areas of the brain. An enhanced statistical framework is needed for assessing the parameters of more complex diffusion models. Our dual tensor model and atlas may contribute to performing these analyses.

Using fiber tracts extracted from volumetric data as a frame of reference, is a promising new way of acquiring statistics. As of today, a thorough comparison with a voxel-based reference frame has not been made, such that both approaches are regarded as complementary. Comparing both methods is challenging, given the large amount of coregistration and tracking methods that are available. Also, optimal parameters for both approaches need to be found. Even harder is the definition of a performance measure, for which the accumulated volume of discriminating regions may be adopted. Still, an adequate phantom or ground truth is currently not available (i.e. diffusion shapes need to be known in addition to fiber orientations that were reconstructed from a phantom [236]). One might choose to determine the microscopic truth with respect to macroscopic DW-measurements, as was previously found in the mouse optic nerve [227].

Increasing interest is into following disease developments over time. For that, subjects must be scanned multiple times after which the data is subject to pairwise comparisons, such as in chapter 11. In case an entire cohort is thought to sample all disease stages, manifold learning might reveal the disease progress (while intra-class variability is being ignored), as has been done previously in structural data [94]. Healthy aging was recently monitored in the FA by means of Principal Component Analysis [120]. Applying these and other pattern recognition techniques allow for estimating complex processes from the data. Eventually, a time resolved atlas may be built that captures specific pathological processes. If created in prospective epidemiological studies [245], the advent of a disease may become apparent.

Cohort sizes tend to increase and may be scanned at multiple centers. Flexible data organisation and processing on distributed resources is then done in collaboration between multiple partners. A virtual laboratory as proposed in this thesis facilitates such an approach to neuroscience, which is what the grid was designed for [108].

The work performed in this thesis focused on the structure of the brain. The relation between structure and function is yet to be explored in full depth. The default mode network was shown to be structurally connected by combining resting state functional MRI and DTI [122]. A combined Positron Emission Tomography (PET) and DTI study of patients diagnosed with Tourette syndrome, abnormalities of the serotonin neurotransmitter were correlated to microstructural integrity [222]. Modeling these joint processes is a highly challenging task and therefore one of tomorrow's image processing assignments.

Eventually, all research should find its way to the clinic. Relatively few DW-MRI-biomarkers are currently employed. This thesis enabled further quantification of pathologies in DW-MRI, which is encouraging for the future.





## BIBLIOGRAPHY

---

- [1] PRTools, a statistical pattern recognition toolbox. <http://www.prtools.org/>. (Cited on page 149.)
- [2] ALiCE grid computing project. <http://www.comp.nus.edu.sg/teoym/alice.htm>. (Cited on page 146.)
- [3] DIPLib, a MATLAB toolbox for scientific image processing and analysis. <http://www.diplib.org/>. (Cited on page 149.)
- [4] 3D-Slicer. Tool for visualization, registration and segmentation, <http://www.slicer.org>. (Cited on page 16.)
- [5] D. Abramson, R. Sosis, J. Giddy, and B. Hall. Nimrod: A tool for performing parameterised simulations using distributed workstations. In *Proc of the Fourth IEEE Int Symp on High Performance Distributed Computing*, pages 122 – 131, 1995. (Cited on page 149.)
- [6] I. Agartz, J. L. R. Andersson, and S. Skare. Abnormal brain white matter in schizophrenia: a diffusion tensor imaging study. *Neuroreport*, 12:2251–2254, 2001. (Cited on pages 14 and 26.)
- [7] Sudesh Agrawal, Jack Dongarra, Keith Seymour, and Sathish Vadhiyar. Netsolve: past, present, and future; a look at a grid enabled server. (Cited on page 149.)
- [8] S. Aja-Fernandez, M. Niethammer, M. Kubicki, M.E. Shenton, and Carl-Fredrik Westin. Restoration of DWI data using a rician LMMSE estimator. *IEEE Trans Med Im*, 27:1389–1403, 2008. (Cited on pages 65, 79, and 133.)
- [9] E. M. Akkerman. Efficient measurement and calculation of MR diffusion anisotropy images using the platonic variance method. *Magn Reson Med*, 49:599–604, 2003. (Cited on pages 15, 32, 40, 60, 101, and 133.)
- [10] A. L. Alexander, K. Hasan, et al. A geometric analysis of diffusion tensor measurements of the human brain. *Magn Reson Med*, 44:283–291, 2000. (Cited on page 90.)
- [11] A.L. Alexander, K. Hasan, et al. A geometric analysis of diffusion tensor measurements of the human brain. *Magn Reson Med*, 44:283–291, 2000. (Cited on page 17.)
- [12] D. C. Alexander. A general framework for experiment design in diffusion mri and its application in measuring direct tissue-microstructure features. *Magn Reson Med*, 60(2):439–48, 2008. (Cited on page 66.)
- [13] D.C. Alexander and G.J. Barker. Optimal imaging parameters for fiber-orientation estimation in diffusion MRI. *Neuroimage*, 27(2):357–367, 2005. (Cited on pages 48, 64, 66, 81, and 82.)
- [14] D.C. Alexander, C. Pierpaoli, et al. Spatial transformations of diffusion tensor magnetic resonance images. *IEEE Trans. Med. Imaging*, 20(11):1131–1139, 2001. (Cited on page 90.)
- [15] Cesare Alippi and Pietro Braione. Classification methods and inductive learning rules: What we may learn from theory. *IEEE Trans Systems Man Cybernetics*, 36:649–655, 2006. (Cited on page 114.)
- [16] American Psychiatric Association (APA). *DSM-IV: Diagnostic and statistic manual of mental disorders (fourth ed.)*. APA, Washington DC, 1994. (Cited on pages 15, 32, and 40.)
- [17] Adam W. Anderson. Theoretical analysis of the effects of noise on diffusion tensor imaging. *Magnetic Resonance in Medicine*, 46:1174–1188, 2001. (Cited on page 49.)
- [18] J. L. Andersson. Maximum a posteriori estimation of diffusion tensor parameters using a rician noise model: Why, how and but. *Neuroimage*, 42:1340–56, 2008. (Cited on pages 65 and 71.)
- [19] M. Annett. A classification of hand preference by association analysis. *Br J Psychol*, 61(3):303–321, 1970. (Cited on pages 15, 32, and 40.)
- [20] Applied Linguistics. Brain lobes image after Gray's Anatomy of the Human Body. <http://danama.org/images>. (Cited on page 2.)
- [21] B.A. Ardekani, J. Nierenberg, et al. MRI study of white matter diffusion anisotropy in schizophrenia. *Neuroreport*, 14:2025–2029, 2003. (Cited on pages 14 and 26.)
- [22] S. Ardekani and U. Sinha. Statistical representation of mean diffusivity and fractional anisotropy brain maps of normal subjects. *JMRI*, 24:1243–51, 2006. (Cited on page 56.)
- [23] V. Arsigny, P. Fillard, et al. Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic Resonance in Medicine*, 56(2):411–421, 2006. (Cited on page 97.)
- [24] J. Ashburner and K.J. Friston. Nonlinear spatial normalization using basis functions. *Human Brain Mapping*, 7:254–266, 1999. (Cited on pages 104 and 134.)
- [25] Y. Assaf and P.J. Basser. Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain. *NeuroImage*, 27:48–58, 2005. (Cited on pages 66 and 76.)
- [26] Y. Assaf, D. Ben-Bashat, J. Chapman, S. Peled, I.E. Biton, M. Kafri, Y. Segev, T. Hendler, A.D. Korczyn, M. Graif, and Y. Cohen. High b-value q-space analyzed diffusion-weighted MRI: Application to Multiple Sclerosis. *Magn Reson Med*, 47:115–126, 2002. (Cited on page 66.)
- [27] Yaniv Assaf, Raisa Z. Freidlin, Gustavo K. Rohde, and Peter J. Basser. New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. *Magnetic Resonance in Medicine*, 52:965–978, 2004. (Cited on page 48.)
- [28] H. E. Assemlal, D. Tschumperle, and L. Brun. Efficient and robust computation of pdf features from diffusion mr signal. *Med Image Anal*, 13(5):715–29, 2009. (Cited on page 81.)
- [29] E. J. Aukema, M. W. A. Caan, N. Oudhuis, C. B. L. M. Majoie, F. M. Vos, L. Reneman, B. F. Last, M. A. Grootenhuys, and A. Y. N. Schouten-van Meeteren. White Matter Fractional Anisotropy Correlates With Speed of Processing and Motor Speed in Young Childhood Cancer Survivors. *International Journal of Radiation Oncology, Biology, Physics*, 74(3):837–843, 2009. (Cited on page 119.)

