Basis Pursuit Based Algorithm for Intra-Voxel Recovering Information in DW-MRI.

Alonso Ramírez-Manzanares
Centro de Investigación en Matemáticas A.C.
Apdo. Postal 402, Guanajuato, Gto.
C.P. 3600, México, (52) 473 73 271 55
alram@cimat.mx

Mariano Rivera
Centro de Investigación en Matemáticas A.C.
Apdo. Postal 402, Guanajuato, Gto.
C.P. 3600, México, (52) 473 73 271 55
mrivera@cimat.mx

Abstract

In this work we applied the Basis Pursuit (BP) methodology for recovering the intra-voxel information in Diffusion Weighted MR Images (DW-MRI). We compare the proposed BP approach with the Diffusion Basis Function Estimation (DBFE) algorithm. DBFE approach was previously applied to recover intra–voxel diffusion information in brain DW-MRI. The intra–voxel information is recovered at voxels that contain axon fiber crosses or bifurcations by means of a linear combination of a known diffusion functions. We state the DBFE solution in the signal decomposition context, i.e., the measured DW-MRI signal is decomposed as a linear combination of signals that belongs to a Base of Diffusion Functions (BDF). In such a BDF, each signal is a M-dimensional vector, where each component indicates the water diffusion coefficient in a known three-dimensional orientation. In this work, we analyze and compare the solution given by DBFE method with the BP methodology. The BP methodology is used in order to select the set of base signals (which are taken from a dictionary) that best explain a given arbitrary signal. Moreover, solution strategies used in the BP and DBFE algorithm are compared and discussed. Examples in synthetic and real images are shown in order to demonstrate the performance of the compared methods.

1. Introduction

1.1. The diffusion weighted intra–voxel problem

Amongst the most challenging goals in neuroimaging is the estimation of connectivity patterns in the brain in vivo. For this purpose, a special Magnetic Resonance Imaging (MRI) technique named Diffusion Weighted MRI (DW-MRI) is used. That technique allows one to obtain an estimation of the orientation of water diffusion in the brain. That diffusion is constrained by the direction of axon nerve bundles. Such an information is very useful in neuroscience research, due to the relationship of brain connectivity with several diseases and, in general, with brain development [10].

The physical relationship of the water diffusion process in a given orientation, for each image position, was established by Stejskal–Tanner [13]:

$$s_i = s_0 \exp(-b\mathbf{g}_i^T \mathbf{D} \mathbf{g}_i) + \eta_i, \tag{1}$$

where the unitary vector $\mathbf{g}_i = [g_x, g_y, g_z]^T$ indicates the *i–th* direction in which a directional independent magnetic gradient is applied, s is the measured DW–MRI signal in the $\mathbf{g}_i$ orientation, s_0 is the measured signal magnitude without diffusion gradient, and b is a constant proportional to both the applied time and magnitude of the magnetic gradients. The η_i coefficient represents a residual (produced by an inadequate model as well as for the present noise in the signal), and the diffusion coefficients in all directions are summarized by the positive definite symmetric 3×3 tensor $\mathbf{D}$. This model is used in the Diffusion Tensor MRI methodology (DT-MRI) (see [13] for more details).

A standard acquisition protocol for a single orientation $\mathbf{g}_i$, give us a 3D image, where in each voxel, the signal intensity indicates the grade of attenuation due to the diffusion, a larger attenuation (s_i small) indicates a significant water diffusion in the configured orientation. A 2D schematization of the measured signals are shown in Figure 1.

