Reconstruction of HARDI Data Using a Split Bregman Optimization Approach

Iván C. Salgado Patarroyo[1], Sudipto Dolui[1],
Oleg V. Michailovich[1], and Edward R. Vrscay[2]

[1] Department of Electrical & Computer Engineering, University of Waterloo, Canada
[2] Department of Applied Mathematics, University of Waterloo, Canada
{icsalgad,sdolui,olegm,ervrscay}@uwaterloo.ca

Abstract. Among the current tools of diffusion Magnetic Resonance Imaging (dMRI), *high-angular resolution diffusion imaging* (HARDI) excels in its ability to delineate complex directional patterns of water diffusion at any predefined location within the brain. It is known that HARDI signals present a practical trade-off between their directional resolution and their signal-to-noise ratio (SNR), suggesting the need for effective denoising algorithms for HARDI measurements. The most effective approaches to alleviate this problem have been proven to be those which exploit both the directional and the spatial regularity of HARDI signals. Unfortunately, many of these algorithms entail substantial computational burdens. Accordingly, we propose a formulation of the problem of reconstruction of HARDI signals which leads to a particularly simple numerical implementation. The proposed algorithm allows for the separation of the original problem into procedural steps which can be executed in parallel, which suggests its computational advantage.

Keywords: Diffusion Weighted MRI, HARDI, total variation image denoising, alternative direction method of multipliers.

1 Introduction

The advent of dMRI has provided a way to measure the diffusivity of brain tissue – the progress which has resulted in a major development of new methods for studying the brain and its connectivity, in particular [1]. Among various methods of dMRI, Diffusion Tensor Imaging (DTI) [1] is the most widespread due to the clinical value of its associated contrasts as well as to the relative simplicity of its acquisition requirements. Intrinsic in the DTI model, however, is its inability to resolve multiple diffusion flows through a voxel of interest [2]. To overcome this deficiency of DTI, HARDI was introduced in [2]. At a technical level, HARDI signals are acquired in the q-space by distributing the sampling points over a spherical shell, the radius of which is controlled by the so-called b-value (typically, in the range $b \in [1000, 3000]$ s/mm^2). While a typical number of sampling points (aka diffusion-encoding directions) in DTI is in the range between 20 and 30, between 60 to 100 points are required for a standard HARDI

M. Kamel and A. Campilho (Eds.): ICIAR 2013, LNCS 7950, pp. 589–596, 2013.

reconstruction [2]. Thus, the improvement in the modelling capacity of HARDI comes at the expense of a greater number of diffusion measurements required, and hence longer data acquisition times. Moreover, the directional resolution of HARDI, and therefore of its associated q-ball imaging (QBI) [3], is known to improve when using higher b-values [4]. Unfortunately, the physics of HARDI allows such an improvement only at the expense of lower values of SNR. Hence, the only way to avoid sacrificing the directional resolution of HARDI, while preserving the informational integrity of the data, is through the use of methods for reconstruction of HARDI signals from their noisy measurements.

The existing approaches to reconstruction of HARDI signals from their observations are vast and diverse. Such methods can be broadly categorized into three groups, which differ in how they impose regularity constraints on the data. Thus, the methods of the first group exploit the assumption that HARDI signals are regular in the directional (i.e., diffusion) domain, while ignoring their spatial regularity [4] [5]. On the other hand, the methods of the second group exploit the spatial regularity of HARDI signals, while ignoring their regularity in the diffusion direction [6] [7] [8]. Naturally, more accurate results have been reported using the methods of the third group, which take into consideration both the spatial and the diffusion-domain behaviour of the diffusion signal [10].

Despite the promising performance of the methods of the third group, many of them remain subject to the disadvantage of relatively high computational complexity, which is partially a result of the high dimensionality of HARDI data. To alleviate the above difficulty, we propose a different formulation of the problem of reconstruction of HARDI data, which leads to a particularly convenient and computationally efficient numerical scheme.