- [30] S. Bagnasco, F. Beltrame, et al. Early diagnosis of alzheimer's disease using a grid implementation of statistical parametric analysis. In *Proceedings of Healthgrid*, pages 69 – 81, 2006. (Cited on page 146.)
- [31] N. Barnea-Goraly, V. Menon, M. Eckert, et al. White matter development during childhood and adolescence: a cross-sectional diffusion tensor imaging study. *Cereb Cortex*, 15:1848–54, 2005. (Cited on page 120.)
- [32] P. J. Basser and C. Pierpaoli. Microstructural and psychological features of tissues elucidated by quantitative-diffusion-tensor MRI. *J Magn Reson B*, 111:209–219, 1996. (Cited on pages 4, 16, 32, 40, 41, 67, 92, 110, 120, 127, and 161.)
- [33] P. J. Basser, J. Mattiello, and D. Le Bihan. Estimation of the effective self-diffusion tensor from the NMR spin echo. *J Magn Reson*, B(103):247–254, 1994. (Cited on pages 14, 36, 56, and 90.)
- [34] P. J. Basser, J. Mattiello, and D. Le Bihan. MR diffusion tensor spectroscopy and imaging. *Biophys J*, 66:259–267, 1994. (Cited on pages 14, 28, 36, and 90.)
- [35] P.J. Basser. Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. *NMR Biomed*, 8:333–334, 1995. (Cited on page 48.)
- [36] P.J. Basser and S.P. Pajevic. Statistical artifacts in diffusion tensor MRI (DT-MRI) caused by background noise. *Magnetic Resonance in Medicine*, 44:41–50, 2000. (Cited on page 93.)
- [37] P.J. Basser, J. Mattiello, and D. Le Bihan. Estimation of the effective self-diffusion tensor from the nmr spin echo. *Journal of Magnetic Resonance*, B(103):247–254, 1994. (Cited on page 48.)
- [38] P.J. Basser, J. Mattiello, and D. Le Bihan. Estimation of the effective self-diffusion tensor from the NMR spin echo. *J Magn Reson*, B(103):247–254, 1994. (Cited on page 146.)
- [39] P.J. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi. In vivo fiber tractography using DT-MRI data. *Magn Reson Med*, 44:625–632, 2000. (Cited on pages 4 and 64.)
- [40] S. Basu, T. Fletcher, and R. Whitaker. Rician noise removal in diffusion tensor MRI. In *Proc. Miccai '06*, LNCS 4190:117–125, 2006. (Cited on page 65.)
- [41] T.E.J. Behrens, H.J. Berg, S. Jbabdi, M.F.S. Rushworth, and M.W. Woolricha. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage*, 34:144–155, 2007. (Cited on pages 7, 66, 81, and 82.)
- [42] D. Ben Bashat, L. Ben Sira, M. Graif, et al. Normal white matter development from infancy to adulthood: comparing diffusion tensor and high b value diffusion weighted MR images. *J Magn Reson Imaging*, 21:503–511, 2005. (Cited on page 128.)
- [43] Hareesh S. Bhatt, V. H. Patel, and A. K. Aggarwal. Web enabled client-server model for development environment of distributed image processing. *LNCS 1971, GRID 2000 R. Buyya and M. Baker (Eds.:)*, pages 135 – 145, 2000. (Cited on page 150.)
- [44] D. Le Bihan. Looking into the functional architecture of the brain with diffusion MRI. *Nature Reviews Neuroscience*, 4:469–480, 2003. (Cited on pages 3 and 6.)
- [45] Y. Bito, S. Hirata, and E. Yamamoto. Optimal gradient factors for ADC measurements. *Proc ISMRM*, page 913, 1995. (Cited on page 66.)
- [46] J. Blaaas, C.P. Botha, B.D. Peters, F.M. Vos, and F.H. Post. Fast and reproducible fiber bundle selection in dti visualization. *Proc. IEEE Visualization*, pages 59–64, 2005. (Cited on pages 48 and 52.)
- [47] C. R. Blain, V. C. Williams, C. Johnston, B. R. Stanton, J. Ganesalingam, J. M. Jarosz, D. K. Jones, G. J. Barker, S. C. Williams, N. P. Leigh, and A. Simmons. A longitudinal study of diffusion tensor MRI in ALS. *Amyotroph.Lateral.Scler.*, pages 1–8, 2007. (Cited on page 140.)
- [48] F.L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE PAMI*, 11:567–585, 1989. (Cited on page 94.)
- [49] James Henry Breasted. *The Edwin Smith Surgical Papyrus*. University of Chicago Press, Chicago, 1930. (Cited on page 2.)
- [50] B. R. Brooks, R. G. Miller, M. Swash, and T. L. Munsat. El escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotroph.Lateral.Scler.Other Motor Neuron Disord.*, 1:293–299, 2000. (Cited on page 132.)
- [51] F. Brugman, J. H. Wokke, H. Scheffer, M. H. Versteeg, E. A. Sistermans, and L. H. Van den Berg. Spastin mutations in sporadic adult-onset upper motor neuron syndromes. *Ann.Neurol.*, 58:865–69, 2005a. (Cited on page 132.)
- [52] F. Brugman, J. H. Wokke, Vianney de Jong, H. J. M., Franssen, C. G. Faber, and L. H. Van den Berg. Primary lateral sclerosis as a phenotypic manifestation of familial als. *Neurology*, 64:1778–79, 2005b. (Cited on page 132.)
- [53] F. Brugman, H. Scheffer, Nillesen W. M. Wokke, J. H., E. Aronica, J. H. Veldink, Van den Berg, and L. H. Paraplegin mutations in sporadic adult-onset upper motor neuron syndromes. *Neurology*, 71:1500–1505, 2008. (Cited on page 132.)
- [54] M. D. Budde, M. Xie, A. H. Cross, and S. K. Song. Axial diffusivity is the primary correlate of axonal injury in the experimental autoimmune encephalomyelitis spinal cord: a quantitative pixelwise analysis. *J Neurosci*, 29:2805–13, 2009. (Cited on page 64.)
- [55] A. I. Buizer, L. M. J. de Sonnevill, M. M. Heuvel-Eibrink, et al. Chemotherapy and attentional dysfunction in survivors of childhood Acute Lymphoblastic Leukemia: effect of treatment intensity. *Pediatr Blood Cancer*, 45:281–290, 2005. (Cited on page 121.)
- [56] A. I. Buizer, L. M. J. de Sonnevill, M. M. Heuvel-Eibrink, et al. Visuomotor control in survivors of childhood Acute Lymphoblastic Leukemia treated with chemotherapy only. *J Int Neuropsychol Soc*, 11:554–565, 2005. (Cited on page 121.)
- [57] Daniel Bunford-Jones, Omer F. Rana, David W. Walker, MatthewAddis, Mike Surridge, and Ken Hawick. Resource discovery for dynamic clusters in computational grids. In *Proceedings of The 15th International Parallel and Distributed Processing Symposium*, pages 759 – 767, 2001. (Cited on page 147.)

- [58] R. W. Butler and J. K. Haser. Neurocognitive effects of treatment for childhood cancer. *Ment Retard Dev Disabil Res Rev*, 12:184–191, 2006. (Cited on page 120.)
- [59] R. W. Butler and R. K. Mulhern. Neurocognitive interventions for children and adolescents surviving cancer. *J Pediatr Psychol*, 30:65–78, 2005. (Cited on page 120.)
- [60] Rajkumar Buyya, Manzur Murshed, David Abramson, and Srikumar Venugopal. Scheduling parameter sweep applications on global grids: a deadline and budget constrained cost-time optimization algorithm. *Software - Practice And Experience*, 35:491 – 512, 2003. (Cited on page 150.)
- [61] M. W. A. Caan, A. W. De Vries, H. G. Khedoe, E. M. Akkerman, L. J. Van Vliet, C. A. Grimbergen, and F. M. Vos. Generating fiber crossing phantoms out of experimental DWIs. *Lecture Notes in Computer Science*, 4791:169–176, 2007. (Cited on page 47.)
- [62] M. W. A. Caan, K. A. Vermeer, L. J. Van Vliet, C. A. Grimbergen, and F. M. Vos. Pruning Datasets in Discriminant Analysis: a DTI Study to Schizophrenia. In *4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007*, pages 1340–1344, 2007. (Cited on page 27.)
- [63] M. W. A. Caan, L. J. van Vliet, C. B. L. M. Majoie, E. J. Aukema, C. A. Grimbergen, and F. M. Vos. Spatial Consistency in 3D Tract-Based Clustering Statistics. In *Proceedings of the 11th international conference on Medical Image Computing and Computer-Assisted Intervention-Part I*, pages 542–549. Springer, 2008. (Cited on pages 109 and 128.)
- [64] M. W. A. Caan, C. A. Sage, M. M. van der Graaf, C. A. Grimbergen, S. Sunaert, L. J van Vliet, and F. M. Vos. Dual tensor atlas generation based on a cohort of coregistered non-HARDI datasets. In *In Proc. MICCAI*, volume LNCS 5761, pages 869–876, 2009. (Cited on pages 55 and 162.)
- [65] M. W. A. Caan, F. M. Vos, L. J. van Vliet, A. H. C. van Kampen, and S. D. Olabarriaga. Gridifying a Diffusion Tensor Imaging Pipeline. *Workshop on Challenges for the Application of Grids in Healthcare, CCGrid'2010*, 2010. (Cited on page 159.)
- [66] M.W.A. Caan, K.A. Vermeer, L.J. van Vliet, C.B.L.M. Majoie, B.D. Peters, G.J. den Heeten, and F.M. Vos. Shaving diffusion tensor images in discriminant analysis: A study into schizophrenia. *Medical Image Analysis*, 10:841–849, 2006. (Cited on pages 13, 57, 64, 110, 146, 147, 148, 158, and 162.)
- [67] M.W.A. Caan, L.J. van Vliet, C.B.L.M. Majoie, E.J. Aukema, C.A. Grimbergen, and F.M. Vos. Spatial Consistency in 3D Tract-Based Clustering Statistics. In *MICCAI 2008*, volume LNCS 5241, pages 535–542, 2008. (Cited on page 91.)
- [68] M.W.A. Caan et al. Adaptive noise filtering for accurate and precise diffusion estimation in fiber crossings. *Miccai 2010*, submitted. (Cited on page 79.)
- [69] B. Cambazoglu, O. Sertel, et al. Efficient processing of pathological images using the grid: Computer-aided prognosis of neuroblastoma. In *Proc 5th IEEE workshop on Challenges of large applications in distributed environments*, pages 35 – 41, 2007. (Cited on page 146.)
- [70] L. K. Campbell, M. Scaduto, D. Van Slyke, et al. A meta-analysis of the neurocognitive sequelae of treatment for childhood Acute Lymphocytic Leukemia. *Pediatr Blood Cancer*, 49:65–73, 2007. (Cited on page 120.)
- [71] Y. Cao, M. Miller, R.L. Winslow, and L. Younes. Large deformation diffeomorphic metric mapping of vector fields. *IEEE Trans. on Medical Imaging*, 24:1216–1230, 2005. (Cited on page 90.)
- [72] Y. Cao, M. Miller, S. Mori, R.L. Winslow, and L. Younes. Diffeomorphic matching of diffusion tensor images. In *proceedings of the MMBIA*, 2006. (Cited on page 90.)
- [73] Y. Cao, M. Miller, S. Mori, R.L. Winslow, and L. Younes. Diffeomorphic matching of diffusion tensor images. In *proc. CVPR*, pages 67–76, 2006. (Cited on page 57.)
- [74] J.D. Carew, C.G. Koay, G. Wahba, A.L. Alexander, M.E. Meyerand, and P.J. Basser. The Asymptotic Behavior of the Nonlinear Estimators of the Diffusion Tensor and Tensor-Derived Quantities with Implications for Group Analysis. Technical Report 1132, University of Wisconsin, 2006. (Cited on page 73.)
- [75] S. Caton, M. Caan, S. Olabarriaga, O. Rana, and B. Batchelor. Using Dynamic Condor-based Services for Classifying Schizophrenia in Diffusion Tensor Images. In *Proceedings of the 2008 Eighth IEEE International Symposium on Cluster Computing and the Grid*, pages 234–241. IEEE Computer Society, 2008. (Cited on pages 145, 164, 167, and 168.)
- [76] Simon Caton, Omer Rana, and Bruce Batchelor. Dynamic condor-based services for distributed image analysis. In *CCGRID '07: Proceedings of the Seventh IEEE International Symposium on Cluster Computing and the Grid*, pages 49 – 56, 2007. (Cited on pages 147, 151, and 153.)
- [77] A. Ceccarelli, M. A. Rocca, P. Valsasina, M. Rodegher, E. Pagani, A. Falini, G. Comi, and M. Filippi. A multiparametric evaluation of regional brain damage in patients with primary progressive multiple sclerosis. *Hum Brain Mapp.*, 30:3009–19, 2009. (Cited on page 6.)
- [78] J. M. Cedarbaum, N. Stambler, E. Malta, C. Fuller, D. Hilt, B. Thurmond, and A. Nakanishi. The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III). *J.Neurol.Sci.*, 169:13–21, 1999. (Cited on page 133.)
- [79] L.-F. Chen, H.-Y. Liao, et al. A new LDA-based face recognition system which can solve the small sample size problem. *Pattern Recognition*, 33:1713–1726, 2000. (Cited on page 37.)
- [80] Ying Chen and Suan Fong Tan. Matlab\**g*: A grid-based parallel matlab. URL <http://hdl.handle.net/1721.1/3863>. (Cited on page 146.)
- [81] S. M. Chou and F. H. Norris. Amyotrophic lateral sclerosis: lower motor neuron disease spreading to upper motor neurons. *Muscl. Nerve*, 16:864–869, 1993. (Cited on page 132.)
- [82] Ron Choy and Alan Edelman. Parallel MATLAB: Doing it right. *Proceedings of the IEEE*, VOL. 93, NO. 2: 331 – 341, 2005. (Cited on pages 147 and 149.)
- [83] H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89:114–141, 2003. (Cited on page 96.)
- [84] H. Chui, A. Rangarajan, J. Zhang, and C.M. Leonard. Unsupervised learning of an atlas from unlabeled point-sets. *IEEE Trans. on Pattern analysis and Machine Intelligence*, 26:160–172, 2004. (Cited on pages 91, 93, 94, 96, and 111.)