Could be found several diffusion orientations in the same brain position (and consequently in a single voxel), for instance, in the case of axon fiber crossing or bifurcations, see Figure 1b. For this reason, a more adequate model have been proposed by Tuch et al. [12, 11] called High Angular Resolution Diffusion Imaging (HARDI) method. That method is based on an observation model in which the signal s is built up as a finite mixture of tensors $\mathbf{D}_j$:

$$s_i = s_0 \sum\nolimits_{j=1}^{\kappa} \beta_j \exp(-b\mathbf{g}_i^T \mathbf{D}_j \mathbf{g}_i), \tag{2}$$

Proceedings of the Sixth Mexican International Conference on Computer Science (ENC'05)
0-7695-2454-0/05 $20.00 © 2005 IEEE

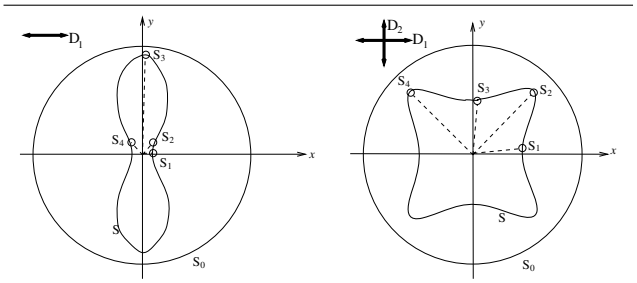


Figure 1. 2D schema of the measured s signals in a voxel. a) The signal s (continuous peanut shape) when there is only one water diffusion orientation D_1 in a voxel. b) The shape of s (continuous concave shape) when there are 2 almost orthogonal diffusion directions D_1 and D_2. The small circles indicates the distance to the center, or the magnitude of s measured at some angles. In both panels, the big circle schematizes the s_0 signal, which is orientation independent.

where κ denotes the number of significant fibers orientations within the voxel, the tensor $\mathbf{D}_j$ explains each one of these diffusion orientations, and the scalar β_j indicates its amount of presence.

If we want to recover the unknown diffusions (i.e. the orientation and magnitude of each one) by using only the measured signal s and s_0, then all unknowns in (2) must be computed (the number κ of orientations, plus six unknowns per tensor, plus the amount β_j of each one), independently for each voxel from a large set of acquired images $\{s\}$. This Diffusion Multi–Tensor Magnetic Resonance Imaging (DMT–MRI) technique allows one to recover the intra–voxel information that is not observed in the standard DT–MRI model (1), see [11]. The drawbacks of the DMT–MRI method are: the large number of additional diffusion images $\{s\}$ required (for instance, 126 diffusion 3D–images are used in [11], and 54 images are used in [6]), resulting in a large acquisition time, and algorithmic problems related to Equation (2); which is highly nonlinear. So that, multiple restarts of the optimization method are required to prevent the algorithm from settling in a local minimum. Furthermore, such an algorithm is not stable for more than two fiber bundles, i.e. for $\kappa > 2$ (see discussion in [11], Chap. 7). There exist another kind of works in this area [7, 8], that have recovered the DT–MRI intra–voxel information. Those works used either a spatial and a intra-model regularization procedure, in order to decompose the DT in a multi-tensor one, though, those works highly depend on the existing voxel neighborhood information.

2. Previous work

2.1. Diffusion basis function estimation

Recently, Diffusion Basis Function Estimation (DBFE) methodology have been proposed in [9], as a solution method for equation (2). This method used a basis of functions called Diffusion Basis Functions (DFB), defined as

$$\Phi_{ij} = \exp(-b\mathbf{g}_i^T \bar{\mathbf{T}}_j \mathbf{g}_i), \tag{3}$$

where each Φ_{ij} should be understood as the decay factor due to a fixed tensor $\bar{\mathbf{T}}_j$ in the *i–th* direction $\mathbf{g}_i$. The set of Φ_{ij} values define a matrix such that each column vector $\Phi_{\cdot j}$ is generated with the signal measured in all the orientations $\mathbf{g}_i$, $i = 1, 2, \ldots, M$, for a fixed tensor $\bar{\mathbf{T}}_j$. By using this DBF Φ, with cardinality N (i.e. composed by N base tensors), the composition of the signal s in the work [9], was established as:

