A Multimodal Regularization Framework. Application to Axon Fiber Orientation Estimation and Transparent Optical Flow Estimation

by

Alonso Ramírez Manzanares

M.Sc., Centro de Investigación en Matemáticas A.C. (2003)

Submitted to the Computer Science Department
in partial fulfillment of the requirements for the degree of

Doctor of Science

at the

CENTRO DE INVESTIGACIÓN EN MATEMÁTICAS A.C.

June 2007

Signature of Author .....
Computer Science Department
June 1st, 2007

Certified by .....
Dr. Mariano José Juan Rivera Meraz
Senior Research Scientist
Thesis Supervisor

Certified by .....
Dr. José Luis Marroquín Zaleta
Senior Research Scientist
Committee Member

Certified by .....
Dr. Verónica Medina Bañuelos
Senior Research Scientist
Committee Member

Certified by .....
Dr. Jean-Bernard Hayet
Senior Research Scientist
Committee Member

Certified by .....
Dr. Salvador Botello Rionda
Senior Research Scientist
Committee Member

Accepted by .....
Dr. Adolfo Sánchez Valenzuela
Chairman, Departmental Committee on Graduate Students

A Multimodal Regularization Framework. Application to Axon Fiber Orientation Estimation and Transparent Optical Flow Estimation

by

Alonso Ramírez Manzanares

Submitted to the Computer Science Department
on June 1st, 2007, in partial fulfillment of the
requirements for the degree of
Doctor of Science

A Multimodal Regularization Framework. Application to Axon Fiber Orientation Estimation and Transparent Optical Flow Estimation

by

Alonso Ramírez Manzanares

Submitted to the Computer Science Department
on June 1st, 2007, in partial fulfillment of the
requirements for the degree of
Doctor of Science

Abstract

Despite the fact that regularization has a long tradition in early vision, most methods tackle the monomodal pixel–labeling task, i.e. those applications compute only one model per pixel (the most plausible one). On the other hand, the multimodal regularization problem, in which two or more labels per pixel are computed, has captured less attention to research groups due to the fact that most labeling problems are formulated in order to recover a single model per image position. Although, there are early vision problems that must be solved by specialized multimodal regularization methods, as for instance, problems related with transparencies in images or in cases where the *partial volume problem* considerably affects the model–estimation task. In this thesis, we provide a multimodal regularization framework which is capable of detecting several models at a single image position. Our proposal points out the corresponding labels by means of a set of real–valued memberships. Thus, our solution framework is based on regularization cost functions composed by three terms named *data*, *oriented spatial regularization* and *intermodel competition*. Since the unknown is a field of real–valued variables, it is possible to apply gradient–based minimization methods with the well–known algorithmic advantages with respect to *hard*–minimization approaches. Moreover, we introduce the Basis Pursuit method in the context of multi–modal regularization problem. Such an approach is convenient since: it permits us to recover sparse solutions (an important feature in many early vision applications), it is implemented in an efficient way and it is robust to outliers (due to noise) in the data. We implement our framework in two state–of–the–art challenging problems: the axon fiber estimation in DW–MRI and the transparent optical flow estimation. We provide solving methods for both problems with algorithmic advantages with respect to state–of–the–art proposals. In particular, we present an efficient method that overcomes the drawbacks of fitting a Gaussian mixture model to the diffusion weighted images for the axon fiber estimation. Our spatial and intermodel regularization allows one to eliminate noise and to recover good solutions with a reduced number of images and with low-requirement data. Our formulation has shown a superior performance in experimental comparison with the state of the art method named Q-Ball. For the transparent optical flow estimation, we provide a novel multimodal regularization framework that improves previous approaches by avoiding combinatorial optimization methods. Moreover, we propose a variant of such a method that is capable of solving transparent Random Dot Kinematogram sequences. The performance of the proposal is validated by means of synthetic and real data: quantitative and qualitative results are presented.

Resumen

No obstante que los métodos de regularización tienen una gran tradición en el área de visión temprana computacional, la mayoría de los métodos propuestos se han enfocado en resolver el problema de etiquetado monomodal de píxeles. Esto es, dichas aplicaciones asignan un único modelo por píxel (el más adecuado de acuerdo a la regularización). Por otra parte, el problema de regularización multimodal, en el cual es necesario asignar más de una etiqueta o modelo a cada pixel de la imagen, ha recibido menos atención por parte de los grupos de investigación. Lo anterior, en parte porque muchos de los problemas de etiquetado han sido formulados en base a la asignación de un solo modelo. Sin embargo, existen problemas en visión temprana los cuales deben de ser resueltos por métodos desarrollados específicamente para regularización multimodal; como por ejemplo, problemas asociados a transparencias en imágenes o en aquellos casos en los cuales el problema de volumen parcial afecta de una manera considerable la tarea de ajuste de modelos.

En esta tesis presentamos un esquema sobre regularización multimodal, el cual detecta de manera simultánea varios modelos en un píxel. Nuestra propuesta indica las etiquetas asignadas por medio de un conjunto de membresías o variables indicadoras con valores reales. En este sentido, nuestro esquema de solución está basado en la minimización de funciones de costo regularizadas. Dado que el parámetro a ser calculado en nuestro esquema es un campo de variables reales, es posible solucionar el problema por medio de métodos de optimización basados en gradientes, los cuales presentan ventajas muy bien conocidas en comparación con métodos de optimización combinatoria. Adicionalmente, introducimos el método de solución denominado Basis Pursuit en el contexto de regularización multimodal. Este método presenta ventajas tales como: recupera soluciones dispersas (lo cual es una característica deseable en varios problemas de visión temprana), es posible implementarlo de una manera eficiente, y es robusto a valores atípicos (que se introducen debido al ruido en los datos observados).

El buen desempeño de nuestra propuesta es validada mediante la aplicación del mismo a la solución de dos problemas muy interesantes: la estimación de las orientaciones de las fibras de axones en el cerebro, y la estimación de flujo óptico transparente. Para dichos problemas, presentamos métodos de solución con ventajas operativas en comparación con propuestas existentes en el estado del arte. En particular, en el caso de estimación de fibras de axones, nuestra propuesta evita los problemas que se presentan al ajustar un modelo de mezcla de Gaussianas a los datos *pesados en difusión* de resonancia magnética. La regularización espacial propuesta permite eliminar el ruido de adquisición y reducir el número de imágenes necesarias sin disminuir la calidad de las estimaciones. Nuestro método de solución muestra un mejor desempeño en comparación con el bien conocido método de estimación denominado Q–Ball. Por el lado de la estimación de flujo óptico transparente, nuestro esquema presenta mejoras, ya que evita el uso de métodos de optimización combinatoria. Adicionalmente, presentamos un método de análisis de secuencias transparentes conocidas como Random Dot Kinematogram (usadas en experimentos de percepción), el cual está basado en detectores de un solo movimiento en una región y en la integración de información a base de difusión. El desempeño de ambas aplicaciones es validado por medio de experimentos en datos reales y sintéticos, con validaciones cuantitativas para el segundo caso.

Thesis Supervisor: Dr. Mariano José Juan Rivera Meraz
Title: Senior Research Scientist

Committee Member: Dr. José Luis Marroquín Zaleta
Title: Senior Research Scientist

Committee Member: Dr. Verónica Medina Bañuelos
Title: Senior Research Scientist

Committee Member: Dr. Jean-Bernard Hayet
Title: Senior Research Scientist

Committee Member: Dr. Salvador Botello Rionda
Title: Senior Research Scientist

Acknowledgments

Agradecimientos

Dedico esta tesis a Ivete por su amor y ternura, que han sido mi motor en estos años.

A mis padres Jovita y Fausto, por su ejemplo y apoyo. A mis cuatro hermanos y a mis amigos que me han brindado una amistad auténtica.

En especial a Mariano J.J. Rivera Meraz, asesor y sobre todo amigo, por su respaldo incondicional y por todo el tiempo dedicado a mi formación. A los doctores José Luis Marroquín Zaleta y Johan Van Horebeek tanto por su amistad como por su apoyo y ayuda en mi investigación. A mi buenos amigos y compañeros Joaquín Peña Acevedo y Arturo González Vega que siempre estuvieron dispuestos a ayudarme a dilucidar problemas en amenas discusiones.

Agradezco mucho a los sinodales y revisores de mi tesis, ya que gracias a su minuciosa revisión y sus comentarios se logró que la calidad de la tesis mejorara significativamente.

I have people to thank for their contribution to our research work: I would like to express my gratitude to Pierre Kornprobst, François Lauze, Baba Vemuri, Paul Carney and Thomas Mareci for their effort and dedication.

Finalmente quiero agradecer a las instituciones que hicieron posible la realización de esta tesis: primeramente al Consejo Nacional de Ciencia y Tecnología (CONACYT) que me otorgó la beca de manutención y colegiatura, y al Centro de Investigación en Matemáticas (CIMAT) que con su agradable ambiente de trabajo y su excelente planta docente hicieron de mi doctorado una experiencia muy placentera.

Glossary

3D – Three dimensional, 19

ADC – Apparent Diffusion Coefficient, 49

BP – Basis Pursuit, 41

CS – Continuous DBF Solution, 69

dB – decibels, 69

DBF – Diffusion Basis Functions, 55

DFD – Displayed Frame Difference equation, 86

DS – Discrete DBF Solution, 69

DSI – Diffusion Spectrum Imaging, 52

DT – Diffusion Tensor, 46

DW – Diffusion Weighted, 45

EAP – Ensemble Average Probability, 45

EM – Expectation Maximization algorithm, 24

EV – Early Vision, 19

FA – Fractional Anisotropy, 47

FOD – Fiber Orientation Distribution, 53

FT – Fourier Transform, 46

GA – Generalized Anisotropy, 61

GMM – Gaussian Mixture Model, 51

GS – Gauss–Seidel, 60

HARDI – High Angular Resolution Diffusion Images, 49

ICM – Iterated Conditional Modes, 25

LP – Linear Program, 41

LS – Least Squares, 47

MP – Matching Pursuit, 41

MR – Magnetic Resonance, 45

MRF – Markov Random Fields, 24

ODF – Orientation Distribution Function, 53

OF – Optical Flow, 19, 85

OFC – Optical Flow Constraint equation, 86

PCA – Principal Component Analysis, 40

PDD – Principal Diffusion Direction, 47

QP – Quadratic Program, 42

RDK – Random Dots Kinematogram, 86, 109

SHD – Spherical Harmonic Decomposition, 51

SNR – Signal to Noise Ratio, 69

List of Notation

α – Field of indicator variables or memberships, 24

$\hat{\alpha}$ – Optimal α solution, 29

b – Constant directly prop. to magnitude of DW gradients and DW time, 47

d_t – Direction of a particle in a random walk at time t, 65

$d(u_i, r)$ – Local distance of OF velocity u_i at r, 93

$d^*_{1:t}$ – Sequence of previous particle displacements, 66

$\mathbf{D}$ – Simmetric positive definite Diffusion Tensor, 47

$\hat{e}$ – A vector with all its components equal to one, 41

$\hat{e}_i$ – i-th eigen-vector, 47

f – An image, 24

$f(r)$ – Pixel/Voxel value or spatiotemporal value, 24

$\mathcal{F}$ – Fourier Transform, 53

λ_i – i-th eigen-value, 47

μ_c – User-defined positive contrast regularization parameter, 30

μ_s – User-defined positive spatial regularization parameter, 30

$\mathcal{N}_r$ – Neighborhood of r position, 32

∇f_r – Image Gradient Intensity at r, 110

$P(\alpha)$ – Prior distribution, 29

$P(f(r)|\Theta, \alpha)$ – Likelihood, 29

$p_k(f(r)|\Theta)$ – Individual likelihood for model Θ_k, 31

P_n – Noise distribution, 30

$p(\mathbf{x})$ – Ensemble-Average Probability, 45

ϕ_j – The j-th Diffusion Basis Functions, 55

$\mathbf{\Phi}$ – Matrix of DBFs as columns (dictionary), 57

r – Image position, 24

$\mathcal{R}_i$ – Compact image region, 26

$S(\mathbf{q}_k, \tau)$ – DWMRI signal, 46

S_0 – MRI signal when no diffusion magnetic field gradient is applied, 46

$\bar{\mathbf{T}}_j$ – The j-th base tensor, 55

Θ_j – A generic possible model for an image position, 24

$\mathbf{\Theta}$ – A matrix composed by models Θ_j (as columns), 40

$U(\alpha, \Theta)$ – Regularized cost function, 30

$U_d(f, \Theta, \alpha, r)$ – Data term, 30

$U_s(\alpha, r)$ – Oriented spatial regularization term, 30

$U_c(\alpha, r)$ – Inter-model competition or contrast term, 30

u_i – i-th OF velocity from a basis, 93

v_i – i-th unitary vector from a basis, 65

w_{krs} – Diffusion weights, 32

W_r – 3×3 spatial window center at r, 95

Publications Arising From This Work

The following are the peer-reviewed publications that have arisen from this thesis research. Chapter 4 of this thesis is closely based on articles #1,#2,#3 and #5. Chapter 5 was partially reported on papers #4 and #6.

1. Alonso Ramírez-Manzanares, Mariano Rivera, Baba. C. Vemuri and Thomas Mareci. *Basis Functions for Estimating Intravoxel Structure in DWMRI.* Procc. In Proc. IEEE Medical Imaging Conference 2004, Rome, Italy, Pp 4207- 4211, October 2004. Digital Object Identifier 10.1109/NSSMIC.2004.1466819

2. Alonso Ramírez-Manzanares and Mariano Rivera, *Basis Pursuit based algorithm for intra-voxel recovering information in DW-MR*, In Proc. IEEE Sixth Mexican International Conference on Computer Science, Pp 152-157, Puebla, México, 2005.

3. Alonso Ramírez-Manzanares and Mariano Rivera. *Basis Tensor Decomposition for Restoring Intra-Voxel Structure and Stochastic Walks for inferring Brain Connectivity DT-MRI.* International Journal of Computer Vision 69(1), 77-92, 2006.

4. Alonso Ramírez-Manzanares, Mariano Rivera, Baba C. Vemuri, Paul Carney and Thomas Mareci. *Diffusion Basis Functions Decomposition for Estimating White Matter Intravoxel Fiber Geometry.* Accepted in IEEE Transactions on Medical Imaging, To be Published.

5. Alonso Ramírez-Manzanares, Mariano Rivera, Pierre Kornprobst and François Lauze.

A Variational Approach for Multi-Valued Velocity Field Estimation in Transparent Sequences. In Proc. 1st International Conference on Scale Space and Variational Methods in Computer Vision. LNCS 4485, Pp. 227–238. May. Ischia, Italy, 2007.

Additionally, the following technical reports were published

1. Alonso Ramez-Manzanares, Mariano Rivera, Baba C. Vemuri and Thomas Mareci. *Basis functions for estimating intra–voxel structure in DW-MRI.* Reporte Técnico del CIMAT, number 18.10.2004, I-04-10 (CC).

2. Alonso Ramírez-Manzanares, Mariano Rivera, Pierre Kornprobst and François Lauze. *Multi-Valued Motion Fields Estimation for Transparent Sequences with a Variational Approach.* Rapport De Recherche Inria, number RR-5920. 2006.

3. Alonso Ramírez-Manzanares, Mariano Rivera, Pierre Kornprobst and François Lauze. *Multi-Valued Motion Fields Estimation for Transparent Sequences with a Variational Approach.* Reporte Técnico del CIMAT, number 22.06.2006, I-06-12 (CC).

Mariano Rivera, who is coauthor in all articles and technical reports, is the advisor of this thesis and he contributed both in ideas and edition of such reports. Pierre Kornprobst and François Lauze contributed with ideas and discussions about the formulations and also with the reports edition. Baba C. Vemuri, Paul Carney and Thomas Mareci kindly provided us the DW-MRI data, gave us feedback about the experiment results and contribute by reviewing the manuscripts.

Contents

1 Introduction **19**

1.1 Motivation 19

1.2 Our Approach 20

1.3 Contribution 21

1.4 Thesis Outline 21

2 Preliminary Work and Statement of the Problem **23**

2.1 Regularization in Image Processing 23

2.2 Model Indicator Variables in a Regularization Framework 24

2.3 Unimodal and Multimodal Regularization 25

2.4 Multi-modal Approaches 26

3 A Multi-Modal Regularization Framework **29**

3.1 Multimodal Observation Model 30

3.2 Multi-Modal Spatial Regularization 32

3.3 Sparsity in Multi-Modal Solutions 33

3.3.1 Our Proposed Inter–Model Quadratic Regularization for Sparse Multi–Modal Solutions 35

3.4 Basis Pursuit Framework for Recovering Multimodal Sparse Solutions 39

3.5 Alternative Multimodal Formulations and Solution Methods 42

3.6 Chapter Summary 43

4 Application I: Diffusion Basis Functions for Axon Fiber Orientation Estimation 45

4.1 A Review of State–of–the–Art in White Matter Intra–voxel Fiber Geometry Recovery . . . 45

4.1.1 DW–MRI and Diffusion Tensor Imaging . . . 46

4.1.2 The Apparent Diffusion Coefficient based on HARDI . . . 49

4.1.3 Plausibility of the Gaussian Mixture Model . . . 51

4.1.4 Review of Multiple Fiber Orientation Methods . . . 52

4.2 The Diffusion Basis Functions . . . 55

4.3 Numerical Solutions For DBF Model . . . 57

4.3.1 Basis Pursuit Algorithm . . . 58

4.3.2 Spatial and Coefficient-Contrast Regularization . . . 59

4.4 Implementation Details . . . 60

4.4.1 Designing the Tensor Basis . . . 60

4.4.2 Computation of a Continuous Solution . . . 62

4.4.3 Avoiding Ill-conditioning in Mehrotra's Algorithm . . . 63

4.4.4 Fast Convergence for the GS Solver . . . 64

4.5 Stochastic Walks for Estimating Fiber Pathways . . . 64

4.5.1 Computation of the Displacement Direction . . . 65

4.5.2 Stochastic Walk Implementation . . . 68

4.6 Results on Synthetic data . . . 68

4.7 Results on Rat Brain DW-MR data . . . 73

4.8 Comparisons With Q–Ball methodology . . . 76

4.9 Results for Particle Random Walks . . . 78

4.10 Chapter Discussion and Conclusions . . . 81

5 Application II: Multi-Valued Motion Fields Estimation for Transparent Sequences with a Variational Approach 85

5.1 Related Work On Multiple Motion Estimation for Transparent Sequences . . . 86

5.2 Problem Statement: From Local to Global . . . 93

5.2.1 Computing Local Velocity Information . . . 94

5.2.2 Objective: Motion Detection Variables 96
5.3 Global Motion Integration via a Variational Approach 96
5.3.1 Attach Term 97
5.3.2 Spatial Regularization 97
5.3.3 Inter-Model Competition 98
5.4 Experiments 99
5.4.1 Algorithmic Details 99
5.4.2 Global Coherent Motion Estimation for Non Transparent Motion Sequences 100
5.4.3 Global Coherent Motion Estimation for Transparent Motion Sequences . 101
5.5 Random Dot Kinematogram Sequences 109
5.5.1 Computation of local motions 109
5.5.2 Spatial integration 110
5.5.3 Results on RDK Sequences 111
5.6 Chapter Conclusions 112

6 Concluding Remarks **115**
6.1 Contributions 115

Chapter 1

Introduction

This chapter presents an overview of the thesis work. We present a motivation based on a general Early Vision (EV) point of view in computer science. Therefore, a general survey about our approach is provided and the main contributions of this research work are listed.

1.1 Motivation

EV is a research area closely related with mathematic ill-posed inverse problems: Optical Flow (OF) estimation, image segmentation, shape from X, stereoscopy, etc. A common strategy is the well known problem regularization for computing suitable solutions [90]. In most cases, the regularization involves a spatial or spatiotemporal smoothness constraints which can be naturally expressed by regularized cost functions.

A common strategy for tackling the above problems is to select the solution from a set of models *a priori* pre-fixed. In this sense, a set of memberships is computed for each image position. Such memberships indicate a degree of association among model and pixel (or voxel for three-dimensional (3D) volumes). One can find in the literature several works that, based on this paradigm, compute the most plausible model for each image position, i.e. those strategies recover monomodal distributions for the memberships [133, 72, 71, 98]. Although such works have presented an impressive performance, some practical problems in EV require to recover a subset of models within an image position and not only "the most plausible one", as for instance, problems that deal with transparent objects in images.

In a Bayesian regularization approach, the data (pixels) are assumed to be produced by a mathematical model (named generation model). Then a common approach consists on proposing procedures to choose among the models the most plausible one, and in a more complete approach the model parameters.

Our methods are developed for analyzing (solving) EV (or image processing) problems in which the data are generated by more than one model. We developed approaches that explicitly introduce a bias for recovering multimodal solutions.

The aim of this research work is to provide a general multimodal regularization scheme and to demonstrate how it can be applied for solving practical problems.

1.2 Our Approach

Our method is composed by the following stages:

1. First, we define a problem-dependent general observation model in the form of a combination of basic solutions (basic models) .

2. Then we associate the presence/absence of basic models in a data term by means of real–valued indicator variables (or memberships).

3. Thus, based on the Markov Random Field framework, we propose a spatial regularization term that promotes oriented–smoothness in the solution.

4. Finally, we introduce prior knowledge about the desired number of models at each position by means of a model–competition cost term.

5. Additionally, we show how to apply the Basis Pursuit formulation for promoting to compute suitable multimodal sparse solutions for cases where the spatial regularization is not required.

The above procedure can be applied to circumstances in which it is necessary to recover several models at each position, and thus it can be seen as a general framework for different multimodal EV problems.

1.3 Contribution

The main product of this thesis is a general framework for solving pixel–wise the multi–labeling problem. This framework includes an anisotropic filtering scheme that promotes an oriented regularization. Such regularization is useful for problems where the models are associated to orientations/directions, as for instance optical flow. Besides, we introduce a novel multimodal model–competition scheme which promotes high contrasted (sparse) multimodal solutions.

We implement our approach in two EV challenging problems. The advantages of our formulations are explained below.

For the first application of this thesis, we attain improvements with respect to state–of–the–art approaches for the axon multi–fiber estimation in Diffusion Weighted Magnetic Resonance (DW–MR) images. We formulate an observation model capable of representing more than one fiber in a voxel. This model describes the DW–MR signal by means of linear combination of discrete prefixed DW–MR signals taken from a proposed basis dubbed Diffusion Basis Functions. Based on the Basis Pursuit approach and our multimodal regularization framework, we present an efficient solution scheme that promotes sparsity in the number of axon bundles required for describing the measured DW–MR signal.

In the second application we tackle the transparent optical flow estimation problem. We developing a new observation model that incorporates a set of measurements for detecting several optical flow velocities in a single pixel. Therefore, we develop a gradient based solution scheme that presents algorithmic advantages with respect to previous state–of–the–art approaches.

1.4 Thesis Outline

In Chapter 2 a review of EV regularization is given, and based on this, we introduce the need for developing multimodal regularization approaches. In Chapter 3 we present the general framework that illustrates the main contributions of the present work. A deep insight of our regularization terms is illustrated and discussed. Additionally, we present and discuss some alternative formulations for multimodal regularization and optional solution methods. Chapters 4 and 5 show the applications of our proposal via real applications to EV problems: a) the estimation of brain water diffusion in axon fiber crossings/bifurcations and b) estimation of

transparent optical flow. We introduce in each application a state–of–the–art review and we illustrate the capabilities of our approach in both synthetic an real images. Finally, in Chapter 6 we discuss our conclusions and propose future work for these lines of research.

Chapter 2

Preliminary Work and Statement of the Problem

In this chapter we review preliminar work related with EV regularization problems. Therefore, we explain the difference between monomodal and multimodal solutions and we point out cases in which a multimodal solution is required.

2.1 Regularization in Image Processing

EV is a research area with a long tradition in regularization of ill-posed problems [90, 66]. An ill-posed problem has no a solution or the solution is not unique or it is not stable under small data perturbations. A clear example in EV is the image interpolation task, i.e. given a discrete set of measurements in a regular grid (pixels), one can compute an infinite number of continuous functions that explain the sampled points. In order to overcome such an ambiguous situation one must regularize the problem [110] by introducing prior knowledge about the desired solution. In the EV context, the most common regularization is to constrain the spatial smoothness. In the interpolation problem, spatial regularization can be introduced by selecting from all possible continuous functions with identical error values the smoothest one. Thus, by means of the regularization, the interpolation task becomes a well–posed problem.

We refer the reader to several works in which a smoothness regularization allows one to recover correct solutions: [50, 68, 42, 133, 1, 72, 71, 31, 128, 98]. Such a regularization is

naturally coded by regularized cost functions which are minimized in order to compute the optimal estimator for the posterior marginal, see [90] and Chapter 3. In addition to smoothness, different prior knowledge has been introduced as regularization terms. For instance, we can find in the literature terms that promote low entropy in the solution [133, 98], a term that promotes the solution to be a positive–definite matrix [127], or a term that promotes to recover small–magnitude optical flow velocities [137].

It is important to note that a regularized functional could have several local minima (especially, when non-convex potentials are used, as for instance, in the well–known discontinuity–preserving problem [96]). Thus, the use of suitable minimization procedures is essential in order to compute high quality solutions with a reasonable computational effort. As a consequence, a lot of the current research work is focused on the development of appropriate optimization techniques for solving this sort of problems [77].

2.2 Model Indicator Variables in a Regularization Framework

First, we will define some notation. The pixel position of an image f is denoted by $r = [x, y, z]$ for 3D data, or $r = [x, y, t]$ for spatiotemporal data. A labeling problem in EV could be seen in a general form as: to associate each $f(r)$ to the most suitable model from a set of N possible ones. The models are denoted as $\Theta_j, j = 1, \ldots, N$. Such models could be pre–fixed (prior models) or simultaneously estimated along with the solution, as in the well–known segmentation/model–estimation Expectation Maximization (EM) algorithm [44].

A possibility for computing a regularized solutions is to use the Bayesian estimation theory jointly with Markov Random Fields (MRF) models [66]. The *a posteriori* or regularized solution can be set as a field of indicator variables

$$\alpha_k(r), k = 1, \ldots, N, \tag{2.1}$$

such that $\alpha_k(r)$ is the *k–th* indicator variable at r and denotes if a pixel/voxel r belongs to model k or not. Based on previous notation, we denote as a *k–layer* the entire volume α_k, for all r.

In many approaches, the indicator variables are defined as binary, i.e. $\alpha_j(r) \in \{0, 1\}$. In

this way, expensive combinatorial optimization methods that spend exponential time must be applied for computing the α field. Examples of previous approaches are, simulated annealing, the Gibbs Sampler approach, the Iterated Conditional Modes (ICM) method (which, in despite of being lighter than the previous two, does not guarantee to converge to global minima), etc. (see [19, 42, 109, 11]). Recently, the Graph Cuts method (a low polynomial algorithm) has demonstrated a good performance for solving discrete pixel–labeling tasks, although because of the method's formulation, the problem should be decomposed into energy functions involving only binary variables [59]. For all previous approaches the computational burden considerably increases when $N > 2$; this is an important drawback for problems in which computing a reasonable accurate solution implies to work with, for instance, tens of models.

On the other hand, a family of approaches define $\alpha_j(r) \in [0, 1]$, as a probability or *membership*. In this sense, such methods constrain vector $\alpha(r)$ to be a probability measure, i.e. $\sum_j \alpha_j(r) = 1$ (with $\alpha_j(r) \geq 0$) as in the *fuzzy* membership approaches [20, 1]. Then, gradient based minimization procedures can be applied with the well-known algorithmic advantages, see [133, 72, 71, 98]. By taking into account the forementioned advantages, the present thesis work is based on this sort of membership approaches instead of *hard* indicator variables.

2.3 Unimodal and Multimodal Regularization

Given the vector solution (2.1) at r, we name monomodal solution the case when a single $\alpha_j(r) \approx 1$ and a multimodal solution when a subset of $K > 1$ coefficients $\alpha_{\{j_1,\cdots,j_K\}}(r) \approx 1$.

Approaches which enforce (2.1) to be a monomodal solution detect in an efficient way the more likely regularized solution in each position, generating fast and robust model detectors [133, 71].

Although previous monomodal approaches are powerful methods in EV, some problems involve the computation of several models at a given position (multimodal solutions). Figure 2-1 illustrates such a case by a general scheme with regions associated to models Θ_j and Θ_k, but also with a region with a mixture of models denoted by $\Theta_j \cap \Theta_k$ (a region where the models overlap). Such a phenomenon is present, for instance, in the problem of transparent optical flow estimation [80, 18]. Also, note that when $\Theta_j \cap \Theta_k = \emptyset$, there are pixels with the well–

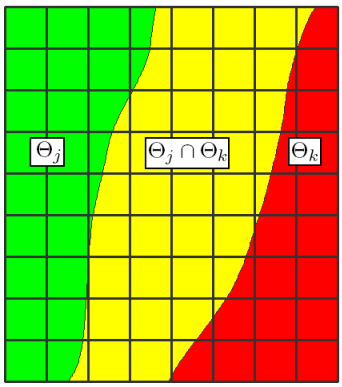


Figure 2-1: General scheme of image volumes that require to compute multimodal solutions: Several pixels must be associated to both Θ_j and Θ_k models.

known *partial volume* problem, i.e. because the discrete feature of the imaging system, parts of discrete–integrated pixels are associated to a model and the rest to another. In the general case, a region could be composed by a mixture of N models or a single pixel could be composed by N portions of different models. In this sense, the solution of a problem that requires multimodal labeling, differs from the classic segmentation directive in which the image is decomposed into a set of model–associated regions $\mathcal{R}_i$, such that $\mathcal{R}_j \cap \mathcal{R}_k = \emptyset$ is always true for any j, k.

Note that for solving this sort of problems it is necessary to develop methods that are explicitly biased for recovering multimodal solutions; this is addressed by this thesis work.

2.4 Multi-modal Approaches

To the best of our knowledge, approaches in literature that are capable of recovering more than one model per pixel are in practice restricted to computing the presence of at most two models per pixel/voxel [109, 11, 117, 85]. On the other hand, some of them [109, 11], involve the computation of a field of binary variables (indicating one or two models in a position), so that the solution must be computed by *hard minimization* procedures (as methods that we point out in section 2.2), which results into an important computational burden.

Moreover, previous approaches suffer the model–selection problem, i.e. it is necessary to choose between fitting a single/multi model (by performing a pre–processing), or to fit both and then choose the one which explains better the measurement $f(r)$.

Other situations in which multimodel approaches can be applied is when the model basis

is too big, however, such model basis can be generated by combinations of elements of a small basis. For instance, in the case of two Transparent Optical Flows (TOF). This can be better understood with the following example: For regular optical flow a possible motion model basis can be expressed as

$$\mathbf{u} = \{u_1, u_2, \ldots, u_n\}$$

where u_k is a motion vector, then the solution corresponds to finding at each pixel the motion vector u_k that better explains the local motion. On the other hand, in the two TOF problem, the motion model basis can be seen as the union of the single motion basis and the possible combinations of two motions:

$$\mathbf{w} = \{\mathbf{u}\} \cup \{w_1, w_2, \ldots, w_N\}$$

with $N = \binom{n}{2}$ and w_k is a pair of simple motions u_i and u_j. In this thesis we avoid such unwanted polynomial growth of the simple model combination by allowing multimodal solutions.

Next chapter presents the modifications and proposals we have developed during this work that result in a novel method for recovering multimodal solutions.

Chapter 3

A Multi-Modal Regularization Framework

In this chapter we present a framework for a multimodal regularization scheme. First we set the basic theory about regularization, then we explain the required modifications for dealing with the multimodal requirements. We introduce the Basis Pursuit approach in the multimodal regularization context where spatial regularization is no need, and finally alternative multimodal formulations and solving methods are given.

In the general Bayesian estimation approach, the Maximum at Posteriori (MAP) estimator for the indicator variables α in (2.1) is given by

$$\hat{\alpha} = \max_{\alpha} P(\alpha|f(r), \Theta) = \max_{\alpha} P(f(r)|\Theta, \alpha)P(\alpha), \tag{3.1}$$

where probability $P(f(r)|\Theta, \alpha)$ is named the likelihood, which indicates, given the observation model, how likely $f(r)$ could be generated from models $\Theta_k, k = 1, \ldots, N$. The prior distribution term $P(\alpha)$ introduces the regularization by codifying our prior knowledge about the desired α solutions.

In the Markov Random Field (MRF) framework the prior distribution $P(\alpha)$ is expressed as a Gibbs distribution with MRF models. Then the posterior probability takes the form

$$P(\alpha|f(r), \Theta) = \frac{1}{Z} \exp\left(-U(\alpha, \Theta)\right)$$

where $U(\alpha, \Theta)$ is named the posterior energy. Then, maximizing the probabilistic formulation (3.1) is analog to minimizing $U(\alpha, \Theta)$. In our approach, the general regularized cost functional is given by three terms. This process leads us to minimize a regularized cost function (see [19, 90, 66]):

$$U(\alpha, \Theta) = \sum_r \Big[U_d(f, \Theta, \alpha, r) + \mu_s U_s(\alpha, r) + \mu_c U_c(\alpha, r) \Big], \tag{3.2}$$

where μ_s and μ_c are positive user–defined regularization parameters, term $U_d(f, \Theta, \alpha, r)$ is the one associated to the likelihood and terms $U_s(\alpha, r)$ and $U_c(\alpha, r)$ are associated to the prior distribution as is explained in the next subsections. In order to apply suitable minimization approaches, it is always desirable to propose appropriate probability distributions in (3.1) such that it results in quadratic potentials for terms in (3.2). Therefore, the solution (MAP estimator) is computed by solving linear systems of equations. In following sections we give a deep insight of each proposed term in (3.2) that follows previous required feature.