- [85] O. Ciccarelli, T. E. Behrens, H. Johansen-Berg, K. Talbot, R. W. Orrell, R. S. Howard, R. G. Nunes, D. H. Miller, P. M. Matthews, A. J. Thompson, and S. M. Smith. Investigation of white matter pathology in ALS and PLS using tract-based spatial statistics. *Hum.Brai. Mapp.*, 30:615–624, 2009. (Cited on pages 139 and 140.)
- [86] V. Cima, G. Logroscino, C. D’Ascenzo, A. Palmieri, M. Volpe, C. Briani, E. Pegoraro, C. Angelini, and G. Soraru. Epidemiology of ALS in Padova district, Italy, from 1992 to 2005. *Eur.J.Neurol*, 16:920–924, 2009. (Cited on page 139.)
- [87] J. Cohen. *Statistical power analysis for the behavioral sciences*. Academy Press, New York. (Cited on page 124.)
- [88] T.F. Coleman and Y. Li. An interior trust region approach for nonlinear minimization subject to bounds. *SIAM J Optim*, 6:418–445, 1996. (Cited on page 52.)
- [89] I. Corouge, P.T. Fletcher, S. Joshi, S. Gouttard, and G. Gerig. Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis. *Med Im Anal*, 10:786–798, 2006. (Cited on pages 7, 48, 90, and 110.)
- [90] M. Cosottini, M. Giannelli, G. Siciliano, G. Lazzarotti, M. C. Michelassi, C. A. Del, C. Bartolozzi, and L. Murri. Diffusion-tensor MR imaging of corticospinal tract in amyotrophic lateral sclerosis and progressive muscular atrophy. *Radiology*, 237:258–264, 2005. (Cited on page 141.)
- [91] A. Czaplinski, A. A. Yen, and S. H. Appel. Amyotrophic lateral sclerosis: early predictors of prolonged survival. *J.Neurol.*, 253:1428–1436, 2006. (Cited on pages 6, 132, and 139.)
- [92] A. Czaplinski, A. A. Yen, E. P. Simpson, and S. H. Appel. Slower disease progression and prolonged survival in contemporary patients with amyotrophic lateral sclerosis: is the natural history of amyotrophic lateral sclerosis changing. *Arch.Neurol.*, 63:1139–1143, 2006. (Cited on page 139.)
- [93] Pawel Czarnul, Andrzej Ciereszko, and Marcin Frcaczak. Towards efficient parallel image processing on cluster grids usinggimp. *M. Bubak et al. (Eds.): ICCS 2004, LNCS 3037*, pages 451 – 458, 2004. (Cited on page 149.)
- [94] B. C. Davis, P. T. Fletcher, E. Bullitt, and S. Joshi. Population Shape Regression From Random Design Data. In *IEEE 11th International Conference on Computer Vision, 2007. ICCV 2007*, pages 1–7, 2007. (Cited on page 176.)
- [95] Ewa Deelman, Tevfik Kosar, Carl Kesselman, and Miron Livny. What makes workflows work in an opportunistic environment? *Concurrency and Computation.: Practice and Experience*, 18:1187 – 1199, 2006. (Cited on pages 149 and 150.)
- [96] A. J. den Dekker and J. Sijbers. *Advanced Image Processing in Magnetic Resonance Imaging*, volume 27 of *Series: Signal Processing and Communications*, chapter 4, pages 85–143. CRC Press, Taylor&Francis Group, LLC, Boca Raton, FL, October 2005. (Cited on pages 65, 68, and 69.)
- [97] A. J. den Dekker, S. van Aert, A. van den Bos, and D. van Dyck. Maximum likelihood estimation of structure parameters from high resolution electron microscopy images. Part I: A theoretical framework. *Ultra-microscopy*, 104:83–106, 2005. (Cited on page 69.)
- [98] M. Descoteaux, R. Deriche, T.R. Knosche, and A. Anwander. Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE Trans Med Im*, 28:269–286, 2009. (Cited on pages 56 and 65.)
- [99] R. Duda and P. Hart. *Pattern Classification*. Wiley, New York, 2001. (Cited on pages 17, 18, 20, 23, 29, 30, 38, and 39.)
- [100] I. Eckstein, D.W. Shattuck, J.L. Stein, K.L. McMahon, G. de Zubicaray, M.J. Wright, P.M. Thompson, and A.W. Toga. Active fibers: matching deformable tract templates to diffusion tensor images. *NeuroImage*, 47:T82–T89, 2009. (Cited on page 90.)
- [101] A. A. Eisen. Comment on the lower motor neuron hypothesis. *Muscl. Nerve*, 16:870–871, 1993. (Cited on page 132.)
- [102] C. M. Ellis, A. Simmons, D. K. Jones, J. Bland, J. M. Dawson, M. A. Horsfield, S. C. Williams, and P.N. Leigh. Diffusion tensor MRI assesses corticospinal tract damage in ALS. *Neurology*, 53:1051–1058, 1999. (Cited on page 133.)
- [103] D. Krefting et al. Simplified implementation of medical image processing algorithms into a grid using a workflow management system. *Future Generation Computer Systems*, 2009. (Cited on page 171.)
- [104] R. P. Feldman and J. T. Goodrich. The Edwin Smith Surgical Papyrus. *Child’s Nerv Syst*, 15:281–284, 1999. (Cited on pages 1 and 2.)
- [105] P. Fillard, X. Pennec, V. Arsigny, and N. Ayache. Clinical DT-MRI estimation, smoothing, and fiber tracking with log-euclidean metrics. *IEEE Transactions on Medical Imaging*, 26:1472–1482, 2007. (Cited on page 65.)
- [106] J. Foong, M. Maier, et al. Neuropathological abnormalities of the corpus callosum in schizophrenia: a diffusion tensor imaging study. *J Neurol Neurosurg Psychiatry*, 68:242–244, 2000. (Cited on pages 14 and 26.)
- [107] R. B. Forbes, S. Colville, G. W. Cran, and R. J. Swingler. Unexpected decline in survival from amyotrophic lateral sclerosis/motor neurone disease. *J.Neurol.Neurosurg.Psychiatry*, 75:1753–1755, 2004. (Cited on page 139.)
- [108] I. Foster, C. Kesselman, and S. Tuecke. The Anatomy of the Grid: Enabling Scalable Virtual Organizations. *Int. J. of Supercomputer Applications*, 15, 2001. (Cited on pages 160 and 176.)
- [109] Lawrence R. Frank. Anisotropy in high angular resolution diffusion-weighted mri. *Magn. Reson. Med.*, 45: 935–939, 2001. (Cited on pages 48, 65, and 81.)
- [110] Lawrence R. Frank. Characterization of anisotropy in high angular resolution diffusion-weighted mri. *Magn. Reson. Med.*, 47:1083–1099, 2002. (Cited on page 48.)
- [111] K.J. Friston, A.P. Holmes, et al. Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp*, 2:189, 1995. (Cited on pages 22, 23, and 176.)

- [112] J. Fritschy, L. Horesh, et al. Using the GRID to improve the computation speed of electrical impedance tomography (EIT) reconstruction algorithms. *Physiol. Meas.*, 26:209–215, 2005. (Cited on pages 146 and 171.)
- [113] K. Fukunaga. *Introduction to statistical Pattern Recognition*. Academic Press, New York, 1990. (Cited on pages 17, 29, and 37.)
- [114] R. Gardner and M. Broman. The Purdue Pegboard. Normative data on 1334 School Children. *J of Clin Neuropsychol*, 1:156–162, 1979. (Cited on pages 122 and 123.)
- [115] G. Gerig, S. Gouttard, and I. Corouge. Analysis of brain white matter via fiber tract modeling. *Engineering in Medicine and Biology Society*, 426:4421–4424, 2004. (Cited on page 110.)
- [116] J. N. Giedd, J. Blumenthal, N. O. Jeffries, et al. Development of the human corpus callosum during childhood and adolescence: a longitudinal MRI study. *Prog Neuropsychopharmacol Biol Psychiatry*, 23:571–588, 1999. (Cited on pages 120 and 128.)
- [117] Tristan Glatard, Johan Montagnat, Diane Lingrand, and Xavier Pennec. Flexible and efficient workflow deployment of data-intensive applications on grids with MOTEUR. *International Journal of High Performance Computing and Applications (IJHPCA)*, 22(3):347–360, August 2008. (Cited on page 164.)
- [118] C.B. Goodlett, P.T. Fletcher, J.H. Gilmore, and G. Gerig. Group analysis of DTI fiber tract statistics with application to neurodevelopment. *NeuroImage*, 45:S133–S142, 2009. (Cited on page 56.)
- [119] P. H. Gordon, B. Cheng, I. B. Katz, M. Pinto, A. P. Hays, H. Mitsumoto, and L. P. Rowland. The natural history of primary lateral sclerosis. *Neurology*, 66:647–653, 2006. (Cited on page 132.)
- [120] S. Gouttard, M. Prastawa, E. Bullitt, W. Lin, C. Goodlett, and G. Gerig. Constrained Data Decomposition and Regression for Analyzing Healthy Aging from Fiber Tract Diffusion Properties. *MICCAI 2009*, pages 321–328. (Cited on page 176.)
- [121] J. M. Graham, N. Papadakis, J. Evans, E. Widjaja, C. A. Romanowski, M. Paley, L. I. Wallis, I. D. Wilkinson, P. J. Shaw, and P. D. Griffiths. Diffusion tensor imaging for the assessment of upper motor neuron integrity in ALS. *Neurology*, 63:2111–2119, 2004. (Cited on page 141.)
- [122] M. D. Greicius, K. Supekar, V. Menon, and R. F. Dougherty. Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cerebral Cortex*, 2008. (Cited on page 177.)
- [123] H. Gudbjartsson and S. Patz. The Rician distribution of noise MRI data. *Magn Res Med*, pages 910–914, 1995. (Cited on pages 50 and 67.)
- [124] A. Guimond and A. Roche. Multimodal brain warping using the demons algorithm and adaptive intensity corrections. *Institut National de Recherche en Informatique et en Automatique, Sophia Antipolis, France*, 1999. (Cited on page 90.)
- [125] T. Hastie, R. Tibshirani, et al. 'Gene shaving' as a method for identifying distinct sets of genes with similar expression patterns. *Genome Biology*, 1(2):research0003.1–0003.21, 2000. (Cited on pages 14 and 19.)
- [126] W. Van Hecke, A. Leemans, E. D'Agostino, S. de Backer, E. Vandervliet, P.M. Parizel, and J. Sijbers. Nonrigid coregistration of diffusion tensor images using a viscous fluid model and mutual information. *IEEE TMI*, 26(11):1598–1612, 2007. (Cited on page 90.)
- [127] B. Hermann, R. Hansen, et al. Neurodevelopmental vulnerability of the corpus callosum to childhood onset localization-related epilepsy. *NeuroImage*, 18:284–292, 2003. (Cited on pages 23 and 34.)
- [128] T. Hosey, G. Williams, and R. Ansoerge. Inference of multiple fiber orientations in high angular resolution diffusion imaging. *Magn Reson Med*, 54:1480–89, 2005. (Cited on pages 64, 66, 67, 73, and 81.)
- [129] T. P. Hosey, S. G. Harding, T. A. Carpenter, R. E. Ansoerge, and G. B. Williams. Application of a probabilistic double-fibre structure model to diffusion-weighted MR images of the human brain. *Magn Reson Imaging*, 26:236–45, 2008. (Cited on page 82.)
- [130] C.Q. Howard, C.H. Hansen, and A.C. Zander. Optimisation of design and location of acoustic and vibration absorbers using a distributed computing network. In *Proc. of ACOUSTICS*, pages 173–178, 2005. (Cited on pages 146 and 150.)
- [131] R. Huang, Q. Liu, et al. Solving the small sample size problem of LDA. *Proc Int Conf Pattern Recognition*, 3:29–32, 2002. (Cited on pages 17 and 29.)
- [132] D. Hubl, T. Koenig, et al. Pathways that make voices. *Arch Gen Psychiatry*, 61:658–668, 2004. (Cited on pages 14 and 26.)
- [133] S. Hunsche, M.E. Mosely, et al. Diffusion-tensor MR imaging at 1.5 and 3.0 T: Initial observations. (Cited on page 21.)
- [134] P. G. Ince, J. Evans, M. Knopp, G. Forster, H. H. Hamdalla, S. B. Wharton, and P. J. Shaw. Corticospinal tract degeneration in the progressive muscular atrophy variant of ALS. *Neurology*, 60:1252–1258, 2003. (Cited on page 132.)
- [135] M. Inglese, S. Makani, G. Johnson, et al. Diffuse axonal injury in mild traumatic brain injury: a diffusion tensor imaging study. *J Neurosurg*, 103:298–303, 2005. (Cited on page 128.)
- [136] L. Iuvone, P. Mariotti, C. Colosimo, et al. Long-term cognitive outcome, brain computed tomography scan, and magnetic resonance imaging in children cured for Acute Lymphoblastic Leukemia. *Cancer*, 95:2562–70, 2002. (Cited on page 128.)
- [137] N. K. Iwata, S. Aoki, S. Okabe, N. Arai, Y. Terao, S. Kwak, O. Abe, I. Kanazawa, S. Tsuji, and Y. Ugawa. Evaluation of corticospinal tracts in ALS with diffusion tensor MRI and brainstem stimulation. *Neurology*, 70:528–532, 2008. (Cited on page 132.)
- [138] B. J. Jellison, A. S. Field, J. Medow, M. Lazarand, M. S. Salamat, and A.L. Alexander. Diffusion tensor imaging of cerebral white matter: A pictorial review of physics, fiber tract anatomy, and tumor imaging patterns. (Cited on page 3.)
- [139] H. Jiang, P. C. van Zijl, J. Kim, G. D. Pearlson, and S. Mori. DtiStudio: resource program for diffusion tensor computation and fiber bundle tracking. *Comput.Methods Program. Biomed.*, 81:106–116, 2006. (Cited on page 134.)