$$s_i = s_0 \sum_{j=1}^{N} \alpha_j \Phi_{ij} + \eta_i \tag{4}$$

were $\alpha = [\alpha_1, \alpha_2, ..., \alpha_j, ..., \alpha_N]^T$ is a vector, such that the scalar $\alpha_j \geq 0$ denotes the contribution of the j-th DBF $\Phi_{\cdot j}$ to all the s_i measures in different gradient directions, i.e. the contribution to the linear mixture in a particular voxel. Note that the basis is not complete, due to the fact that the contained orientations $j = 1, \ldots, N$ are a discretization of the 3D space. This feature carry us to observe a remanent scalar η_i for each measurement. By increasing the cardinality in the basis (consequently the 3D angular resolution is increased), this remainder is diminished, so that, it could be attributed to the noise in the signal.

The use of the model (4), with a large basis, allow us to estimate the diffusion orientations with a high angular resolution, only by computing the α vector in each brain position. The solution proposed in [9] was derived by the minimization of the following error function

$$\begin{aligned} E(\alpha) &= \|\Phi\alpha - s\|_2^2 \\ \text{subject to} &\quad \alpha_j \geq 0\,, \forall j \end{aligned} \tag{5}$$

The minimization was achieved in [9] by the derivation of (5) with respect to a each α_k, and by solving the obtained liner system equalized to zero. The solution of the linear system was obtained by the use of a Gauss-Seidel approach, were the non–negativity constraint was accomplished by projecting to zero the negative α_j values in each iteration. That minimization approach has the drawback that requires several iterations to achieve the best α value.

3. Theoretic framework: Pursuit techniques

A compact signal representation have been largely used in signal processing problems [1, 2, 3] . In this context, is desirable to represent a given signal by a set of coefficients associated with elements of a dictionary of functions. The elements of such a dictionary are called *atoms*. Very often, the used dictionary is redundant (or over-complete), in the sense that the signal atoms are not orthogonal. In the signal processing context it is common to select atoms as Wavelets, Gabor dictionaries, Cosine Packets, Chirplets and Warplets, among others. The idea is to select from the dictionary the atoms that best match the signal structures, using a criteria for choosing among equivalent decompositions.

A common using criteria is to accomplish with the basic feature of sparsity, i.e., we want to recover a representation with the fewest significant coefficients. Additionally, a desirable feature is to spend a reasonable time in to achieve the decomposition.

The mathematical model that represents the decomposition of a signal $s \in \Re^M$ as a linear combination of atoms ϕ's, each one of them belonging to a dictionary Γ is

$$s = \sum_{i \in \Gamma} \alpha_i \phi_i + \eta = \Phi\alpha + \eta, \tag{6}$$

where Φ is a $M \times N$ matrix composed by N atoms, each atom ϕ_i is a column of Φ. The $\alpha \in \Re^N$ vector contains the linear combination coefficients. The residual η grasp the signal energy that can not be explained (or fitted) by the dictionary. When the dictionary was selected such that it is capable to represent all possible signals in the work domain, we expect that η captures the energy of the present noise in the signal.

Commonly $N >> M$, that makes the problem ill-posed: we have only a few measures $s_i, i = 1, \ldots, M$ and we want to select between a (normally) huge dictionary, the atoms that best represent the signal as a linear combination. The solution of (6) for α, given by the inversion of the pseudo-inverse of Φ is prohibitive in such a case, and it does not allow us neither to introduce prior information about α, nor lead the algorithm to a solution with some desired properties (sparsity for instance).