3.1 Multimodal Observation Model

The first term in (3.2) is the data term, it attaches the model indicator variables $\alpha(r)$ to the measured data. Based on Bayesian estimation theory, the first step is to compute the likelihood field. In some cases it is possible to formulate an observation model that is implicity based on the indicator variables (as for instance, in the linear combination case $f(r) = \sum_j \alpha_j \Theta_j$), so that, the likelihood depends on α as

$$P(f(r)|\Theta, \alpha) = P_n \left(f(r) - \sum_j \alpha_j \Theta_j \right), \tag{3.3}$$

for an assumed known noise distribution P_n . In this case, minimizing the data term

$$U_d(f, \Theta, \alpha, r) = V \left(f(r) - \sum_j \alpha_j \Theta_j \right), \tag{3.4}$$

w.r.t. the α vector, corresponds to maximizing the likelihood of the proposed model, where V is generally some convex potential (e.g. when P_n is supposed Gaussian).

When our observation model does not explicitly include the indicator variables one can

Table 3.1: Data terms that link the likelihood vector $p(f(r)|\Theta)$ to the vector α.

$\mathbf{U_d}(\mathbf{f}, \boldsymbol{\Theta}, \boldsymbol{\alpha}, \mathbf{r})$	**Used in**
$\sum_k (\alpha(r)_k - p_k(f(r)\|\Theta))^2$	[72]
$-\log\left(\alpha(r)^T p(f(r)\|\Theta)\right)$	[71]
$\sum_k \alpha_k(r) \left[-\log\left(p_k(f(r)\|\Theta)\right)\right]$	[133]
$\sum_k \alpha_k^2(r) \left[-\log\left(p_k(f(r)\|\Theta)\right)\right]$	[98, 97]

compute the individual likelihoods for each model as

$$p_k(f(r)|\Theta) = P_n(f(r) - \Theta_k). \tag{3.5}$$

Thus, it is possible to link the data likelihood to the indicator variables (2.1) by means of a data term. For instance, several terms for this aim are shown in Table 3.1 (a modification of the approach presented on the first row is proposed in [43]). Note that in all cases, as required, the data term is minimized when the α vector is similar to the likelihood one.

For the applications of this thesis, we use the data term in (3.4) (see Chapter 4) and the one shown in last row of Table 3.1 (see Chapter 5).

We note that an hypothetic method that introduces indicator variables for all combinations of models, i.e. α_i, $\alpha_{i,j}$, $\alpha_{i,j,k}$, ..., etc. presents an important drawback: the huge number of required variables and the consequent increase in the computational burden. So that, for practical purposes, we should diminish the number of combinations, eliminating in this way the capability of detecting an arbitrary number of models. Differently, we use the indicator variables in (2.1) and we avoid the constraint $\sum_j \alpha_j(r) = 1$ since we detect multimodal solutions (see Section 2.3). Thus, for instance, we expect to recover $\alpha_i(r) \approx \alpha_j(r) \approx \alpha_k(r) \approx 1$, $\alpha_l(r) \approx 0, l \neq i, j, k$ when three models Θ_i, Θ_j and Θ_k explain the observed data at r. In this way, if the additional constraint $0 \leq \alpha_k(r) \leq 1$ is provided, our indicator variables can be interpreted as probabilities but the $\alpha(r)$ vector is not necessarily a probability measure. In this sense, the $\alpha(r)$ vector is a vector of probability measures, where each entry indicates the probability of the presence ($\alpha_k(r)$) and the probability of absence ($1 - \alpha_k(r)$) of model Θ_j.

3.2 Multi-Modal Spatial Regularization

The second term in (3.2) performs a spatial/spatiotemporal regularization. Based on membership approaches [133, 72, 71] we use a functional that exploits the MRF model and minimizes the differences between variable indicator vectors within a neighborhood. Let $\mathcal{N}_r$ be the second order spatial neighborhood of r:

$$\mathcal{N}_r = \{s : \|r - s\|_2 < 2\}. \tag{3.6}$$

Thus we define the general form of our spatial regularization term as

$$U_s(\alpha, r) = \sum_{s:s\in\mathcal{N}_r} \sum_k w_{krs} \left(\alpha_k(r) - \alpha_k(s)\right)^2, \tag{3.7}$$

where w_{krs} are spatial–oriented weights explained in the sequel. Such a regularization term is derived from the relationship between the Gibbs distribution and the MRF prior given the neighborhoods $\mathcal{N}_r$, as was stated by the Hammersley and Clifford theorem, see [19].

Term (3.7) integrates the local information (likelihood vector) in order to estimate global solutions, see [66]. Such a spatial integration allows the method to deal with corruption caused by ill-posed observation models, acquisition noise and incomplete data.

The normalized w_{krs} weights, constrain different level of smoothness for $\{k, r, s\}$ configurations [71]. For instance, in a 2D image, an x–orientation filtering is achieved for all α layers by setting

$$w_{krs} = \begin{cases} 1 - |\mathcal{N}_r|\epsilon & |s_x - r_x| = 1, s_y - r_y = 0 \\ \epsilon & \text{else} \end{cases},$$

where $|\cdot|$ denotes set cardinality and $\epsilon \ll 1$.

Moreover, when the models are intrinsically associated to directions/orientations the weights could be determined by such a feature of each model: one can introduce prior knowledge about the orientation in the spatial data integration, see Figure 5-2. In the applications of this thesis (in Chapters 4 and 5), two practical examples are given: the weights are oriented along the axon fiber and along the optical flow velocity associated to each model, respectively. Figure 3-1 illustrates such a prior information by an scheme in which two objects are moving with

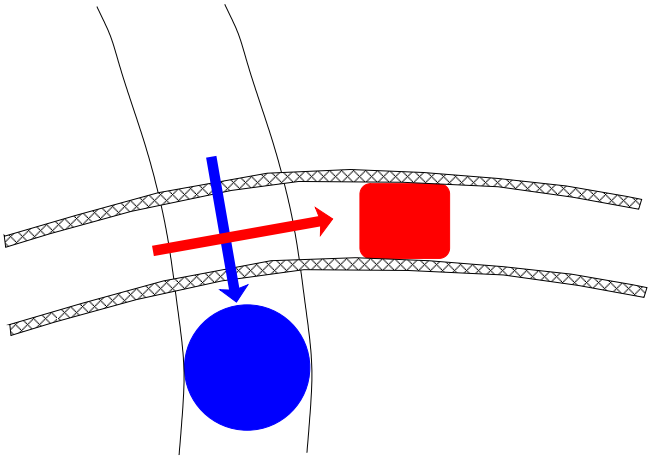


Figure 3-1: The anisotropic filtering effect on the spatial regularization. See text for details.

the velocities denoted by the colored arrows. For instance, for the red object, one wants to recover the red velocity only within the region of its trajectory, but, if one applies a strong spatial regularization, then the displacement boundaries could be blurred, invading in this way the shaded regions. Thus, by using the proposed anisotropic smoothness one diminishes the influence of each model along its associated transversal orientation, and promotes the influence only along the associated orientation (where it is more likely to detect the same model when the velocity is smooth). Therefore, this anisotropic filtering can be seen as a sort of edge–preserving strategy.

Therefore, in this work, we promote an anisotropic filtering, so that, our weights are more related with the ones proposed in [122].

3.3 Sparsity in Multi-Modal Solutions

Because the problem is ill-posed or because of the noise in the data, a simple maximum likelihood estimator could detect spurious models. In order to discern among prominent and spurious models, we introduce an ad–hoc regularization that promotes the recovery of sparse solutions.

Sparsity or "compact coding" scheme has been biologically motivated: In the context of biological vision models it is desirable to minimize the number of models that respond to any particular event. The prior information is that the probability of any model to be present is equal, but such probability is low for any given model. Such coding scheme is like the

one presumed by experiments in retina and primary visual cortex for natural images analysis [34, 35, 79], where a suitable response may be described in terms of structural primitives (as lines , for instance), so that, only a subset from all detectors should indicate contribution (for instance, only a few log-Gabor filters from among a huge bank of them).

In the work in [133] is proposed to penalize the Shannon's entropy (to promote sparsity) for the probability measure vector $\alpha(r)$ as

$$\hat{U}_c(\alpha, r) = -\sum_k \alpha_k(r) \log(\alpha_k(r)),$$

however, the minimization scheme becomes complex, since the derivative of this cost function is not linear. Thus, works in [98, 97] penalize the Gini's entropy by means of a quadratic potential as

$$\hat{U}_c(\alpha, r) = 1 - \sum_k \alpha_k^2(r).$$

However, note that in both cases the minimum entropy is reached for monomodal solutions [49], thus, such functionals are not suitable for the aims of this thesis.

On the other hand, multimodal non-linear cost functions that promote sparsity have been proposed [78] as

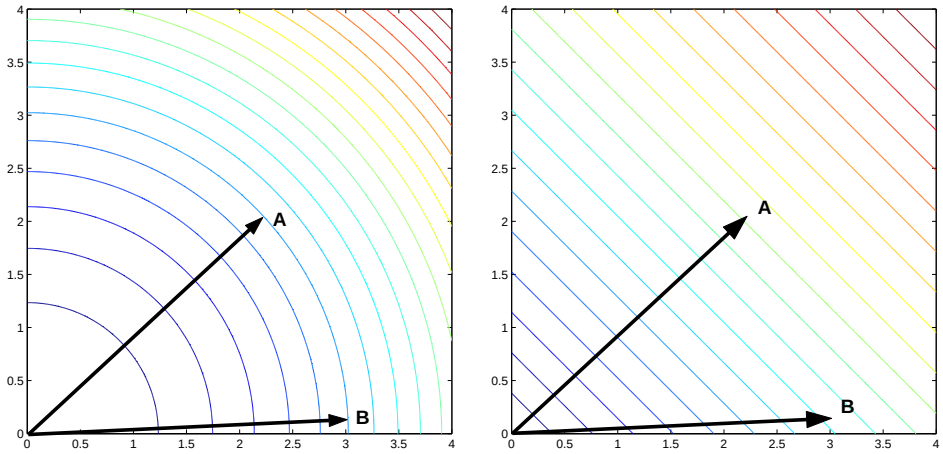
$$\hat{U}_c(\alpha, r) = -\sum_k \exp(-\alpha_k^2(r)),$$

$$\hat{U}_c(\alpha, r) = \sum_k \log(1 + \alpha_k^2(r))$$

and

$$\hat{U}_c(\alpha, r) = \sum_k |\alpha_k(r)| = \|\alpha(r)\|_1. \tag{3.8}$$

Figure 3-2 illustrates how the sparsity is promoted by functional (3.8) (as was reported in [78], no meaningful differences in implementation are found between the former three terms). For Panel 3-2(a) a solution that contains significant changes in the proportion of the values for both variables α_1 and α_2 does not involve a major cost. Differently, for (3.8) one can see that solutions that include only one variable (point **B**, for instance) minimizes the cost. Therefore, the three previous potentials promote the recovery of sparse solutions, although, their non-linear feature results in the well-known minimization drawbacks.



(a) (b)

Figure 3-2: L–2 norm vs. L–1 norm for prior of positive α coefficients. Level curves of U_c for a 2D case with α_1 is laying on the X-axis and α_2 on Y-Axis. The vectors A and B show cases for $\sqrt{\alpha_1^2 + \alpha_2^2} = 3$. (a) $P(\alpha_k) \propto \exp(-\alpha_k^2)$ and (b) $P(\alpha_k) \propto \exp(-|\alpha_k|)$.

3.3.1 Our Proposed Inter–Model Quadratic Regularization for Sparse Multi–Modal Solutions

For overcoming the shortcomings of the regularization terms explained above, in this section we propose to perform an *inter–model* regularization over the $\alpha(r)$ vector by means of a quadratic term defined in general form as

$$U_c(\alpha, r) = -\sum_k \big(\alpha_k(r) - C_{\alpha(r)}\big)^2 \tag{3.9}$$

where $C_{\alpha(r)}$ is an intermediate value laying between zero and the maximum value expected, according to our model, by any α_k coefficient. This term promotes large contrast in the $\alpha_k(r)$ coefficients, since it is minimized as each $\alpha_k(r)$ is as far as possible from the value $C_{\alpha(r)}$. We can understand $C_{\alpha(r)}$ as a parameter (maybe constant) that acts as a threshold between the significant models and the non–significant ones.

Denoting the mean value of $\alpha(r)$ vector by $\bar{\alpha}(r) = \frac{1}{N}\sum_k \alpha_k$, let's suppose $C_{\alpha(r)} = \bar{\alpha}(r)$. Thus, one can obtain interesting information by expanding term (3.9) as

$$U_c(\alpha, r) = -\sum_k (\alpha_k(r) - \bar{\alpha}(r))^2 \tag{3.10}$$

$$= -\sum_k \alpha_k^2(r) + \frac{2}{N}\left(\sum_k \alpha_k(r)\right)^2 - N\bar{\alpha}^2(r)$$

$$= -\sum_k \alpha_k^2(r) + N\bar{\alpha}^2(r). \tag{3.11}$$

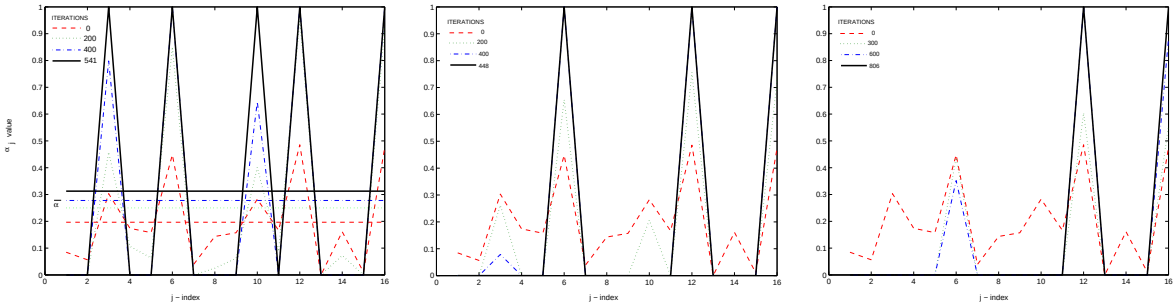
In this sense, to understand the potential's behavior we note that the first term in (3.11) promotes the "switching on" of models and avoids the trivial null solution $\alpha(r) = 0$, while the second term penalizes the number of switched–on models. Hence for a fixed mean value (controlled by the second term) the first term (that acts in a similar way to Gini's entropy term) prefers highly contrasted solutions.

According to our experiments and inspired from the previous analysis, we propose a more versatile inter–model regularization term as

$$U_c(\alpha, r) = -\sum_k \alpha_k^2(r) + \kappa N\bar{\alpha}^2(r) \tag{3.12}$$

where $\kappa > 0$. Note that for $\kappa = 1$, this term is equal to (3.10). In this way, we can tune the κ parameter depending on the number of models that we want to detect at each pixel. Figure 3-3 shows the regularization behavior of this term, by itself, along the iterations. We show the cases $\kappa = 1$, $\kappa = 2$ and $\kappa = 4$ in Panels (a), (b) and (c) respectively. Note that only the prominent coefficients (that are present since the first iteration) are enhanced and the others are completely attenuated. As can be seen, potentials (3.10) and (3.12) are suitable for recovering multimodal solutions. For this example we minimized (3.12) by a fixed-step gradient minimization [77], with step size $h = 0.005$.

It is important to note that potential (3.12) can be tuned so that for a given κ value a multimodal solution (with two or more detected motions) has lower energy than a unimodal one or viceversa: for instance, the final solution (in the convergence) shown in Panel 3-3(c) has an energy $U_c(\alpha, r) = -1$, if we turn–off arbitrarily $\alpha_{12}(r) = 0$ then we get a higher energy $U_c(\alpha, r) = -0.75$ for this monomodal solution. Moreover, instead of setting a monomodal



(a) (b) (c)

Figure 3-3: Regularization effect along the iterations of the inter–model competition term (3.11). At iteration zero several models have significant values; at minimization convergence the non–prominent coefficients are completely attenuated. (a) result for $\kappa = 1$, i.e. term (3.10), where the horizontal line marks the mean vector value $\bar{\alpha}$, (b) for $\kappa = 2$ and (c) for $\kappa = 4$.

solution, by arbitrarily turning–on a third coefficient, for instance $\alpha_6(r) = 1$, we get again a higher energy $U_c(\alpha, r) = -0.75$. This behavior is plotted in Fig. 3-4 for several κ values and for a different number of modes in the vector solution. This behavior makes an important distinction with respect to entropy that always has lower energy for unimodal solutions [133, 98]. Additionally, our proposed potential, based on quadratic terms, is easily differentiable and therefore simple minimization algorithms can be used.

By analyzing term (3.10), one can see that such a term promotes a large variance in the α-values since it promotes (by the negative sign) deviation with respect to the mean value. This is illustrated by an scheme in Figure 3-5, where we show that it is clear that high–contrasted solutions have higher variance than the low–contrasted ones.

It is important to state that term (3.9) could lead us to a non-convex regularized function (3.2) for the case $\mu_c \gg \mu_s$. Figure 3-6 illustrates this for a simple 1D example (i.e. $r \in \Re$) with $f(r) = 0.75$ and $N = 2$ models $\{\Theta_1 = 0, \Theta_2 = 1\}$. For illustrative aims, we fixed the neighborhood α–vectors as $\alpha(r-1) = [0.2, 0.75]^T$, $\alpha(r+1) = [0.15, 0.85]^T$ and we define $C_{\alpha(r)} = 0.5$. Thus the unknown is $\alpha(r) \in [0,1] \times [0,1]$. It is clear that, given $f(r)$ and the two α-vectors in $\mathcal{N}_r$, the model θ_2 is the most plausible one. Panel 3-6(a) shows the functional with the inter–model regularization term turning–off, i.e. $\mu_c = 0$. Then, by increasing the μ_c value, we drive the solution to the desired one (almost binary): $\alpha(r) = [0,1]^T$, Panel 3-6(b). In the same way, an increment in the selection of μ_c parameters generates a non-convex cost function with several local minima, as is illustrated in Panel 3-6(c). We note,

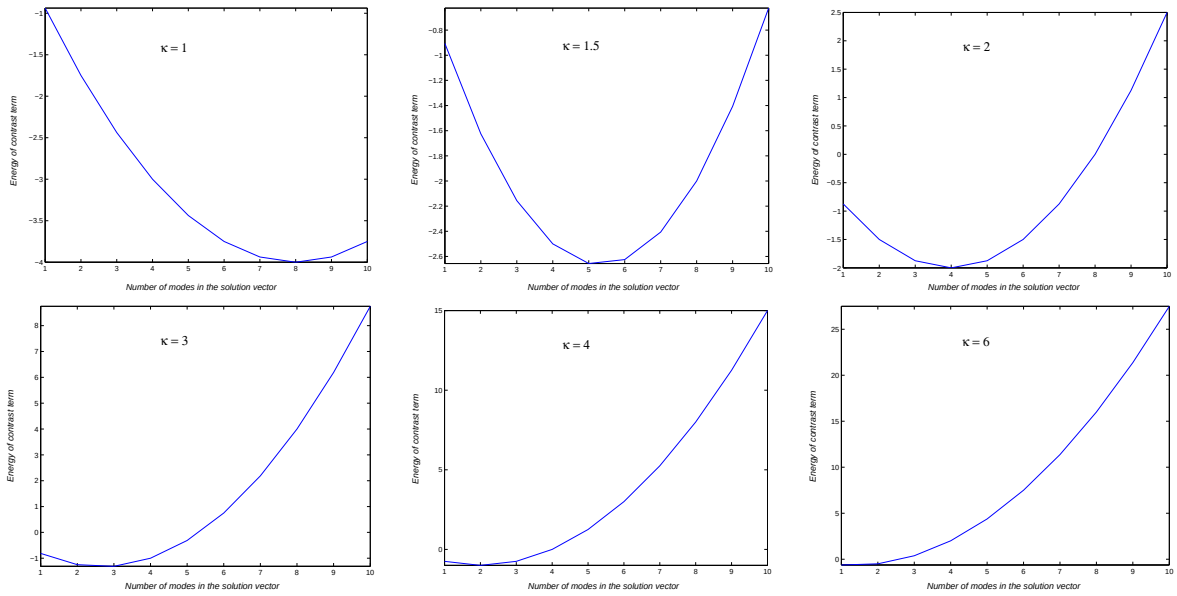
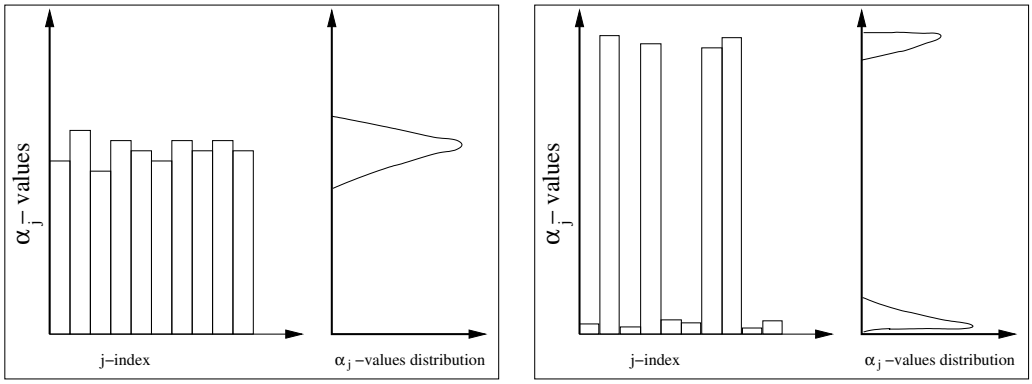
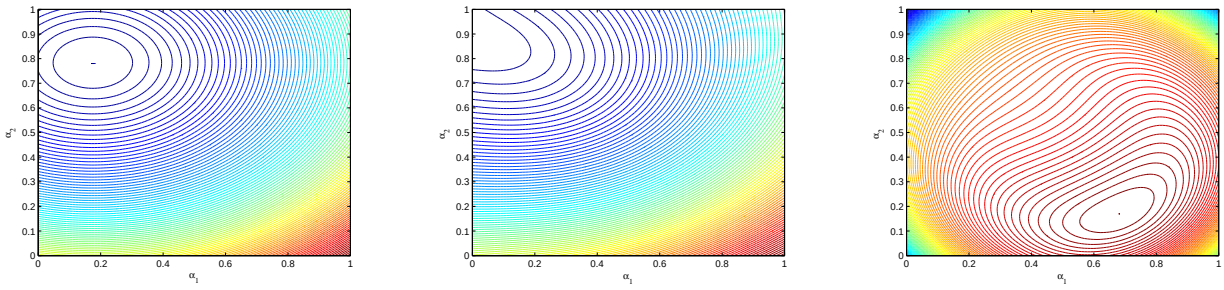


Figure 3-4: Number of modes (i.e. $\alpha_i = 1$) that minimize the energy of term (3.12) with $N = 16$ and different κ values.



(a) (b)

Figure 3-5: Variance in the α-coefficients value. (a) A low-contrasted solution is associated to low variance and (b) conversely, high variance is given by high-contrasted solutions.



(a) (b) (c)

Figure 3-6: Convex or non-convex regularized cost functions (3.2) depending on parameter selection. Shape of the functional for parameters $[\mu_s, \mu_c]$ equal to (a) $[1, 0]$, (b) $[1, 1]$ and (c) $[1, 20]$ respectively.

however, that in the region close to the global solution the gradient of the functional correctly indicates the desired minimization direction. For this reason, the μ_c value must be gradually introduced since it is important to perform the coefficient–contrast regularization until we have an intermediate regularized solution (as for instance the one presented in Panel 3-6(a)). Thus, we perform a deterministic annealing over this parameter that gradually introduces the inter–model regularization along the minimization, as is explained in the application Chapters 4 and 5. Such a deterministic annealing is inspired in the work [25], where *graduated non-convexity* was used for approximating the global solution for nonconvex minimizations.

Finally, according to our experiments, by using (3.12) it is possible to avoid any other regularization term to promote the "switching on" of the indicator variables. However, such a scheme is difficult to tune; thus in our applications (for instance, in Chapter 5) we prefer to factorize such a term.

3.4 Basis Pursuit Framework for Recovering Multimodal Sparse Solutions

In this section, we show how to introduce the Basis Pursuit formulation [30] for recovering multimodal sparse solutions when there is no need to perform data spatial integration.

Compact signal representation is a well studied problem by signal processing researchers [30, 45, 69]. In this context, it is convenient to represent a given signal by a set of coeffi-

cients associated with elements of a dictionary (or basis) of functions. The elements of such a dictionary are called *atoms* (or basis functions). Very often, the dictionary is redundant (over-complete) and the signal atoms could be non-orthogonal. That can offer advantages: one can define a large set of specialized basis function, then only a small subset of the dictionary represents with significant detail the important structures of the observed data [65]. In this sense, setting the basis is by itself a regularization process that biases the observation representation to a given structure.

In some interesting approaches [78, 79, 65] the basis functions are estimated by the method itself. An intuitive example of this situation is the Principal Component Analysis (PCA) in which we explain the observed data by means of the eigen–vectors of the covariance matrix (basis functions) of a multivariate Gaussian (by assuming that the data have Gaussian structure and by constraining the basis to be orthogonal).

In the signal processing context it is common to select atoms as Wavelets, Gabor filter banks, Cosine Packets, Chirplets and Warplets, among others [30]. The idea is to select from the dictionary the atoms that best match the signal structures, using a criterion for choosing among equivalent decompositions. A commonly used criterion is to accomplish with the basic principle of sparsity, i.e., to represent the signal with a fewest coefficients as possible. Additionally, a desirable feature is to achieve the decomposition in a computationally efficient way.

In our notation, the mathematical model that represents the decomposition of a signal $f(r) \in \Re^M$ as a linear combination of atoms Θ_j, each of them belonging to a dictionary, is given by

$$f(r) = \mathbf{\Theta}\alpha(r) + \eta(r); \tag{3.13}$$

$\mathbf{\Theta}$ is a $M \times N$ matrix where the j–th column corresponds to the j–th model Θ_j and the non–negative weights of the linear combination are given by the $\alpha(r)$ vector.

Note that $\mathbf{\Theta}$ is rectangular with $M < N$ for the cases where one acquire less $f_k(r)$ measurements than the number of *atoms* (a convenient and common situation). Consequently, we have more unknowns α's than data measurements; thus the problem (3.13) is ill-posed and should be constrained or regularized in order to compute a meaningful solution.

In [30] was noted that trying to solve it by using the pseudo-inverse of $\mathbf{\Theta}$ leads to a major

disadvantage: it does not allow us to introduce prior information about α, preventing us to lead the method to a solution with some desired properties (sparsity for instance).

Recently Basis Pursuit (BP) [30] technique has been proposed for solving the problem (3.13) (and consequently, for minimizing the data term (3.4) when no spatial regularization is need), by introducing in an efficient–algorithmic way the sparsity constraint presented in (3.8). Such a method proposes to compute a solution by means of a Linear Program (LP) problem as:

$$\begin{aligned} \min \quad & \|\alpha\|_1 = \textstyle\sum_j \alpha_j = \hat{e}^T \alpha \\ \text{subject to} \quad & \boldsymbol{\Theta}\alpha = f(r), \\ & \alpha_j \geq 0 \ , \ \forall j, \end{aligned} \tag{3.14}$$

where $\hat{e}$ is a vector with all its components equal to one (we can use just $\hat{e}^T\alpha$ since the sign of the α components is already constrained). In this sense, it was shown in [65] that the formulation in (3.14) is equivalent to solve (3.13) with the Laplacian prior $P(\alpha_k) \propto \exp(\lambda|\alpha_k|)$.

By constraining the L–1 norm of α vector, one guarantee that a global minimum will be attained. Because of noise, and for cases when the dictionary $\boldsymbol{\Theta}$ is not complete, the reconstruction constraint in (3.14) could not to be accomplished, resulting in an over-constrained LP problem. Therefore an appropriate minimization procedure is used: specifically an interior-point method which also tries to minimize the magnitude of the residual vector $\eta_\alpha = \boldsymbol{\Theta}\alpha - f(r)$ (see [77]). In our experiments we used the powerful primal-dual predictor-corrector Mehrotra's algorithm [74, 77] that computes the results in a fraction of the computational effort required by other less-sophisticated interior point methods. Note that formulation (3.14) is capable for dealing with complete, uncomplete and overcomplete dictionaries, which allows one to adapt the formulation to different problems.

The BP method has shown a better performance with respect to other pursuit techniques, for instance Matching Pursuit (MP) [69], which uses a non-linear iterated algorithm for solving (3.13). Despite the fact that MP is a simple and easy-to-implement algorithm, it has an important drawback: very often the first selected atom globally fits several signal structures, but, this atom could not be well adapted to represent the signal local structures (see [45]).

Differently, BP indicates with few α coefficients the atoms that best fit the local structures, with the drawback that it requires a more sophisticated minimization tool.

We note that the solution of the Mehrotra's algorithm is achieved by computing the roots of a function which includes the gradient of the augmented Lagrangian of the constrained system (3.14) and the constraint $\mathbf{\Theta}\alpha - f(r) = 0$, see [77]. Such a solution is computed by means of the Newton's method for non–linear equations. In this sense, there is not a user–defined parameter that weights the importance of each one of the two goals in (3.14) (sparsity and measurement reconstruction). According to our experiments, the method gives a reconstruction–error as small as possible then it minimizes the L1–norm of the α vector.

In the standard LP formulation of the Mehrotra's algorithm either primal (the α coefficients) or dual (not used in our case) variables are constrained to be positive, so that, the only user–defined constraint is the system $\mathbf{\Theta}\alpha = f(r)$, see [77].

3.5 Alternative Multimodal Formulations and Solution Methods

It is important to note that it is possible to fit the model (3.13) by means of several minimization methods. This subsection shows alternative minimization approaches and formulations, then we indicate advantages and drawbacks of each one.

The basic idea is to minimize the least–square error in the observation model (3.13) as

$$\hat{\alpha} = \min_{\alpha} \|\mathbf{\Theta}\alpha - f\|_2^2 = \min_{\alpha} \alpha^T \mathbf{G}\alpha + d^T\alpha, \tag{3.15}$$

s.t. $\alpha_k \geq 0$. Second expression in (3.15) is the well-known Quadratic Program (QP) formulation where $G = \mathbf{\Theta}^T\mathbf{\Theta}$ and $d^T = -2f^T\mathbf{\Theta}$. The previous problem can be solved with methods as Non-Negative Least-Squares, Quadratic Penalization, Logarithmic Barrier, Augmented Lagrangian, Active Set, Gradient Projection and QP Inner Point, see [77]. For all these methods it is not necessary to indicate complicated starting minimization points and the results are competitive. However, such formulations do not penalize sparsity (they do not introduce any other regularization neither) thus the computed solutions could be not suitable for most applications.

In order to introduce sparsity in (3.15), we can formulate the following QP problem:

$$\hat{\alpha} = \min_{\alpha} \alpha^T \mathbf{G}\alpha + [d + \lambda_c \hat{e}]^T \alpha \tag{3.16}$$

s.t. $\alpha_k \geq 0$, where the $\lambda_c \hat{e}^T \alpha$ term performs the sparsity constraint as in (3.14). According to our experiments, the main disadvantage in previous formulation is that it needs of selecting a suitable λ_c value, since the abuse on this value carry on to an incorrect model fitting, i.e. α entries could be significantly attenuated.

Finally, we note that it is possible to pixel–wise integrate the neighborhood information in a BP scheme as

$$\begin{aligned} \min \quad & \hat{e}^T \alpha(r) \\ \text{s.t.} \quad & \mathbf{\Theta}\alpha(r) = f(r), \\ & \mathbf{I}\alpha(r) = F_{\mathcal{N}_r}, \\ & \alpha_j \geq 0 \ \mathbf{,} \ \forall j, \end{aligned} \tag{3.17}$$

where $F_{\mathcal{N}_r} = \frac{1}{|\mathcal{N}_r|} \sum_{s \in \mathcal{N}_r} f(s)$. In this case we should solve problem (3.17) for the entire image iteratively until convergence. Previous procedure could require significative computational burden.

Because of the drawbacks for the formulations given in this section, we recommend to use the methods described in sections 3.1–3.3.

3.6 Chapter Summary

In this chapter we provide a general framework that is capable of dealing with the regularization of multimodal problems. We presented a quadratic regularization cost function that integrates the data measurements, performs a spatial data integration and promotes sparse solutions. Because of the sparsity constraint, our method promotes almost binary solutions in the recovered memberships. In this sense, we provide a method that uses derivative–based minimization algorithms for the multi–labeling task problem. Thus, our quadratic potentials are easily differentiable. Therefore the energy minimization is obtained by solving linear systems

of equations with non–negativity constraints.

Moreover, we show how to integrate the BP approach in a multimodal regularization framework when no spatial regularization is need, which gives algorithmic advantages and promotes the recovering of sparse solutions in a natural way.

Finally we provide alternative formulations and their minimization procedures for this sort of problems where we remark advantages and drawbacks of such formulations.

The good behavior of our formulation is demonstrated in the following two chapters (4 and 5) where we apply the proposed multimodal regularization method to two challenging problems that are naturally associated with multimodal labeling tasks.