- [140] H. Jiang, P.C.M. van Zijl, et al. DtiStudio: Resource program for diffusion tensor computation and fiber bundle tracking. *Computer Methods and Programs in Biomedicine*, 81:106–116, 2006. (Cited on pages 93 and 111.)
- [141] D. K. Jones and P. J. Basser. "Squashing peanuts and smashing pumpkins": How noise distorts diffusion-weighted MR data. *Magnetic Resonance in Medicine*, 52(5):979–993, 2004. (Cited on pages 66 and 82.)
- [142] D.K. Jones. Determining and visualizing uncertainty in estimates of fiber orientation from Diffusion Tensor MRI. *Magnetic Resonance in Medicine*, 49:7–12, 2003. (Cited on page 92.)
- [143] D.K. Jones, M.A. Horsfield, and A. Simmons. Optimal strategies for measuring diffusion in anisotropic systems by Magnetic Resonance Imaging. *Magn Reson Med*, 42:515–525, 1999. (Cited on page 66.)
- [144] D.K. Jones, L.D. Griffin, D.C. Alexander, M. Catani, M.A. Horsfield, R. Howard, and S.C.R. Williams. Spatial normalization and averaging of diffusion tensor MRI data sets. *NeuroImage*, 17:592–617, 2002. (Cited on page 90.)
- [145] Z. Jun and Y. Umetani. A problem solving environment for automatic matlab 3d finite element code generation and simplified grid computing. In *Proceedings of the Second IEEE International Conference on e-Science and Grid Computing (e-Science'06)*, pages 99 – 104, 2006. (Cited on page 146.)
- [146] V. Kalogeraki, P. M. Melliari-Smith, and L. E. Moser. Using multiple feedback loops for object profiling, scheduling and migration in soft real-time distributed object systems. *IEEE Int Symp on Object-Oriented Real-Time Distributed Computing*, pages 291 – 300, 1999. (Cited on page 149.)
- [147] R. A. Kanaan, J. S. Kim, et al. Diffusion tensor imaging in schizophrenia. *Biol Psychiatry*, 58(12):921–929, 2005. (Cited on pages 14, 16, 26, 28, 34, 146, and 173.)
- [148] Ole T. Karlsen, Rieko Verhagen, and Wim M.M.J. Bovée. Definiton of displacement probability and diffusion time in q-space magnetic resonance measurements that use finite-duration diffusion encoding gradients. *Magnetic Resonance in Medicine*, 165:185–195, 2003. (Cited on page 70.)
- [149] J. A. Kent-Braun, C. H. Walker, M. W. Weiner, and R. G. Miller. Functional significance of upper and lower motor neuron impairment in amyotrophic lateral sclerosis. *Muscl. Nerve*, 21:762–768, 1998. (Cited on page 133.)
- [150] P. L. Khong, D. L. W. Kwong, G. C. F. Chan, et al. Diffusion-tensor imaging for the detection and quantification of treatment-induced white matter injury in children with Medulloblastoma: A Pilot Study. *Am J Neuroradiol*, 24:734–740, 2003. (Cited on page 120.)
- [151] P. L. Khong, L. H. T. Leung, G. C. F. Chan, et al. White matter anisotropy in childhood medulloblastoma survivors: association with neurotoxicity risk factors. *Radiology*, 236:647–652, 2005. (Cited on page 128.)
- [152] P. L. Khong, L. H. T. Leung, A. S. M. Fung, et al. White matter anisotropy in post-treatment childhood cancer survivors: preliminary evidence of association with neurocognitive function. *J Clin Oncol*, 24:884–890, 2006. (Cited on page 120.)
- [153] W. K. Kim, X. Liu, J. Sandner, et al. Study of 962 patients indicates progressive muscular atrophy is a form of ALS. *Neurology*, 73:1686–1692, 2009. (Cited on page 132.)
- [154] G. Kindlmann. Teem. <http://teem.sourceforge.net>. (Cited on page 92.)
- [155] P.B. Kingsley and W.G. Monahan. Selection of the optimum b-factor for diffusion-weighted magnetic resonance imaging assessment of ischemic stroke. *Magn. Reson. Med.*, 51:996–1001, 2004. (Cited on page 66.)
- [156] Stephen D. Kleban and Scott H. Clearwater. Fair share on high performance computing systems: What does fair really mean? In *Proceedings of the 3rd IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGRID'03)*, pages 146 – 153, 2003. (Cited on page 155.)
- [157] C. G. Koay, L. C. Chang, C. Pierpaoli, and P. J. Basser. Error propagation framework for diffusion tensor imaging via diffusion tensor representations. *IEEE Trans Med Im*, 26(8):1017–1034, 2007. (Cited on page 65.)
- [158] W. Kort, M. Schittekatte, and P. H. Dekker. WISC-III, 3rd ed. Wechsler D. Handleiding en verantwoording (technical manual). *Amsterdam: Harcourt Test Publishers*, 2005. (Cited on pages 122 and 123.)
- [159] B.W. Kreher, J.F. Schneider, et al. Multitensor approach for analysis and tracking of complex fiber configurations. *Magn Res Med*, 54:1216–1225, 2005. (Cited on pages 48, 56, 65, and 81.)
- [160] M. D. Kubicki, C. F. Westin, et al. Cingulate fasciculus integrity disruption in schizophrenia: a magnetic resonance diffusion tensor imaging study. *Biol Psychiatry*, 54:1171–1180, 2003. (Cited on page 14.)
- [161] B. Landman, P. L. Bazin, and J. Prince. Diffusion tensor estimation by maximizing Rician likelihood. *IEEE 11th ICCV*, pages 2433–40, 2007. (Cited on page 65.)
- [162] Adam Lathers, Mei-Hui Su, Alex Kulungowski, Abel W. Lin, Gaurang Mehta, Steven T. Peltier, Ewa Deelman, and Mark H. Ellisman. Enabling parallel scientific applications with workflow tools. *IEEE Challenges of Large Applications in Distributed Environments*, pages 55–60, 2006. (Cited on pages 146, 147, and 171.)
- [163] D. LeBihan, E. Breton, D. Lallemand, P. Grenier, E. Cabanis, and M. Laval-Jeantet. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology*, 161:401–407, 1986. (Cited on page 66.)
- [164] A. Leemans, J. Sijbers, S. De Backer, E. Vandervliet, and P. Parizel. Multiscale white matter fiber tract coregistration: A new feature-based approach to align diffusion tensor data. *Magnetic Resonance in Medicine*, 55: 1414–1423, 2006. (Cited on page 90.)
- [165] Mike Lewis and Andrew Grimshaw. The core legion object model. In *Proceedings of HPDC-5*, pages 551 – 561, 1996. (Cited on page 149.)
- [166] R. Lützkendorf, J. Bernarding, et al. Enabling of grid based diffusion tensor imaging using a workflow implementation of FSL. *Stud Health Technol Inform*, 147:72–81, 2009. (Cited on page 171.)
- [167] D. J. Mabbott, M. Noseworthy, E. Bouffet, et al. White matter growth as a mechanism of cognitive development in children. *NeuroImage*, 33:936–946, 2006. (Cited on pages 123 and 128.)
- [168] M. Maddah, W.M. Wells III, S.K. Warfield, C.-F. Westin, and W.E.L. Grimson. Probabilistic clustering and quantitative analysis of white matter fiber tracts. In *IPMI 2007*, volume 20, pages 372–383, 2007. (Cited on pages 90 and 110.)