Recently, a Basis Pursuit (BP) [1] technique have been proposed for solving the problem (6), i.e. for computing the α coefficients. It was formulated as:

$$\begin{aligned} &\min \|\alpha\|_1 \\ &\text{subject to } \Phi\alpha = s \end{aligned} \tag{7}$$

where the constraint over the L-1 norm of the α vector guaranties that the solution is unique. In this way, BP has a interesting relationship to areas related to ill-posed problems. The problem above, can be transformed in a linear programming one (see [1]), and solved by the simplex method. Although, in the presence of noise and depending on the chosen dictionary, the constraint in (7) could not to be accomplished, resulting in a over-constrained linear programming problem. In those cases, we need to use a more adequate minimization procedure, like a interior-point method which tries to minimize the residual vector $r_b = \Phi\alpha - s$ (see [5]). The BP method have shown a better performance with respect to other pursuit techniques, for instance Matching Pursuit (MP) [3], which uses a non-linear iterated algorithm for solving (6). Although MP is a simple and easy-to-implement algorithm, it presents the drawback that very often the first selected atom globally fits several signal structures, but, this atom could be not well adapted to represent the signal local structures (see [2]). Differently BP indicates with a few α coefficients the atoms that best fits the local structures, with the drawback that it requires a more sophisticated minimization tool.

4. Basis pursuit based algorithm for DW–MRI

Now, we will formulate the intra–voxel structure estimation as a selection of atoms from a given dictionary, i.e. in the pursuit framework, particularly in a Basis Pursuit framework.

We propose to solve the problem stated by (4), by the modification of the the minimization problem in (5) in a BP methodology, i.e. we propose to solve the problem:

$$\begin{aligned} &\min \quad \|\alpha\|_1 \\ &\text{subject to} \quad \Phi\alpha = s, \quad \alpha_j \geq 0\,, \forall j \end{aligned} \tag{8}$$

Since we have constrained already the sign of the α components, (8) can be formulated in the following linear program

$$\begin{aligned} &\min \quad \textstyle\sum_j \alpha_j = \hat{e}^T\alpha \\ &\text{subject to} \quad \Phi\alpha = s, \qquad \alpha_j \geq 0\,, \forall j \end{aligned} \tag{9}$$

where $\hat{e}$ is a vector with all its components equal to one.

The previous linear program could be solved, in theory, with the simplex method, but, because in most cases the MRI signal s contains noise and because the dictionary Φ is not complete, the linear problem in (9) is an over–constrained one. That means that we expect that the remanent η_i [defined in (4)] will be different to zero. For this reason we solve it by using a powerful interior-point minimization method. In our experiments we used the primal-dual predictor-corrector Merhotra's algorithm (see [4, 5]) that computes the results in a fraction of the computational

effort that requires the DBFE method described in Section 2.1. Since the Merhotra's algorithm formulation assumes that either primal (the α coefficients) and dual (not used in our case) variables are positives, the constraints are fulfilled only by the system $\Phi\alpha = s$. Note that the main differences between (8) and (5) are: the use of a L-1 norm vs. the use of a L-2 norm, and the constraint over the size of the α vector vs. the constraint over the error $\Phi\alpha - s$. In this sense, we known that the L-1 norm belongs to the robust potential category, differently to the L2-norm. A nice property of the problem stated in (8) (a linear programming one), is that, in the noise absence, the obtained solution is a global minimizer .

5. Results

This section shows experimental results and a performance comparison between the DBFE method and the proposed BP method, either in synthetic and in real DW-MRI data.

5.1. Synthetic experiments

In order to test the performance of the proposed method for selecting the diffusion signals ϕ_j that best composed the observed s, we generate several synthetic examples. In the first example we generated a signal s which is exactly composed by a linear combination of two atom signals, $s = \frac{1}{2}\phi_i + \frac{1}{2}\phi_j$. Then, we used a different number of measures M and we compare the norm of the error between the obtained solution and the real one. The dictionary was composed by $N = 36$ atoms, we used $b = 1000$ in (3). The comparison between the method DBFE reported in [9] and the BP one proposed here in Section 4 is shown in Figure 2. Note that the error diminishes in both methods by increasing the number of measures s_i. However, note that the error obtained by the BP method is smaller and requires less measures (only 4) for obtaining the exact result.