Chapter 4

Application I: Diffusion Basis Functions for Axon Fiber Orientation Estimation

Water diffusion estimation has extensively been used in recent years as an indirect way to infer axonal fiber pathways and this in turn has made the in vivo estimation of fiber connectivity patterns one of the most challenging goals in neuroimaging. For this purpose, a special Magnetic Resonance Imaging (MRI) technique named Diffusion Weighted (DW) MR is used. This imaging technique allows one to estimate the preferential orientation of the water diffusion phenomenon in the brain, which in the white matter case, is usually constrained along the orientation of axon fibers. This information is very useful in neuroscience research due to the relationship of brain connectivity with several diseases and, in general, with brain development [29][91].

4.1 A Review of State–of–the–Art in White Matter Intra–voxel Fiber Geometry Recovery

The DW–MRI acquisition protocol measures (indirectly) the Probability Density Function (PDF) or Ensemble-Average Probability (EAP) $p(\mathbf{x})$ for the displacement vector $\mathbf{x} = \mathbf{x}_e - \mathbf{x}_0$

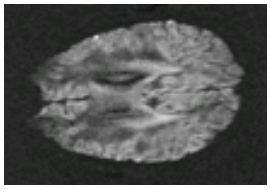


Figure 4-1: An axial slice from a DW–MR human brain scan for a single 3D orientation $\mathbf{g}_k$. The dark regions correspond to sites where the water diffusion is significant along $\mathbf{g}_k$.

corresponding to the particle displacement located at $\mathbf{x}_0$ at the beginning and at $\mathbf{x}_e$ at the end of the experiment for a fixed time τ.

The Fourier Transform (FT) relationship between the PDF and the DW–MR signal for a voxel is given by

$$E(\mathbf{q}_k) = \frac{S(\mathbf{q}_k, \tau)}{S_0} = \int_{\Re^3} p(\mathbf{x}) \exp(-i\mathbf{q}_k^T \mathbf{x}) d\mathbf{x} = \int_{\Re^3} p(\mathbf{x}) \cos(\mathbf{q}_k^T \mathbf{x}) d\mathbf{x} \tag{4.1}$$

where $E(\mathbf{q})$ is the normalized DW-MR signal, S_0 is the measured signal when no diffusion magnetic field gradient is applied (a standard $T2$ image [29]), $S(\mathbf{q}_k, \tau)$ is the observed attenuated signal value given the diffusion vector $\mathbf{q}_k = \gamma \delta G \mathbf{g}_k$ where γ is the gyromagnetic ratio, δ is the applied time for the directional magnetic gradient, G is the magnitude of the applied diffusion magnetic field gradient and the unitary vector $\mathbf{g}_k = [g_{kx}, g_{ky}, g_{kz}]^T_{k=1,\ldots,M}$ indicates the k–th orientation of the diffusion–encoding gradients, finally τ is the effective diffusion time. Note that antipodal symmetry in (4.1), i.e. $p(\mathbf{x}) = p(-\mathbf{x})$, is assumed, so that the cosine expansion in the FT is the only one that is used.

Thus, the axon fiber orientation problem can be stated as the estimation of $p(\mathbf{x})$ based on as few as posible measurements of its FT.

4.1.1 DW–MRI and Diffusion Tensor Imaging

In most medical applications, the angular variation of water diffusion has been extensively summarized by Diffusion Tensor (DT) MRI [14][17]. In the recent past, there has been a flurry of research activity based on DT–MRI: denoising [120, 131, 127, 128, 113, 36], field segmentation [140, 125, 124, 126, 63, 62] and field registration [99, 41] among others. In [107], Stejskal-Tanner presented the mono-exponential model of the DW–MR decayed signal and in [14], Basser et al.,

developed the DT model which is given by the following expression:

$$S(\mathbf{q}_k, \tau) = S_0 \exp(-\mathbf{q}_k^T \mathbf{D} \mathbf{q}_k \tau) + \varepsilon_k, \tag{4.2}$$

where the anisotropic diffusion coefficients (with units equal to mm/s^2) are summarized by the positive definite symmetric 3×3 tensor $\mathbf{D}$. This model implicitly assumes that $p(\mathbf{x})$ presents Gaussian distribution as:

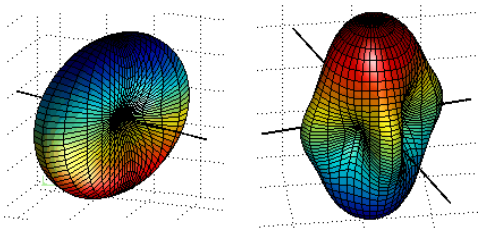
$$p(\mathbf{x}) = \frac{1}{\sqrt{(4\pi\tau)^3 \det(\mathbf{D})}} \exp\left(-\frac{\mathbf{x}^T \mathbf{D}^{-1} \mathbf{x}}{4\tau}\right). \tag{4.3}$$

Signal noise in 4.2 has Rician distribution and is represented by ε_k, see [47] and the Appendix. A standard acquisition protocol consist of M 3D-images, where the voxel intensity is attenuated by the fiber orientation (modeled by the anisotropic tensor $\mathbf{D}$) and the orientation of the diffusion magnetic field gradient $\mathbf{g}_k$. Thus lower gray values indicate larger attenuation, revealing a large water diffusion in the tissue along the orientation $\mathbf{g}_k$, see Figure 4-1. A commonly used convention is to let $b = (\gamma \delta G)^2 \tau$ and thus making b (denoted in s^2/mm) a constant directly proportional to the magnitude of the directional gradients and the acquisition time. If S_0 and at least six measurements $S(\mathbf{q}_k, \tau)_{k=1,\ldots,6}$ are provided (taken in independent non–coplanar orientations), then the DT can be estimated by a Least Squares (LS) procedure or a nonlinear LS [127, 128]. The DT can be visualized as a 3D ellipsoid, with the principal axis aligned with the eigenvectors, $[\hat{e}_1, \hat{e}_2, \hat{e}_3]$, and scaled by the eigen–values, $\lambda_1 \geq \lambda_2 \geq \lambda_3$, defining the diffusivity along each axis. Thus, $\hat{e}_1$ is named the Principal Diffusion Direction (PDD) , and in the case of a single fiber bundle, it is associated with the orientation of the fibers.

The *fractional anisotropy* (FA) is the 3D anisotropy measure most commonly used for DT [17][16]:

$$FA(\mathbf{D}) = \sqrt{\left[(\lambda_1 - \lambda_2)^2 + (\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_3)^2\right] / 2\left(\lambda_1^2 + \lambda_2^2 + \lambda_3^2\right)}.$$

Note that for a DT fitted to a highly oriented diffusion within the voxel ($\lambda_1 >> \lambda_2 \simeq \lambda_3$) FA is close to one, while FA is close to zero for a DT fitted to the isotropic diffusion case ($\lambda_1 \simeq \lambda_2 \simeq \lambda_3$).

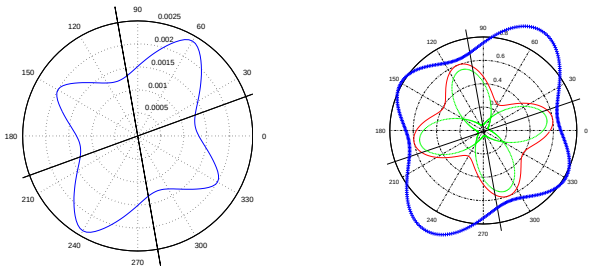


(a) (b)

Figure 4-2: $S(\mathbf{q}_k, \tau)$ for (a) a single fiber, (b) two fibers. The black axis denotes the fiber orientations. To simulate them we used $b = 5000 s^2/mm$.

Noise is not the only factor that affects the fiber orientation estimation in DW-MR images; voxels with partial volume effect have a more pronounced impact. For instance, the observed DT at voxels where two or more fibers cross, split, or merge is the average of the diffusion in the constituent fiber orientations, and thus the fitted DT inadequately represents such an intra–voxel information [29, 134, 135, 2, 95]. In order to compute a good estimation of the underlying fiber pathways, it is necessary to use methods that solve the limitations of the mono–exponential tensor model and recover the intra–voxel information. In order to visualize this phenomenon, we refer to the example illustrated in Figure 4-2. The one fiber case is illustrated in Panel 4-2(a); the $S(\mathbf{q}_k, \tau)$ surface has a unique extremum in the orientation with maximum diffusion (the biggest signal attenuation orientation). In this case, the fitted DT, capable of representing a single PDD, can explain such a phenomenon quite well. On the other hand, panel 4-2(b) shows the two fibers case, that can be verified by the multiple observed extrema. For this case, a single DT can not explain the measured DW–MRI signal. Thus, the presence of two (or more) well defined fiber orientations results in a low–anisotropy fitted DT. This fact increases the uncertainty of the tissue orientation (i.e., the inverse problem is not well-defined).

The loss of directional information in fiber crossings affects the estimation of fiber pathways since all the DT–based methods in literature follow the PDD field in order to infer the fiber pathways [92, 120, 16, 15, 139, 61]. In order to compute a good estimation of the underlying fiber pathways, it is necessary to use methods that solve the limitations of the mono–exponential tensor model and, in this way, recover the intra–voxel information.



(a) (b)

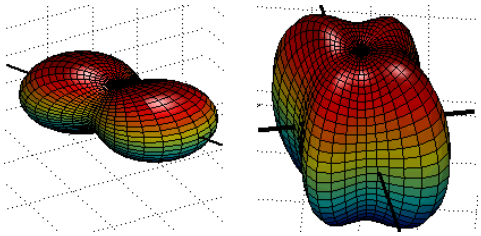
Figure 4-3: (a) Simulated ADC map for 2 fibers oriented with angles $\theta = \{20^o, 100^o\}$ respectively, computed from synthetic DW–MRI signal by applying the Söderman model [106] (b=5000). We show in blue the ADC map and in black the correct fiber orientations. (b) Free–diffusion case: in green the signal due to the first fiber and the signal due to the second fiber. In red, the addition of both independent signals. In blue the ADC map of red signal.

4.1.2 The Apparent Diffusion Coefficient based on HARDI

Because the limitations of the parametric DT model, research groups drew attention to compute non-parametric diffusivity coefficients with High Angular Resolution Diffusion Images (HARDI) [118]. One possibility is to compute the non–parametric coefficients $D(\mathbf{q}_k)$ as

$$D(\mathbf{q}_k) = -\log[S(\mathbf{q}_k, \tau)/S_0]/b, \tag{4.4}$$

with $k = 1, \ldots, M$, for a large set of orientations (M large). However, in the two (or more) fibers case, the maxima of this Apparent Diffusion Coefficient (ADC) Map do not correspond with the fiber orientations [118, 117, 114, 82]. Figure 4-3 illustrates such a behavior; for two fibers lying in X-Y plane and along directions $\theta = \{20^o, 100^o\}$, panel 4-3(a) shows in blue the ADC Map, and in black lines the real fiber orientations. Note the discrepancies between the maxima of the ADC Map and the fiber orientation axes. The DW-MRI signals in Panel 4-3(a) were generated with the cylinder confined diffusion model derived by Söderman et al [106], see also [123, 82] (we used the same set of parameters proposed in [82]). In the panel 4-3(b), we show the signals (computed with the free diffusion model) for each fiber (in green), i.e. the *peanut* shapes, the ADC map is shown in blue, and the addition of both signals is shown in red. As one can see, for both cases, the minima of the addition (red line) correspond (as expected) with the maxima of the ADC map (blue line), but not with the fiber orientations. Because the joint effect of the diffusions in the DW-MR signal does not present the maximal decay along



(a) (b)

Figure 4-4: Panel (a) and (b) depict the ADC Maps for the $S(\mathbf{q}_k, \tau)$ signals shown in Panels 4-2(a) and 4-2(b) respectively. The black axis denotes the fiber orientations.

the orientations of the fibers. For instance a synthetic case for 3 fibers is reported in [82] and a real ADC map that shows the above effect for 2 fibers is reported in [117]. See Panel 4-4(a) and Panel 4-4(b) for a 3D synthetic case.

For the non-free diffusion MR signal, the expression for diffusion restricted within a cylinder was derived by Söderman [106]. Such a model, for standard diffusion parameters, generates almost the same signal as the one obtained by the addition of the exponential decays in Figure 4-3, as can be seen by a qualitative examination. More recently Von Dem Hagen et al. [123] validated the Söderman's model by comparing the synthetic signals with the observed in a physical phantom.

Since the ADC Map computed with HARDI data does not indicate the fiber orientations, the observation models that propose the computation of the fiber orientations by decomposing the standard DT as an addition of tensors [95]:

$$S(\mathbf{q}_k, \tau) = S_0 \exp\left(-\mathbf{q}_k^T \left\{\sum_{j=1}^{L} D_j\right\} \mathbf{q}_k \tau\right) \tag{4.5}$$

or to decompose it as a high rank tensor [81]:

$$\log \frac{S(\mathbf{q}_k, \tau)}{S_0} = -b \left\{\sum_{i1=1}^{3} \sum_{i2=1}^{3} \cdots \sum_{il=1}^{3} D_{i1,i2,\ldots,il} g_{i1} g_{i2} \ldots g_{il}\right\} \tag{4.6}$$

fail to detect the correct fiber orientation because the diffusion maxima may not correspond to the fiber orientations. To avoid such an inconvenience and based on the observation model (4.5), in [95] regularization terms that integrate the information around the voxel's neighborhood are

introduced.

4.1.3 Plausibility of the Gaussian Mixture Model

Figure 4-3 illustrates a more plausible model for the decayed signal phenomenon. The Gaussian Mixture Model (GMM) [117]:

$$S(\mathbf{q}_k, \tau) = S_0 \sum_{j=1}^{L} \beta_j \exp(-\mathbf{q}_k^T \mathbf{D}_j \mathbf{q}_k \tau) + \varepsilon_k; \tag{4.7}$$

where the real coefficients $\beta_j \in [0, 1]$ indicate the portion in the total diffusion for each independent tensor $\mathbf{D}_j$, i.e. $\left(\sum_j \beta_j = 1\right)$. Such a model implicitly assumes that the PDF $p(\mathbf{x})$ presents a GMM distribution (because of the linear feature of the FT) as:

$$p(\mathbf{x}) = \frac{1}{\sqrt{(4\pi\tau)^3}} \sum_{j=1}^{L} \beta_j \frac{1}{\sqrt{\det(\mathbf{D_j})}} \exp\left(-\frac{\mathbf{x}^T \mathbf{D_j}^{-1} \mathbf{x}}{4\tau}\right). \tag{4.8}$$

The GMM explains quite well the diffusion phenomenon for two or more fibers within a voxel (by assuming no exchange between fibers, i.e. the signals are independently added). The GMM was explored by Basser et al. in [13]; and they concluded that its solution requires a large number of measurements $S(\mathbf{q}_k, \tau)$, and that it would be problematic because of the non-linearity. Tuch et al. [117] proposed to solve (4.7) with a multi–start gradient descent algorithm with HARDI data. Such a state–of–the–art technique allows one to recover intra–voxel information that is not observed in the standard DT-MRI neither in the HARDI ADC Map. Frank [39] develops the *Spherical Harmonic Decomposition* (SHD) method by expanding his model to the N–fiber case based on model (4.7). Parker and Alexander [85] used a Levenberg-Marquardt algorithm to fit the GMM. Recently Ozarslan et al. [82] use the GMM to perform an important refinement in their Diffusion Orientation Transform (DOT) for computing the displacement probability. See [29] for more details about model (4.7).

4.1.4 Review of Multiple Fiber Orientation Methods

Tuch et al. proposed a non–linear LS method for solving (4.7) by using the following simplified model [117, 114]:

$$S(\mathbf{q}_k, \tau) = S_0 \sum\nolimits_{j=1}^{L} \beta_j \exp(-\mathbf{q}_k^T R_j^T \Lambda R_j \mathbf{q}_k \tau), \tag{4.9}$$

where $\Lambda = \text{diag}[\lambda_1, \lambda_2, \lambda_3]$ is a constant diagonal matrix and R_j are the associated rotation matrix of each DT. Thus, the unknowns in (4.9) are the number of fibers L, the mixture coefficients β and the director cosines of R_j. The drawbacks of the method are: the large number of required diffusion images $\{S\}$ that notably increases the acquisition time (for instance, 126 diffusion 3D–images are used in [117, 114] and 54 in [85]) and the algorithmic problems related to the non-linearity of (4.7). So that, multiple restarts of the optimization method are required for preventing the algorithm from settling in local minima. Note that, in that formulation, it is necessary to choose between a single Gaussian model or a GMM by choosing the model which explains better the DW–MRI signals (for instance, in work [85] they used the SHD as a model selection preprocessing). Furthermore, no stable solution has been reported for more than 2 fiber bundles, i.e. for $L > 2$ (see discussion on Ref. [117], Chap. 7).

Work [8] presents an interesting model–based approach by assuming that the DW signal is composed by the addition of the signal due to the hindered water diffusion (extra–axonal space) and the one related to the restricted water diffusion (intra axonal space). Such a model was extended to a multi-fiber case, though, the above explained model-selection problem is present.

Frank [39] analyzes HARDI DW–MR data by representing it into the SHD, see [5] for implementation details. Frank's method eliminates artifacts contained in the DW data, and gives information for discerning between the one and multi-fiber case. However, the model is described only in a qualitative way: for the multi–fiber case a method for estimating the exact number of fibers within the voxel and the orientations was not addressed.

A successful non-parametric representation is Q–space methods. Such a line of research is based on Diffusion Spectrum Imaging (DSI) [130, 119, 129]. Q–space methods exploits the FT relationship in (4.1). The non–parametric EAP is defined as

$$p(\mathbf{x}) = S_0^{-1} \int S(\mathbf{q}, \tau) e^{-i\mathbf{q}^T\mathbf{x}} d\mathbf{q} = \mathcal{F}^{-1}\left[E(\mathbf{q}, \boldsymbol{\tau})\right], \tag{4.10}$$

where $\mathcal{F}^{-1}$ denotes the inverse FT with respect to the diffusion vector $\mathbf{q}$ (the transform evaluates an integral over spherical $\mathbf{q}$ domain, i.e. for several orientations and magnitudes, which requires a significantly large DW data set). Once $p(\mathbf{x})$ is approximated, the fiber's orientation is estimate by computing the Orientation Distribution Function (ODF) as the radial projection of p onto the unit sphere

$$\Psi(\hat{\mathbf{x}}) = \int_0^\infty p(\alpha\hat{\mathbf{x}})d\alpha, \tag{4.11}$$

and by looking for the orientations in which this projection presents peaks.

DSI successfully represent, in a non-parametric way, several fibers within the voxel, although the large number of required DW images makes the method unpractical for medical purposes. Modern methods for recovering EAP, as Q–Ball [115] (method that collapse the probability function in a sphere of constant radius), Persistent Angular Structure (PAS) [54] (that computes the PAS as a spherical function with minimum entropy and which FT best fits the measurements) or Diffusion Orientation Transform (DOT) [82] (that use a Laplace–Series intermediate representation for S that simplifies the computation of (4.10)) have demonstrated good results with a smaller number of diffusion images. However, in our opinion, the number of DW images is still large in order to obtain a good EAP. A drawback in the previous methods is the need of a post-processing in order to locate the maxima of the EAP [114, 115, 82, 4].

Recently, spherical deconvolution techniques successfully computes the Fiber Orientation Distribution (FOD), F, by modeling the DW–MR signal as the convolution over the unit sphere of the response of a single fiber, $E_f(\mathbf{q}, \hat{\mathbf{x}})$, with the FOD as:

$$E(\mathbf{q}) = \int E_f(\mathbf{q}, \hat{\mathbf{x}})F(\hat{\mathbf{x}})d\hat{\mathbf{x}}. \tag{4.12}$$

In order to linearize the model, we can represent F by using a linear basis as

$$F(\hat{\mathbf{x}}) = \sum_{k=1}^{K} \beta_k \gamma_k(\hat{\mathbf{x}}). \tag{4.13}$$

By substituting (4.13) in (4.12) we have

$$E(\mathbf{q}) = \sum_{k=1}^{K} \left(\beta_k \int E_f(\mathbf{q}, \hat{\mathbf{x}}) \gamma_k(\hat{\mathbf{x}}) \right), \tag{4.14}$$

which is a linear system $\mathbf{E} = \mathbf{XB}$, where $\mathbf{E} = (E(\mathbf{q_1}), \ldots, E(\mathbf{q_M}))^T$ is the vector of normalized measurements $\mathbf{B} = (\beta_1, \ldots, \beta_K)$ is the vector of basis weights and the entries of matrix $\mathbf{X}$ are

$$X_{i,k} = \int E_f(\mathbf{q_i}, \hat{\mathbf{x}}) \gamma_k(\hat{\mathbf{x}}) d\hat{\mathbf{x}}. \tag{4.15}$$

FOD was represented with a linear basis for spherical functions in [6, 112] and in [3] was proposed a maximum-entropy formulation of the spherical deconvolution (MESD) problem with a non–linear deconvolution kernel (a generalization of PAS method). Although the method in [3] presents good results (the non-linear formulation allows one to recover smaller errors than in the methods reported in [6, 112]) the required non-linear minimization does not guarantee to get the global minimum and requires a significant computational effort. Similarly, in [7] it is proposed a simple axial symmetric model of diffusion, where the angular distribution of fibers is computed by a deconvolution process and by assuming constant, both, mean diffusivity and perpendicular diffusivity in all the white matter (a similar assumption was used in [112]).

We note that, as methods that use the standard DT, the recovered non-parametric EAP have been used for tracking fibers. In this case the particles follow the orientations of peaks in the probability mass, see [84, 116].

In most previous works [117, 6, 115, 112, 82], large b-values (larger than $2000 s^2/mm$) or large datasets are required in order to have high contrast signals or for recovering good angular resolution, i.e. an undesirable situation. For this reason it is necessary to develop methods that reduce such requirements. For instance, we propose in a recent article [95], a regularization-based approach for recovering the underlying fiber geometry within a voxel. This method assumed that the observed tensors are a linear combination of a given tensor basis. Therefore, the aim of this approach was to compute the unknown coefficients of this linear combination. Regularization of this ill-posed problem required prior assumption on the piecewise smoothness of nerve bundle orientations. The results were satisfactory, however, the multi-tensor model

was decomposed based on diffusivity computed from a previously fitted DT (instead of the raw measurements $S(\mathbf{q}_k, \tau)$) and, thereby, leading to an information loss.

4.2 The Diffusion Basis Functions

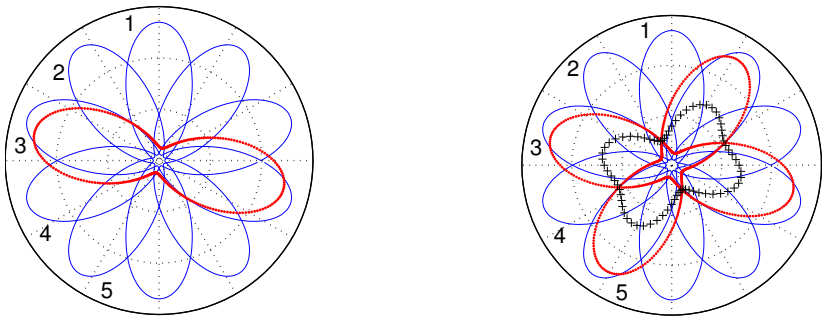
In this section we propose a discrete diffusion model based on the GMM in (4.7). In order to simplify the solution of such a model, we propose to use a set of Diffusion Basis Functions (DBF) $\{\phi\}$, which are generated from a tensorial basis as the one used in [95]. Such a tensor basis is defined as a fixed set of tensors $\bar{\mathbf{T}}$ with cardinality equal to N. The individual basis tensors $\bar{\mathbf{T}}_j$ are chosen, such that, they are distributed as uniformly as possible in the 3D space of orientations, and their anisotropy is chosen according to prior information about longitudinal and transversal fiber diffusion. For the human brain, it is reasonable to assume that the anisotropy and magnitude of the water diffusion for a single fiber in white matter is almost constant in all the volume [117, 112, 7], we will discuss this topic in section 4.10. For instance, one could expect that the longitudinal fiber diffusion is about five times the transversal one: $[\lambda_1, \lambda_2, \lambda_3] \approx [1 \times 10^{-3} mm^2/s, 2 \times 10^{-4} mm^2/s, 2 \times 10^{-4} mm^2/s]$ [29, 114]. However, these values could change among patients, so that, we instead recommend setting the basis eigen–values according to the procedure described in section 4.4.1. By fixing the basis eigen–values, we reduce the degrees of freedom for the problem. Thus, we propose to model the DW–MR signal, at each voxel, with:

$$S(\mathbf{q}_k, \tau) = \sum\nolimits_{j=1}^{N} \alpha_j \phi_j(\mathbf{q}_k, \tau) + \eta_{\mathbf{q}_k} + \varepsilon_k; \tag{4.16}$$

with $\alpha_j \geq 0$; where we define the *j–th* DBF by:

$$\phi_j(\mathbf{q}_k, \tau) = S_0 \exp(\mathbf{q}_k^T \bar{\mathbf{T}}_j \mathbf{q}_k \tau), \tag{4.17}$$

thus, $\phi_j(\mathbf{q}_k, \tau)$ is understood as the coefficient of the DW signal for the diffusion vector $\mathbf{q}_k$ due to a single fiber modeled by the basis tensor $\bar{\mathbf{T}}_j$. The non-negative α_j denotes the contribution of the *j–th* DBF $\{\phi_j(\mathbf{q}_k, \tau)\}_{k=1,\ldots,M}$. Note that the basis $\{\phi\}$ is incomplete, because the available orientations are a discretization of the 3D space (see section 4.10). So that, a residual $\eta_{\mathbf{q}_k}$ in the



(a) (b)

Figure 4-5: 2D scheme of DBFs. (a) Continuous-blue line shows the DBFs generated by an uniformly distributed tensor basis with cardinality $N = 5$; the doted-red line shows the signal $S(\mathbf{q})$ generated by an arbitrary DT. (b) Scheme for a two fiber case, the dotted-red line shows the $S(\mathbf{q}, \tau)$ measured signals for the two arbitrary tensors, the half-addition of both signals are shown in the cross-marked black line. See text for details.

signal representation is observed. By choosing a basis with a large cardinality we can diminish $\eta_{\mathbf{q}_k}$, until it becomes insignificant enough, and then neglected for practical purposes. *As can be noted, an advantage in our model (4.16) is that the unknowns are the α–coefficients because the $\phi_j(\mathbf{q}_k, \tau)$ coefficients can be pre-computed.* In fact, we need to compute the best linear combination of DBFs that reproduce the signal S. This is illustrated in the 2D scheme shown in Figure 4-5. Panel 4-5(a) shows a single fiber case where we compute the α_j values that, given a set of 5 DBFs (continuous-blue lines) reproduce the $S(\mathbf{q}_k, \tau)_{k=1,\ldots,M}$ measurements (dotted-red line) as accurately as possible; for this particular case, we expect $\alpha_3 \approx 1$. On the other hand, for the two fiber case (Panel 4-5(b)), the α coefficients should reproduce the addition (cross-marked black line); in this case, we expect $\alpha_3 \approx \alpha_5 \approx 0.5$. Note that in our approach, we do not work with the schematized continuous measurements in Figure 4-5, but with a discrete set of M samples (measurements).

Although in this work we use the free–diffusion model in (4.17) for setting the DBF, it is possible to use another diffusion model as the cylinder restricted diffusion model proposed in [106], see discussion in section 4.10.

By substituting our observation model (4.16) in the EAP equation (4.10), we obtain:

$$P(\mathbf{x}) = S_0^{-1} \sum_{j=1}^{N} \alpha_j \mathcal{F}^{-1} \left[\phi_j(\mathbf{q}, \tau) \right]. \tag{4.18}$$

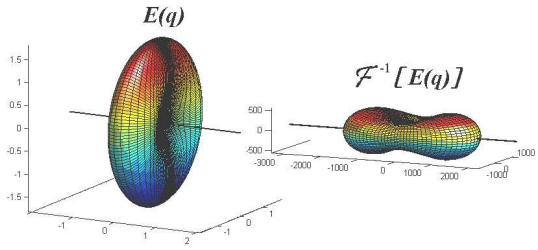


Figure 4-6: Normalized diffusion weighted signal $E(q)$ and the EAP for a basis Tensor $\bar{\mathbf{T}}_j$. The black axis denotes the PDD.

As the DBF is a Gaussian (according to the free–diffusion model) and the FT of a Gaussian results in a Gaussian $\left(\mathcal{F}[g((x),\Sigma)](\mathbf{w}) \propto g(\mathbf{w},\Sigma^{-1})\right)$, then in our case, the EAP is a GMM with peaks oriented along the PDDs of the corresponding basis tensors. Moreover, the peaks in $P(\mathbf{x})$ are determined by the largest α_j and therefore also the fiber orientations. This is illustrated in Figure 4-6 where we show for a given basis tensor the synthetic DW signal and the EAP computed with (4.10). As can be seen, the maxima of a single EAP in the GMM corresponds with the PDD of the associated basis tensor.

4.3 Numerical Solutions For DBF Model

In this section we present two procedures for estimating the coefficients α in (4.16). The first one is suitable for noise–free cases and the second one allows one to eliminate the pernicious effect of the noise by performing a spatial integration of the data.

We first introduce the notation that will be useful in the following. The observation model (4.16) can be written in matrix form as

$$S = \mathbf{\Phi}\alpha + \eta; \tag{4.19}$$

with $\alpha_j \geq 0, \forall j$; $\mathbf{\Phi}$ is an $M \times N$ matrix where the j–th column corresponds to the j–th DBF ($\mathbf{\Phi}_j = [\phi_j(\mathbf{q}_k,\tau)]_{k=1,\ldots,M}$) and $S \in \Re^M$ is the vector composed by all the DW signals. Note that because of our requirements, the matrix $\mathbf{\Phi}$ is rectangular ($M < N$): we want to acquire as few as possible S signals and to recover solutions with a high angular resolution; for instance we use $N = 129$ DBFs and $M = 23$ DW-MR images. Consequently, we have more unknowns, α's, than data, S's; the problem (4.19) is ill-posed and should be constrained or regularized in

order to compute a meaningful solution.

In the next subsections we present two algorithms for computing the best α vector by means of introducing prior information about desired features in the solution.

4.3.1 Basis Pursuit Algorithm

The development of algorithms for compact signal representation has been an important research area, we refer the reader to Section 3.4 for a survey.

In our notation, (4.19) is the mathematical model for representing the decomposition of the signal $S \in \Re^M$ as a linear combination of *atoms* $\mathbf{\Phi}_j$ in the dictionary $\mathbf{\Phi}$.

In [30], the Basis Pursuit (BP) technique was proposed for solving the problem (4.19), i.e., for computing the α coefficients. Based on the BP framework, we propose to compute a solution to (4.19) by means of an LP problem of the form:

$$\begin{aligned} \min \quad & \|\alpha\|_1 = \textstyle\sum_j \alpha_j = \hat{e}^T \alpha \\ \text{subject to} \quad & \mathbf{\Phi}\alpha = S, \\ & \alpha_j \geq 0 \ , \forall j, \end{aligned} \tag{4.20}$$

where $\hat{e}$ is a vector with all its components equal to one (we can use just $\hat{e}^T\alpha$ instead of $\|\alpha\|_1$ since the sign of the components of α is already constrained). Because of the noise, and given that $\mathbf{\Phi}$ is an uncomplete dictionary, the signal reconstruction constraint in (4.20) could not to be fully accomplished, resulting in an over-constrained LP problem. Therefore an appropriate minimization procedure is required: an interior-point method which tries to minimize the magnitude of the residual vector $\eta_\alpha = \mathbf{\Phi}\alpha - S$, see [77]. In our experiments we used the powerful Mehrotra's primal-dual predictor-corrector algorithm (see [74, 77]) that computes the results in a fraction of the computational effort required by other less-sophisticated interior point methods. The BP method has shown, in general, a better performance with respect to other pursuit techniques as, for instance, Matching Pursuit (MP) [69]. The BP method represents with few α coefficients the *atoms* that best fit the local structures.

4.3.2 Spatial and Coefficient-Contrast Regularization

The adverse effect of noise or a limited number of measurements $S(\mathbf{q}_k, \tau)$ could possibly lead most methods to miss the original fiber directions. In such situations, the BP method could erroneously estimate the α coefficients: they may not correspond to the correct axonal fiber orientations or may not indicate the right number of fibers within each voxel. In such a case, the use of a spatial regularization diminishes the noxious noise effect [31, 86]. In this work, in order to reduce such an adverse effect, we propose to filter the α coefficients and therefore to introduce prior knowledge about the piecewise smoothness assumption on the axon fibers orientation and for promoting high contrast in the α–coefficients.