- [169] J.B.A. Maintz and M.A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2:1–36, 1998. (Cited on pages 6 and 56.)
- [170] J.G. Malcolm, M.E. Shenton, and Y. Rathi. Neural Tractography Using An Unscented Kalman Filter. *Information Processing in Medical Imaging (IPMI)*, pages 126–138, 2009. (Cited on page 65.)
- [171] J.G. Malcolm, M.E. Shenton, and Y. Rathi. Two-Tensor Tractography Using a Constrained Filter. *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pages 894–902, 2009. (Cited on page 65.)
- [172] J. Mandrioli, P. Faglioni, P. Nichelli, and P. Sola. Amyotrophic lateral sclerosis: prognostic indicators of survival. *Amyotroph.Lateral.Scler.*, 7:211–220, 2006. (Cited on page 132.)
- [173] J.-F. Mangin, C. Poupon, C. Clark, D. Le Bihan, and I. Bloch. Eddy-current distortion correction and robust tensor estimation for mr diffusion imaging. *In Proc. MICCAI 2001*, 2208/2001:186–193, 2001. (Cited on pages 60, 77, 101, 122, and 133.)
- [174] Elias S. Manolakos, Demetris G. Galatopoulos, and Andrew P. Funk. Distributed matlab based signal and image processing using javaports. *In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'04)*, pages 217 – 220, 2004. (Cited on page 147.)
- [175] S. Mika, G. Rätsch, J. Weston, B. Schölkopf, and K.-R. Müller. Fisher discriminant analysis with kernels. In Y.-H. Hu, J. Larsen, E. Wilson, and S. Douglas, editors, *Neural Networks for Signal Processing IX*, pages 41–48. IEEE, 1999. (Cited on pages 38 and 39.)
- [176] H. Mitsumoto, A. M. Ulug, S.L. Pullman, C. L. Gooch, S. Chan, M. X. Tang, X. Mao, A. P. Hays, A. G. Floyd, V. Battista, J. Montes, S. Hayes, S. Dashnaw, P. Kaufmann, P. H. Gordon, J. Hirsch, B. Levin, L. P. Rowland, and D. C. Shungu. Quantitative objective markers for upper and lower motor neuron dysfunction in ALS. *Neurology*, 68:1402–1410, 2007. (Cited on pages 140 and 141.)
- [177] M. Moleski. Neuropsychological, neuroanatomical, and neurophysiological consequences of CNS chemotherapy for acute lymphoblastic leukemia. *Arch Clin Neuropsychol*, 15:603–630, 2000. (Cited on page 120.)
- [178] J. Montagnat, B. Isnard, et al. A data-driven workflow language for grids based on array programming principles. *WORKS'09*, 2009. (Cited on page 165.)
- [179] J. Montagnat et al. Medical images simulation, storage, and processing on the european datagrid testbed. *Journal of Grid Computing*, 2:387–400, 2004. (Cited on page 171.)
- [180] B. D. Moore. Neurocognitive outcomes in survivors of childhood cancer. *J Pediatr Psychol*, 30:51–63, 2005. (Cited on page 120.)
- [181] S. Mori, K. Oishi, et al. Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *Neuroimage*, 40:570–82, 2008. (Cited on pages 6, 56, 58, 134, and 162.)
- [182] M. E. Moseley, Y. Cohen, J. Kucharczyk, J. Mintorovitch, H. S. Asgari, M. F. Wendland, J. Tsuruda, and D. Norman. Diffusion-weighted MR imaging of anisotropic water diffusion in cat central nervous system. *Radiology*, 176:439–445, 1990. (Cited on page 132.)
- [183] R. K. Mulhern, S. L. Palmer, W. E. Reddick, et al. Risks of young age for selected neurocognitive deficits in medulloblastoma are associated with white matter loss. *J Clin Oncol*, 19:472–479, 2001. (Cited on page 120.)
- [184] D. A. Mulrooney, J. P. Neglia, and M. M. Hudson. Caring for adult survivors of childhood cancer. *Curr Treat Options Oncol*, 9:51–66, 2008. (Cited on page 120.)
- [185] Z. Nagy, H. Westerberg, and T. Klingberg. Maturation of white matter is associated with the development of cognitive functions during childhood. *J Cogn Neurosci*, 16:1227–33, 2004. (Cited on page 120.)
- [186] P. C. Nathan, S. K. Patel, K. Dilley, and others. Guidelines for identification of, advocacy for, and intervention in neurocognitive problems in survivors of childhood cancer: a report from the Children's Oncology Group. *Arch Pediatr Adolesc Med*, 161:798–806, 2007. (Cited on page 120.)
- [187] O. O. Brihuega-Moreno, F. P. Heese, and L. D. Hall. Optimization of diffusion measurements using cramer-rao lower bound theory and its application to articular cartilage. *Magn Reson Med*, 50(5):1069–76, 2003. (Cited on page 66.)
- [188] L.J. O'Donnell, C.-F. Westin, and A.J. Golby. Tract-based morphometry. In *MICCAI '07*, volume 4792, pages 161–168, 2007. (Cited on page 110.)
- [189] T. Oinn et al. Taverna: A tool for the composition and enactment of bioinformatics workflows. *Bioinformatics journal*, 17(20):3045–3054, 2004. (Cited on page 165.)
- [190] S.D. Olabarriaga and J. Montagnat. Special section on medical imaging on grids. *Future Generation Computer Systems*, page in press, 2009. (Cited on page 171.)
- [191] S.D. Olabarriaga, P.T. de Boer, et al. Virtual Lab for fMRI: Bridging the Usability Gap. <http://www.vl-e.nl/vbrowser>, 2006. (Cited on page 164.)
- [192] Silvia Olabarriaga, Tristan Glatard, Kamel Boulebar, and Piter T. de Boer. From 'low-hanging' to 'user-ready': initial steps into a healthgrid. In *HealthGrid*, Studies in Health Technology and Informatics, pages 70–79, Chicago, June 2008. IOS Press. (Cited on page 79.)
- [193] Silvia Olabarriaga, Tristan Glatard, and Piter T. de Boer. A virtual laboratory for medical image analysis. *IEEE Transactions on Information Technology In Biomedicine (TITB)*, pages (in-press), 2010. (Cited on page 164.)
- [194] Evren Ozarslan and Thomas H. Mareci. Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. *Magn. Reson. Med.*, 50: 955–965, 2003. (Cited on pages 48, 65, and 81.)
- [195] Cherri M. Pancake and Curtis Cook. What users need in parallel tool support: Survey results and analysis. *In Proc IEEE Scalable High-Performance Computing Conference*, pages 40 – 47, 1994. (Cited on page 149.)
- [196] H.J. Park, M. Kubicki, M.E. Shenton, A. Guimond, R.W. McCarley, S.E. Maier, R. Kikinis, F.A. Jolesz, and C.-F. Westin. Spatial normalization of diffusion tensor MRI using multiple channels. *NeuroImage*, 20: 1995–2009, 2003. (Cited on page 90.)

- [197] O. Pasternak, Y. Assaf, N. Intrator, and N. Sochen. Variational multiple-tensor fitting of fiber-ambiguous diffusion-weighted magnetic resonance imaging voxels. *Magn Reson Im*, 26:1133–44, 2008. (Cited on page 81.)
- [198] S. Peled. New perspectives on the sources of white matter DTI signal. *IEEE Trans Med Im*, 26(11):1448–55, 2007. (Cited on page 66.)
- [199] S. Peled and C.F. Westin. Geometric extraction of two crossing tracts in dwi. *Proc. Intl. Soc. Mag. Reson. Med*, 13:1, 2005. (Cited on page 50.)
- [200] S. Peled, O. Friman, F. Jolesz, and C. F. Westin. Geometrically constrained two-tensor model for crossing tracts in DWI. *Magnetic Resonance Imaging*, 24(9):1263–1270, 2006. (Cited on page 161.)
- [201] S. Peled, O. Friman, et al. Geometrically constrained two-tensor model for crossing tracts in DWI. *Magn Reson Med*, 24:1263–1270, 2006. (Cited on pages 48, 65, and 81.)
- [202] C. Pierpaoli and P. J. Basser. Toward a quantitative assessment of diffusion anisotropy. *Magn Reson Med*, 36:893–906., 1996. (Cited on pages 132 and 134.)
- [203] C. Pierpaoli and D. K. Jones. Removing CSF Contamination in Brain DT-MRIs by Using a Two-Compartment Tensor Model. *Proc ISMRM*, page 1215, 2005. (Cited on page 70.)
- [204] C. Pierpaoli, P. Jezzard, P. J. Basser, A. Barnett, and G. Di Chrio. Diffusion tensor mr imaging of the human brain. *Radiology*, 201:637–648, 1996. (Cited on pages 51, 60, and 66.)
- [205] V. Prckovska, A. F. Roebroek, Wlpm Pullens, A. Vilanova, and B. M. T. Romeny. Optimal acquisition schemes in high angular resolution diffusion weighted imaging. *Miccai LNCS 5242*, pages 9–17, 2008. (Cited on page 81.)
- [206] A. A. Qazi, A. Radmanesh, L. O’Donnell, G. Kindlmann, S. Peled, S. Whalen, C. F. Westin, and A. J. Golby. Resolving crossings in the corticospinal tract by two-tensor streamline tractography: Method and clinical assessment using fMRI. *Neuroimage*, 2008. (Cited on page 65.)
- [207] D. Qui, D. L. W. Kwong, G. C. F. Chan, et al. Diffusion tensor magnetic resonance imaging finding of discrepant fractional anisotropy between the frontal lobes and parietal lobes after whole-brain irradiation in childhood medulloblastoma survivors; reflection of regional white matter radiosensitivity? *Int J Radiation Oncology Biol Phys*, 69:846–851, 2007. (Cited on page 128.)
- [208] Raihan Ur Rasool and Qingping Guo. A pro-middleware for grid computing. *I. Stojmenovic et al. (Eds.): ISPA 2007, LNCS 4742*, pages 556 – 562, 2007. (Cited on page 146.)
- [209] Raihan Ur Rasool and Guo Qingping. Users-grid matlab plug-in: Enabling matlab for the grid. *IEEE International Conference on e-Business Engineering (ICEBE’06)*, pages 473 – 478, 2006. (Cited on pages 146 and 147.)
- [210] W. E. Reddick, H. A. White, J. O. Glass, et al. Developmental model relating white matter volume to neurocognitive deficits in pediatric brain tumor survivors. *Cancer*, 97:2512–19, 2003. (Cited on page 120.)
- [211] W. E. Reddick, J. O. Glass, S. L. Palmer, et al. Atypical white matter volume development in children following craniospinal irradiation. *J Neurooncol*, 7:12–19, 2005. (Cited on pages 120 and 128.)
- [212] W. E. Reddick, Z. Y. Shan, J. O. Glas, et al. Smaller white-matter volumes are associated with larger deficits in attention and learning among long-term survivors of Acute Lymphoblastic Leukemia. *Cancer*, 106:941–949, 2006. (Cited on page 120.)
- [213] T. G. Reese, O. Heid, et al. Reduction of eddy-current-induced distortion in diffusion MRI using twice-refocused spin echo. *Magn Reson Med*, 49:177–182, 2003. (Cited on pages 15, 32, 40, and 101.)
- [214] G.K. Rohde, S. Pajevic, and C. Pierpaoli. Multi-channel registration of diffusion tensor images using directional information. *In proceedings of 2004 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2004. (Cited on page 90.)
- [215] C. Rønning, K. Sundet, B. Due-Tønnessen, et al. Persistent cognitive dysfunction secondary to cerebellar injury in patients treated for posterior fossa tumors in Childhood. *Pediatr Neurosurg*, 41:15–21, 2005. (Cited on page 128.)
- [216] J. Ruiz-Alzola, C.F. Westin, S.K. Warfield, C. Alberola, S. Maier, and R. Kikinis. Nonrigid registration of 3d tensor medical data. *Medical Image Analysis*, 6(2):143–161, June 2002. (Cited on page 90.)
- [217] M. Sach, G. Winkler, V. Glauche, J. Liepert, B. Heimbach, M. A. Koch, C. Buchel, and C. Weiller. Diffusion tensor MRI of early upper motor neuron involvement in amyotrophic lateral sclerosis. *Brain*, 127:340–350, 2004. (Cited on page 141.)
- [218] C. A. Sage, M. M. van der Graaff, M. W. A. Caan, et al. White matter involvement in phenotypes of MND: a longitudinal diffusion tensor imaging study. *Submitted to Neurology*. (Cited on page 132.)
- [219] C. A. Sage, W. Van Hecke, R. Peeters, J. Sijbers, W. Robberecht, P. Parizel, G. Marchal, A. Leemans, and S. Sunaert. Quantitative diffusion tensor imaging in amyotrophic lateral sclerosis: Revisited. *Hum.Brai. Human Brain Mapping*, 2009. (Cited on page 132.)
- [220] C.A. Sage, R. R. Peeters, A. Gorner, W. Robberecht, and S. Sunaert. Quantitative diffusion tensor imaging in amyotrophic lateral sclerosis. *Neuroimage*., 34:486–499, 2007. (Cited on pages 134 and 140.)
- [221] C.A. Sage, R.R. Peeters, Astrid Gorner, Wim Robberecht, and Stefan Sunaert. Quantitative diffusion tensor imaging in amyotrophic lateral sclerosis. *Neuroimage*, 34:486–499, 2007. (Cited on pages 64, 66, and 110.)
- [222] A. S. D. Saporta, H. T. Chugani, C. Juhasz, M. I. Makki, O. Muzik, B. J. Wilson, and M. E. Behen. Multi-modality Neuroimaging in Tourette Syndrome: Alpha-<sup>11</sup>C Methyl-L-Tryptophan Positron Emission Tomography and Diffusion Tensor Imaging Studies. *Journal of Child Neurology*, page In press, 2009. (Cited on page 177.)
- [223] V. J. Schmithorst, M. Wilke, B. J. Dardzinski, et al. Cognitive functions correlate with white matter architecture in a normal pediatric population: a diffusion tensor MRI study. *Hum Brain Mapp*, 26:139–147, 2005. (Cited on page 120.)
- [224] M. E. Shenton, M. E. Dickey, et al. A review of MRI findings in schizophrenia. *Schizophr Res*, 49:1–52, 2001. (Cited on pages 6, 14, 17, 26, 28, and 36.)