The proposed formulation of the reconstruction problem is based on a standard Bayesian approach, in which an optimal solution is obtained as a global minimizer of an associated cost function. What renders the proposed method computationally efficient is the way in which the cost function is minimized. In particular, we employ a split Bregman procedure which, in fact, is equivalent to the alternative direction method of multipliers (ADMM) [11]. Using this method allows replacing the original problem (that involves composite, cross-domain regularization) by a sequence of simpler (single-domain) problems, which admit either closed-form or easily computable solutions. It is worth noting that the method proposed herein and that introduced in [9] share multiple characteristics, but their relative effectiveness differ, being dependent on a specific application at hand. Thus, the method of [9] was intended to be used for reconstructing the diffusion signals from their sub-critical measurements. In such a case, the sparseness of signal representation (as provided through the use of spherical ridgelets) is essential for attaining useful reconstruction results. The method proposed in the present work, on the other hand, requires fully sampled data, while allowing one to bypass the computationally extensive procedure of sparse approximation, thereby substantially improving the efficiency of the reconstruction procedure. The results comparing the performance of the proposed method with a number of alternative methods are presented in the final section of this note.

2 Proposed Method

Let $\Omega \subseteq \mathbb{R}^3$ be the spatial region over which the HARDI signal is measured. For simplicity, assume Ω is a uniform discrete rectangular lattice, formally given by $\Omega = \{(n,m,l) : 0 \leq n < N, 0 \leq m < M, 0 \leq k < L\}$. On the other hand, let $\{\mathbf{u}_k\}_{k=1}^K$ be a set of K diffusion-encoding directions, along which the diffusion signal is measured, with $\mathbf{u}_k \in \mathbb{S}^2 \triangleq \{\mathbf{u} \in \mathbb{R}^3 \mid \|\mathbf{u}\|_2 = \sqrt{\mathbf{u} \cdot \mathbf{u}} = 1\}$. Subsequently, the normalized HARDI signal E at orientation $\mathbf{u}_k$ and location $\mathbf{r} \in \Omega$ can be expressed as $E(\mathbf{u}_k, \mathbf{r}) = s(\mathbf{u}_k, \mathbf{r})/s_0(\mathbf{r})$, where $s(\mathbf{u}_k, \mathbf{r})$ is the k-th diffusion-encoded image (related to $\mathbf{u}_k$) and $s_0(\mathbf{r})$ is the corresponding b_0-image.

Practically, it is convenient to think of the whole signal E as a function of $\mathbf{u}_k$ and $\mathbf{r}$, in which case $E \in \mathbb{R}_+^{\Omega \times K}$. It is also possible to "parse" E in either the diffusion or the spatial direction. This leads to the following two ways to interpret the data structure. Specifically, E can be considered to be a "stack" of K diffusion-encoded images E_k, in which case one has $E = \{E_k\}_{k=1}^K$, with $E_k \in \mathbb{R}_+^{\Omega}$, for $k = 1, 2, \ldots, K$. Alternatively, E can be considered as a collection of K-dimensional vectors E^i of the diffusion measurements acquired at a total of $|\Omega|$ voxels within Ω. In this case, $E = \{E^i\}_{i=1}^{|\Omega|}$, with $E^i \in \mathbb{R}_+^K$, for $i = 1, 2, \ldots, |\Omega|$.