In our notation, a voxel position is denoted by $r = [x, y, z]$, such that α_{jr} is the α_j–*th* coefficient at the r voxel position and $\mathcal{N}_r$ denotes the second order spatial neighborhood of r: $\mathcal{N}_r = \{s : |r - s| < 2\}$. Therefore the α_{jr} coefficient is implicitly associated with the fraction diffusion in a given orientation (i.e. along the PDD of the associated basis tensor $\bar{\mathbf{T}}_j$), and the spatial smoothness of the α_j layer ($\forall r$) is intimately related with the fiber's spatial smoothness. Moreover, if an axon bundle has a trajectory closely parallel to the *j-th* PDD, then we expect a large value for the α_j coefficient. Thus, by smoothness, the neighbor voxels along the orientation of the fiber should have its α_j coefficient large too. Similarly, such a behavior is expected for the close-to-zero coefficients too: if a fiber is not present in a position, then it is not likely to detect its prolongation along its orientation. The above prior knowledge is coded in the regularization term [95]:

$$U_s(\alpha, r) = \sum_{s:s\in\mathcal{N}_r} \sum_j w_{jrs} \left(\alpha_j(r) - \alpha_j(s)\right)^2;$$

which penalizes the difference between neighboring coefficients along the underlying fiber orientations. Such a process is controlled with the anisotropic weight factors

$$w_{jrs} = (s - r)^T \bar{\mathbf{T}}_j (s - r) / \|s - r\|^4,$$

where the power in the denominator is a empirically fixed parameter that limites the influence of the weights [122].

Additionally, we promote high contrast in the α–coefficients for distinguishing the repre-

sentative α–coefficients (orientations) from the noisy ones and for computing a sparse solution. Thus we force each α_{jr} coefficient to be different from the arithmetic mean $\bar{\alpha}_r = \sum_j \alpha_{jr}/N$, by minimizing $U_c(\alpha, r) = -\sum_j (\alpha_{jr} - \bar{\alpha}_r)^2$, see [95]. Finally, the cost function that we propose to minimize is:

$$U(\alpha, r) = \|S - \mathbf{\Phi}\alpha_r\|_2^2 + \mu_s U_s(\alpha, r) + \mu_c U_c(\alpha, r), \tag{4.21}$$

subject to $\alpha_{jr} \geq 0$, where the non–negative control parameters μ_s and μ_c allow us to tune the amount of regularization. The potentials were chosen in order to keep the cost function (4.21) quadratic. Thus by equaling to zero the partial derivative with respect to each α_{jr} results in a constrained linear system. It can easily be solved by using a Gauss–Seidel (GS) scheme [27] (used in our experiments because of its efficient use of memory) or a conjugate gradient technique which is time efficient [77]. The non-negativity constraint over the α coefficients is accomplished along with the minimization with a particular case of the well known Gradient Projection: the negative α_{jr} values are projected to zero in each iteration [77]. The tuning of the spatial regularization parameter is quite simple: the large μ_s value eliminates noise but a too large value over–smooths the recovered solution. We found that, in our experiments, $\mu_s \in [0.5, 3.0]$ produces an adequate noise reduction. As was explained in [95], the μ_c value is gradually introduced because it is important to perform the coefficient–contrast regularization once we have an intermediate regularized solution: for each iteration $k = 1, 2, \ldots, n$, we set $\mu_c^{(k)} = \mu_c \left(1.0 - 0.95^{100k/n}\right)$ that increases to μ_c in the approximately 90% of the total number of iterations n, with $\mu_c \in [0.1, 0.5]$ for all our experiments.

4.4 Implementation Details

In this section we describe important implementation details to be taken into account for obtaining high quality results.

4.4.1 Designing the Tensor Basis

There are two main aspects: (i) the eigen–vector orientations and (ii) the eigen–values.

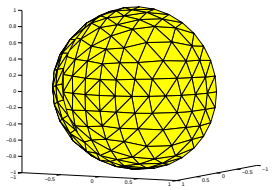


Figure 4-7: Tessellated half–sphere that generates the 129 basis orientations (vertex of the structure) used for designing the DBFs in our experiments, see text.

Orientations

The basis orientation set depends on a compromise between the desired resolution of the results and the computational effort. The procedure for obtaining the 3D balanced orientations is exactly the same as that of selecting the acquisition DW orientations in the MR machine [105], i.e. one can use the almost uniformly distributed directions given by the n-fold tessellated icosahedron hemispheres, or by using an electrostatic repulsion model [57], or by using the recent algorithm that partitions the unit sphere into regions of equal area an small diameter [64]. In particular, we used recursive tessellations of a square pyramid (having equilateral triangles as sides) that results in $\{3, 9, 33, 129, 513, 2049, \dots\}$ almost uniform orientations for $\{0, 1, 2, 3, 4, 5, \dots\}$ successive tessellations, respectively. We used the $N = 129$ orientations shown in Figure 4-7 in all our experiments. *Note that in our approach, the high angular resolution is in the tensor basis but not in the acquired signals* $S(\mathbf{q}, \boldsymbol{\tau})$.

Eigen–Values

As it was mentioned in section 4.2, we can make use of prior information about longitudinal and transversal diffusion. As the diffusion parameters may change between patients or by scale-factor effects in the signal, then it is important to determine the best set of parameters for each experiment. In present work we perform experiments using rat brain data. The optimal parameters were determined by fitting the standard DT model to the voxels in the *corpus callosum*, a well known region with high Generalized Anisotropy (GA) [83] and relatively free of crossing fibers. Then, the mean values of the fitted DTs are used for designing the basis; in particular we found $[\lambda_1, \lambda_2, \lambda_3] = [6 \times 10^{-4} mm^2/s, 2 \times 10^{-4} mm^2/s, 2 \times 10^{-4} mm^2/s]$.

Therefore, given the azimuthal (φ_j) and elevation (ψ_j) angles of the j–th basis direction,

computed according to subsection 4.4.1, the j–th basis tensor is computed as

$$\bar{\mathbf{T}}_j = R_{\varphi_j\psi_j}\text{diag}[\lambda_1, \lambda_2, \lambda_3]R^T_{\varphi_j\psi_j},$$

where $R_{\varphi_j\psi_j}$ is the corresponding rotation matrix.

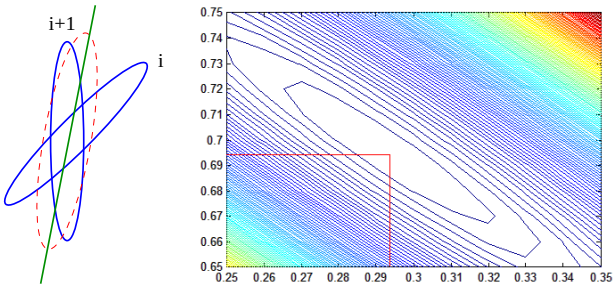
We do not constrain $\sum_j \alpha_j = 1$ because (assuming that S_0 is accurate enough) a well designed basis will automatically satisfy it. A summation different enough from 1 indicates error in the DBF design; in our experiments, for such a voxelwise summation, we obtained a mean value equal to 0.96.

4.4.2 Computation of a Continuous Solution

We note that the formulation presented in section 4.2 produces a discrete set of PDDs, that can be conveniently post-processed at each voxel for obtaining refined continuous orientations with smaller angular errors. Assuming the 2D example shown in Figure 4-8, for instance, the BP approach gives us a solution with minimum $\|\alpha\|_1$ and minimum magnitude of the error $r_\alpha = \mathbf{\Phi}\alpha - S$ (as shown in the Panel 4-8(b)), the maximum diffusion orientation (plotted in a dotted-red line), lies at an intermediate value between the orientation of the two closest basis tensors (continuous–lines) $\bar{\mathbf{T}}_i$ and $\bar{\mathbf{T}}_{i+1}$. The same situation could be present in the solutions given by the procedure in subsection 4.3.2. For computing the continuous solution we group the orientations in *clusters* and we assign an unique orientation to each *cluster*. A *cluster* $\mathbf{\Omega} = \{v_l\}$ is a set of vectors associated with the basis tensors that contribute to the GMM (i.e. their corresponding coefficients are $\alpha_l > 1 \times 10^{-2}$) and with a transitive neighborhood relationship. Thus, $v_l = PDD(\bar{\mathbf{T}}_l)$, where $PDD(\bar{\mathbf{T}}_l)$ is the first eigen-vector of the basis tensor $\bar{\mathbf{T}}_l$ and we denote by

$$\mathcal{N}_{v_l} = \{v_j = PDD(\bar{\mathbf{T}}_j) : e_{lj} \in \mathcal{E}\} \cup \{v_j = -PDD(\bar{\mathbf{T}}_j) : v_j^T v_l \leq \max_{e_{lk}\in\mathcal{E}} v_k^T v_l\} \tag{4.22}$$

the set of neighbor vectors to v_l; where $\mathcal{E}$ is the set of edges e_{ij} in the tessellation structure, so that e_{ij} is the edge that connects vectors v_i and v_j. This is schematized in a 2D example in figure 4-9, where we show the semi-sphere tessellation structure of orientations in green (i.e. the edges e_{ij}) and four orientations with representative α_j values (the red ones and the blue



(a) (b)

Figure 4-8: Example of a single fiber case. (a) Discrete solution (*cluster*) composed with two DBFs. (b) Level curves of $\|\mathbf{\Phi}\alpha - S\|_2$ for α_i (X-axis) and α_{i+1} (Y-axis) coefficients, with $\alpha_k = 0, \forall k \neq i, j$. See text for details.

ones).

The *cluster* centroid $\bar{Q} \in \Re^3$ is then computed by the weighted summation:

$$\bar{Q} = \sum_{v_l \in \mathbf{\Omega}} \alpha_l v_l. \tag{4.23}$$

Therefore we obtain a new GMM with continuous DTs (with eigen–values $[\lambda_1, \lambda_2, \lambda_3]$ and oriented along $\bar{Q}$) and mixture coefficients equal to $\left|\bar{Q}\right|$.

Due to the high sparsity in the α vector for the discrete solutions, the processed *clusters* were composed in most cases of two or three vectors. We note that, for cases with more than one fiber, the correct estimation of the number of fibers is limited by the angular resolution of the DBFs and the angle between such fibers: the estimation could fail in cases where the orientation of the fibers are closer than approximately three times the angle between two neighboring DBFs: in those cases, our method could erroneously collapse such orientations in a single cluster.

4.4.3 Avoiding Ill-conditioning in Mehrotra's Algorithm

Last iterations of Mehrotra's algorithm could involve the solution of an ill-conditioned problem of the form $\mathbf{Ax} = \mathbf{b}$, see [77]. For avoiding such a problem, we modify $\mathbf{A}$ by adding the value $\kappa = 5 \times 10^{-6}$ to the diagonal when its smallest eigen–value is less than 1×10^{-3}: we solve instead $(\mathbf{A} + \boldsymbol{\kappa}\mathbf{I})\mathbf{x} = \mathbf{b}$.

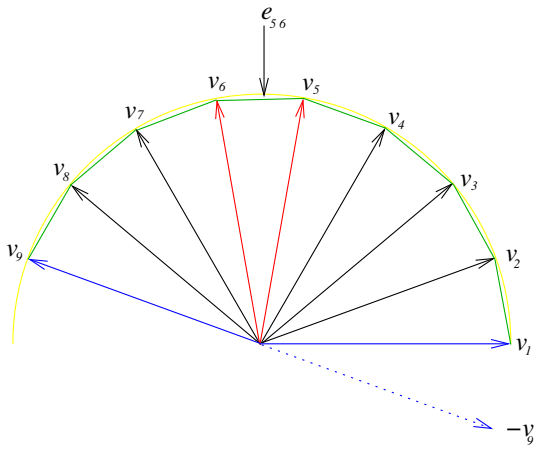


Figure 4-9: 2D scheme of the neighborhood of orientations defined in (4.22). For this case we have 2 clusters, the one composed by the vectors v_5 and v_6 (which are neighbors) and the one composed by v_1 and $-v_9$ (where v_1 and v_9 are also neighbors). See text for details.

4.4.4 Fast Convergence for the GS Solver

For cases where we need to compute a robust solution by means of spatial integration, and in order to speed up the GS solver for (4.21), we use the BP solution as the initial guess for α. Then we eliminate noise by means of a spatial integration performed by a small number of iterations of the GS approach.

4.5 Stochastic Walks for Estimating Fiber Pathways

The second stage of the method for recovering fiber pathways is the computation of virtual particles walks through the multi–tensor field. In opposition to deterministic walk methods, reported in [16, 92, 139], our approach consists of stochastic walks.

Before we present the stochastic walk method, we establish some definitions. $\mathbf{x}_t$ denotes the position vector of a particle at iteration t, d^*_{t+1} is the direction (unitary vector) that leads the particle from the position $\mathbf{x}_t$ to the next step, $\mathbf{x}_{t+1}$, and δ is the step size, i.e. $\delta = |\mathbf{x}_{t+1} - \mathbf{x}_t|$, then:

$$\mathbf{x}_{t+1} = \mathbf{x}_t + \delta d^*_{t+1}. \tag{4.24}$$

In our case, the particles pathways are closely related with the fiber structure. Therefore, we use the DBF decomposition introduced in section 4.3 in order to infer the fiber pathways.

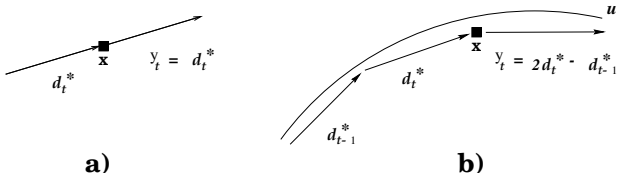


Figure 4-10: Definition of $\mathbf{y}_t$ for a) the first order walk and b)second order walk.

Thus, such decomposition coefficients control the stochastic walks of the virtual particles. One can expect that in voxels where only one fiber is present, a single coefficient of the DBFs has a significantly large value and the PDD of the associated basis tensor indicates the fiber orientation. On the other hand, in a bifurcation two coefficients have significant values. The particle should choose any of those paths corresponding to the PDD's.

In the simple case of one predominant displacement orientation, $\boldsymbol{\mu}$ (note that the PDD is associated with both $\hat{e_1}$ and $-\hat{e_1}$), we model the likelihood of a new position of a particle, given the predominant orientation and the last position, with the Dimroth–Watson distribution [70]:

$$P\left(\mathbf{x}_{t+1}\middle|\mathbf{x}_t,\boldsymbol{\mu}\right) = \frac{1}{K}\exp\left(\nu\left[\boldsymbol{\mu}^T\frac{(\mathbf{x}_{t+1}-\mathbf{x}_t)}{\delta}\right]^2\right) = \frac{1}{K}\exp\left(\nu|\boldsymbol{\mu}^T d_{t+1}^*|^2\right); \tag{4.25}$$

where the parameter, ν, regulates the concentration around $\pm\boldsymbol{\mu}$, and K is a normalization constant.

In the next subsection we discuss how, based on the DBF decomposition, the direction d_{t+1}^* and the step size δ are chosen, i.e., we expand the model given by equation (4.25) for the case of several PDD's per voxel.

4.5.1 Computation of the Displacement Direction

For computing the displacement direction, d_{t+1}^* in (4.24), we use a Bayesian estimation principle based on Markov processes, as it is now explained. Let $\mathbf{v} = \{v_1, v_2, ..., v_N\}$ be a set of fixed unitary orientation vectors (v_i and $-v_i$ are not distinguished) each one aligned with the corresponding PDD's of the basis tensors, $\bar{\mathbf{T}}$. Then, by using the Bayes Rule, we compute the probability of choosing a particular v_i orientation, as the orientation of the vector d_{t+1}, given the α–coefficient vector at the position r and the sequence of previous displacements:

Figure 4-11: Calculation of the term $P\left(d_{t+1} \parallel v_i \middle| d^*_{1:t}\right)$.

$d^*_1, d^*_2, ..., d^*_t$, denoted by $d^*_{1:t}$, with[1] :

$$P\left(d_{t+1} \parallel v_i \middle| \alpha(r), d^*_{1:t}\right) = \frac{1}{Z} P\left(\alpha(r) \middle| d_{t+1} \parallel v_i, d^*_{1:t}\right) P\left(d_{t+1} \parallel v_i \middle| d^*_{1:t}\right), \tag{4.26}$$

where Z is a normalization constant, and $x \parallel y$ denotes that x is parallel to y. The vector d^*_{t+1} is computed by performing a sampling of the posterior probability distribution (4.26). This Montecarlo procedure can be understood as a stochastic tournament. The ambiguity in the sign of the orientation is solved by choosing from d_{t+1} and $-d_{t+1}$, the one with positive inner product with the past direction, i.e. the closest direction with the past trajectory.

The first term in (4.26), the likelihood term, can be simplified by using the independence between the coefficients α, at the current position, and the sequence $d^*_{1:t}$:

$$P\left(\alpha(r) \middle| d_{t+1} \parallel v_i, d^*_{1:t}\right) = P\left(\alpha(r) \middle| d_{t+1} \parallel v_i\right). \tag{4.27}$$

We model this distribution with a probability mixture model of the form [48] [compare with (4.25)]:

$$P\left(\alpha(r) \middle| d_{t+1} \parallel v_i\right) = \frac{1}{Z_M} \sum_j^N \beta_{jr} \exp(\nu(v_i^T v_j)^2), \tag{4.28}$$

where the mixing proportion parameters β_{jr} are computed by normalizing the $\alpha_j(r)$ coefficients, i.e.

$$\beta_{jr} = \frac{\alpha_j(r)}{\sum_k \alpha_k(r)}, \tag{4.29}$$

[1] Due that: $P(A|B,C) = P(B|A,C)\, P(A|C) \,/ P(B|C)$

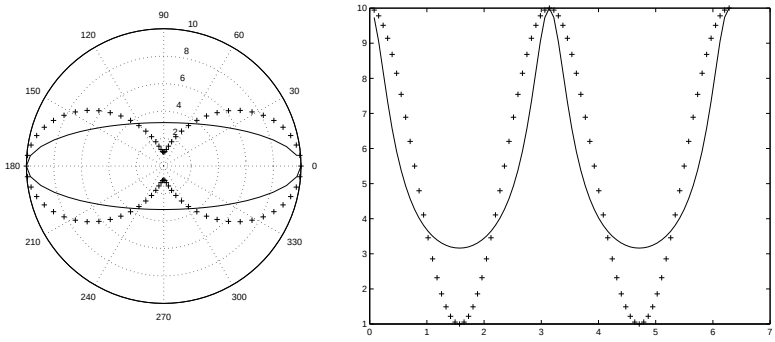


Figure 4-12: Functions for prior probability distributions: $\exp(-\mathbf{y}_t^T \bar{\mathbf{T}}_i \mathbf{y}_t)$ [in (+)] and $1/\sqrt{\mathbf{y}_t^T \hat{\mathbf{T}}_i \mathbf{y}_t}$ (in continue line). Polar (left) and rectangular (right) plots.

this satisfy $\sum_j^N \beta_{jr} = 1$. Given that the particle positions take continuous values ($\mathbf{x}_t \in \Omega \subset \mathbb{R}^3$), then, mixture coefficients $\beta_{j\mathbf{x}_t}$ are computed with a trilinear interpolation. Note that for a very large value of the ν parameter, one obtains sharper distributions [i.e. $\exp(\nu(v_i^T v_j)^2) = 1$ if $i = j$ and it is equal to zero otherwise]. In a such case the computation of the likelihood is simplified:

$$P\left(\alpha(\mathbf{x}_t) \middle| d_{t+1} \parallel v_i\right) = \beta_{i\mathbf{x}_t}. \tag{4.30}$$

This significantly reduces the computational time, specially for a large numbers of particles and displacement steps.

The prior probability in (4.26) codifies our prior knowledge about smooth fiber trajectories. Specifically, this prior promotes that d_{t+1}^* extends the sequence $d_{1:t}^*$ in a smooth way. If the stochastic walk is modeled as a Markovian process, then the smoothness constraint, on d_{t+1}^*, depends on a few past steps. For instance, we define the unitary vector: $\mathbf{y}_t = d_t^*/|d_t^*|$ for a first order Markov processes that preserves the last tendency, or

$$\mathbf{y}_t = (2d_t^* - d_{t-1}^*)/|2d_t^* - d_{t-1}^*| \tag{4.31}$$

for a second order Markov processes that preserves the last curvature, see Figure 4-10. Then, we use

$$P\left(d_{t+1} \parallel v_i \middle| d_{1:t}^*\right) = \frac{1}{Z_2} \frac{1}{\sqrt{\mathbf{y}_t^T \hat{\mathbf{T}}_i \mathbf{y}_t}} \tag{4.32}$$

as prior probability for $d_{t+1} \parallel v_i$; where Z_2 is a normalization constant, $\hat{\mathbf{T}}_i = trace[\bar{\mathbf{T}}_i]\mathbf{I} - \bar{\mathbf{T}}_i$ is the inertia tensor associated to the i^{th} basis tensor, $\bar{\mathbf{T}}_i$. A geometric interpretation shows that (4.32) corresponds to the radius, ρ, in the $\mathbf{y}$–direction of the 3D–ellipsoid defined by the level set: $(\rho\mathbf{y}_t)^T\hat{\mathbf{T}}_i(\rho\mathbf{y}_t) = Z_2$. This is illustrated by Figure 4-11; the radius in the direction $\mathbf{y}_t$ is measured for the basis tensors. In this way, it is clear that the largest value is computed with the tensor B. Consequently, in this case, it is more probable that the next walk direction is parallel to the PDD of the basis tensor B. We use (4.32) instead of the (apparently natural) Bingham's distribution [70]:

$$P\left(d_{t+1} \parallel v_i \middle| d^*_{1:t}\right) \propto \exp(-\mathbf{y}_t^T\bar{\mathbf{T}}_i\mathbf{y}_t), \tag{4.33}$$

because (4.32) is sharper for high probabilities and allows us a clearer distinction between close orientations, see Figure 4-12.

Finally, we note that the above stochastic Bayesian algorithm is related to the well–known Kalman filter [38], where our likelihood term is associated with the predictor step and our prior term is associated with the corrector step.

4.5.2 Stochastic Walk Implementation

Given that the computed walk have coarse trajectories because of the discrete nature of the tensor basis, we refine the vector d^*_{t+1} in order to obtain smooth trajectories. Such refinement consists on the use as displacement vector the resultant addition of two vectors: the previous direction step d^*_t, and the winner of the stochastic tournament, d^*_{t+1}. Note that, the refined orientation does not necessary belong to the set $\mathbf{v}$.

4.6 Results on Synthetic data

To show important features of the signals, all previous figures were generated using the diffusion–weighting parameter b=5000 s/mm^2 and a high angular resolution [so that, the $S(\mathbf{q}_k, \tau)$ signal generates contrasted plots]. In real imaging protocols lower values for previous parameters are preferred. In this section, we show synthetic results obtained using only 23 diffusion encoding orientations, relatively small b values and small ratios between the longitudinal diffusion (D_{al})

and the transversal diffusion (D_{tr}), see Figure 4-13. An example of a realistic $S(\mathbf{q}_k, \tau)$ is shown in Panel 4-13(a). Same figure shows synthetic noise-free experiments and demonstrates the capability of our method for resolving multiple fiber orientations (in yellow parallelepiped) with a small error. We show the Discrete Solution (DS) and the Continuous Solution (CS) computed according to the procedure described in subsection 4.4.2. The real axes for the maximum diffusion orientations are plotted as black lines. In Panel 4-13(f) we show, for illustrative aims, the EAP for the recovered multi-DT in Panel 4-13(e), computed with inverse FT of the GMM as indicated in [8]. We note that the peaks of the EAP (aligned, as expected, with the PDDs of the recovered multiDTs) correspond with the axes for the maximum diffusion orientations. Such EAP peaks are directly determined by the orientation of DBF with significant α values. Thus, in our approach, for bunch fiber detection we look for large α values and the computation of the EAP is not needed. For computing previous solutions, the BP solver requires approximately 35 ms per voxel, implemented in C language, on a modest PC Pentium IV, 2.8Mhz.

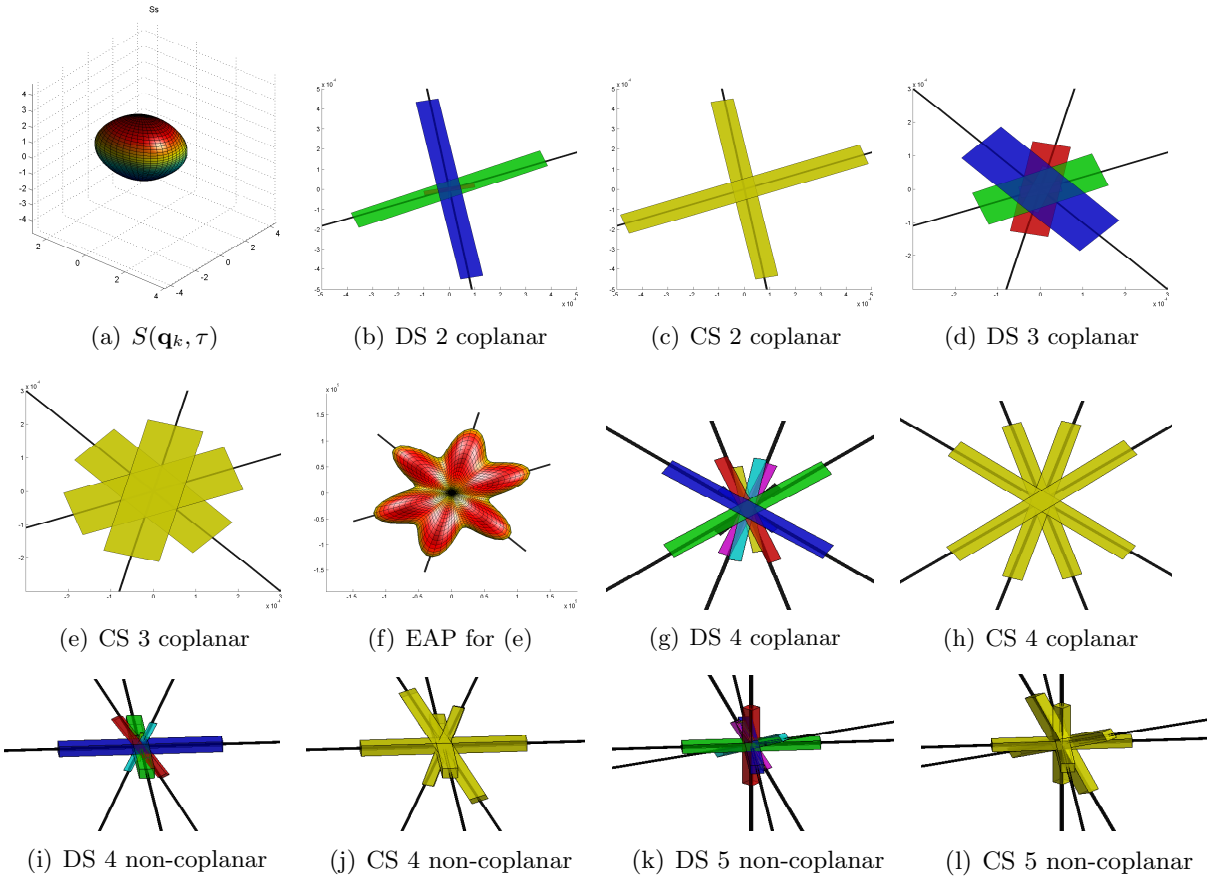
In order to analyze the expected error in real conditions, we performed 3D synthetic experiments simulating 3 non-coplanar fibers within the voxel, oriented with azimuthal and elevation angles equal to $[\pi/4, \pi/4]$, $[3\pi/4, \pi/4]$ and $[3\pi/2, \pi/4]$, respectively. In Tables 4.1, 4.2, 4.3 and Figure 4-14, we show the computed mean angular error, $\bar{\theta}$, of 100 experimental outcomes taking into account 4 important variables that directly affect the solution quality:

1. Noise robustness. The $S(\mathbf{q}, \tau)$ signals were corrupted with Rician noise with a Signal–to–Noise Ratio (SNR) (See the Appendix for the SNR definition) range from 2 (6.02 decibels (dB)) to 16 (24.08 dB), see Table 4.1.

2. Error in the diffusion basis with respect to the diffusion parameters in the data. The purpose of this set of experiments is to evaluate the sensitivity of the method to deviations in the pre–fixed DBFs with respect to the real diffusion parameters which change between voxels, see Table 4.2.

3. Method capability for recovering intra–voxel geometry with different b-values, see Figure 4-14.

4. Sensitivity to changes in the fibers compartment size, see Table 4.3.



(a) $S(\mathbf{q}_k, \tau)$ (b) DS 2 coplanar (c) CS 2 coplanar (d) DS 3 coplanar

(e) CS 3 coplanar (f) EAP for (e) (g) DS 4 coplanar (h) CS 4 coplanar

(i) DS 4 non-coplanar (j) CS 4 non-coplanar (k) DS 5 non-coplanar (l) CS 5 non-coplanar

Figure 4-13: Results of noise-free synthetic experiments. The fiber axis are plotted in black. (a) DW signal for 4 non–coplanar fiber orientations, b=1200 s/mm^2, $D_{al} = 1{\times}10^{-3}$, $D_{tr} = 2{\times}10^{-4}$. (b) and (c) Discrete Solution (DS) and refined Continuous Solution (CS) for 2 coplanar fibers, b=800 s/mm^2, $D_{al} = 6 \times 10^{-4}$, $D_{tr} = 2 \times 10^{-4}$ (these diffusion parameters are similar to those obtained from the rat brain white matter). (d), (e) and (f) DS, CS and EAP for 3 coplanar fibers, b=800 s/mm^2, $D_{al} = 6 \times 10^{-4}$, $D_{tr} = 2 \times 10^{-4}$. (g) and (h) DS and CS for 4 coplanar fibers, b=1350 s/mm^2, $D_{al} = 1 \times 10^{-3}$, $D_{tr} = 2 \times 10^{-4}$. (i) and (j) DS and CS for 4 non-coplanar fibers, b=1200 s/mm^2, $D_{al} = 1 \times 10^{-3}$, $D_{tr} = 2 \times 10^{-4}$. (k) and (l) DS and CS for 5 non-coplanar fibers, b=1000 s/mm^2, $D_{al} = 1 \times 10^{-3}$, $D_{tr} = 2 \times 10^{-4}$.

As one can see, the mean angular error, $\bar{\theta}$, is small enough for a large set of parameter variations. These results improve the methods of the state-of-the-art. The multi–tensor recovering method [60] is restricted to recover only one or two fibers orientations within a voxel, and reports a mean angular error smaller than 10 degrees for simulated fibers with an $\overline{SNR}$=80 (We note that the $\overline{SNR}$ is not defined in [60], so it can not be directly compared with our SNR definition). In our work, we obtained $\bar{\theta} \approx 5$ degrees for SNR=6 (15.56 dB), for the 3 fibers case, see Table 4.1. For SNR ≥ 6, our algorithm is capable of yielding high quality results ($\bar{\theta} \leq 6$ degrees) with realistic b values, see Figure 4-14. Additionally, we note that Figure 4-14 shows that an increment in the b values does not diminishes the $\bar{\theta}$ (differently than one would expect at first sight). In our experiments, such a phenomena is merely due to the used SNR definition (see the Appendix): an increase in the b value involves an increase in the fluctuations of the DW signal (i.e. one observes a more contrasted signal). So that, the numerator of 4.34 increases too, thus, we must increase the amount of noise in order to keep the same SNR. A similar behavior is found in the DW MR machines by chance, where an increment in the b value (that involves the use of stronger magnetic fields) yields an increment of noise.

Figure 4-15 demonstrate the spatial and contrast regularization performance, introduced in section 4.3.2. We simulate a crossing of two fibers with b=1250 s/mm^2, $D_{al} = 1 \times 10^{-3}$ mm^2/s, $D_{tr} = 2 \times 10^{-4}$ mm^2/s, SNR = 2 (6.02 dB) and a 2D tensor basis composed of $N = 30$ orientations. Panel 4-15(a) shows the noise corrupted recovered solution with the BP procedure (i.e. without regularization). The resultant orientation errors are similar to the ones reported by Perrin et al. [88], where $\bar{\theta} \approx 30$ degrees was reported in a crossing zone for a realistic phantom and, in our opinion, reveals the need of introducing a regularization mechanism for dealing with high noise data. Panels 4-15(b),4-15(c) and 4-15(d) show the noise removal effect when our proposed quadratic regularized method is used. The regularization parameters used in the experiments were $[\mu_s, \mu_c] = [1.0, 0.5]$, $[\mu_s, \mu_c] = [2.0, 0.5]$ and $[\mu_s, \mu_c] = [3.0, 0.5]$, for in Panels 4-15(b), 4-15(c) and 4-15(d), respectively.

Table 4.1: Mean angular error $\bar{\theta}$ Vs. SNR. $M = 23$ measurements, $b = 1250\ s/mm^2$, DBF parameters $[\lambda_1, \lambda_{2,3}, N] = [9\times10^{-4}, 1\times10^{-4}, 129]$, diffusion parameters $[D_{al}, D_{tr}] = [1\times10^{-3}, 2\times 10^{-4}]$, compartment sizes $\varsigma_i = 1/3, (i = 1, 2, 3)$.

SNR	$\bar{\theta}$
2 (6.02 dB)	15.21
4 (12.04 dB)	7.75
6 (15.56 dB)	5.29
8 (18.06 dB)	3.68
10 (20.00 dB)	3.66
12 (21.58 dB)	2.74
14 (22.92 dB)	2.15
16 (24.08 dB)	1.85

Table 4.2: Mean angular error $\bar{\theta}$ Vs. Basis Parameters. $N = 129$, $M = 23$ diffusion measurements, $b = 1250\ s/mm^2$, $SNR = 6$ (15.56 dB), diffusion parameters $[D_{al}, D_{tr}] = [1 \times 10^{-3}, 2 \times 10^{-4}]$, compartment sizes $\varsigma_i = 1/3, (i = 1, 2, 3)$.