- [225] J. Sijbers and A. J. den Dekker. Maximum Likelihood Estimation of Signal Amplitude and Noise Variance From MR Data. *Magnetic Resonance in Medicine*, 51:586–594, 2004. (Cited on page 69.)
- [226] S.M. Smith, M. Jenkinson, et al. Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *NeuroImage*, 31:1487–1505, 2006. (Cited on page 110.)
- [227] S-K Song, S-W Sun, Won-Kyu Ju, Shio-W-Juan Lin, A. H. Cross, and A. H. Neufeld. Diffusion tensor imaging detects and differentiates axon and myelin degeneration in mouse optic nerve after retinal ischemia. *NeuroImage*, 20:1714–1722, 2003. (Cited on pages 2, 6, 64, 81, and 176.)
- [228] E.O. Stejskal and J.E. Tanner. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *J. Chem. Phys.*, 42:288–292, 1965. (Cited on page 4.)
- [229] Z. Sun, F. Wang, et al. Abnormal anterior cingulum in patients with schizophrenia: a diffusion tensor imaging study. *Neuroreport*, 14:1833–1836, 2003. (Cited on page 14.)
- [230] Douglas Thain, Todd Tannenbaum, and Miron Livny. Distributed computing in practice: The condor experience. *Concurrency and Computation: Practice and Experience*, 17:323–359, 2005. (Cited on page 150.)
- [231] P. Thevenaz, T. Blu, and M. Unser. Interpolation revisited. *IEEE TMI*, 19:739–758, 2000. (Cited on page 97.)
- [232] J.-P. Thirion. Image matching as a diffusion process: an analogy with Maxwell's demons. *Med Image Anal*, 2(3):243–260, 1998. (Cited on pages 16, 33, and 40.)
- [233] C. E. Thomaz, J. P. Boardman, et al. Using a maximum uncertainty l<sub>1</sub>-based approach to classify and analyse mr brain images. In *Proc. MICCAI'04*, LNCS 3216:291–300, 2004. (Cited on pages 28 and 36.)
- [234] G. J. Tortora. *Principals of human anatomy*. Harper&Row publishers, New York, USA, 1983. (Cited on page 1.)
- [235] J. D. Tournier, F. Calamante, and A. Connelly. Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *Neuroimage*, 35(4):1459–1472, 2007. (Cited on pages 65 and 81.)
- [236] J. D. Tournier, C. H. Yeh, F. Calamante, K. H. Cho, A. Connelly, and C. P. Lin. Resolving crossing fibres using constrained spherical deconvolution: Validation using diffusion-weighted imaging phantom data. *Neuroimage*, 42(2):617–625, 2008. (Cited on pages 77 and 176.)
- [237] N. Toussaint, A. van Muiswinkel, F. G. Hoogenraad, R. Holthuisen, and S. Sunaert. Resolving fiber crossings: a two fiber model simulation result. *Proc ISMRM*, 13:1339, 2005. (Cited on pages 59, 79, and 169.)
- [238] David S. Tuch, R.M. Weisskoff, J.W. Belliveau, and V.J. Wedeen. High angular resolution diffusion imaging of the human brain. In *In: Proceedings of the 8th Annual meeting of ISMRM, Philadelphia*, page 321, 1999. (Cited on pages 56 and 64.)
- [239] D.S. Tuch. Diffusion MRI of complex tissue structure. *PhD thesis*, 2002. (Cited on pages 48 and 65.)
- [240] M. Turk and A. Pentland. Eigenfaces for recognition. *J Cognitive Neuroscience*, 3:71–86, 1991. (Cited on pages 17, 29, and 37.)
- [241] A. M. Ulug, T. Grunewald, M. T. Lin, A. K. Kamal, C. G. Filippi, R. D. Zimmerman, and M. F. Beal. Diffusion tensor imaging in the diagnosis of primary lateral sclerosis. *J.Mag. Reson.Imaging*, 19:34–39, 2004. (Cited on page 139.)
- [242] A. van den Bos. *Parameter estimation for scientists and engineers*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2007. (Cited on pages 68 and 69.)
- [243] W. Van Hecke, A. Leemans, E. D'Agostino, S. De Backer, E. Vandervliet, P.M. Parizel, and J. Sijbers. Nonrigid coregistration of diffusion tensor images using a viscous fluid model and mutual information. *IEEE Trans Med Im*, 26(11):1598–1612, November 2007. (Cited on pages 57 and 58.)
- [244] W. van Hecke, J. Sijbers, E. D'Agostino, F. Maes, S. De Backer, E. Vandervliet, P.M. Parizel, and A. Leemans. On the construction of an inter-subject diffusion tensor magnetic resonance atlas of the healthy human brain. *NeuroImage*, 43:69–80, 2008. (Cited on pages 56, 58, and 162.)
- [245] M.W. Vernooij, M. De Groot, A. van der Lugt A, M.A. Ikram, G.P. Krestin, A. Hofman, W.J. Niessen, and M.M. Breteler. White matter atrophy and lesion formation explain the loss of structural integrity of white matter in aging. *NeuroImage*, 43:470–7, 2008. (Cited on pages 165 and 176.)
- [246] G. Villarreal, D. A. Hamilton, D. P. Graham, et al. Reduced area of the corpus callosum in posttraumatic stress disorder. *Psychiatry Res*, 131:227–235, 2004. (Cited on page 128.)
- [247] J. Visser, van den Berg-Vos R. M., H. Franssen, L. H. van den Berg, O. J. Vogels, J. H. Wokke, and de Visser M. de Jong, J. M. Mimic syndromes in sporadic cases of progressive spinal muscular atrophy. *Neurology*, 58:1593–1596, 2002. (Cited on page 132.)
- [248] J. Visser, R.M. van den Berg-Vos, H. Franssen, L. H. Van den Berg, J. H. Wokke, J. M. de Jong, R. Holman, R. J. de Haan, and M. de Visser. Disease course and prognostic factors of progressive muscular atrophy. *Arch.Neurol.*, 64:522–528, 2007. (Cited on page 132.)
- [249] F. M. Vos, M. W. A. Caan, K. A. Vermeer, C. Majoie, G. J. Den Heeten, and L. J. Van Vliet. Linear and Kernel Fisher Discriminant Analysis for Studying Diffusion Tensor Images In Schizophrenia. In *4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007*, pages 764–767, 2007. (Cited on page 35.)
- [250] G. Wahba. *Spline models for observational data*. Society for Industrial and Applied Mathematics, Philadelphia, 1990. (Cited on page 94.)
- [251] S. Wakana, H. Jiang, Nagae-Poetscher, P. C. L. M., van Zijl, and S. Mori. Fiber tract-based atlas of human white matter anatomy. *Radiology*, 230:77–87, 2004. (Cited on pages 132 and 134.)
- [252] S. Wakana, H. Jiang, et al. Fiber tract-based atlas of human white matter anatomy. *Radiology*, 230:77–87, 2004. (Cited on pages 16, 33, 40, 56, and 104.)
- [253] F. Wang, Z. Sun, et al. Anterior cingulum abnormalities in male patients with schizophrenia determined through diffusion tensor imaging. *Am J Psychiatry*, 161:573–575, 2003. (Cited on page 14.)
- [254] S. Wang and E. R. Melhem. Amyotrophic Lateral Sclerosis and Primary Lateral Sclerosis: The Role of Diffusion Tensor Imaging and Other Advanced MR-Based Techniques as Objective Upper Motor Neuron