In fact, Figure 3 give us an insight about the reason of the large error in the solution obtained with the DBFE method. In this experiment we generated a synthetic signal s that is composed by 2 diffusion orientations in a 2D plane, 63^o and 157^o respectively. The dictionary is composed by diffusion signals that splits the 2D orientation space in equidistant intervals of 10^o, i.e., the orientation set was $[0^o, 10^o, 20^o, \ldots, 170^o]$. Note that such a dictionary can not exactly represents the signal s. The result obtained with the DBFE algorithm is shown in panel 3a, and the result for the BP method is shown in panel 3b. In both figures the x axis corresponds to the orientation (in degrees) of the $N = 18$ atoms (or diffusion signals in the dictionary) and the y axis indicates the value of the associated α_j value. As can be seen, the result obtained by the DBFE involves more

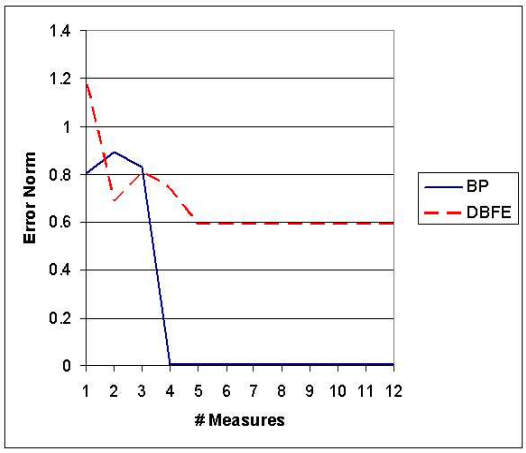


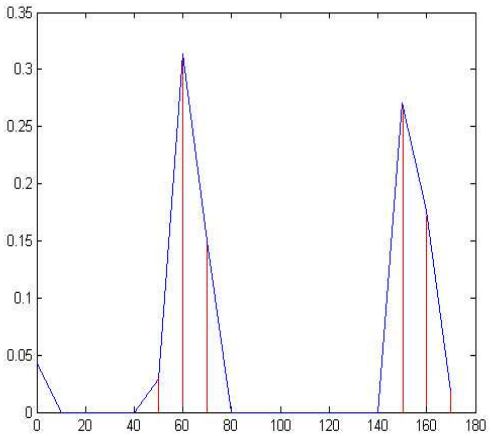
Figure 2. Comparison for both methods. We show the obtained error vs. the number of measures M, see text.

coefficients than the needed ones, and the result given by the BP method uses only the necessary number of coefficients, i.e., in the case of BP, the signal ϕ_7 associated with 60^o have the biggest α coefficient, and jointly with the coefficient associated to 70^o (a small one) fits good enough the diffusion measured in the 63^o orientation. In the same way, the second diffusion is fitted only by the combination of the atom associated to 160^o (biggest) and 150^o (smallest), and all the others coefficients are equal to zero. The BP method required 12 iterations of Mehrotra's minimization method. On the other hand, the method DBFE was iterated 1300 times, and the obtained result contains more coefficients that the needed ones for the s representation. Note that the solution given by DBFE is almost right, i.e., it indicates with the biggest α coefficient the diffusion function that best explains the signal. Although, the solution given by BP is sharper, thus, better defined, indicating accurately the magnitude of the coefficients of the linear combination. We can stated that, for the medical purposes, the BP solution is better.

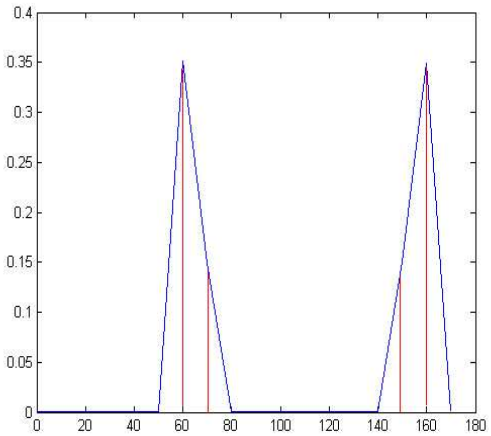
Additionally, we tested the robustness of both methods to the present noise in the measured signal. Figure 4 shows the obtained results. We generate a synthetic signal and then we added independent Gaussian noise with zero mean and different standard deviations. The x axis indicates the SNR ratio and the y axis shows the norm of the difference between the obtained solution vector α and the real one. As can be seen, the error of the method BP is the smallest one.