λ_1	$\lambda_{2,3}$	$\left\|\bar{\mathbf{T}} - \mathbf{D}_{D_{al},D_{tr}}\right\|_F$	$\bar{\theta}$
8.50×10^{-4}	0.5×10^{-4}	1.5×10^{-4}	5.45
9.00×10^{-4}	1.0×10^{-4}	1.0×10^{-4}	5.77
9.50×10^{-4}	1.5×10^{-4}	5.0×10^{-5}	5.02
1.00×10^{-3}	2.0×10^{-4}	0.0	5.46
1.05×10^{-3}	2.5×10^{-4}	5.0×10^{-5}	5.11
1.10×10^{-3}	3.0×10^{-4}	1.0×10^{-4}	4.96
1.15×10^{-3}	3.5×10^{-4}	1.5×10^{-4}	5.60

Table 4.3: Mean angular error $\bar{\theta}$ Vs. Compartment sizes ($[\varsigma_1, \varsigma_2, \varsigma_3]$). $M = 23$ diffusion measurements, $b = 1250\ s/mm^2$, $SNR = 8$ (18.06 dB), tensor basis parameters $[\lambda_1, \lambda_{2,3}, N] = [9 \times 10^{-4}, 1 \times 10^{-4}, 129]$, diffusion parameters $[D_{al}, D_{tr}] = [1 \times 10^{-3}, 2 \times 10^{-4}]$.

compartment sizes	$\bar{\theta}$	Mean Recovered $[\bar{\varsigma}_1, \bar{\varsigma}_2, \bar{\varsigma}_3]$
$[0.333, 0.333, 0.333]$	3.90	$[0.279, 0.283, 0.279]$
$[0.433, 0.283, 0.283]$	7.15	$[0.363, 0.220, 0.221]$
$[0.533, 0.233, 0.233]$	14.16	$[0.439, 0.186, 0.183]$
$[0.633, 0.183, 0.183]$	19.27	$[0.510, 0.150, 0.159]$

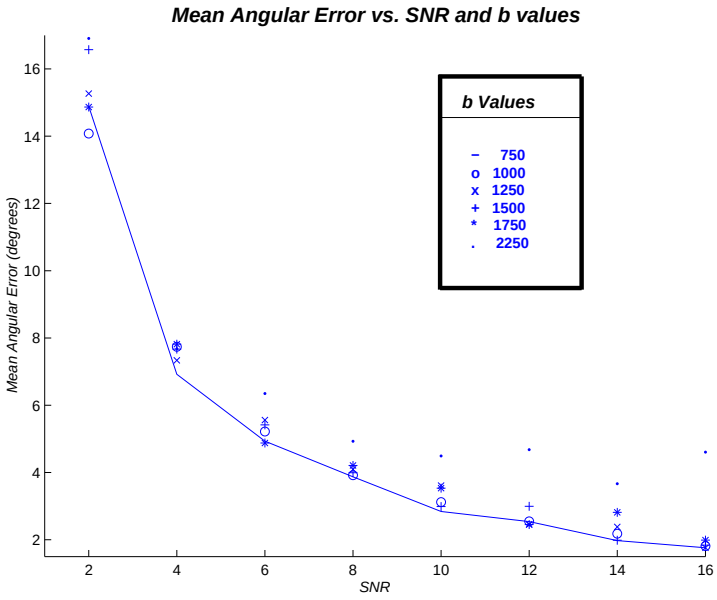


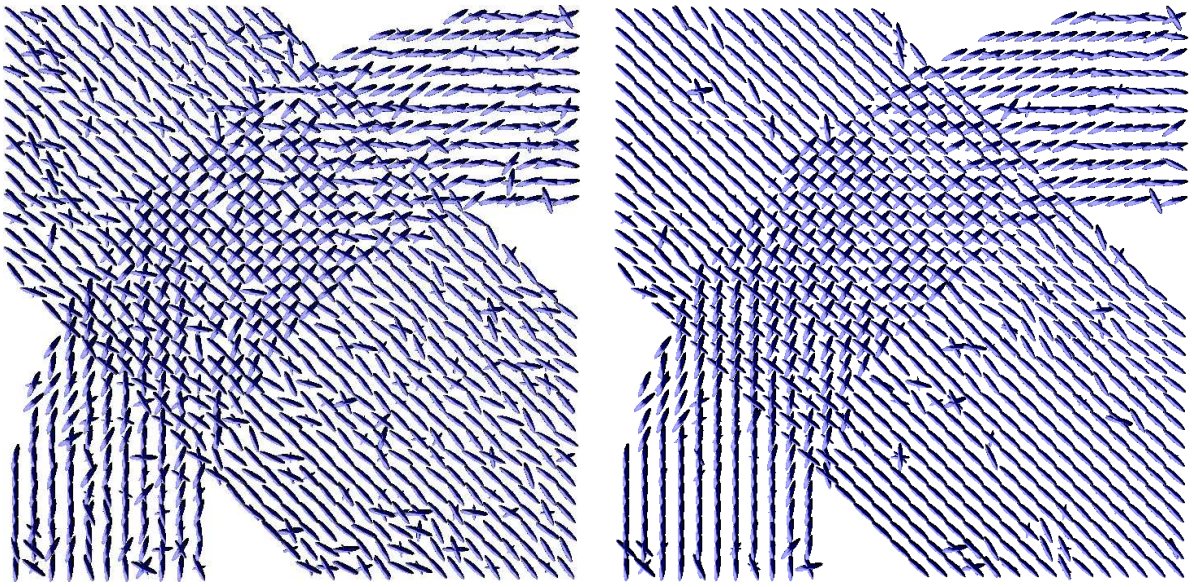
Figure 4-14: Mean angular error $\bar{\theta}$ Vs. SNR and b-values. $M = 23$ diffusion measurements, tensor basis parameters $[\lambda_1, \lambda_{2,3}, N] = [9\times10^{-4}, 1\times10^{-4}, 129]$, diffusion parameters $[D_{al}, D_{tr}] = [1 \times 10^{-3}, 2 \times 10^{-4}]$, compartment sizes $\varsigma_i = 1/3, (i = 1, 2, 3)$. See text for details.

4.7 Results on Rat Brain DW-MR data

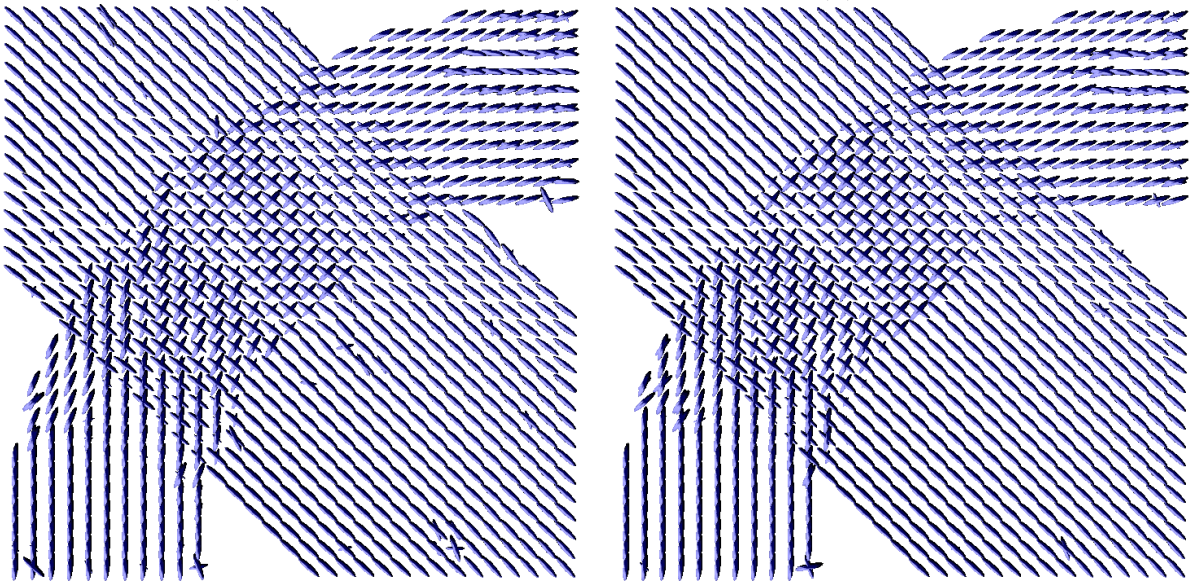
First, we describe the DW–MR acquisition procedure and the image properties. This information and the data were provided by the Department of Biochemistry and Molecular Biology of the University of Florida.

Under deep anesthesia, a Sprague Dawley rat was transcardially exsanguinated then perfused with a fixative solution of 4% paraformaldehyde in phosphate buffered saline (PBS). The corpse is stored in a refrigerator overnight then the brain was extracted and stored in the fixative solution. For MR measurements, the brain was removed from the fixative solution then soaked in PBS, without fixative, for about 12 hours (overnight). Prior to MR imaging, the brain was removed from the saline solution and placed in a 20 mm tube with fluorinated oil (Fluorinert FC-43, 3M Corp., St. Paul, MN) and held in place with plugs. Extra care was taken to remove any air bubbles in the sample preparation.

The multiple-slice diffusion weighted image data were measured at 750 MHz using a 17.6 Tesla, 89 mm bore magnet with Bruker Avance console (Bruker NMR Instruments, Billerica, MA). A spin-echo, pulsed-field-gradient sequence was used for data acquisition with a repetition



(a) $\bar{\theta} = 9.21°$ (b) $\bar{\theta} = 7.76°$



(c) $\bar{\theta} = 6.98°$ (d) $\bar{\theta} = 6.47°$

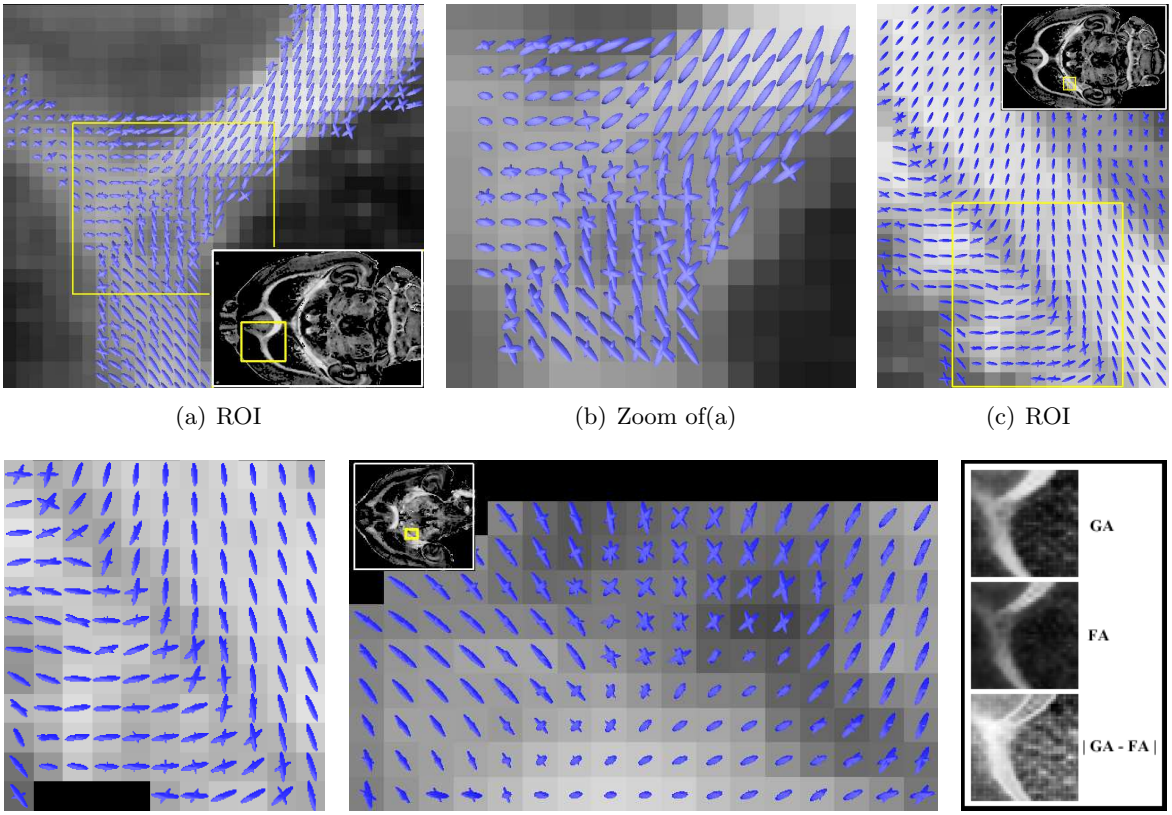
Figure 4-15: Simulated crossing fibers, the signals were corrupted with Rician noise, SNR = 2.0 (6.02 dB). (a) Solution without regularization (BP based method). (b), (c) and (d) noise removal effect with the quadratic formulation and the mean angular errors $\bar{\theta}$. The solution in (d) is over–smoothed because of a too large μ_s value. See text for details.

time of 1400 ms and an echo time of 28 ms. The diffusion weighted gradient pulses were 1.5 ms long and separated by 17.5 ms. A total of 32 slices, with a thickness of 0.3 mm, were measured with an orientation parallel to the long-axis of the brain (slices progressed in the dorsal-ventral direction). These slices have a field-of-view 30 mm x 15 mm in a matrix of 200 x 100. The diffusion weighted images were interpolated to a matrix of 400 x 200 for each slice. Each image was measured with 2 diffusion b weights: 100 and 1250 s/mm^2. Diffusion-weighted images with 100 s/mm^2 were measured in 6 gradient directions determined by a tetrahedral based tessellation on a hemisphere. The images with a diffusion-weighting of 1250 s/mm^2 were measured in 46 gradient-directions, which are also determined by the tessellation on a hemisphere. The 100 s/mm^2 images were acquired with 20 signal averages and the 1250 s/mm^2 images were acquired with 5 signal averages in a total measurement time of approximately 14 hours. In our DBF based reconstruction, we used only the DW images with b=1250 s/mm^2.

Representative results for this rat brain data are shown in Figure 4-16. The GMM model is computed for each position plotted and shown as overlapped ellipsoids. The processed brain regions are indicated by the highlighted boxes in the GA map. The intersecting fibers of cingulum and corpus callosum are seen in Panels (a) and (b) (see Plate 111 and Figure 111, Paxinos and Watson [87]). In Panels (c) and (d), the detailed fiber structure of the fimbria of the hippocampus can be seen, that illustrates the entry of fibers into the fimbria from surrounding structures. This detailed analysis shows that the computed fiber orientations appear to be congruent with the prior anatomical knowledge for those regions. Note that according to Panel 4-16(d), a significant difference between the GA (computed from a 6-rank tensor [81]) and FA map are found in the crossing zone, the same region where we detected more than one fiber per voxel (as noted in [33]).

The capabilities of the regularization presented in section 4.3.2 are shown in Figure 4-17, note how the noise effect is eliminated and the obtained results with $M = 23$ measurements are equivalent to the ones obtained with $M = 46$ measurements without regularization.

Finally, we show in Figure 4-18 a region of decussation in the cerebellum, in which we recovered voxels with three fiber bundles using the BP approach (i.e. without spatial regularization). Note that the region is composed of voxels with two and three maximum diffusion orientations; in particular, in the center we can observe voxels with three spatially congruent



(a) ROI (b) Zoom of(a) (c) ROI

(d) Zoom of(c) (e) ROI (f) GA, FA Maps

Figure 4-16: Computed DTs of the GMM from a real rat brain DW-MR set superimposed over the GA axial map. Note several fiber crossings and splits. (f) GA map, FA map and their difference for the ROI in (a).

fiber orientations.

4.8 Comparisons With Q–Ball methodology

In this section, we compare the performance of the proposal method with respect to Q–Ball, a well known non–parametric method [115]. For all Q–Ball results, we compute the EAP for the 129 orientations defined in section 4.4.1 (the same orientations that we use for building the DBFs) and the integration over the equators was performed over 36 interpolated uniformly spaced points. In the kernel regression stage we used the following parameters: cutoff $\alpha_c = 20$ degrees and $\sigma_{Q-Ball} = 10$ degrees. A peak in the computed EAP was defined as the maximum value in a radius of 20 degrees.

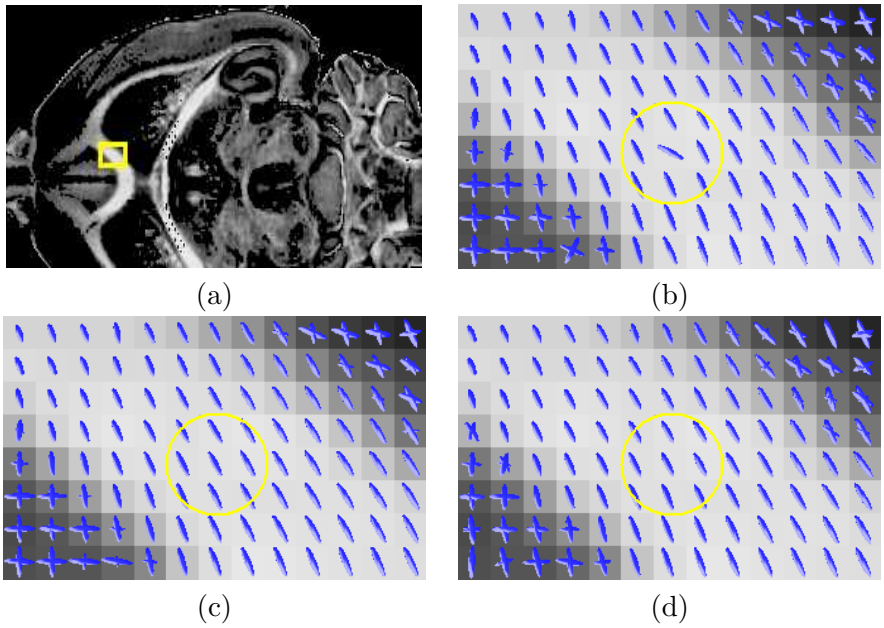


Figure 4-17: Results of regularization in the rat corpus callosum. (a) ROI in axial GA map. (b) Without spatial regularization (BP based method) by using $M = 23$ measurements, the yellow circle indicates a voxel where the noise and the reduced number of measurements produces an inaccurate result. (c) With $M = 23$ measurements and quadratic regularization: $\mu_s = 0.50$, $\mu_c = 0.18$. (d) With $M = 46$ measurements with the BP method. Note that the result obtained in (c) and (d) are equivalent for all practical purposes.

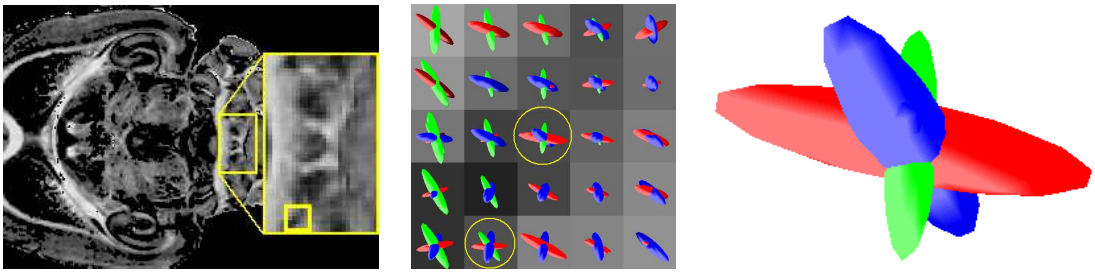
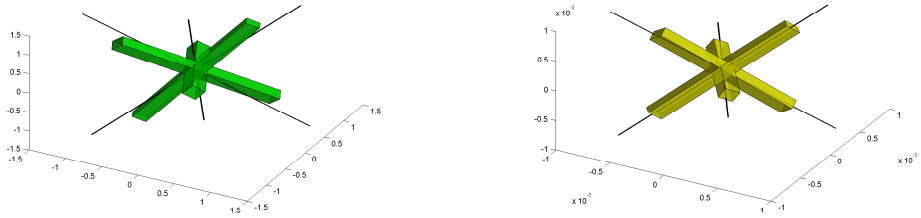


Figure 4-18: Real 3 fiber crossing in a rat cerebellum. (a) ROI GA map. (b) Region in which three fibers are present. The diffusion along **X** axis were plotted in red, along **Y** axis in green and along **Z** axis in blue. Note that the region contains an intersection of 3 fiber bundles. (c) Zoom in a voxel where the 3 bundles are crossing.



(a) (b)

Figure 4-19: Synthetic S signal generated for a three fiber crossing with compartment sizes $\varsigma_i = 1/3, (i = 1, 2, 3)$, tensor basis parameters $[\lambda_1, \lambda_{2,3}, N] = [9 \times 10^{-4}, 3 \times 10^{-4}, 129]$, diffusion parameters $[D_{al}, D_{tr}] = [1 \times 10^{-3}, 2 \times 10^{-4}]$, $M = 46$ measurements, $b = 1250$ s/mm^2 and $SNR = 5$ (13.97 dB) . (a) Result for Q–Ball, mean angle error (for the three fibers) $\bar{\theta} = 10.90$ degrees (b) Result for DBF approach, $\bar{\theta} = 3.78$ degrees.

Figure 4-19 shows a comparison, given the same signal S for a three fiber crossing with Rician noise and in realistic acquisition conditions. Note that our proposed method reports smaller mean angular error, $\bar{\theta}$, than Q–Ball.

In Figure 4-20 we show the Q–Ball solution for the rat DW-MR images. Confronting Panels 4-20(a) and 4-20(b) with Panels 4-16(b) and 4-16(d) respectively (both results without spatial regularization), the Q–Ball results presents poor performance for such conditions, i.e. low spatial coherence in the crossing zone in Panel 4-20(a) and inability in resolving the intra–voxel information (dark region) in the crossing zone in Panel 4-20(b).

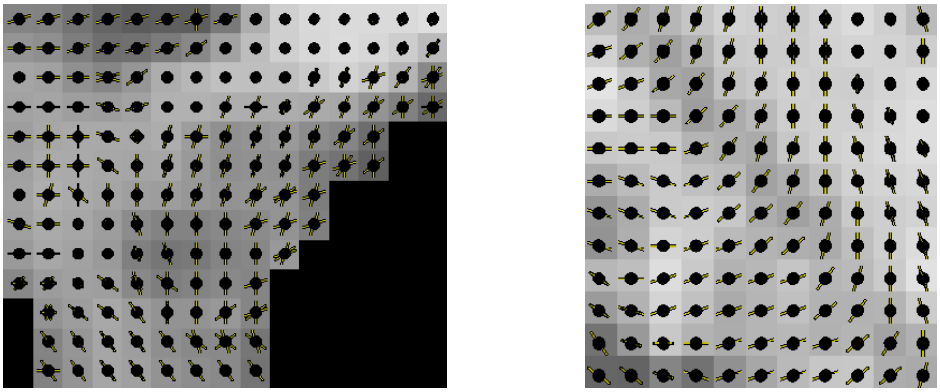
Statistical values for the performance of both methods are shown in Table 4.4. Each experiment consists of 50 Monte–Carlo outcomes with variations of the acquisition parameters. The $\bar{\theta}$ value reported by the DBF method is about half of the one obtained by the Q–Ball approach. This behavior agrees with the results on rat DW–MRI: For $M = 46$ and b=1250 s/mm^2 we expect a significant large value $\bar{\theta}$ for Q–Ball, about twice the one obtained by the DBF approach.

4.9 Results for Particle Random Walks

We present results of our random walk method on Figure 4-21. Panel 4-21(a) shows a 3D synthetic multi–tensor field simulating a fiber crossing with smooth wavy paths. We show in

Table 4.4: Mean angular error $\bar{\theta}$ for DBF and Q–Ball reconstructions. Three fiber crossing with compartment sizes $\varsigma_i = 1/3, (i = 1, 2, 3)$, tensor basis parameters $[\lambda_1, \lambda_{2,3}, N] = [9 \times 10^{-4}, 3 \times 10^{-4}, 129]$, diffusion parameters $[D_{al}, D_{tr}] = [1 \times 10^{-3}, 2 \times 10^{-4}]$. When the parameter was not under analysis we set $M = 46$ measurements, $b = 1250\ s/mm^2$ and $SNR = 6$ (15.56 dB).

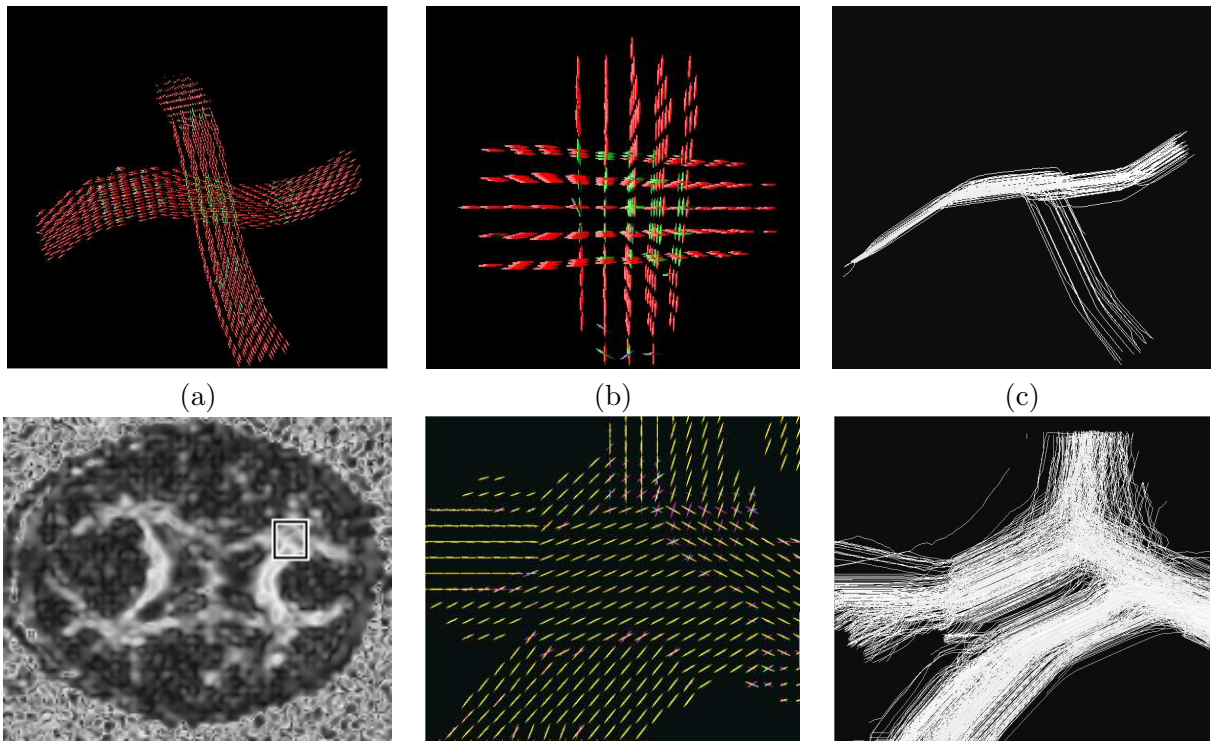
SNR→[$\bar{\theta}_{QBall}$,$\bar{\boldsymbol{\theta}}_{\mathbf{DBF}}$]	M→[$\bar{\theta}_{QBall}$,$\bar{\boldsymbol{\theta}}_{\mathbf{DBF}}$]	b→[$\bar{\theta}_{QBall}$,$\bar{\boldsymbol{\theta}}_{\mathbf{DBF}}$]
10→[8.70 , ***2.48***]	513→[7.57 , ***1.70***]	3000→[9.23 , ***3.37***]
6→[9.41 , ***4.81***]	129→[8.03 , ***3.78***]	2000→[9.48 , ***3.61***]
4→[11.02 , ***5.82***]	46→[9.57 , ***3.97***]	1250→[9.42 , ***3.77***]
2→[24.24 , ***11.41***]	23→[27.57 , ***5.43***]	900→[9.18 , ***3.48***]



(a) (b)

Figure 4-20: Q–Ball results for the rat brain DW–MRI, confront with the DBF results in 4-16(b) and 4-16(d).

Panel 4-21(b) the detail of the intersection. Panel 4-21(c) shows the particle paths of a set of 100 particles with starting point in the left part of the horizontal fiber. We note that approximately 15% of the particles are deviated to the other simulated fiber bundle. This feature is product of the stochastic nature of the particle walks and allows us to explore possible bifurcations in fiber bundles. In comparison, deterministic walk methods will recover the same trajectory for all the particles that were started at the same point. Additionally, Panels 4-21(d), 4-21(e) and 4-21(f) show the results of an experiment over a real recovered multi–tensor field. Panel 4-21(d) shows the FA of a human brain axial slice and the small square indicate the region of detail. Panel 4-21(e) shows the detail of the recovered multi–tensor field and Panel 4-21(f) the computed trajectories of the particles. In this case, the displacement of the particles was constrained to lay in the axial slice shown in Panel 4-21(d).



(a) (b) (c)

d) e) f)

Figure 4-21: Results for the proposed random walks on synthetic (a), (b) and (c), and human DW–MRI data (d), (e) and (f), see text for an explanation.

4.10 Chapter Discussion and Conclusions

The use of Basis Functions (for instance radial or kernel basis functions) that span a subspace of smooth functions is a common and successful strategy for noise reduction in signal and image processing problems. Such a strategy can be seen as an implicit regularization procedure where prior knowledge is introduced by selecting the right form of the basis function. In our case, the chosen basis functions are directly related with the signal observation model. Thus, besides promoting noise reduction, our formulation reconstructs the signal by estimating the control parameters of the diffusion process (the $\alpha-$coefficients). An important characteristic of the proposed DBFs is that they are over-complete for spanning the subspace of smooth functions: Some reconstructions can be computed with different combination of α-coefficients; for instance, because of the sparsity constraint, an isotropic diffusion can be approximated with several triads of DBFs with self-orthogonal PDD; similarly a flat (2D-isotropic) diffusion with different possible pairs of DBFs. This could be seen as a limitation of our model that makes the restoration process ill-posed. However, it only means that if the $S(q_k)$ signal does not exhibit preferential diffusion directions, then our proposal (like others methods as DT, Q–space or deconvolution methods) will be unable to recover the intra–voxel geometry.

Undefined diffusion directions can be caused by noise or tissue properties, as in gray matter or Cerebral Spinal Fluid. In this work we assume that white matter has previously been segmented from other tissues and thus the proposed model can recover the intra–voxel fiber structure for the case of low level of noise. In other cases, for relatively high level noise, a regularization process that codifies the prior knowledge about smooth fiber trajectories is proposed. Subsection 4.3.1 and 4.3.2 presented our approaches for the two noise level cases.

The present work is based on the assumption that the MR signals for a single fiber orientation are sufficiently homogeneous in the white matter tissue (as in [117, 112, 7]), so that, for each voxel, the MR signal could be explained as a linear combination of DBFs that takes into account changes only in orientation. In [112] it was noted that if the diffusion parameters change by different myelination levels, axonal diameters and axonal densities, then the diffusion parameters violate the homogeneity assumption and the relative volume fractions will not be exactly recovered. However, such errors are small and do not significantly alter the estimated fiber orientations (the most important data in axon fiber tracking). The later conclusion is

congruent with our experimental results shown in Table 4.2.

We have presented a new representation for directly obtaining the local nerve fiber geometry from DW–MR measurements. Our proposal, by means of a discrete approximation of the GMM dubbed DBF, overcomes well known difficulties of fitting a GMM to DW–MR data; i.e. our proposal:

1. Automatically computes the number of fibers and the compartment sizes within each voxel, avoiding the need of prior knowledge about the number of Gaussians, contrary to [117, 85].

2. Is capable of detecting more than 2 fibers within a voxel, which improves the state-of-the-art methods based on parametric GMM [117, 85].

3. Allows us to infer complicated local fiber geometry with DWIs collected along a sparse set of diffusion encoding directions (46, or 23 by using quadratic regularization) as opposed to techniques that use a large number of directions in HARDI data sets [117, 85, 115, 54, 82].

4. Yields small angular errors for relatively small b values (1250 s/mm^2), demonstrated by experiments.

5. Has the additional advantage of being formulated as a constrained LP or constrained quadratic optimization problem, that are solved efficiently by a parallelizable interior point method or by the solution of a bounded linear system, respectively.

To the best of our knowledge, the aforementioned properties considerably advance the state-of-the-art.

It is important to note that (4.20) uses an L-1 norm instead of an L-2 norm. In this sense, we know that the L-1 norm belongs to the robust potential category, distinct from the L2-norm [51, 96]. From an ill-posed problem the BP scheme allows us to introduce prior information about the desired solution namely: to select the sparsest solution among all the possible that minimizes the magnitude of the residual vector $r_\alpha = \mathbf{\Phi}\alpha - S$. This could be translated in the DW-MR framework as, "to explain the voxel's DW signal with as few as possible DBFs."

Because the solution is given in a parametric form, the fiber orientations are computed by basis PDDs weighted by the recovered α coefficients, so that the probability of displacement

is achieved without the need of looking for peaks in non-parametric models as in [84, 116]. Moreover, in our case, for fiber pathway tracking one can use the simple method reported in Section 4.5.

Distinct from the model–free methods (as Q-Ball, DOT, etc.), our method implicitly incorporates prior knowledge on axonal water diffusion models for the reconstruction of the diffusion signals. In particular, we use the free diffusion model because the parameters (the DT) can be easily estimated from the corpus callosum for each patient (see subsection 4.4.1). However, the proposed method can be adapted to use others axonal water diffusion models, as the cylindrical confined diffusion model [106]. In such a case it is necessary to compute the diffusion coefficient, cylindrical radius and length for the Söderman's et al. model.