Next, let $\mathcal{Y} \in \mathbb{R}^{K \times P}$ be a $K \times P$ matrix, with its columns formed by the values of real-valued, even-degree spherical harmonics (SHs) of degree n_{max} inclusive, evaluated at $\{\mathbf{u}_k\}_{k=1}^K$ [4]. Note that the total number of such SHs is given by $P = 0.5(n_{max}+1)(n_{max}+2)$, and hence n_{max} has to be chosen so as to obey $P \leq K$. Then, congruent with the assumption in [4], we assume that for each vector of diffusion measurements E^i, there exist a vector $c^i \in \mathbb{R}^P$ of SH coefficients such that $E^i \approx \mathcal{Y}c^i$. Further, we require $\Delta^* E^i$ to have a relatively small norm, with Δ^* denoting the Laplace-Beltrami operator. Note that the fact that SHs constitute the eigenfunctions of Δ^* in conjunction with Parseval's theorem suggests that $\|\Delta^* E^i\|_2^2$ can be evaluated in terms of the SH coefficients c^i of E^i as $\|\Lambda c^i\|_2^2$, where Λ is a diagonal matrix with its i-th diagonal element $\Lambda(i,i)$ equal to $-n(n+1)$ if the i-th column of $\mathcal{Y}$ corresponds to a SH of degree n. Note that the above regularity assumptions imply that the functional

$$c^i \mapsto \|\mathcal{Y}c^i - E^i\|_2^2 + \lambda \|\Lambda c^i\|_2^2 \tag{1}$$

is likely to be minimized at an optimal c^i [4].

To account for the spatial-domain regularity of the diffusion signal, each of the diffusion-encoded images E_k is assumed to be of bounded variation [12], which suggests relatively small values of their total-variation semi-norms

$$\|E_k\|_{\mathrm{TV}} = \sum_{(n,m,l) \in \Omega} |\nabla E_k(n,m,l)|, \tag{2}$$

with ∇ standing for the operator of discrete differencing and $|\cdot|$ denoting the Euclidean norm of a vector.

Our final step is to combine both the spatial- and diffusion-domain regularity constraints into a single estimation procedure. To this end, let $c \in \mathbb{R}^{\Omega \times P}$ be the set of the SH coefficients associated with the whole HARDI signal E. Moreover,

let the operator $\mathsf{Y} : \mathbb{R}^{\Omega \times P} \to \mathbb{R}^{\Omega \times K} : c \mapsto \mathsf{Y}(c) \approx E$ be a map between the space of SH coefficients and their respective HARDI signals. Then, given E, the optimal c should solve

$$\min_{c} \left\{ \frac{1}{2} \|\mathsf{Y}(c) - E\|^2 + \frac{\lambda}{2} \|\mathit{\Lambda}(c)\|^2 + \mu \|\mathsf{Y}(c)\|_{TV} \right\}, \tag{3}$$

where $\lambda, \mu > 0$ are predefined regularization parameters, and the norms are defined as $\|\mathsf{Y}(c) - E\|^2 = \sum_{i=1}^{|\Omega|} \|\mathcal{Y} c^i - E^i\|_2^2$, $\|\mathit{\Lambda}(c)\|^2 = \sum_{i=1}^{|\Omega|} \|\mathit{\Lambda} c^i\|_2^2$, and

$$\|\mathsf{Y}(c)\|_{TV} = \Big[\sum_{k=1}^{K} \|E_k\|_{TV}^{\alpha} \Big]^{1/\alpha}, \tag{4}$$

with, e.g., $\alpha = 1$ or $\alpha = 2$ [13].

The optimization problem (3) can be applied either to the measured signal E or to its related apparent diffusion coefficients (ADC) which are equal to $\tilde{E} \triangleq -(1/b) \log E$, with b being the b-value in use. In either case, however, solving (3) involves composite, cross-domain optimization. A particularly efficient solution to this problem is proposed below.

3 Solution Steps

The original problem (3) can be simplified by means of the ADMM approach of [11]. To this end, we replace the unconstrained optimization in (3) by a constrained optimization of the form

$$\begin{gathered} \min_{c,u} \left\{ \frac{1}{2} \|\mathsf{Y}(c) - E\|^2 + \frac{\lambda}{2} \|\mathit{\Lambda}(c)\|^2 + \mu \|u\|_{TV} \right\} \\ \text{s.t.} \quad \mathsf{Y}(c) - u = 0, \end{gathered} \tag{5}$$

where $u \in \mathbb{R}^{\Omega \times K}$ is an auxiliary optimization variable. The above problem can be readily solved using the method of augmented Lagrangian multipliers, which suggests an iterative update of c and u according to