- Markers. *Ann.N.Y.Acad.Sci.*, 1064:61–77, 2005. (Cited on page 134.)
- [255] Web Books. Nervous system. <http://www.web-books.com>. (Cited on page 2.)
- [256] C.-F. Westin, S. E. Maier, et al. Processing and visualization for diffusion tensor MRI. *Med Image Anal*, 6: 93–108, 2002. (Cited on pages 4, 40, 41, 91, 92, and 161.)
- [257] M.R. Wiegell, H.B.W. Larsson, et al. Fiber crossing in human brain depicted with diffusion tensor MR imaging. *Radiology*, 217:897–903, 2000. (Cited on page 26.)
- [258] J. C. Wong, L. Concha, C. Beaulieu, W. Johnston, P. S. Allen, and S. Kalra. Spatial profiling of the corticospinal tract in amyotrophic lateral sclerosis using diffusion tensor imaging. *J.Neuroimaging*, 17:234–240, 2007. (Cited on pages 135, 140, and 141.)
- [259] M. W. Woolrich, S. Jbabdi, B. Patenaude, M. Chappell, S. Makni, T. Behrens, C. Beckmann, M. Jenkinson, and S.M. Smith. Bayesian analysis of neuroimaging data in FSL. *NeuroImage*, 45:S173–186, 2009. (Cited on page 176.)
- [260] J. R. Wozniak, L. Krach, E. Ward, et al. Neurocognitive and neuroimaging correlates of pediatric traumatic brain injury: a diffusion tensor imaging (DTI) study. *Arch Clin Neuropsychol*, 22:555–568, 2007. (Cited on page 120.)
- [261] B.T.T. Yeo, T. Vercauteren, P. Fillard, J.M. Peyrat, X. Pennec, P. Golland, N. Ayache, and O. Clatz. DT-REFIND: Diffusion Tensor Registration with Exact Finite-Strain Differential. *IEEE Transactions on Medical Imaging*, 28(12):1914–1928, 2009. (Cited on page 90.)
- [262] L. Younes. Combining geodesic interpolating splines and affine transformations. *IEEE Trans. on Image Processing*, 15:1111–1119, 2006. (Cited on page 107.)
- [263] W. Yuan, S. K. Holland, V. J. Schmithorst, et al. Diffusion tensor MR imaging reveals persistent white matter alteration after traumatic brain injury experienced during early childhood. *Amer J Neuroradiol*, 28:1919–25, 2007. (Cited on page 123.)
- [264] H Zhang, P A Yushkevich, D C Alexander, and J C Gee. Deformable registration of diffusion tensor MR images with explicit orientation optimization. *Medical Image Analysis*, 10(5):764–785, 2006. (Cited on pages 90, 100, and 104.)
- [265] H. Zhang, P. Yushkevich, D. Rueckert, and J. Gee. A computational DTI template for aging studies. *Proc ISMRM*, page 3230, 2009. (Cited on page 104.)
- [266] C. Z. Zhu, Y. F. Zang, et al. Discriminative analysis of brain function at resting-state for Attention-Deficit/Hyperactivity Disorder. In *Proc. MICCAI'05*, LNCS 3750:468–475, 2005. (Cited on pages 14, 28, and 36.)
- [267] S. Zoccollella, E. Beghi, G. Palagano, A. Fraddosio, V. Guerra, V. Samarelli, V. Lepore, I. L. Simone, P. Lamberti, L. Serlenga, and G. Logroscino. Analysis of survival and prognostic factors in amyotrophic lateral sclerosis: a population based study. *J.Neurol.Neurosurg.Psychiatry*, 79:33–37, 2008. (Cited on pages 6 and 132.)

## LIST OF ABBREVIATIONS

---

ADC	Apparent Diffusion Coefficient
ALL	Acute Lymphoblastic Leukaemia
ALS(-B/L)	Amyotrophic Lateral Sclerosis (bulbair/limb onset)
ALSFRS-R	Amyotrophic Lateral Sclerosis Functional Rating Scale–revised
AMC	Academic Medical Center
ANOVA	Analysis Of Variance
(b/g/s)CC	(Body/Genu/Splenium of the) Corpus Callosum
CNS	Central Nervous System
CP	Cerebral Peduncle
CPU	Central Processing Unit
CRM	Condor Resource Manager
CSRT	Craniospinal Radiotherapy
CSM	Central Service Manager
CR	Corona Radiata
(r)CRLB	(relative) Cramér-Rao Lower Bound
CSF	Cerebrospinal Fluid
CST	Corticospinal Tract
DCT	Discrete Cosine Transform
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders, fourth edition
DTI	Diffusion Tensor Imaging
DTI-TK	Diffusion Tensor Imaging Toolkit
DTT	Diffusion Tensor Tractography
DW-MRI	Diffusion Weighted Magnetic Resonance Imaging
DWI	Diffusion Weighted Image
EPI	Echo Planar Imaging
FA	Fractional Anisotropy
FACT	Fiber Assignment by Continuous Tracking
FOV	Field Of View
FSL	Functional Software Library
GM	Grey Matter
GWENDIA	Grid Workflow Efficient Enacment for Data Intensive Applications
HAFT	High Angular Fiber Tracking
HARDI	High Angular Resolution Diffusion Imaging
HTTP	Hypertext Transfer Protocol
IAS	Image Analysis Service
IOF/IFO	Inferior Occipitofrontal/Fronto-occipital Fasciculus
IP	Internet Protocol
IPE	Image Processing Engine

K(F)DA	Kernel (Fisher) Discriminant Analysis
KS	Kolmogorov-Smirnov
LDA	Linear Discriminant Analysis
LMMSE	Linear Minimum Mean Square Error
LMN	Lower Motor Neurons
MCR	Matlab Component Runtime
MD	Mean Diffusivity
MED	Medulloblastoma
MI	Meta-Interpreter / Mutual Information
ML	Maximum Likelihood
MND	Motor Neuron Disease
MO	Medulla Oblongata
MOTEUR	hoMe-made OpTimisEd scUfl enactoR
MRI	Magnetic Resonance Imaging
MS	Multiple Sclerosis
MTX	MethoTrexate
P	Pons
PCA	Principal Component Analysis
PET	Positron Emission Tomography
PLIC	Posterior Limb of the Internal Capsule
PLS	Primary Lateral Sclerosis
PMA	Primary Motor Atrophy
PSM	Point Set Matching
RMS(E)	Root Mean Squared (Error)
ROI	Region Of Interest
SCUFL	Simple Conceptual Unified Flow Language
SNR	Signal-to-Noise Ratio
SPM	Statistical Parametric Mapping
S(T)D	Standard Deviation
TE	Echo Time
TID	Task Identifier
TR	Repetition time
TSO	Task Specification Object
UMN	Upper Motor Neurons
URL	Uniform Resource Locator
VBA	Voxel-Based Analysis
VBrowser	Virtual Resource Browser
VC	Vital Capacity
VL-e	Virtual Laboratory for Enhanced Science
(S)WM	(Subcortical) White Matter
WMFA	Mean FA in the White Matter

## LIST OF PUBLICATIONS

---

### *Journal papers*

M.W.A. Caan, K.A. Vermeer, L.J. van Vliet, C.B.L.M. Majoie, B.D. Peters, G.J. den Heeten and F.M. Vos, (2006), Shaving diffusion tensor images in discriminant analysis: A study into schizophrenia, In *Medical Image Analysis*, 10 (5), 1-9.

P.G. Barth, C.B.L.M. Majoie, M.W.A. Caan et al., Pontine tegmental cap dysplasia: a novel brain malformation with a defect in axonal guidance, in: *Brain* (2007), 130, 2258-2266.

E.J. Aukema, M.W.A. Caan, N. Oudhuis, et al., White Matter Fractional Anisotropy Correlates With Speed of Processing and Motor Speed in Young Childhood Cancer Survivors, *Int J Radiat Oncol Biol Phys.* 2009;74(3):837-43.

M.W.A. Caan, H.G. Khedoe, D.H.J. Poot, A.J. den Dekker, S.D. Olabbarriaga, C.A. Grimbergen, L.J. van Vliet, F.M. Vos, Estimation of diffusion properties in crossing fiber bundles, *IEEE Transactions on Medical Imaging*, 2010 (in press)

### *Conference papers*

M.W.A. Caan, K.A. Vermeer, L.J. van Vliet, C.A. Grimbergen, F.M. Vos, Pruning datasets in discriminant analysis: a DTI study to Schizophrenia, in: *ISBI 2007, Proc. 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2007*, 1340-1344.

F.M. Vos, M.W.A. Caan, K.A. Vermeer, C.B.L.M. Majoie, G.J. den Heeten, L.J. van Vliet, Linear and kernel Fisher discriminant analysis for studying diffusion tensor images in Schizophrenia, in: *ISBI 2007, Proc. 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2007*, 764-767.

M.W.A. Caan, A.W. de Vries, H.G. Khedoe, E.M. Akkerman, L.J. van Vliet, C.A. Grimbergen and F.M. Vos, Generating fiber crossing phantoms out of experimental DWIs, in: N. Ayache, S. Ourselin, A. Maeder (eds.), *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2007, LNCS 4791, 169-176.

S. Caton, M.W.A. Caan, S.D. Olabbarriaga, O.F. Rana and B. Batchelor, Using Dynamic Condor-based Services for Classifying Schizophrenia in Diffusion Tensor Images, in: *2008 Eighth IEEE International Symposium on Cluster Computing and the Grid (ccgrid)* pp. 234-241, 2008.

M.W.A. Caan, L.J. van Vliet, C.B.L.M. Majoie, E.J. Aukema, C.A. Grimbergen and F.M. Vos, Spatial Consistency in 3D Tract-Based Clustering Statistics, in D. Metaxas et al. (Eds.): MICCAI 2008, Part I, LNCS 5241, pp. 535–542, 2008.

M.W.A. Caan, C.A. Sage, M.M. van der Graaf, C.A. Grimbergen, S. Sunaert, L.J. van Vliet, F.M. Vos, Dual Tensor Atlas Generation based on a Cohort of Coregistered non-HARDI Datasets, MICCAI 2009, Part I, LNCS 5761, pp. 869–876, 2009.

M.W.A. Caan, F.M. Vos, L.J. van Vliet, A.H.C. van Kampen, S.D. Olabbarriaga, Gridifying a Diffusion Tensor Imaging Analysis Pipeline, CCgrid-health 2010 workshop

M.W.A. Caan, H.G. Khedoe, D.H.J. Poot, A.J. den Dekker, S.D. Olabbarriaga, C.A. Grimbergen, L.J. van Vliet, F.M. Vos, Adaptive noise filtering for accurate and precise diffusion estimation in fiber crossings, MICCAI, 2010 (in press)

S.R. van Noorden, M.W.A. Caan, M.M. van der Graaff, L.J. van Vliet, and F.M. Vos, A Comparison of the Cingulum Tract in ALS-B Patients and Controls Using Kernel Matching, MICCAI 2010 (in press)

#### *Conference abstracts*

S.D. Olabbarriaga, M.W.A. Caan, C. Byrman, Grid-enabled Medical Image Analysis: From Jobs to Workflow-based Service, EGEE User Forum, 2010

M.W.A. Caan, C.A. Sage, M.M. van der Graaf, C.A. Grimbergen, S. Sunaert, L.J. van Vliet, F.M. Vos, Dual Tensor Tracking in Low Angular Resolution Diffusion Weighted MRI, ISMRM 2010

## SUMMARY

---

### **Quantifying pathology in Diffusion Weighted MRI**

In this thesis algorithms are proposed for quantification of pathology in Diffusion Weighted Magnetic Resonance Imaging (DW-MRI) data. Functional evidence for brain diseases can be explained by specific structural loss in the white matter of the brain. That is, certain biomarkers may exist where the disease inhibits improper functioning. Axonal and myelin sheath damage hamper neural connectivity. This can be assessed *in vivo* by measuring a change in the diffusion of water molecules. DW-MRI may deliver such biomarkers by capturing subtle changes in the diffusion process in an early disease stage. Diffusion tensor derived scalar measures such as the Mean Diffusivity and Fractional Anisotropy (FA) quantify this process. When comparing such measures on group level, patients may be found to significantly differ from healthy controls.

Conventional analysis treats measurements in each voxel independently. Due to high inter-brain connectivity, we hypothesize that multiple brain regions are involved in complex brain diseases such as schizophrenia. We introduce a new machine learning framework for performing an analysis of these regions in comparative studies on volumetric data. By ‘shaving’ the mapping computed by discriminant analysis, a characteristic set of regions is automatically extracted (chapter 2). We opt to ‘prune’ the dataset by iteratively discarding misclassified subjects from the cohort, such that the mapping is based merely on representative prototypes (chapter 3). Then a comparison is made between a linear and non-linear kernel discriminant analysis, to identify the dimensionality of the problem (chapter 4).