Model based methods (as presented here) have the additional advantage over model–free methods of being more robust to noise because one can discard unreasonable fiber topologies; see experimental comparisons for an unique fiber region of DT-MRI versus Q–Ball results in the fiber phantom by Perrin et al. [88]. In many cases, the selection among different mathematical models is based on algorithmic (numerical and algebraic) advantages. This is the case with our approach.

The proposed method is very efficient as the DBF used in the GMM can be pre-computed by using the acquisition parameters. We demonstrated via experiments, the performance of our algorithm on synthetic and real data sets, and in the former case, the results were validated.

Finally, we presented a novel stochastic particle walk procedure based on Bayesian estimation theory and a second order Markov random process model. The procedure allows us to estimate the fiber pathways and therefore deducts the brain connectivity. The stochastic movement of the particle allows one to explore possible bifurcations on fiber bundles.

Appendix. Noise generation and SNR definition

For the MR images, the Rician noise distribution results in the magnitude of the complex number such that the real and imaginary parts were corrupted with additive independent Gaussian noise with $\mathcal{N}(0, \sigma^2)$. Thus one can simulate signals $S_\sigma(\mathbf{q}_k, \tau)$ corrupted with Rician

noise [136] as:

$$S_\sigma(\mathbf{q}_k, \tau) = \sqrt{(S(\mathbf{q}_k, \tau) + \varepsilon_1)^2 + \varepsilon_2^2};$$

where $\varepsilon_1 \sim \mathcal{N}(0, \sigma)$, $\varepsilon_2 \sim \mathcal{N}(0, \sigma)$. Signal–To–Noise–Ratio (SNR) was computed according to the ratio of the peak–to–peak distance in the signal to the Root Mean Square of the noise signal (that as convention is equal to σ [104]) as:

$$SNR(S, \sigma) = \frac{\max(S) - \min(S)}{\sigma}. \tag{4.34}$$

To the aim of correct experiment reproducibility, we prefer the above SNR convention that avoids dependency on the Direct Current (DC) component in the signal (differently to one that depends on the mean value of S). For the decibel standard, we use

$$SNR_{dB}(S, \sigma) = 20 \log_{10}(SNR(S, \sigma)).$$

Chapter 5

Application II: Multi-Valued Motion Fields Estimation for Transparent Sequences with a Variational Approach

There exists a very wide literature on apparent motion estimation, also called Optical Flow (OF) . Such a craze for optical flow is notably due to the number of applications that require some motion estimation to perform their tasks. We refer the reader to [40, 9, 12] for some reviews on this topic. Although less models are proposed concerning multiple motions, it is our conviction that considering more complex stimuli will also bring some new solutions and ideas for simple optical flow estimation.

In this chapter we propose a framework based on a finite sampling of the space of velocities. Having chosen a finite set of admissible velocities, our goal is to recover a coherent spatio-temporal field that encodes at each location the presence of one or more velocities from our sample set. To recover such a field, we start with local velocity measurements, and then we minimize an energy function that encodes our prior knowledge about the optical flow smoothness and the expected number of motions (a relatively small number, say, one or two) at a particular site of the image. A simple adaptation allows the algorithm to deal with the Random Dots

Kinematogram sequences, in both single and transparent OFs.

The chapter is organized as follows. Section 5.1 reviews and comments on some related works on multiple motion estimations. Then Section 5.2 describes the proposed framework and related notations are introduced. Section 5.3 states a discrete variational model to handle multiple motions, and the role of each term in the resulting energy is discussed. The performance of the approach is illustrated in Section 5.4, on synthetic, synthesized realistic and real sequences. A modification of our formulation is proposed in Section 5.5; such a method is capable of solving Random Dot Kinematogram sequences which are used in transparent and non–transparent motion perception experiments. We conclude and present future work in Section 5.6.

5.1 Related Work On Multiple Motion Estimation for Transparent Sequences

Motion estimation methods rely on a form for data conservation along motion trajectories and some spatial or spatiotemporal regularity. Regularity in that context corresponds to some local smoothness assumption of the motion field. The most elementary form of data conservation and probably the most used is the Lambertian assumption, or brightness constancy, which states that intensities remain constant. Given a sequence $f(\mathbf{x}, t) = f(x_1, x_2, t)$, then the conservation can be stated as the *Displayed Frame Difference Equation* (DFD)

$$f(\mathbf{x} - u, t - 1) = f(\mathbf{x}, t), \tag{5.1}$$

or its linearization, the *Optical Flow Constraint Equation* (OFC):

$$\left(u^1 \frac{\partial}{\partial x_1} + u^2 \frac{\partial}{\partial x_2} + \frac{\partial}{\partial t} \right) f(\mathbf{x}, r) = (\nabla f(\mathbf{x}, t))^T \begin{pmatrix} u^1 \\ u^2 \\ 1 \end{pmatrix} = 0; \tag{5.2}$$

where $\nabla f = (f_{x_1}, f_{x_2}, f_t)^T$ and (u^1, u^2) are the spatial components of the velocity vector u. The gradient ∇f provides an affine constraint on the velocity space and is sometimes referred to as a "motion constraint vector."

Although widely used, this model has a well known limited validity, intensities do not always remain constant due to, among other causes, changing lighting conditions, specularities and clearly it cannot cope with multiple motions, especially in the case of transparency.

Transparency can be modeled as a superposition of moving layers, a linear superposition meaning addition of layer intensities, or a generalized one [80] where intensity addition is replaced by an operation with similar formal algebraic properties such as multiplication in reflection. A simple superposition model was introduced by Burt *et al.* in [18] for the case of two motions. The observed image sequence f is assumed to come from the combination $f = P_1 \oplus P_2$ of two moving patterns P_1 and P_2 with respective motions u_1 and u_2, such that brightness constancy holds for each (P_i, u_i):

$$P_i(\mathbf{x} - u_i, t-1) = P_i(\mathbf{x}, t) \tag{5.3}$$

or its linearization

$$(\nabla P_i(\mathbf{x}, t))^T \begin{pmatrix} u_i^1 \\ u_i^2 \\ 1 \end{pmatrix} = 0.$$

In the mere case of linear superposition, i.e. when the combination operation $\oplus$ is just a pixelwise addition

$$f(\mathbf{x}, t) = P_1(\mathbf{x}, t) + P_2(\mathbf{x}, t) \tag{5.4}$$

and

$$f(\mathbf{x} - u_1, t-1) = P_1(\mathbf{x} - u_1, t-1) + P_2(\mathbf{x} - u_1, t-1). \tag{5.5}$$

From (5.3) we have that

$$P_1(\mathbf{x}, t) = P_1(\mathbf{x} - u_1, t-1). \tag{5.6}$$

By subtracting (5.4) from (5.5) and by substituting (5.6) we have

$$f(\mathbf{x} - u_1, t-1) - f(\mathbf{x}, t) = P_2(\mathbf{x} - u_1, t-1) - P_2(\mathbf{x}, t).$$

So that, the general form is

$$f(\mathbf{x}-u_i,t-1)-f(\mathbf{x},t)=D_j(\mathbf{x},t)\stackrel{def}{=}P_j(\mathbf{x}-u_i,t-1)-P_j(\mathbf{x},t) \tag{5.7}$$

where $(i,j)=(1,2)$ or $(2,1)$ and the displaced frame difference is non zero, but one of the patterns has been eliminated. In the case that the motion of each pattern P_i is constant on at least three frames at times $t-2,t-1$ and t then the "difference pattern" D_j satisfies the DFD $D_j(\mathbf{x}-u_j,t-1)-D_j(\mathbf{x},t)=0$ and assuming that u_i is known, u_j can be computed by a single motion estimation technique. Burt *et al.* then derive from this fact a three frames algorithm for estimating u_1 and u_2. They start, in a multiresolution setting, with a coarse estimate of u_1 (for instance), and use a single motion algorithm on the resulting difference pattern D_2 in order to compute an estimate of u_2. This estimate is then used to form the difference pattern D_1 and get a new estimate of u_1 from it. This process is iterated until convergence.

A more thorough study and extension of this idea is proposed in a subsequent paper [28], where a frequency domain interpretation, including multiresolution pyramid effects, is provided. In particular a "dominant velocity extraction" mechanism is explained, and is then used by Irani and Peleg in [52] (see also [53]).

Starting from the linear superposition principle, Shizawa and Mase explore in a series of papers [101, 102, 103] a frequency domain, total least squares formulation of the multiple motion problems. They start from the single motion case, the OFC constraint equation (5.2) is replaced by the spatiotemporal *linear* homogeneous one

$$(\nabla f(\mathbf{x},t))^T\begin{pmatrix}u^1\\u^2\\u^3\end{pmatrix}=0,\quad \vec{u}=(u^1,u^2,u^3)\neq 0 \tag{5.8}$$

or its frequency domain counterpart

$$\vec{u}^T\omega\hat{f}(\omega)=0$$

where $\omega=(\omega_1,\omega_2,\omega_3)$ are the spatial and temporal frequencies and $\hat{f}$ is the FT of f (the $2\pi\sqrt{-1}$ multiplicative constant has been dropped). In that case, assuming constant motion,

the best $\vec{u}$ can be retrieved as the minimizer of the energy

$$E_{single}(\vec{u}) = \frac{\vec{v}^t \left(\int \omega\omega^t |\hat{f}(\omega)^2| d\omega \right) \vec{u}}{\vec{v}^t \vec{v} \int |\hat{f}(\omega)^2| d\omega}.$$

This is a total least squares problem whose solution is given as the (an) eigenvector corresponding to the smallest eigenvalue of the 3×3 symmetric, positive (semi-)definite matrix

$$A = \int \omega\omega^t |\hat{f}(\omega)^2| d\omega$$

which is a *Structure Tensor* (see [37, 21, 132] for instance). For the recovery of n motions at a given location, the linear, first order, constraint (5.8) is replaced by a n-th order, n-th multilinear one obtained by "cascading" the linear first order ones. For example, in the two motions case, the pair $(\vec{u_1}, \vec{u_2})$ would be a zero of the bilinear symmetric map

$$(\vec{u_1}, \vec{u_2}) \mapsto \vec{u_2}^T \mathcal{H}(f) \vec{u_1} = 0 \tag{5.9}$$

where $\mathcal{H}$ is the spatiotemporal Hessian operator. Multilinear maps can be factored through linear ones using the Tensor Product construction, and this leads to a two–stage formulation of the multiple motion recovery as: first a total least squares computation on the n-th tensor power of the velocity space, which singles out one element (in fact a line) of that space, and secondly a decomposition of this element into a tensor product of n velocities. A closed-form formula for that decomposition is provided in the case $n = 2$, but becomes rapidly more complicated for higher orders. Very recently, Mota *et al.* have extended these ideas in [75] and Mühlich and Aach have proposed an algebraic framework based on homogeneous parts of symmetric algebras in [76].

The very algebraic nature of the motion constraint in frequency domain has lead Vernon to propose in [121] an algorithm for the decoupling of moving patterns, for both transparency and occlusion models. An algorithm for the specific problem of reflections is proposed by Zou and Kambhamettu in [138].

The simplest homogeneous form (with $u_i^3 = 1$) of equation (5.9) provides the *2-fold optical*

flow constraint equation as introduced by Shizawa and Mase in [102]:

$$\left(u_1^1\frac{\partial}{\partial x_1}+u_1^2\frac{\partial}{\partial x_2}+\frac{\partial}{\partial t}\right)\left(u_2^1\frac{\partial}{\partial x_1}+u_2^2\frac{\partial}{\partial x_2}+\frac{\partial}{\partial t}\right)f(\mathbf{x},t)=0 \tag{5.10}$$

This form is used by Liu *et al.* in [67] with Hermite polynomial based differentiation filters and specific checks for the presence of single or multiple motions. Darell and Simoncelli "dualize" this constraint in [32] in order to construct some Fourier "donuts" used to respond to one or more velocities. The 2-fold optical flow constraint is used in the present work, in order to build local multiple motions "probes".

The nonlinear form of this constraint provides, what one may call, the *2-fold displaced frame difference equation* used in this present work,

$$f(\mathbf{x}+\vec{u_1}+\vec{u_2},t+2)-f(\mathbf{x}+\vec{u_1},t+1)-f(\mathbf{x}+\vec{u_2},t+1)+f(\mathbf{x},t+2)=0. \tag{5.11}$$

It can be extended to more than two motions and has been used as starting point by several authors. For instance, Stuke *et al.* use it in [108] to derive a block-matching approach to the multiple motion problem. In their subsequent work [109], the authors regularized spatially the block matching solution by promoting smooth solutions with a Markov Random Field (MRF) framework, improving the noise robustness of the method. However, finding a solution results in a *computationally heavy minimization* (because of the use of a field of binary indicator variables) and complex (due to a statistical confidence test used to discern the number of motions at each pixel).

Starting with equation (5.11), Pingault *et al.* in [89] perform a N-th order Taylor expansion around velocity values. A multi-resolution non linear least squares estimation is performed, using a Levenberg-Marquardt algorithm. Recently, Auvray *et al.* proposed in [10] an algorithm based on equation (5.11). The method is also multi-resolution, uses a simplex algorithm for its initialization and adds a postprocessing step, especially efficient when the two velocities are close to the real solution.

The approaches described above are based on a single higher order constraint designed to "react" to multiple motions. On the other hand, a series of methods have been developed by

incorporating several single, low order, motion constraints. When dealing with transparency, they all use an essentially unmentioned idea of a local dominance of one of the layers in some spatiotemporal neighborhood of the image sequence. These local dominances are scattered in the image plane/volume and are associated with different layers at different positions. We will now describe a few of these approaches.

In the robust statistics approach of Black and Anandan [22], the transparency is treated through a segmentation approach. The image plane is assumed to be partitioned into regions, each but one corresponding to a parametric motion model $u = u(a)$, a being the parameter vector for the region. This is done by iteratively estimating on a region $\mathcal{O}_i$ a dominant motion $u(a_i)$, the inlier pixel region $\mathcal{R}_i$ for that motion and the outlier region $\mathcal{O}_{i+1} = \mathcal{O}_i - \mathcal{R}_i$, providing, after n iterations the decomposition

$$\mathcal{R}_1 \cup \mathcal{R}_2 \cup \ldots \mathcal{R}_n \cup \mathcal{O}_n$$

where $\mathcal{R}_i$ moves with velocity $u(a_i)$, $i = 1 \ldots n$ and $\mathcal{O}_n$ is the final outlier region. In their paper, they apply the strategy to a image pair (I_1, I_2) with $n = 2$. The motion parameters a_1 and a_2 are then assumed to represent the motion of two layers that cover the entire image plane. These layers are recovered by a nulling process

$$L_2 = I_2(\mathbf{x} - u(a_1)) - I_1$$
$$L_1 = I_2(\mathbf{x} - u(a_2)) - I_1.$$

The authors describe the process as a "no-model" one (with respect to transparency) and as a limited one. It is however clear that the authors assume that transparency is due to superposition of moving patterns, through the ways these patterns are recovered, via a nulling process similar to the one of [18].

Mixture models for multiple motion computation have been introduced by Jepson and Black in [55, 56]. A parametric layered flow model is considered. One assumes that the motion can be explained by up to N parametric motion fields with parameter vectors $\vec{a}_n$. As input, one uses motion vector constraints as mentioned above, and the probability of observing constraint $\vec{c}_r$ at location r, assuming velocity layer n is given by $p_n(\vec{c}_r | r, \vec{a}_n)$. In order to take outlier

measurements into account, a model for it is added in a "zero-th layer" $p_0(\vec{c}_r)$. Assuming then that each layer is given the probability α_i of being selected (including $i = 0$ and thus requesting that $\sum_{i=0}^{N} \alpha_i = 1$), then one can write the mixture model for a constraint $\vec{c}_k$

$$p\left(\vec{c}_r | r, (\vec{a}_i)_{i=1\ldots N}, (\alpha_i)_{i=0\ldots N}\right) = \sum_{i=0}^{N} \alpha_i p_i(\vec{c}_r | r, \vec{a}_i).$$

The problem is then to compute the best mixture and motion parameters $(\alpha_i, \vec{a_i})_i$. This is usually done using EM-like algorithms.

Ju, Black *et al* [58] proposed the "Skin and Bones" model, in which multi-layered affine models are defined on small rectangular patches of the image (bones), then an ownership field defines the likelihood that each pixel comes from a particular layer. The goal is to solve for the affine model parameters and the ownership field. This is done within a robust estimation framework using an EM-algorithm. An inter-patch regularization (skin) term introduces a regularization effect in the model parameters estimation. Another layered representation is proposed by Black *et al.* in [23]. In that approach, they consider that multiple motions may appear due to occlusions and limited forms of transparency. The method introduces models for illumination changes and specular reflections, and allows one to eliminate them, improving the computation of the optical flow of the scene. In this formulation, a set of membership weights are computed in order to indicate which layer is more likely to belong to a region. Although the method captures the changes in illumination, it does not allow to compute the optical flow of moving transparencies. Weiss and Adelson [133] and recently Rivera et al. [98] proposed EM-based approaches for computing different layered motion models in an image sequence and its segmentation based on these models. They use as prior knowledge the smooth feature of the velocities. The solution in such cases is given by a field of probabilities measures that indicates layers ownership. Last methods produce pixel-wise unimodal solutions (single motions) because of the use of a distance measure for single motions as well as their entropy controls.

This section has focused on optical flow recovery and does not include some of the related questions, and the especially important one of layer recovery, at the exception of [121, 138]. We mention here the work of Toro *et al.* [111] where the knowledge of motion is fundamental, as

opposed to the work of Sarel and Irani [100], where such a separation is performed by optimizing some correlation measures. Also non mentioned here are the perceptual/neurophysiological aspects of transparency.

5.2 Problem Statement: From Local to Global

Let us assume that we have an estimation of the likelihood of a set of velocities at each spatio-temporal position. Then our goal is to propose an approach for integrating this local information in order to get a more global and robust velocity information. This integration is necessary for dealing with complex motions (such as transparent motion sequences) and with noise, as we will see in the sequel.

First, let us define a finite sampling of the velocity space, i.e. we consider N vectors

$$\{u_1, \ldots, u_N\},$$

describing the set of possible velocities (such a predefined finite sample of the velocity space is inspired in the human visual cortex where the different cells are tuned to a specific velocity).

Given a gray-scale image sequence $f : (x,t) \in \Omega \times [0,T] \rightarrow \Re$, the input is the set of functions $d(u_i, r) \in \Re^+|_{i=1\ldots N}$ describing at each position if the velocity u_i is likely to be observable, at a **local** scale, at r position (where $r = (x,t)$ are the spatiotemporal coordinates). We show in Figure 5-1 a velocity space representation. We refer the reader to Section 5.2.1 for the estimation of $d(u_i, r)$.

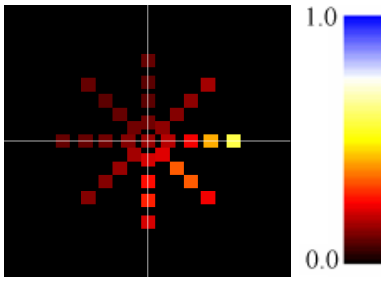


Figure 5-1: Example of a velocity space composed by 33 velocity vectors, specified through their magnitudes and orientations, respectively $\{0, 1, 2, 3, 4\}$ pixels and $\{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3}{4}\pi, \pi, \frac{5}{4}\pi, \frac{3}{2}\pi, \frac{7}{4}\pi\}$ radians. The color indicates the likelihood of each velocity.

As we have said, the goal is to compute the velocities (one or more) at each spatiotemporal position r. For this, we will associate to each possible motion u_i a variable $\alpha_i(r)$ which indicates if such a model (motion) is or not present at r. In order to obtain a robust solution, we perform an integration process over the local information provided by $d(u_i, r)$ and we introduce prior constraints about the possible number of simultaneous motions in r. This will be explained in next subsections.

5.2.1 Computing Local Velocity Information

As was explained above, one needs an initial **local** estimation of the data likelihood, a function $d(u_i, r) \in \mathbb{R}^+|_{i=1...N}$ which describes at each position $r = (\mathbf{x}, t)$ whether the velocity u_i can explain locally the apparent motion (characterized by $d(u_i, r) \approx 0$) or not (characterized by $d(u_i, r) \gg 0$). In general, d can be implemented as a norm (or quasi-norm) over a similarity operator.

Two well known similarity operators which satisfy those requirements for the single motion case are: a) The non-linear difference (Correlation-based)

$$M_{\mathcal{C}}^{(1)}(u_i)f(\mathbf{x}, t) \stackrel{def}{=} f(\mathbf{x}, t) - f(\mathbf{x} - u_i, t - 1). \tag{5.12}$$

b) The linearized version (Differential-based)

$$M_{\mathcal{D}}^{(1)}(u_i)f(x_1, x_2, t) \stackrel{def}{=} \left(u_{i1}\frac{\partial}{\partial x_1} + u_{i2}\frac{\partial}{\partial x_2} + \frac{\partial}{\partial t}\right) f(x_1, x_2, t).$$

The superscript indicates the number of displacements that these operators (and the following ones) take into account. Following Shizawa and Mase work [103], one can define an operator for two velocities as the composition:

$$M_{\mathcal{D}}^{(2)}(u_i, u_j)f(\mathbf{x}, t) \stackrel{def}{=} M_{\mathcal{D}}^{(1)}(u_i)M_{\mathcal{D}}^{(1)}(u_j)f(\mathbf{x}, t);$$

where products $\frac{\partial}{\partial r}\frac{\partial}{\partial s}$ are naturally expanded as $\frac{\partial^2}{\partial r \partial s}$. Composing instead the nonlinear corre-

lation operators $M_{\mathcal{C}}^{(1)}$ provides the nonlinear operator for two velocities:

$$M_{\mathcal{C}}^{(2)}(u_i,u_j)f(\mathbf{x},t) \stackrel{def}{=} f(\mathbf{x},t)-f(\mathbf{x}-u_i,t-1)-f(\mathbf{x}-u_j,t-1)+f(\mathbf{x}-u_i-u_j,t-2),$$

that corresponds to the distance reported in [108]. As one can see, this operator expands the non–linear single velocity one (5.12), and in this case, it imposes the constraint that velocities u_i, u_j must be constant during last three frames: t, $t-1$ and $t-2$.

We introduce here the general mechanism we have used in order to select the local velocity descriptors $d(u_i,r)$ from multiple motion operators. Given an integer $k \in [1,N]$ (in the case we concern here $k=1,2$), let us assume that we have a family of "k velocities probe" operators

$$\mathbf{M} = \{M^{(k)}(u_{i_1},\ldots,u_{i_k}), \quad 1 \leq i_1 < \cdots < i_k \leq N\}$$

where $M^{(k)}(u_{i_1},\ldots,u_{i_k})f(r) \approx 0$ if the velocity vectors $u_{i_1},\ldots,u_{i_k}$ explain the motion of the image sequence f at the position r. We build them by either cascading k correlation based filters. Such a probe will be denoted in general by $M_{\mathcal{C}}$, or k differential-based ones, these ones will be denoted $M_{\mathcal{D}}$. Then for each vector u_i, let us consider the subset $\mathbf{M}_{u_i}$ of all the operators involving u_i and define

$$d_{\mathcal{C}}(u_i,r) = \min_{M_{\mathcal{C}} \in \mathbf{M}_{u_i}} \frac{1}{k} \sum_{s \in W_r} \left(M_{\mathcal{C}} f(s)\right)^2 \tag{5.13}$$

and

$$d_{\mathcal{D}}(u_i,r) = \min_{M_{\mathcal{D}} \in \mathbf{M}_{u_i}} \frac{1}{k} \sum_{s \in W_r} \left(M_{\mathcal{D}} f(s)\right)^2, \tag{5.14}$$

where W_r is a 3×3 spatial window centered at r. In the present work, we used as input distances (5.13) and (5.14) for different experiments, showing the general framework feature of our proposal as is explained in the following. Note that because of the Taylor's series approximation, the distance (5.13) is more suitable than (5.14) for long displacements.

5.2.2 Objective: Motion Detection Variables

Based on the previous discussion, we define the problem unknowns as the vector valued field α, which is a vector at position r:

$$\alpha(r) = (\alpha_1(r), \ldots, \alpha_N(r)), \quad \alpha_i(r) \in [0,1] \quad \forall r \in \Omega \times [0,T],$$

therefore $\alpha_i(r)$ can be interpreted as the probability of observing velocity u_i at position r. Note that although the entries of $\alpha(r)$ are probabilities, $\alpha(r)$ is not a probability measure (as in [133, 98]) in the sense that it is not constrained to sum one. In other words, if two motions u_i and u_j are present at a particular pixel position, r, then we expect that both associated probabilities $\alpha_i(r)$ and $\alpha_j(r)$ will be close to one.

In Section 5.3, we propose an approach for computing the α vector field by means of a variational integration process of the local information $d(u_i, r)$ (equations (5.13) and (5.14)).

5.3 Global Motion Integration via a Variational Approach

Let $d(u_i, r)$ defined as in Section 5.2.1, we look for the velocity distribution minimizing the energy

$$\begin{aligned}
\mathcal{E}(\alpha) = \sum_r \Bigg\{ & \sum_i d(u_i, r)\alpha_i^2(r) + \mu_a(1 - \alpha_i(r))^2 && (5.15)\\
& + \frac{\mu_s}{2} \sum_{s:s\in\mathcal{N}_r} \sum_i w_i(r,s)[\alpha_i(r) - \alpha_i(s)]^2 && (5.16)\\
& - \mu_c \sum_i \left(\alpha_i^2(r) - \kappa\bar{\alpha}^2(r)\right) \Bigg\}, && (5.17)\\
& \text{subject to } \alpha_i(r) \in [0,1], \forall i;
\end{aligned}$$

with $\bar{\alpha}(r) \stackrel{def}{=} \frac{1}{N}\sum_i \alpha_i(r)$, where κ is a positive scalar and μ_a, μ_s and μ_c are some positive constants, the weights $w_i(r,s)$ will be defined in the sequel and $\mathcal{N}_r \stackrel{def}{=} \{s : r, s \in \Omega \times [0,T], \|r - s\| < 2\}$ is the spatiotemporal neighborhood of the r position.

Before going more into details, let us give a general idea on the meaning of each term. The first term (5.15) is called the attach term since it links the input (the functions d's) to the

unknown α (see Section 5.3.1). The second term (5.16), see Section 5.3.2, is a smoothing term and its role is to integrate local to global motion estimation. The last term (5.17), see Section 5.3.3, gives a prior that controls the number of active motion layers. The compromise between the last term and the attach term introduces a motion model competition mechanism.

5.3.1 Attach Term

In order to compute the presence of the i-th model, we use an approach related with the outlier rejection method [24] and with the EM formulation [23, 133, 98, 58]. In particular, using the notation in [98], this term can be expressed as $\alpha_k^2(r)\log(v_k(r)) = \alpha_k^2(r)d(u_i, r)$, where $v_k(r)$ is the likelihood of observing the motion k at the pixel r. Recall that function $d(u_i, r)$ is close to zero when the velocity u_i explains correctly the motion at position r, and is a positive large value otherwise. Minimizing first term in (5.15) with respect to $\alpha_i(r)$ produces $\alpha_i(r)$ close to 0 for high $d(u_i, r)$ values, indicating in this way that such a motion model is not likely at position r. Furthermore, second quadratic term in (5.15) avoids the null trivial solution $\alpha(r) = 0$.

5.3.2 Spatial Regularization

Term (5.16) allows us to integrate the local information by regularization, in order to obtain a more global estimation. In the previous work of Stuke *et al.*[109] was noted the necessity of a spatial regularization process. However, given that their approach is based on the computation of categorical variables, hard (combinatorial) optimization methods are required, for instance the computationally-expensive Gibbs Sampler algorithm.

Differently, in our approach, spatial–temporal smoothness means that we want to diminish the difference between the real valued vector $\alpha(r)$ and the ones $\alpha(s)$ in its neighborhood, $\mathcal{N}_r$. Given that our indicator variables are real valued, we can use differentiable potentials with the well-known algorithmic advantages. We use the approach presented in [94] for achieving such a regularization, i.e. a directional one. Therefore the smoothing process is controlled by directional fixed weights,

$$w_{irs} = \frac{(s-r)^T \bar{\mathbf{I}}_i (s-r)}{\|s-r\|^4},$$

generated from the i^{th} tensor associated to the i^{th} velocity model: $\bar{\mathbf{I}}_i = \gamma I_d + U_i U_i{}^T$, where I_d is

the identity matrix, $\gamma = 0.1$ and $U_i = [u_{i1}, u_{i2}, 1]^T / \|[u_{i1}, u_{i2}, 1]\|$ is a homogeneous-coordinate unitary vector. For small γ values these weights, w_{irs}, promote a strong smoothness along the i^{th} velocity direction, see [94]. This is illustrated in Figure 5-2. As a consequence, piece-wise smooth optical flows are recovered and the boundaries are well-defined along the velocity model, see results in Figure 5-6 and the discussion in Subsection 3.2.

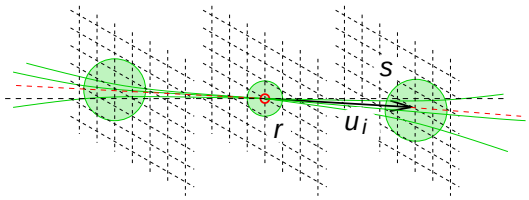


Figure 5-2: The diffusion coefficients w_i. The diffusion process is performed in the spatio-temporal neighborhood of a given point r, according to the associated velocity u_i. The domain of influence is schematically represented by the circles. The strength of the influence of the point r to the point s then depends on the spatiotemporal distance between r and s, taking into account trajectories leaving r with speed u_i.

5.3.3 Inter-Model Competition

To introduce the inter-model competition prior, fundamental in our approach, we first remind the expected behavior of the attach term (see Section 5.3.1): if velocity u_i explains locally the motion at position r, then $d(u_i, r)$ is small and consequently the corresponding α_i value is not penalized. Since our aim is to detect multiple simultaneous motions (transparent motions) we may have several α's switched-on at a given position. Thus we may have problems at sites where multiple spurious matches are locally detected, for example in homogeneous regions, where $d(u_i, r)$ is small for many (maybe all) the velocities. For this reason we need a mechanism for eliminating spurious models (to switch-off α's) and to promote the valid ones, i.e., to recover almost binary solutions. So that, our inter-model competition term should behave similarly to entropy-control potentials (as the Shannon's or Gini's used respectively in [133, 98]) in the sense of removing spurious models. In our case, however, we need a suitable term that can handle vectors that are not a probability measures and multi–modal solutions (see subsection 5.2.2).

Thus, we use the contrast term (5.17) that depends on the $\alpha(r)$ mean value and the parameter κ. The κ parameter is very useful for controlling the number of switched–on models, we

refer the reader to subsection 3.3.1 for a deep insight of this term.

5.4 Experiments

5.4.1 Algorithmic Details

The cost function $\mathcal{E}(\alpha)$ is quadratic and can be minimized by solving the linear system that results from $\frac{\partial \mathcal{E}(\alpha)}{\partial \alpha_i(r)} = 0, \quad \forall i, \forall r$, with the constraints $\alpha_i(r) \in [0, 1]$. We use the low memory requirement GS algorithm. Given an estimate α_i^n, we iterate until convergence:

$$\alpha_i^{n+1}(r) = \frac{\mu_s \sum_{s \in \mathcal{N}_r} w_i(r,s)\alpha_i^n(s) - c\mu_c \bar{\alpha}^n(r)}{d(u_i, r) + \mu_s \sum_{s \in \mathcal{N}_r} w_i(r,s) - \mu_c}.$$

The bound constraints on $\alpha_i(r)$ are then enforced by projecting non-feasible values to bounds at each iteration. We note that for obtaining a smooth algorithm convergence, it was important to keep fixed the mean of the previous iteration, $\bar{\alpha}^n(r)$, for updating the current $\alpha(r)$ vector. This can be seen as an over-relaxation strategy. We initially set $\alpha_i^0(r) = 0.5, \forall i, r$.

A Deterministic Annealing strategy in the μ_c parameter introduces the *inter–model competition* only until an approximate solution with valid representative models have predominant $\alpha_j(r)$ values: For each iteration $k = 1, 2, \ldots, n$, we set $\mu_c^{(k)} = \mu_c a_k$, where μ_c is the chosen contrast level and $a_k = 0.1 + 0.9(1 - 0.95^{(100k/n)})$ is a factor that increases to 1 in approximately 80% of the total iterations. Results are sensitive to the annealing speed of μ_c: Premature increment could lead to an incorrect solution. Nevertheless, we used the same annealing scheduling in all our experiments.

The large value μ_s eliminates noise but one that is too large blurs the motion boundaries. We used $\mu_s \in [50, 100]$ for an adequate noise reduction. Parameter μ_a increases/decreases the number of models that are present at some position, we use $\mu_a \in [1, 4]$ in our experiments. Parameter $c = 1$ performs well for most noise-free synthetic sequences. For noise-contaminated, real sequences or when the number of basis velocities is increased (so that several spurious models may be present) the prominent models are obtained by increasing this parameter, $c \in [1, 4]$.