5.2. Results obtained in real DW-MRI data

The results obtained with brain DW–MRI are shown in figure 5. In this case, we used the DW images that are in-



a)



b)

Figure 3. α coefficients obtained in the case where the diffusion orientation was not in the dictionary. a) Result of DBFE method, b) Result of the BP method, see text.

cluded in the BioTensor software, provided by the web site of the Scientific Computing and Imaging Institute of the University of Utah (`http://software.sci.utah.edu/archive/archive_main.html`). Such a data is constructed by $75 \times 109 \times 29$ voxels, and with voxel dimensions $2.0 \times 2.0 \times 2.0 mm^3$. The number of measurements per voxel in independent directions $\mathbf{g}_i$ is $M = 12$, and was acquired with $b = 1000$. The dictionary of diffusion functions was composed by $N = 33$ equidistant orientations. Panel 5a depicts the brain map (a sagital slice), where the region of interest is marked by a white square. Panel 5b shows the classical DT fitted in (1) (see [13] for details) that give us a reference for the orientations. We show in panel 5c the

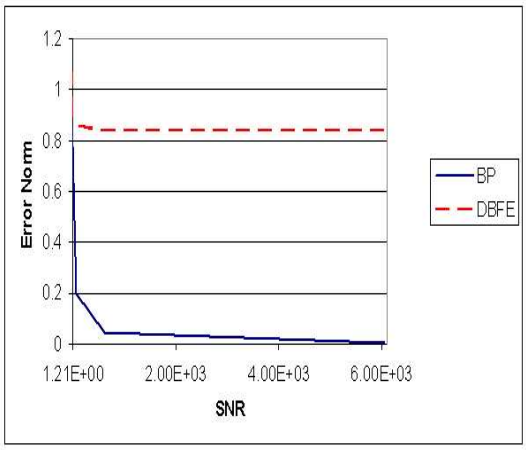


Figure 4. Variation for both methods of the error in the results by the SNR variation in the synthetic signal, see text.

computed orientations, weighted by its associated α coefficient, recovered by the DBFE method. Finally, the computed orientations with the BP method are shown in Panel 5d, weighted by its associated α coefficient. In both 5c and 5d panels, the tensor diffusion associated with the largest α_i coefficient is shown in red color, the second in green and the third in blue. It is important to note that this results are congruent with the a priori anatomical knowledge for this region which corresponds to the brain *corona radiata*. Note that the center of the slice contains a fiber crossing (depicted in the DT–MRI field in panel 5b with no predominant orientation tensors, i.e. ball shaped tensors) and note that the intra–voxel orientations recovered, depicted in panel 5c and panel 5d indicates well the crossing by means of almost orthogonal diffusion orientations. Additionally, note that, in some regions, the recovered orientations by the BP method are more defined and contrasted, i.e. there are some brain positions where only one diffusion direction was detected instead of two, differently to the DBFE result.

In order to compare the computational effort, we measured the computational time for both methods, they were implemented in the same language and were executed in the same computer. BP approach is about 40 times faster than the DBFE approach.