Methods are proposed for modeling fiber crossings, that constitute 30% of white matter, which cannot be adequately described by a single tensor. These regions are currently disregarded in comparative studies. More complex diffusion models that are introduced need to be adequately evaluated. We propose a new method for creating experimental phantom data of fiber crossings, by mixing the DWI-signals from high FA-regions (5). Complex models demand HARDI (High Angular Resolution Diffusion Imaging) data acquired with higher SNR, diffusion weighting and angular resolution. In comparative studies, scanning time may be insufficient for meeting the requirements. We propose a method to create a dual tensor atlas from multiple coregistered non-HARDI datasets. The random fluctuations in the pose of subjects in the scanner as well as anatomical heterogeneity contribute to an increased angular resolution (chapter 6). Finally, we build an optimization framework for estimating both diffusion shape and orientation in fiber crossings. This work sets fundamental limits for comparative studies to correctly

analyze crossing white matter structures. Firstly, it assesses the precision and accuracy with which parameters may be estimated. Secondly, the optimal acquisition parameters are selected in order to do so. This model allows for estimating consistent FA-profiles along crossing tracts (chapter 7).

Fiber tracts provide a specific frame of reference for computing statistics. We perform 3D tract-based comparison of tensor-derived indices between groups. Inter-subject correspondence is achieved by non-rigid registration based on a joint clustering and matching (chapter 8). The clustering delivers atlas points that serve as a frame of reference for performing the analysis. Spatial consistency is taken to reflect a significant difference between groups (chapter 9).

The potential of our methods is demonstrated in two comparative studies: on Childhood Cancer survivors and Amyotrophic Lateral Sclerosis respectively (chapters 10 and 11).

High throughput analysis of our data is realized by adopting a grid computing approach (chapter 12). The grid provides fast and easy access to shared resources. We present our progress over the past four years in the development and porting the DW-MRI analysis pipeline to grids (chapter 13). By doing so, our algorithms and results can also be accessed by fellow researchers.

Matthan Caan



### **Kwantificeren van ziektebeelden in diffusie-gewogen MRI**

In dit proefschrift worden algoritmes gepresenteerd die ziektebeelden kwantificeren in diffusie-gewogen MRI (DW-MRI) data. Klinisch bewijs voor een hersenziekte kan worden verklaard door een bepaald structureel verlies in de witte stof in de hersenen. *Biomarkers* beschrijven deze plaatsen waar de ziekte aangrijpt, met functieverlies als gevolg. Schade aan axonen en myeline verstoren ter plaatse de neuronale geleiding. Deze schade kan *in vivo* worden bepaald, door een veranderde diffusie van watermoleculen. DW-MRI is in staat in een vroeg stadium van een ziekte subtiele veranderingen in het diffusieproces te meten en zo bepaalde biomarkers te leveren. Dit proces wordt beschreven door de gemiddelde diffusie en fractionele anisotropie (FA), te herleiden uit de diffusie tensor. Door deze maten op groepsniveau te vergelijken, kunnen significante verschillen tussen patiënten en controles vastgesteld worden.

In een conventionele analyse worden alle voxels onafhankelijk beschouwd. Wij verwachten dat, als gevolg van de vele verbindingen in het brein, meerdere gebieden betrokken zijn bij een complexe hersenziekte, zoals schizofrenie. We introduceren daarom een zelflerend systeem dat een vergelijkende studie naar deze gebieden in volumetrische data kan uitvoeren. We ‘schaven’ de projectie die door een discriminatie analyse gevonden wordt, zodat we automatisch die gebieden vinden die samen een onderscheidende verzameling vormen (hoofdstuk 2). Vervolgens ‘snoeien’ we de dataset, door verkeerd geklassificeerde personen iteratief uit de studie te verwijderen (hoofdstuk 3). Dit heeft tot gevolg dat de projectie alleen gebaseerd wordt op representatieve prototypes. Om de dimensionaliteit van ons probleem te bepalen, maken we een vergelijking tussen de uitkomsten van een lineaire en niet-lineaire discriminatie analyse (hoofdstuk 4).

Kruisende fibers komen voor in 30% van de witte stof. Omdat een enkele tensor deze gebieden niet goed kan modelleren, beschrijven wij methodes die dat wél kunnen. Dit doen we, omdat deze gebieden momenteel genegeerd worden in vergelijkende studies. Een complexer diffusie-model moet goed worden geëvalueerd, voor het in een studie toegepast kan worden. Wij hebben een nieuwe methode ontwikkeld om experimentele fantoomdata te genereren. Hiertoe mengen we de DWI-signalen uit gebieden met een hoge FA (hoofdstuk 5). Complexe modellen kunnen alleen geschat worden in datasets met een hoge hoekresolutie en een goede signaal-ruis verhouding (HARDI, *High Angular Resolution Diffusion Imaging*). Echter, de acquisitie van zo’n dataset duurt aanzienlijk langer dan gebruikelijk in een studie. Wij kiezen er daarom voor een twee-tensor atlas te schatten uit meerdere gecoregistreerde datasets die niet aan deze eisen voldoen. Wij

verkrijgen een verhoogde hoekresolutie doordat de ligging van personen in de scanner en de anatomie van de hersenen licht varieert (hoofdstuk 6). Tenslotte ontwikkelen we een optimalisatie-methode voor het schatten van zowel de vorm als de orientatie van de diffusie in kruisingen. In dit werk bepalen we de grenzen waarbinnen een vergelijking van tensor-maten langs kruisende witte-stof-structuren op een juiste wijze kan worden gemaakt. Ook selecteren we de optimale acquisitie-parameters voor het uitrekenen van deze tensor-maten (hoofdstuk 7).

Statistische maten langs gereconstrueerde fiber-banen kunnen namelijk heel specifiek inzicht bieden in een bepaalde ziekte. Waar bestaande methodes fiber-banen als lijnstructuren beschouwen, zijn wij in staat een drie-dimensionale vergelijking van tensor-afgeleide maten te maken. De correspondentie tussen personen bepalen we met een niet-rigide registratie op basis van een gezamenlijke clustering van de fiber-banen (hoofdstuk 8). De clustering levert een verzameling atlaspunten die als basis dient voor de statistische analyse. De spatiële samenhang langs de atlaspunten dient hierin als maat voor significantie (hoofdstuk 9).

In twee vergelijkende studies, naar kinderen die kanker overleefden en naar patiënten met Amyotrofe Laterale Sclerose, laten we zien waartoe de door ons ontwikkelde methodes in staat zijn (chapters 10 and 11).

Om grote hoeveelheden data snel te kunnen verwerken, maken we gebruik van een grid van computers (hoofdstuk 12). Het grid geeft snelle en eenvoudige toegang tot gedeelde bronnen. Wij laten zien hoe wij in de afgelopen vier jaar onze analyse van DW-MRI data hebben gemigreerd naar het grid (chapter 13). Zo zorgen we ervoor dat ook collega-onderzoekers onze algoritmes en resultaten kunnen gebruiken.

Matthan Caan

## ACKNOWLEDGEMENTS

---

On a sunny spring afternoon in 2003, two students entered Frans Vos' room, looking for a Master's project. One was enthusiastic and asked questions while the other was just quietly listening. Surprisingly, the latter returned to apply for a project about 'something to do with orientation space but then different'. The project started in Delft and turned into a journey to a huge hospital with small meeting rooms in Amsterdam, went on to yet another hospital far abroad in Boston USA and ended eventually *Around the world in eighty days* in a Not-So-Broadway theatre. Fortunately, Koen Vermeer kept a close eye on the student, since he saw the potential of a pattern recognition approach to medical imaging.

A few years before, that same quiet student ruined one of the first demonstrations of Lucas van Vliet in the Systems and Signals course, by recognizing Mozart's *Eine kleine Nachtmusik* in the magnitude of the Fourier Transformed signal. Perhaps this event linked them together, such that they could dive into other magnitude data: MR Diffusion Weighted Images. It has been fun to collaborate with a number of student. Other colleagues of QI were also there for support and company. Fortunately, Ganesh Khedoe keeps DTI in Delft alive.

The work in this thesis may be called a joint effort between Amsterdam and Delft. Ard den Heeten and Kees Grimbergen recognized that advanced image analysis of 3.0 Tesla DW-MRI data was relevant. Aart Nederveen continued support from the 3.0 Tesla research facility (and was good company in the train back home). The radiologists Charles Majoie and Liesbeth Reneman were even able to endure a physicist's presentation of results (anterior is left, or right). Cristina Lavini and Erik Akkerman were always willing to share their physicists' point of view in such a situation. Until now, the colleagues from Z-o and 'radiology upstairs' made the AMC feel like home, and they still do.

'Grid' is not about computers, but about people and the connections between them. Silvia Olabarriaga made the integration and implementation of the grid in our project (or vice versa) to a success, long before other people recognized the added value. For example, she sets one to write a paper with people one never met before: an interesting experience. Fortunately, we could get acquainted with Simon Caton and Omer Rana shortly afterwards in the *Auberge* of Paul Bocuse. This work led to the development of a web portal, that could be completed thanks to the support of Antoine van Kampen. In the end, invaluable support was given by Carsten Byrman, Piter de Boer, Jan Just Keijser, Tristan Glatard, Jeroen Engelberts, Tom Visser, Angela Luyf, Barbera van Schaik, Mark Santcroos, Shayan Shahand and Yassene Mohammed.

The ALS-study succeeded due to good collaboration with Maaïke van der Graaff (who knows where to stay in Leuven), Caroline Sage and Stefan Sunaert. Working with Eline Aukema and Netteke Schouten ensured a quick publication of the children's study. In contrast, it took a while before fiber crossings were unraveled. Dirk Poot and Arjan den Dekker were however always patient and willing to help. Wim van Hecke's coregistration method is still state-of-the-art, while Ben Jeurissen helped in building a CSD-atlas.

Over the years, my family has granted unconditional support, they show that love outfavours achievements. Gertrude, you know that this October 11<sup>th</sup> is not the most important one in my life, although from now on I'm no longer just your 'ir... By the way, proposition 10 is yours.

I'd like to acknowledge all these people for their contribution to this work. Maybe we will work together in the future, maybe we won't. For now, we achieved our goal of further quantification of pathology in DW-MRI. Thank you!

## CURRICULUM VITAE

---

Matthan Willem Abraham Caan was born in Kapelle, The Netherlands on July 24, 1980. After completing secondary school (V.W.O., Calvin College, Goes) in 1998, he went on to study Applied Physics at Delft University of Technology. He received his Master's degree in 2005 on the subject 'Methods for Classification of Magnetic Resonance Diffusion Tensor Images'. Part of this thesis was an internship in the Brigham and Women's hospital at Harvard Medical School (Boston, MA, USA), on the topic of registration of Diffusion Tensor Images.

He then pursued a Ph.D. on this topic in a joint collaboration of the Radiology department of the Academic Medical Center in Amsterdam and the Quantitative Imaging group at Delft University of Technology, under the supervision of prof. Kees Grimbergen and prof. Lucas van Vliet.

Following completion of his Ph.D. he is currently employed as post-doctoral research fellow at the Radiology department of the Academic Medical Center in Amsterdam, where he conducts research in the field of Magnetic Resonance image analysis.