One example of the used velocity basis is the one composed by 33 velocity vectors, specified

through their magnitudes and orientations, respectively $\{0, 1, 2, 3, 4\}$ pixels and $\{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3}{4}\pi, \pi, \frac{5}{4}\pi, \frac{3}{2}\pi, \frac{7}{4}\pi\}$ radians, see Figure 5-3. We chose it according to the present displacements in our test sequences, but a different basis can be chosen depending on the problem. This change does not affect significantly the previous parameter selection.

5.4.2 Global Coherent Motion Estimation for Non Transparent Motion Sequences

In this subsection, we experiment with non transparent motion sequences. The first example deals with the aperture problem and motion integration, while the second illustrates the performance of our algorithm on a real sequence.

Minimization Procedure Performs Motion Integration

The first experiment concerns single motion sequences and it shows how a correct global estimation is obtained based on local velocity estimations. This integration is illustrated with a synthetic sequence that consists of an oblique bar translating in the horizontal direction (Figure 5-3).

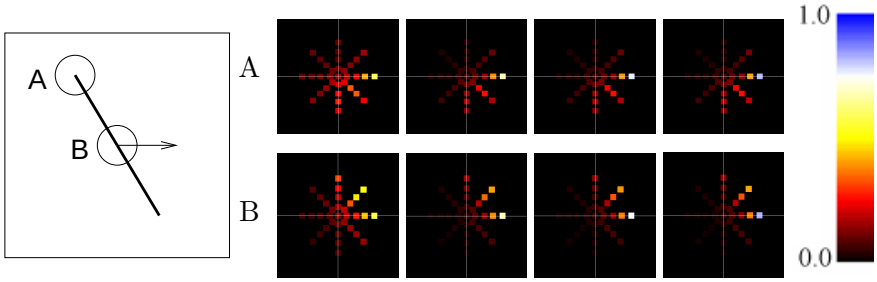


Figure 5-3: Translating bar example. We show the α's evolution (small squares denote the associated velocities) for the iteration number 1, 3, 5 and 10, for the 2 points marked in the figure on the left. The pseudocolor scale for the range in the alpha values [0,1] is shown to the right-hand side.

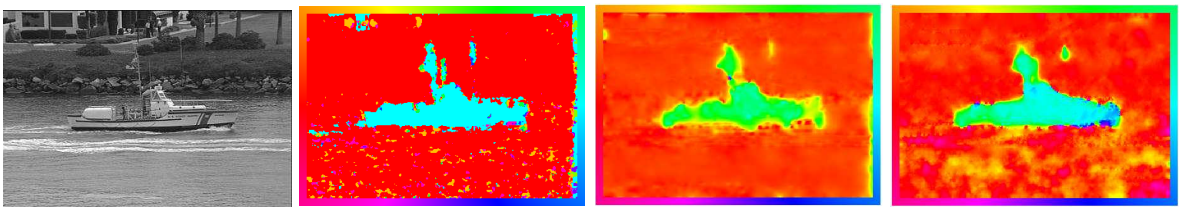
Interestingly, some psychophysics studies show how motion integration is performed in tracking tasks. In [73] is shown that the eyes will first follow the normal direction of the bar, according to 1D motion detectors. Then, after few milliseconds, there is a correction of the pursuit toward the horizontal direction, once 2D cues from ending points are integrated (That is illustrated by Figure 5-3 region B). These kind of experiments suggest that there is a parallel processing

between 1D and 2D motion signal with different temporal dynamics and that some time is needed to extract from them a stable response.

Figure 5-3 shows the evolution of the probabilities α at two given spatiotemporal location, depending on the convergence of the energy minimization. The integration phenomenon can be observed, so that the iterations of the optimization procedure can be interpreted as time evolution in real experiments.

Non Transparent Multiple Motion Sequence.

Figure 5-4 shows an example of the computed result with our algorithm on a real sequence, called `coastguards`. The background moves roughly horizontally to the left, while the foreground object, a coastguard boat, moves roughly to the right. For comparison purposes some results computed with standard variational approaches [50, 9] are displayed. Figure 5-4 (b) shows the flow corresponding to the most probable velocity at every position. The orientation of the solution is color-coded: the color associated to a each orientation is shown in the border of the image, for instance, a red pixel in the image indicates a motion to the left. In this case, as expected, unimodal solutions are obtained. As one can see, our method is capable of recovering the correct single OF, which is smooth and presents well defined boundaries.



(a) (b) (c) (d)

Figure 5-4: Optical flow in real scene (single motion case). (a) One frame of the coastguard sequence. Results computed with (b) our approach, (c) Horn and Schunck method [50] and (d) Aubert et al. approach [9].

5.4.3 Global Coherent Motion Estimation for Transparent Motion Sequences

In this subsection, we present several experiments for transparent motions. The first sequence we used is a synthetic one, the next three ones were artificially created from real photographs

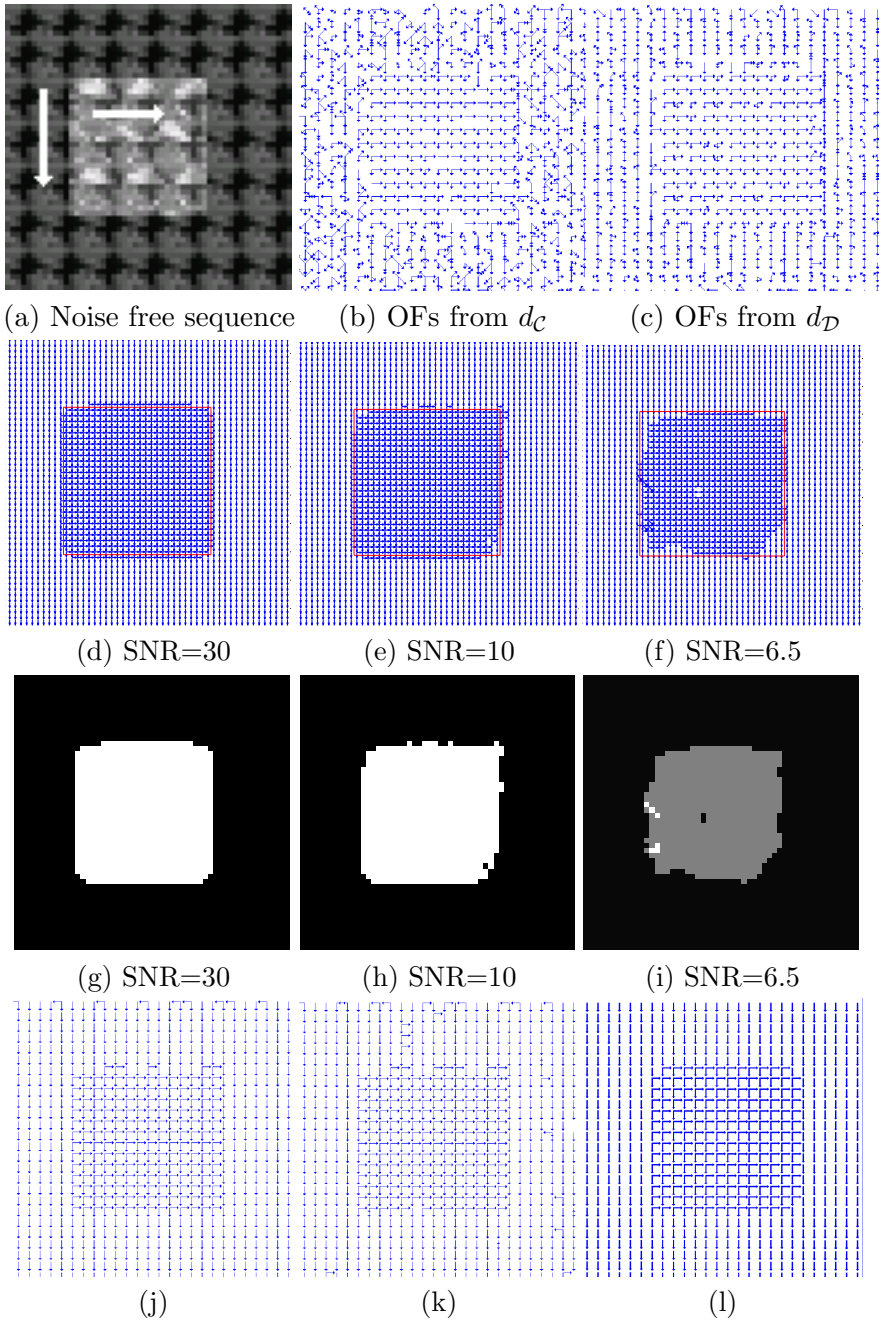
while the last one is a natural sequence with transparency and occlusions..

Illustration of the Local Measurements.

Figure 5-5 (a) shows a synthetic sequence with transparent motion similar to the one used in [109]. The sequence dimensions are 54×54×16 and it is composed by a moving background (with velocity $\hat{u} = [0, -1]$) and an overlapped moving transparent square (with velocity $\hat{v} = [1, 0]$). We now demonstrate how local measures could be highly disturbed by noise in the acquisition process. By looking only at the distance measures, several incorrect movements could be considered as valid candidates in each position. For instance, Figures 5-5(b) and 5-5(c) show the present movements associated to the minimum distance value, for the $d_{\mathcal{C}}$ distance [108] and the $d_{\mathcal{D}}$ distance [103], see Section 5.2.1. The data were noise corrupted with a Signal to Noise Ratio (SNR) equal to 30. As one can see, the quality of this first approximation is poor i.e., several incorrect movements are detected, so that a regularization process is required, which is discussed in the next paragraph.

Regularization of Local Measurements.

As mentioned above, consider the synthetic sequence shown in Figure 5-5(a). Gaussian noise has been added in order to evaluate the robustness of our proposal. Figures 5-5(d)-(i) shows the results for a frame. The percentages of pixels with a wrong estimation are 2.12%, 2.29% and 2.40% respectively. Note that the method can deal with a strong noise corruption, as a SNR=10, and shows better performance than the approach reported in [109] (see Figures 5-5 (d) and 5-5 (e), and results in [109]). Also, note that our method produces relatively good results even for an extremely corrupted sequence, as the one shown in Figures 5-5(f) and 5-5(i). For comparison purposes, Figures 5-5 (j) and 5-5 (k) show the computed OF with the computationally expensive Gibbs sampler approach for minimizing the discrete energy function reported in [109]. In [109] a deterministic relaxation ICM algorithm was used, which, differently to Gibbs sampler approach, is prone to converge to local minima. The noise-free case is shown in Figure 5-5 (j), and the SNR=30 case in Figure 5-5 (k). The shown results correspond to the computed solution after 150,000 iterations (about 2.5 hours, in a PC Pentium IV, 3.0 GHz), that results 150 times slower than our approach. For the Gibbs sampler results, the quality decreases



(a) Noise free sequence (b) OFs from $d_{\mathcal{C}}$ (c) OFs from $d_{\mathcal{D}}$

(d) SNR=30 (e) SNR=10 (f) SNR=6.5

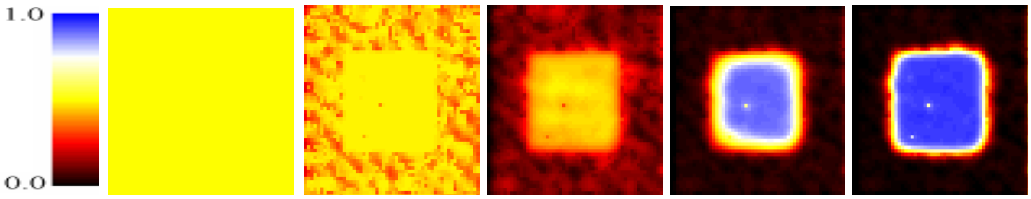
(g) SNR=30 (h) SNR=10 (i) SNR=6.5

(j) (k) (l)

Figure 5-5: First row: performance of the two local motion estimators. (a) noise-free frame. (b) Non-regularized velocities for $d_{\mathcal{C}}$ distance (SNR=30) and (c) for $d_{\mathcal{D}}$ distance. Second and third rows: Our results for different amounts of Gaussian noise with $d_{\mathcal{C}}$ distance as input. Figures (d)-(f) velocity fields and Figures (g)-(i) number of motions per pixel: in (g),(h) white = 2 velocities, and black = 1 velocity. In (i) white = 3 velocities, gray = 2 and black = 1. Last row: the Gibbs Sampler scheme: (j) results in noise-free sequence and (k) in a noise-corrupted one (SNR=30). (l) same than (d), but sampled for comparison with (j) and (k).

for noise corrupted sequences (see Figure 5-5 (k) and compare it with the one computed with the proposed method in about 1 minute in Figure 5-5 (l)).

We can verify that our spatial regularization, jointly with our inter-model competition, develops well in order to separate the velocities that are present in a region. Figure 5-6 shows the evolution of the layer associated with velocity [1,0]. Note that the layer takes a significantly large value [by growing from small values (red-yellow) to 1 (blue)] in the square region and that the contribution of this layer is completely eliminated in the background region.



It. 0 It. 1 It. 11 It. 31 It. 200

Figure 5-6: Evolution for the layer associated with velocity [1,0]. For this experiment the sequence was strongly corrupted noise (SNR=15). We show the layer values in the pseudo–color scale shown in the left. Note that the presence of the movement [1,0] in the background is pushed to zero because of the spatial regularization and the inter-model competition mechanism of the algorithm.

Realistic Texture Sequences

It is important to note that high textured sequences are relatively easy to solve using local motion measures. The real performance of a method for transparent motion should be evaluated in realistic textured scenes: recovering transparent motion in sequences with homogeneous regions presents difficulties because several models may locally explain the data. We have tested our approach using a series of experiments, where we use both synthetic and real image sequences.

To the aim of comparison, we tested the method in the sequence shown in figure 5-7, which is similar to the one presented in [32]. In this sequence, the left image is moving with velocity $\hat{u} = [1, 0]$ and the right image is moving with velocity $\hat{v} = [-1, 0]$. The transparent region corresponds to the area where the two images overlap. The images of the sequence have dimensions $64 \times 64 \times 20$. The results for frame 8 are shown in Figure 5-7. Figure 5-7(b) shows the recovered map for 1 or 2 movements, the white regions indicates the presence of 2

movements and the black ones indicates the presence of 1 movement. Figure 5-7(c) shows the recovered multi-velocity field. For the sake of clarity, we show separately the recovered field for the velocity [1,0] in Figure 5-7(d) and the recovered field for the velocity [-1,0] in Figure 5-7(e).

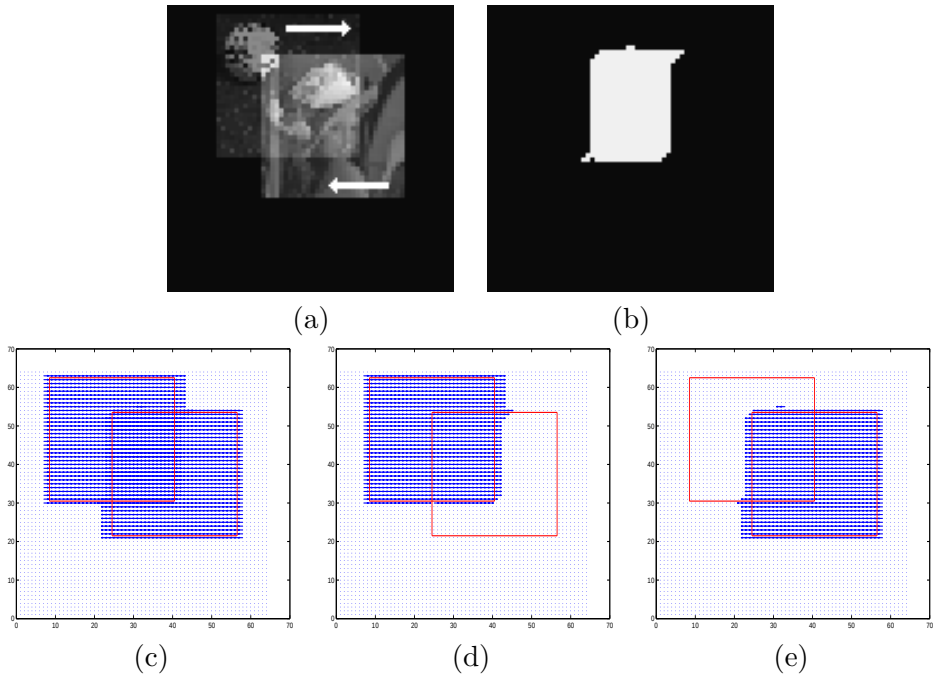
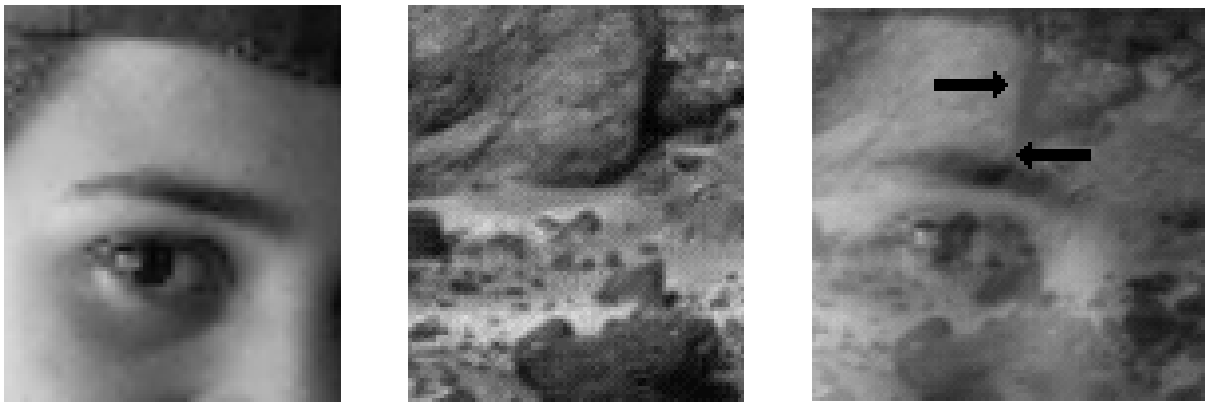


Figure 5-7: Results for the transparent sequence (a)

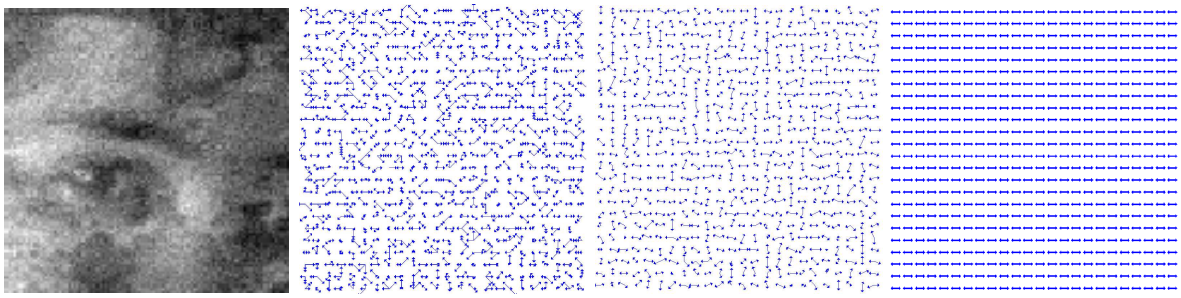
The second example presented here is composed of two photographs: a face (limited textured scene) and a Mars landscape, see Figure 5-8. Figures 5-9(b) and 5-9(c) shown the computed optical flow associated to the minimum distance value for the $d_{\mathcal{C}}$ and $d_{\mathcal{D}}$ distances, respectively. We corrupt the sequence with a strong noise (SNR=8) (Figure 5-9 (a)), and the computed velocities field is shown in Figure 5-9 (d). For this experiment distance $d_{\mathcal{D}}$ was used in the attach term. Note that the right optical flow is recovered in all the pixels regardless the high amount of noise. For the same experiment, Figure 5-10 show the evolution (with the same color–code used in Figure 5-6) of one of the two active layers, in this case the one associated with the velocity [-1,0]. One can observe that gaps corresponding to non-textured regions are correctly filled.

Figure 5-11 shows the method's performance for the case when the transparent region is composed by different combinations of velocities across time and when the image layers contains non-textured regions (realistic ones). Figure 5-11(a) shows a frame of the transparent sequence.



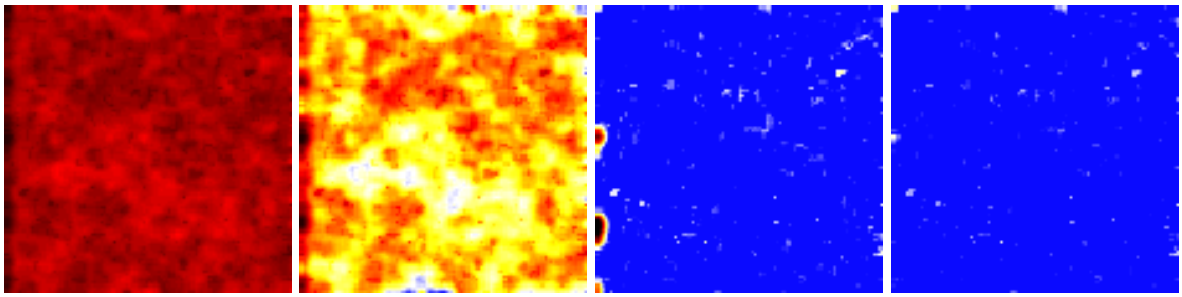
(a) (b) (c)

Figure 5-8: Realistic synthetic sequence. (a) Limited-texture image, I_1, with motion $u = [1, 0]$. (b) Rocky Martian landscape, I_2, with motion $v = [-1, 0]$. (c) Transparent generated sequence with $f = \frac{3}{5}I_1 + \frac{2}{5}I_2$.



(a) (b) (c) (d)

Figure 5-9: Transparent motion estimation on a realistic sequence corrupted with Gaussian noise. (a) Central frame highly noise corrupted (SNR=8). Velocities associated with the minimum distance for (b) $d_{\mathcal{C}}$ and (c) $d_{\mathcal{D}}$ measurements (SNR=30). In (d) the result obtained with the proposed method for the high corrupted sequence in (a) (SNR =8), note that we recovered the right velocities in all positions.



Iteration 11 Iteration 21 Iteration 31 Iteration 100

Figure 5-10: Evolution in the values for the layer associated with the velocity [-1,0].

The changing velocities are schemed in Figure 5-11(b): the background is moving with velocity [-1,0] and the airplane is moving with velocity [1,-1], [1,0] and [2,2] in equal time intervals[1]. The obtained multi-velocity vector fields are shown in Figures 5-11(c), 5-11(d) and 5-11(e). For this experiment we used the $d_{\mathcal{C}}$ distance (5.13) in the attach term.

Transparency and Occlusion in a Real Sequence

In order to show the performance of the proposed method in a real situation, we shown the results obtained for a real sequence in Figure 5-12. The sequence is composed by two robots moving slope down, see Figure 5-12(a), 5-12(b) and 5-12(c)[2]. The upper-left robot is located behind a glass, the lower-right one is located in front of the camera and the reflection of the second one is located into the upper-central part. The associated resultant vector fields are shown in Figures 5-12(d), 5-12(e) and 5-12(f) for the 3^{th}, 12^{th} and 22^{th} frame respectively). The recovered velocities were [1.5, 0.4] pixels for upper-left robot and [-1.5,0.5] for both the lower-right robot and its reflection. Note that despite the fact that the lower right-robot is moving a little faster than its reflection (easy to deduct from the projective geometry), both were associated to the same velocity model, because of the discrete nature of the velocity basis.

In this experiment, we perform a spatiotemporal Gaussian smoothing process ($\sigma = 0.5$) of the input sequence and we processed only the regions that contain displacements as is explained below. The static background was removed automatically by thresholding the difference between consecutive frames, and then applying opening-closing morphologic operators. By using this pre-process, we obtain an activity-mask that indicates the pixels where a change in time occurred, i.e. the regions where the optical needs to be computed. We used the $d_{\mathcal{C}}$ distance (5.13) in the attach term.

In all previous experiments, we computed a dense optical flow in at most 200 minimization iterations.

[1]The data sequence and the results can be downloaded at the public web site `http://www.cimat.mx/~mrivera/vision/transparent_sequences/index.html`.

[2]the AVI file and the computed flows can be downloaded at the public web site `http://www.cimat.mx/~mrivera/vision/transparent_sequences/index.html`.

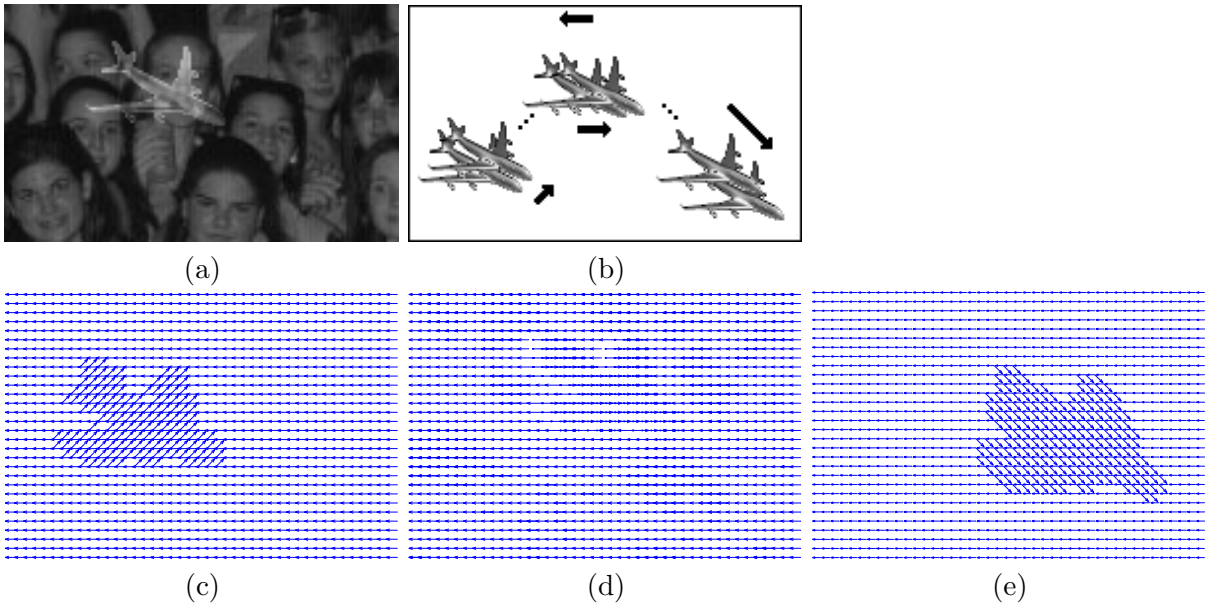


Figure 5-11: Results for a synthetic transparent sequence in which both the velocity of the background and the velocity of the object changes across the time. (a) A frame taken from the sequence. (b) Scheme of velocities: the airplane experiment velocities [1,-1],[1,0] and [2,2], and the background experiment the velocity [-1,0]. (c), (d) and (e) Sampled recovered multi-velocity fields for frames 5, 23 and 39 respectively.

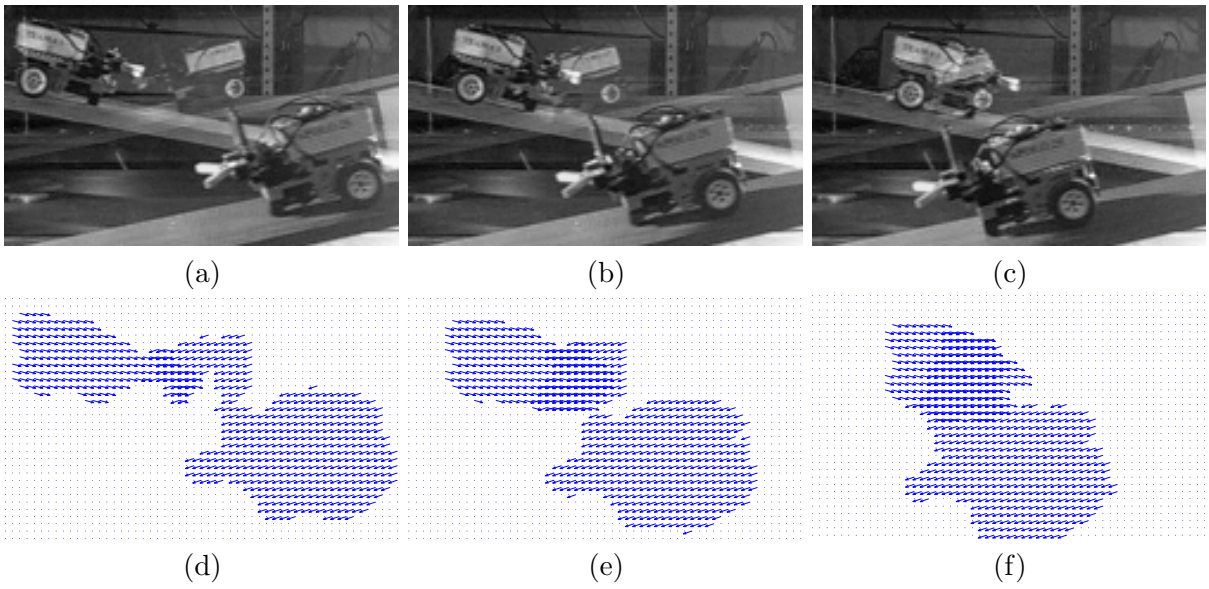


Figure 5-12: Experiment with a real transparent sequence. (a)(b)(c) Frames 3^{th}, 12^{th} and 22^{th} of the real sequence: the upper-left robot is moving slope down behind a glass, the lower-right on is moving slope down in front of camera and its reflection is captured in the upper–central part. (d), (e), (f) Sampled recovered multi-velocity fields for the respective frames.

5.5 Random Dot Kinematogram Sequences

Random Dots Kinematogram (RDK) are sequences commonly used in motion perception experiments [46, 93, 26]. RDK sequences are composed by a set of moving randomly distributed dots (see panels 5-13a and 5-14a), which are moving with different directions and/or speeds. The movement of these points in a direction causes the perception of the movement of the entire display. Moreover, if a part of this dots are moving in a direction while all the others are moving in another direction, the human observer has the perception of transparency. A full representation of these visual situations include both local and global motions. These sequences arise from the experiments in [93], that show that representations of multiple velocities do not coexist at the finest spatial scale of motion analysis. According to the experiment in [26], the transparency detection occurs in two early vision stages: a) local motion detectors which show a winner take all interaction and b) integration of motion signals over a more extended region. Neurophysiologically, there is evidence that such processes (local and integration) are carried out in different networks of the cortical area: neurons V1 encode more local motion information and spatial integrative process seems to be a property of neurons V5, so called MT.

Inspired by the motion perception model reported in [26], we propose to recover in two stages the predominant motions in RDK sequences. We first smooth noisy or sparse motion signals to disambiguate the aperture problem. Then we propagate the information to the entire display, in a integration process. The implementation is explained in the following subsections.

5.5.1 Computation of local motions

In first stage we slightly modify the cost function (5.15)–(5.17) in order to deal with the non-textured regions (areas without dots) that must be integrated to the local motions. The adapted cost function is:

$$\begin{aligned} U_a(\alpha) &= \sum_r \sum_i^N \Big[T_r \left(d^{(1)}(f, r, \mathbf{u}_i)\alpha_i^2(r) + \Psi(\alpha_i(r)) \right) \\ &+ \mu_s \sum_{s:s\in\mathcal{N}_r} w_{irs}(\alpha_i(r) - \alpha_i(s))^2 \\ &- T_r\mu_c \left(\alpha_i(r) - C_{\alpha(r)}\right)^2 + \mu_p \|\mathbf{u}_i\| \alpha_{ir}^2 \Big] \end{aligned} \tag{5.18}$$

where the threshold $C_{\alpha(r)}$ was defined in Section 3.3.1, the term

$$T_r = |\nabla f_r| \Big/ (\lambda_g + |\nabla f_r|) \tag{5.19}$$

indicates a confidence coefficient that depends on the local texture and λ_g is a user–defined parameter that controls the contrast on T_r. Regions where the aperture problem can be solved have a large T_r coefficient (this kind of regions include corners, textured ones and borders), while homogeneous regions without texture, where any movement is equal likely to occur have a T_r close to zero. We use this coefficient for detecting regions where the movement information is poor so that we need to acquire the information from its neighbor. Additionally, we use the term that assumes a prior favoring slow velocities, used in [137]. This term, controlled by the parameter μ_p, promotes that the movement $[0, 0]$ is set as the likely one, when there is no other prior knowledge, and allow us to detect motionless non–textured regions.

For solving (5.18) we use the single–motion distance defined in (5.12), since for these sequences, representations of multiple velocities do no coexist at this spatial scale (see [93]).

5.5.2 Spatial integration

The integration of global motion needs a large interaction area (different scales). In this second stage we use as initial solution the α coefficients obtained in the previous stage (where the local movements have been detected), and then we perform the diffusion–based spatial integration by the minimization of:

$$U_b(\alpha) = \sum_r \sum_i^N \left[\sum_{s:s\in\mathcal{N}_r} \alpha_i(s)(\alpha_i(r) - \alpha_i(s))^2 - \mu_c \left(\alpha_i(r) - C_{\alpha(r)}\right)^2 \right], \tag{5.20}$$

where the diffusion weights are the $\alpha_i(s)$ coefficients. By using these diffusion weights, we promote that each position r to be similar to the neighbor position s, when position s has a large $\alpha_i(s)$ coefficient, i.e., where the neighbor was marked by the previous stage as a position with a predominant α coefficient (close to one).