6. Conclusions

In this paper, we stated the problem of intra-voxel diffusion estimation in DW–MRI in the context of signal decomposition, such a decomposition is obtained by means of atoms that belongs to a dictionary. The BP methodology was applied, demonstrating that the achieved results are

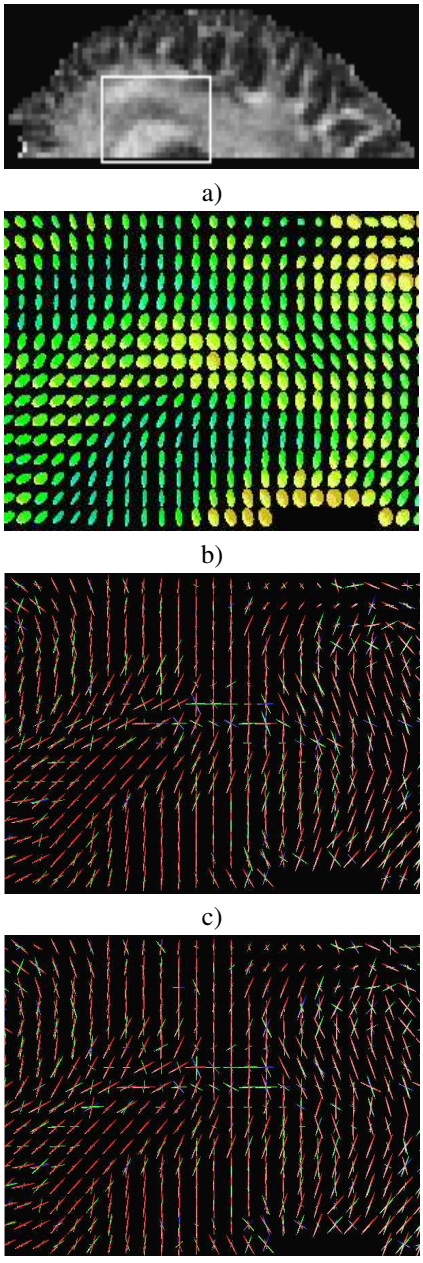


Figure 5. Results obtained in real DW–MR images. a) Brain map of the region of interest in a sagital slice, showing a part of the *corona radiata*. b) Classical DT-MRI field, shown as orientation reference. c) Diffusion orientations weighted by the α coefficients obtained by the DBFE method, shown for comparison. d) Diffusion orientations weighted by the α coefficients obtained by the BP method. See text.

better compared with the DBFE methodology, which represents the state of the art in this approach. The obtained results are better in the sense that the uncertainty in the diffusion orientation was diminished and the required computational effort was reduced. Such a differences can be explained, in part, because the BP-like approach uses a objective function based in the L-1 norm (which is known to be a robust potential) and the DBFE uses a non-robust L-2 based cost norm.

References

[1] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1999.

[2] R. Gribonval, P. Depalle, X. Rodet, E. Bacry, and S. Mallat. Sound signal decomposition using a high resolution matching pursuit. In *ICMC'96*, 1996.

[3] S. M. and Z. Zhang. Matching pursuit with time-frequency dictionaries. *IEEE Transactions in Signal Processing*, 41:3397–3415, 1993.

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

[5] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer Series in Operation research, 2000.

[6] G. Parker and D. Alexander. Probabilistic monte carlo based mapping of cerebral connections utilising whole-brain crossing fibre information. In *IPMI'03*, pages 684–695, 2003.

[7] 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.

[8] A. Ramirez-Manzanares and M. Rivera. Basis tensor decomposition for restoring intra–voxel structure and stochastic walks for inferring brain connectivity dt–mri. *to appear in IJCV*, 2005.

[9] A. Ramírez-Manzanares, M. Rivera, B. C. Vemuri, and T. C. Mareci. Basis functions for estimating intra–voxel structure in dw–mri. In *Medical Imaging Conference, 2004*, 2004.

[10] B. Richard. *Introduction to Functional Magnetic Resonance Imaging Principles and Techniques*. Cambridge: Cambridge University Press, 2002.

[11] D. Tuch. *Diffusion MRI ofcomplex tissue structure*. PhD thesis, Harvard-MIT, Cambridge MA, January 2002.

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

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