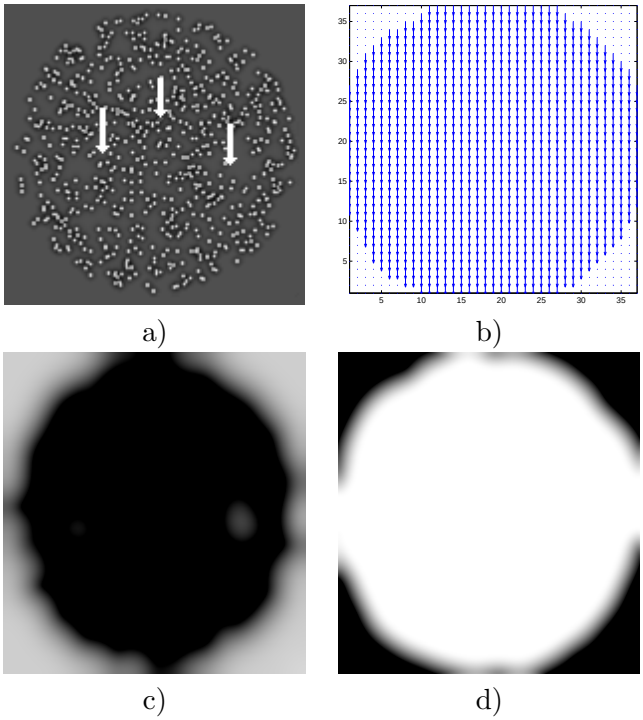


Figure 5-13: Results obtained for a RDK sequence with 1 movement. See text.

5.5.3 Results on RDK Sequences

Figure 5-13 shows the obtained results for a RDK sequence with one movement $\hat{u} = 0, -1$. The original $301 \times 301 \times 9$ sequence is shown in panel 5-13a. The recovered subsampled velocity field is shown in panel 5-13b. We show the recovered velocity layers with values different from zero in order to indicate the regions where each velocity model $\mathbf{u}_i$ is present; panel 5-13c shows the α layer associated to the model $\mathbf{u}_i = [0, 0]$ and panel 5-13d shows the α layer associated with the velocity model $\mathbf{u}_j = [0, -1]$, The white regions indicate areas with coefficients close to one and the dark zones indicates the presence of coefficients close to zero.

In a more complex experiment, Figure 5-14 shows the results for a transparent RDK sequence composed by two opposed movements $\hat{u}_1 = 1, 0$ and $\hat{u}_2 = -1, 0$. The sequence dimensions where $301 \times 301 \times 9$, panel 5-14a. We show the subsampled velocity field in panel 5-14b. The alpha layers for the three present movements $\mathbf{u}_i = 0, 0$ and $\mathbf{u}_j = 1, 0$ and $\mathbf{u}_k = 0, 0$, are shown in panels 5-14c 5-14d and 5-14e respectively, the other α layers are composed by zeros.

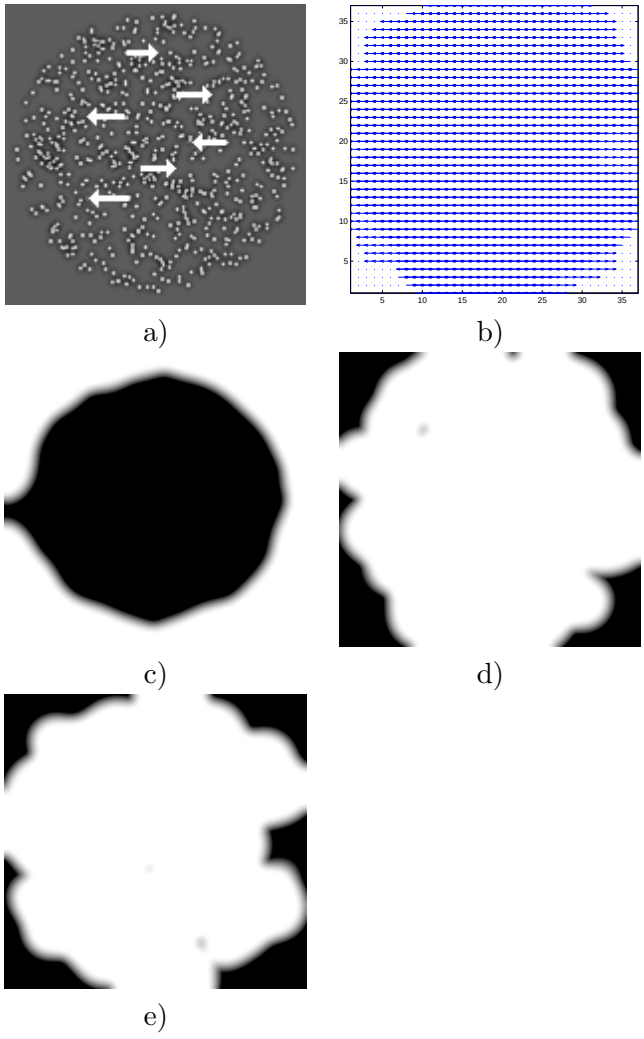


Figure 5-14: A transparent RDK sequence and the results obtained, see text.

5.6 Chapter Conclusions

In this chapter we have presented an energy cost formulation in order to estimate multiple motions. The unknown is a vector valued field that indicates the motions present in a particular spatiotemporal position. Our formulation extends previous works based on layered optical flow computation, by using a distance measure suitable for transparent motions and proposing an inter–model competition mechanism proper for multi–valued solutions. For the multi–motion case, the proposed inter–model competition mechanism behaves like those used for entropy–control in probability–measure–based approaches for single motion. This term is by itself a

novel contribution of this work.

Our formulation allows us to tackle sequences having single or multiple layers moving. The optimization process makes the integration of local velocities information by using suitable diffusion terms. The performance of the presented approach is demonstrated by synthetic experiments in textured and non-textured sequences as well as real sequences.

Additionally, we have investigated RDK sequences and we have proposed a regional based single motion estimation on a diffusion based integration.

In future work it is planned to focus on the diffusion terms and investigate how the different velocity maps may interact together.

Chapter 6

Concluding Remarks

6.1 Contributions

This thesis reports our research work on early vision multimodal regularization; a new general method is provided.

Our work is based on minimization of quadratic cost functions as well as the application of Basis Pursuit approach for multimodal regularization of ill-posed problems. We proposed a suitable regularization scheme that recovers high–quality solutions using a reasonable computational burden. We summarize our contributions as:

1. A general framework for assigning pixel–wise multimodal solutions which is capable of computing almost discrete indicator variables. Thus, this approach allows one to clearly discern which models are present and which are not. In particular:

   (a) We provide a general anisotropic filtering scheme that promotes an oriented regularization. Such an oriented regularization is useful for problems where the models are associated to orientations.

   (b) We introduce a novel cost term that promotes a multimodal inter–model competition. Such a term eliminates the non-relevant solutions and enhances the more plausible ones by promoting high contrasted multimodal solutions.

2. We show improvements in the state–of–the–art solutions for the axon multi–fiber estimation in DW–MR Images by applying the above multimodal scheme:

(a) We provide an observation model capable of representing more than one fiber in a voxel. This model explains the DW–MR signal by means of a linear combination of discrete prefixed DW-MR signals taken from a proposed basis dubbed Diffusion Basis Functions. We propose a procedure for setting these basis by performing a simple analysis of DW data in a convenient brain region.

(b) Based on the Basis Pursuit approach and the Diffusion Basis Functions scheme, we present an efficient solution method that promotes sparsity in the number of axon bundles required for describing the measured DW-MR signal.

(c) We develop a Bayesian framework for reconstructing the fiber pathways given the Diffusion Basis Functions solution by means of a stochastic particle walks.

3. We tackle the transparent Optical Flow estimation problem by applying our general multimodal regularization approach:

   (a) We propose an observation model based on a set of distances to different velocity models. For this aim, we used previously reported measures capable for detecting several optical flow velocities in a single pixel.

   (b) We develop a gradient–based solution scheme for this problem that presents algorithmic advantages with respect to previous state-of-the-art approaches.

   (c) We present a method for analyzing transparent RDK–sequences based on a regional single motion detector and a diffusion–based integration.

4. For both applications, the methods were tested by means of experiments on synthetic and real data. The results were validated quantitatively and qualitatively respectively.

Bibliography

[1] M. N. Ahmed, S. M. Yamany, N. A. Mohamed, A. A. Farag, and T. Moriarty. Bias field estimation and adaptive segmentation of mri data using modified fuzzy c-means algorithm. In *Proc. IEEE Int. Conf. Computer Vision and Pattern Recogn.*, pages 250–255, Fort Collins, CO, 1999.

[2] D. C. Alexander. An introduction to computational diffusion MRI: the diffusion tensor and beyond. In J. Weickert and H. Hagen, editors, *Visualization and Image Processing of Tensor Fields*. Springer, Berlin, 2005.

[3] D. C. Alexander. Maximum entropy spherical deconvolution for diffusion MRI. In *Proc. IPMI*, pages 76–87, 2005.

[4] D. C. Alexander. Multiple-fibre reconstruction algorithms for diffusion MRI. *Annals of the New York Academy of Sciences*, 1046:113–133, 2005.

[5] D. C. Alexander, G. J. Barker, and S. R. Arridge. Detection and modeling of non-gaussian apparent diffusion coefficient profiles in human brain data. *Magn. Reson. Med.*, 48(2):331–340, Jul 2002.

[6] A. W. Anderson. Sub-voxel measurement of fiber orientation using high angular resolution diffusion tensor imaging. In *Proc. 10th Anual Meeting of the ISMRM*, page 440, 2002.

[7] A. W. Anderson. Measurement of fiber orientation distributions using high angular resolution diffusion imaging. *Magn. Reson. Med.*, 54(5):1194–1206, 2005.

[8] Y. Assaf, R. Z. Freidlin, G. K. Rohde, and P. J. Basser. New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. *Magn. Reson. Med.*, 52(5):965–978, 2004.

[9] G. Aubert, R. Deriche, and P. Kornprobst. Computing optical flow via variational techniques. *SIAM Journal of Applied Mathematics*, 60(1):156–182, 1999.

[10] V. Auvray, P. Bouthemy, and J. Lienard. Multiresolution parametric estimation of transparent motions. In *Proc. Int. Conf. on Image Processing (ICIP'05)*, 2005.

[11] V. Auvray, P. Bouthemy, and J. Lienard. Motion estimation in x-ray image sequence with bi-distributed transparency. In *ICIP 06*, Atlanta, USA, 2006.

[12] J. Barron, D. Fleet, and S. Beauchemin. Performance of optical flow techniques. *The International Journal of Computer Vision*, 12(1):43–77, 1994.

[13] P. J. Basser and D. K. Jones. Diffusion-tensor MRI: theory, experimental design, and data analysis. *NMR Biomed.*, 15:456–467, 2002.

[14] P. J. Basser, J. Mattiello, and D. LeBihan. MR diffusion tensor spectroscopy and imaging. *Biophys. J.*, 66:259–267, 1994.

[15] P. J. Basser, S. Pajevic, C. Pierpaoli, and A. Aldroubi. Fiber tract following in the human brain using DT-MRI data. *IEICE Trans on Inf. and Sys.*, E85-D(1):15–21, 2002.

[16] 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.

[17] P. J. Basser and C. Pierpaoli. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *J. Magn. Reson. B*, 111:209–219, 1996.

[18] J. Bergen, P. Burt, R. Hingorani, and S. Peleg. Computing Two Motions From Three Frames. In *Proceedings of the third International Conference on Computer Vision*, pages 27–32, Osaka, Japan, Dec. 1990.

[19] J. Besag. Spatial interaction and the statistical analysis of lattice systems (with discussion). *Journal of Royal Statistical Society*, 2:192–236, 1974.

[20] J. C. Bezdek, J. Keller, R. Krisnapuram, and N. R. Pal. *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing.* Kluwer Academic Publishers, first edition, 1999.

[21] J. Bigun, G. H. Granlund, and J. Wiklund. Multidimensional orientation estimation with applications to texture analysis and optical flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(8):775–790, Aug. 1991. Report LiTH-ISY-I-0828 1986 and Report LiTH-ISY-I-1148 1990, both at Computer Vision Laboratory, Linköping University, Sweden.

[22] M. Black and P. Anandan. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *CVGIP: Image Understanding*, 63(1):75–104, 1996.

[23] M. Black, D. Fleet, and Y. Yacoob. Robustly estimating changes in image appearance. *Computer Vision and Image Understanding*, 78:8–31, 2000.

[24] M. Black and P. Rangarajan. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *The International Journal of Computer Vision*, 19(1):57–91, 1996.

[25] A. Blake and A. Zisserman. *Visual Reconstruction.* MIT Press, Cambridge, MA, first edition, 1987.

[26] O. Braddick. Local and global representations of velocity; transparency opponency, and global direction perception. *Perception*, 26:995–1010, 1997.

[27] R. L. Burden and J. D. Faires. *Numerical Analysis.* Brooks/cole., 7th edition, 2001.

[28] P. Burt, R. Hingorani, and R. Kolczynski. Mechanisms for Isolating Component Patterns in the Sequential Analysis of Multiple Motion. In *Proceedings of the IEEE Workshop on Visual Motion*, pages 187–193, Princeton, NJ, Oct. 1991.

[29] R. Buxton. *Introduction to Functional Magnetic Resonance Imaging Principles and Techniques.* Cambridge University Press, 2002.

[30] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM Review*, 43(1):129–159, 2001.

[31] Y. Chen, W. Guo, Q. Zeng, G. He, B. C. Vemuri, and Y. Liu. Recovery of intra-voxel structure from HARD DWI. In *Proc. International Symposium in Biomedical Imaging*, pages 1028–1031, October 2004.

[32] T. Darrell and E. Simoncelli. Separation of transparent motion into layers using velocity-tuned mechanisms. In *MIT Media Laboratory Vision and Modeling Group Technical Report*, 1993.

[33] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche. Apparent diffusion coefficients from high angular resolution diffusion images: Estimation and applications. *Magn. Reson. Med.*, 56(2):395–410, Aug. 2006.

[34] D. J. Field. Scale-invariance and self-similar 'wavelet' transforms: an analysis of natural scenes and mammalian visual systems. In O. U. Press., editor, *Wavelets, Fractals and Fourier Transforms: New Developments and New Applications.* Oxford University Press., 1993.

[35] D. J. Field. What is the goal of sensory coding? *Neural Computation*, 6:559–601, 1994.

[36] P. Fillard, V. Arsigny, X. Pennec, and N. Ayache. Clinical DT-MRI estimation, smoothing and fiber tracking with log-Euclidean metrics. In *Proc. ISBI*, pages 786–789, 2006.

[37] W. Förstner. A Feature Based Corresponding Algorithm for Image Matching. *International Archives of Photogrammetry and Remote Sensing*, 26:150–166, 1986.

[38] D. Forsyth and J. Ponce, editors. *Computer Vision - A modern approach.* Prentice Hall, 2002.

[39] L. R. Frank. Characterization of anisotropy in high angular resolution diffusion-weighted MRI. *Magn. Reson. Med.*, 47:1083–1099, 2002.

[40] B. Galvin, B. McCane, K. Novins, D. Mason, and S. Mills. Recovering motion fields: an evaluation of eight optical flow algorithms. *British Machine Vision Conference*, pages 195–204, 1998.

[41] J. C. Gee, D. C. Alexander, M. Rivera, and J. T. Duda. Non-rigid registration of diffusion tensor MR images. In I. Press, editor, *Proc. IEEE ISBI*, pages 477–480. IEEE, July 2002.

[42] S. Geman and D. Geman. Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Machine Intell.*, 6(6):721–741, 1984.

[43] L. Grady. Random walks for image segmentation. *IEEE Trans. Pattern Anal. Machine Intell.*, 28(11):1768–1783, 2006.

[44] P. Green. On use of the em algorithm for penalized likelihood estimation. *J. Roy. Statist. Soc*, 52(3):443–452, 1990.

[45] R. Gribonval, P. Depalle, X. Rodet, E. Bacry, and S. Mallat. Sound signal decomposition using a high resolution matching pursuit. In *Proc. ICMC*, pages 293–296, 1996.

[46] S. Grossberg, E. Mingolla, and L. Viswanathan. Neural dynamics of motion integration and segmentation within and across apertures. *Vision Res.*, 41:2521–2553, 2001.

[47] H. Gudbjartsson and S. Patz. The Rician distribution of noisy MRI data. *Magn. Reson. Med.*, 34:910–914, 1995.

[48] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer-Verlag, 2001.

[49] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer-Verlag New York, LLC, 1 edition, 2003.

[50] B. Horn and B. Schunck. Determining Optical Flow. *Artificial Intelligence*, 17:185–203, 1981.

[51] P. Huber. *Robust Statistics*. John Wiley & Sons, New York, 1981.

[52] M. Irani and S. Peleg. Motion analysis for image enhancement: resolution, occlusion, and transparency. *Journal on Visual Communications and Image Representation*, 4(4):324–335, 1993.

[53] M. Irani, B. Rousso, and S. Peleg. Computing Occluding and Transparent motions. *The International Journal of Computer Vision*, 12(1):5–16, Jan. 1994.

[54] K. M. Jansons and D. C. Alexander. Persistent angular structure: new insights from diffusion magnetic resonance imaging data. *Inverse Probl.*, 19:1031–1046, 2003.

[55] A. Jepson and M. Black. Mixture Models for Optical Flow Computation. In *CVPR93*, pages 760–761, 1993.

[56] A. Jepson and M. Black. Mixture Models for Optical Flow Computation. Technical Report RBCV-TR-93-44, University of Toronto, Department of Computer Science, Apr. 1993.

[57] 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(3):515–525, 1999.

[58] S. Ju, M. Black, and A. Jepson. Skin and Bones: Multi-layer, Locally Affine, Optical Flow and Regularization with Transparency. In *Proceedings of CVPR 96*, pages 307–314, San Francisco, CA, June 1996.

[59] V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph cuts? *IEEE Trans. Pattern Anal. Machine Intell.*, 26:147–159, 2004.

[60] B. W. Kreher, J. F. Schneider, I. Mader, E. Martin, J. Hennig, and K. A. Il'yasov. Multitensor approach for analysis and tracking of complex fiber configurations. *Magn. Reson. Med.*, 54:1216–1225, sep 2005.

[61] M. Lazar, D. M. Weinstein, J. S. Tsuruda, K. M. Hasan, K. Arfanakis, M. E. Meyerand, B. Badie, H. A. Rowley, V. Haughton, A. Field, and A. L. Alexander. White matter tractography using diffusion tensor deflection. *Hum. Brain. Mapp.*, 18(4):306–321, 2003.

[62] C. Lenglet, M. Rousson, and R. Deriche. DTI segmentation by statistical surface evolution. *IEEE Trans. Med. Imag.*, 25(6):685–700, Jun 2006.

[63] C. Lenglet, M. Rousson, R. Deriche, O. D. Faugeras, S. Lehericy, and K. Ugurbil. A Riemannian approach to diffusion tensor images segmentation. In *Proc. IPMI*, pages 591–602, 2005.

[64] P. Leopardi. A partition of the unit sphere into regions of equal area and small diameter. *Electronic Transactions on Numerical Analysis*, 25:309–327, 2006.

[65] M. S. Lewicki and T. J. Sejnowski. Learning overcomplete representations. *Neural Computation*, 12:337–365, 2000.

[66] S. Z. Li. *Markov Random Field Modeling in Image Analysis*. Springer Verlag, 2001.

[67] H. Liu, T. Hong, M. Herman, and R. Chellappa. Spatio-temporal filters for transparent motion segmentation. In *Proceedings of the International Conference on Image Processing*, pages 464–468, Washington, USA, 1995.

[68] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *International Joint Conference on Artificial Intelligence*, pages 674–679, 1981.

[69] S. Mallat and Z. Zhang. Matching pursuit with time-frequency dictionaries. *IEEE Trans. Signal Processing*, 41(12):3397–3415, 1993.

[70] K. V. Mardia. *Statistics of Directional Data*. Academic Press, 1972.

[71] J. L. Marroquín, E. A. Santana, and S. Botello. Hidden Markov measure field models for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(11), 2003.

[72] J. L. Marroquin, F. A. Velasco, M. Rivera, and M. Nakamura. Gauss-Markov measure field models for low-level vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23:337–348, 2001.

[73] G. Masson, Y. Rybarzcyk, E. Castet, and D. Mestre. Temporal dynamics of motion integration for the initiation of tracking responses at ultra-short latencies. *Visual Neuroscience*, 17(5):754–767, 2000.

[74] S. Mehrotra. On the implementation of a primal-dual interior point method. *SIAM J. Optimization*, 2(4):575–601, 1992.

[75] C. Mota, I. Stuke, T. Aach, and E. Barth. Divide-and-Conquer Strategies for Estimating Multiple Transparent Motions. In *Proceedings of the 1st International Workshop on Complex Motion, Schloss Reisensburg, Germany. Lecture Notes on Computer Science, LNCS Vol. 3417*, 2005.

[76] M. Mühlich and T. Aach. A Theory of Multiple Orientation Estimation. In *European Conference on Computer Vision*, volume 2, pages 69–82, 2006.

[77] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer Series in Operation Research, second edition, 2000.

[78] B. A. Olshausen and D. J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381:607–609, 1996.

[79] B. A. Olshausen and D. J. Field. Sparse coding with an overcomplete basis set: A strategy employed by v1? *Vision Research*, 37:3311–3325, 1997.

[80] A. V. Oppenheim. Superpostion in a Class of Nonlinear Systems. In *IEEE International Convention*, pages 171–177, New York, USA, 1964.

[81] E. Ozarslan and T. Mareci. Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. *Magn. Reson. Med.*, 50(5):955–965, Nov 2003.

[82] E. Ozarslan, T. Shepherd, B. C. Vemuri, S. Blackband, and T. Mareci. Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT). *Neuroimage*, 31(3):1083–1106, Jul 2006.

[83] E. Ozarslan, B. C. Vemuri, and T. H. Mareci. Generalized scalar measures for diffusion MRI using trace, variance, and entropy. *Magn. Reson. Med.*, 53(4):866–876, 2005.

[84] G. J. M. Parker and D. C. Alexander. Probabilistic anatomical connectivity derived from the microscopic persistent angular structure of cerebral tissue. *Philos. Trans. R. Soc. B Biol. Sci.*, 360(1467):893 – 902, 2005.

[85] J. Parker and D. Alexander. Probabilistic Monte Carlo based mapping of cerebral connections utilising whole-brain crossing fibre information. In *Proc. IPMI*, pages 684–695, July 2003.

[86] O. Pasternak, N. Sochen, and Y. Assaf. Variational regularization of multiple diffusion tensor fields. In J. Weickert and H. Hagen, editors, *Visualization and Image Processing of Tensor Fields*. Springer, Berlin, 2005.

[87] G. Paxinos and C. Watson. *The Rat Brain in Stereotaxic Coordinates*. San Diego: Academic Press, second edition, 1998.

[88] M. Perrin, C. Poupon, B. Rieul, P. Leroux, A. Constantinesco, J. F. Mangin, and D. LeBihan. Validation of Q-Ball imaging with a diffusion fibre-crossing phantom on a clinical scanner. *Philos. Trans. R. Soc. B Biol. Sci.*, 360(1467):881–891, 2005.

[89] M. Pingault, E. Bruno, and D. Pellerin. A Robust Multiscale B-Spline Function Decomposition for Estimating Motion Transparency. *IEEE Transactions on Image Processing*, 12(11):1416–1426, Nov. 2003.

[90] T. Poggio, V. Torre, and C. Koch. Computational vision and regularization theory. *Nature*, 317:314–319, 1985.

[91] R. A. Poldrack. A structural basis for developmental dyslexia: Evidence from diffusion tensor imaging. In M. Wolf, editor, *Dyslexia, Fluency, and the Brain*, pages 213–233. York Press, 2001.

[92] C. Poupon, C. Clark, V. Frouin, J. Regis, I. Bloch, D. LeBihan, and J. Mangin. Regularization of diffusion-based direction maps for the tracking of brain white matter fascicles. *Neuroimage*, 12(2):184–195, 2000.

[93] N. Qian, R. A. Andersen, and E. H. Adelson. Transparent motion perception as detection of unbalanced motion signals. III. Modeling. *Neuroscience*, 14(12):7381–7392, Dec. 1994.

[94] A. Ramirez-Manzanares and M. Rivera. Brain nerve boundless estimation by restoring and filtering intra-voxel information in DT-MRI. In *Second Workshop on Variational and Level Sets Methods*, pages 71–80, Oct. 2003.

[95] A. Ramirez-Manzanares and M. Rivera. Basis tensor decomposition for restoring intra-voxel structure and stochastic walks for inferring brain connectivity in DT-MRI. *Int. Journ. of Comp. Vis.*, 69(1):77–92, 2006.

[96] M. Rivera and J. Marroquin. Adaptive rest condition potentials: first and second order edge-preserving regularization. *Computer Vision and Image Understanding*, 88:76–93, 2002.

[97] M. Rivera and P. P. Mayorga. Image segmentation by means of a entropy-controlled trimap approach. *Internal Report at CIMAT, To be published*, 2007.

[98] M. Rivera, O. Ocegeda, and J. L. Marroquin. Entropy controlled gauss-markov random measure field models for early vision. In *LNCS 3752, Springer-Verlag VLSM 2005*, pages 137–148, 2005.

[99] J. Ruiz-Alzola, C. F. Westin, S. K. Warfield, A. Nabavi, and R. Kikinis. Nonrigid registration of 3D scalar, vector and tensor medical data. In *Proc. MICCAI*, pages 541–550, 2000.

[100] B. Sarel and M. Irani. Separating transparent layers through layer information exchange. In T. Pajdla and J. Matas, editors, *Proceedings of the 8th European Conference on Computer Vision*, pages 328–341, Prague, Czech Republic, 2004. Springer–Verlag.

[101] M. Shizawa and K. Mase. Simultaneous Multiple Optical Flow Estimation. In *International Conference on Pattern Recognition*, volume 1, pages 274–278, 1990.

[102] M. Shizawa and K. Mase. Principle of Superposition: a Common Computational Framework for Analysis of Multiple Motion. In *IEEE Workshop on Visual Motion*, pages 164–172, 1991.

[103] M. Shizawa and K. Mase. A unified computational theory for motion transparency and motion boundaries based on eigenergy analysis. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, pages 289–295, Lahaina, Hawai, June 1991. IEEE.

[104] J. Sijbers, A. den Dekker, J. V. Audekerke, and M. Verhoye. Estimation of the noise in magnitude MR images. *Magnetic Resonance Imaging*, 16(1):87–90, 1998.

[105] S. Skare, M. Hedehus, M. E. Moseley, and T. Q. Li. Condition number as a measure of noise performance of diffusion tensor data acquisition schemes with MRI. *J. Magn. Reson.*, 147:340–352, Dec 2000.

[106] O. Söderman and B. Jönsson. Restricted diffusion in cylindrical geometry. *J. Magn. Reson.*, 117:94–97, Nov 1995.

[107] E. O. Stejskal. Use of spin echoes in a pulsed magnetic-field gradient to study anisotropic restricted diffusion and flow. *J. Chem. Phys.*, 43:3597–3603, 1965.

[108] I. Stuke, T. Aach, E. Barth, and C. Mota. Estimation of multiple motions by block matching. In *4th ACIS International Conference onSoftware Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD 2003)*, pages 358–362, 2003.

[109] I. Stuke, T. Aach, E. Barth, and C. Mota. Multiple-motion-estimation by block matching using MRF. In *ACIS, International Journal of Computer and Information Science*, 2004.

[110] A. N. Tikhonov and V. A. Arsenin. *Solutions of Ill-posed Problems.* Winston & Sons, Washington, 1977.

[111] J. Toro, F. Owens, and R. Medina. Using Known Motion Fields for Image Separation in Transparency. *Pattern Rcognition Letters*, 24:597–605, 2003.

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

[113] D. Tschumperlé and R. Deriche. Vector-valued image regularization with PDE's: A common framework for different applications. *IEEE Trans. Pattern Anal. Machine Intell.*, 27(4):506–517, April 2005.

[114] D. S. Tuch. *Diffusion MRI of complex tissue structure.* PhD thesis, Harvard-MIT, Cambridge MA, January 2002.

[115] D. S. Tuch. Q-Ball imaging. *Magn. Reson. Med.*, 52:1358–1372, 2004.

[116] D. S. Tuch, J. W. Belliveau, and V. Wedeen. A path integral approach to white matter tractography. In *Proc. 8th Annual Meeting of the ISMRM*, page 791, 2000.

[117] D. S. Tuch, T. G. Reese, M. R. Wiegell, N. Makris, J. W. Belliveau, and V. J. Wedeen. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magn. Reson. Med.*, 48(4):577–582, 2002.

[118] D. S. Tuch, R. M. Weisskoff, J. W. Belliveau, and V. J. Wedeen. High angular resolution diffusion imaging of the human brain. In *Proc. 7th Annual Meeting of the ISMRM*, page 321, 1999.

[119] D. S. Tuch, M. R. Wiegell, T. G. Reese, J. W. Belliveau, and V. Wedeen. Measuring cortico-cortical connectivity matrices with diffusion spectrum imaging. In *Proc. 9th Annual Meeting of the ISMRM*, page 502, 2001.

[120] B. C. Vemuri, Y. Chen, M. Rao, T. McGraw, Z. Wang, and T. Mareci. Fiber tract mapping from diffusion tensor MRI. In *Proc IEEE Workshop VLSM*, pages 81–88, 2001.

[121] D. Vernon. Decoupling Fourier Components of Dynamic Image Sequences: a Theory of Signal Separation, Image Segmentation and Optical Flow Estimation. In *European Conference on Computer Vision*, volume 2, pages 69–85, June 1998.

[122] F. Vigueras. Filtrado y segmentacion de imagenes usando difusion anisotropica. Master's thesis, Centro de Investigacion en Matematicas (CIMAT), 2001.

[123] E. von dem Hagen and R. Henkelman. Orientational diffusion reflects fiber structure within a voxel. *Magn. Reson. Med.*, 48(3):454–459, 2002.

[124] Z. Wang and B. C. Vemuri. An affine invariant tensor dissimilarity measure and its applications to tensor-valued image segmentation. In *Proc. CVPR*, pages 228–233, 2004.

[125] Z. Wang and B. C. Vemuri. Tensor field segmentation using region based active contour model. In *Proc. 8th ECCV'04, Springer LNCS*, volume 3024, pages 304–315, 2004.

[126] Z. Wang and B. C. Vemuri. DTI segmentation using an information theoretic tensor dissimilarity measure. *IEEE Trans. Med. Imag.*, 24(10):1267–1277, 2005.

[127] Z. Wang, B. C. Vemuri, Y. Chen, and T. Mareci. A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from DWI. In *Proc. IPMI*, volume 18, pages 660–671, 2003.

[128] Z. Wang, B. C. Vemuri, Y. Chen, and T. H. Mareci. A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex DWI. *IEEE Trans. Med. Imag.*, 23(8):930–939, 2004.

[129] V. J. Wedeen, P. Hagmann, W.-Y. I. Tseng, T. G. Reese, and R. M. Weisskoff. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magn. Reson. Med.*, 54(6):1377–1386, Oct 2005.

[130] V. J. Wedeen, T. G. Reese, D. S. Tuch, M. R. Weigel, J. G. Dou, R. M. Weiskoff, and D. Chessler. Mapping fiber orientation spectra in cerebral white matter with Fourier-transform diffusion MRI. In *Proc. 8th Annual Meeting of the ISMRM*, page 82, 2000.

[131] J. Weickert. Diffusion and regularization methods for tensor-valued images. In *Proc. First SIAM-EMS Conf. AMCW01*, 2001.

[132] J. Weickert and H. Hagen, editors. *Visualization and Processing of Tensor Fields*. Springer Verlag, 2006.

[133] Y. Weiss and E. Adelson. A unified mixture framework for motion segmentation: incorporating spatial coherence and estimating the number of models. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, pages 321–326, San Francisco, CA, June 1996. IEEE.

[134] C. F. Westin, S. E. Maier, H. Mamata, F. A. Jolesz, and R. Kikinis. Processing and visualization for diffusion tensor MRI. *Medical Image Analysis*, 6(2):93–108, 2002.

[135] M. R. Wiegell, M. Henrik, B. W. Larsson, and V. J. Wedeen. Fiber crossing in human brain depicted with diffusion tensor MR imaging. *Radiology*, 217:897–903, 2000.

[136] A. M. Wink and J. Roerdink. BOLD noise assumptions in fMRI. *J. Biomed. Imag.*, 2006(12014):1–11, 2006.

[137] E. P. S. Y. Weiss and E. H. Adelson. Motion illusions as optimal percepts. *Nature Neuroscience*, 5(6):598 – 604, 2002.

[138] W. Zhou and C. Kambhamettu. Separation of Reflection by Fourier Decoupling. In *Asian Conference on Computer Vision*, Jeju Island, Korea, Jan. 2004.

[139] L. Zhukov and A. H. Barr. Oriented tensor reconstruction: Tracing neural pathways from diffusion tensor MRI. In *Proc. 13th IEEE Visualization 2002*, pages 387–394, 2002.

[140] L. Zhukov, K. Museth, D. Breen, R. Whitaker, and A. Barr. Level set modeling and segmentation of DT-MRI brain data. *J. Electronic Imaging*, 12:125–133, January 2003.