$$\begin{aligned} (c^{t+1}, u^{t+1}) = \arg\min_{c,u} \Big\{ & \frac{1}{2} \|\mathsf{Y}(c) - E\|^2 + \frac{\lambda}{2} \|\mathit{\Lambda}(c)\|^2 + \\ & + \mu \|u\|_{TV} + \frac{\delta}{2} \|\mathsf{Y}(c) - u + p^t\|^2 \Big\} \\ p^{t+1} = p^t + (\mathsf{Y}(c^{t+1}) - u^{t+1}), & \end{aligned} \tag{6}$$

where t stands for the iteration index, $\delta > 0$ is a predefined constant (e.g., $\delta = 0.5$) and p^t denotes the vector of (augmented) Lagrange multipliers at iteration t.

Finally, the optimization over c and u can be superseded by minimization w.r.t. c and u sequentially, in which case we can "split" the above problem into two simpler sub-problems, namely

$$\min_{c} \left\{ \frac{1}{2} \|\mathsf{Y}(c) - E\|^2 + \frac{\lambda}{2} \|\mathit{\Lambda}(c)\|^2 + \frac{\delta}{2} \|\mathsf{Y}(c) - (u^t - p^t)\|^2 \right\} \tag{7}$$

and

$$\min_{u} \left\{ \frac{\delta}{2} \|u - (\Upsilon(c^{t+1}) + p^t)\|^2 + \mu \|u\|_{TV} \right\}. \tag{8}$$

It is important to note that the optimization in (7) is *separable in the spatial domain*, as it can be solved on a voxel-by-voxel basis using

$$[c^{t+1}]^i = (\mathcal{Y}^T \mathcal{Y} + \lambda \Lambda^T \Lambda)^{-1} \mathcal{Y}^T \left(\frac{E^i + \delta([u^t - p^t]^i)}{1 + \delta} \right), \tag{9}$$

for each $i = 1, 2, \ldots, |\Omega|$. At the same time, in the case when $\alpha = 1$ in (4), the optimization in (8) is *separable in the diffusion domain*, as it can be solved on a direction-by-direction basis using

$$[u^{t+1}]_k = \arg\min_{u_k} \left\{ \frac{1}{2} \|u_k - (\Upsilon(c^{t+1}) + p^t)_k\|^2 + \frac{\mu}{\delta} \|u_k\|_{\mathrm{TV}} \right\}, \tag{10}$$

for each $k = 1, 2, \ldots, K$. Note that the above problem is known as total-variation denoising, and it can be efficiently solved using, e.g., the method of [12].

In summary, the proposed algorithm alternates the (separable) minimizations in (9) and (10), followed by updating the Lagrange multipliers according to (6). The iterations are terminated once the relative difference between subsequent solutions drops below a predefined tolerance (e.g., $\leq 0.1\%$). In our experiments, an average number of required iterations to reach such a tolerance (with $p^0 = u^0 = 0$ and $\delta = 0.5$) has been observed to be 20.

4 Results

As a first step in quantitatively assessing the performance of our algorithm, we compare its results with those produced by a number of state-of-the-art methods. Specifically, as the reference methods we used: 1) the constrained TV-denoising approach of [7], 2) the unconstrained vectorized TV-denoising approach of [13], and 3) the Tikhonov-type SH fitting procedure of [4]. In our experiments, the reference methods of [7] and [13] have been observed to produce more accurate results when applied to the ADC signals $\tilde{E}$, as opposed to the original HARDI data E. For convenience, the above algorithms will be referred to as TV1 and TV2, respectively. We note that, although using different assumptions regarding the nature of measurement noises, both TV1 and TV2 would be conceptually equivalent to our method if we decided to ignore the signal regularity over the $\mathbf{u}$ coordinate by setting $\lambda = 0$ in (3). On the other hand, the reference method of [4] has been applied to both $\tilde{E}$ and E, with the resulting reconstruction procedures referred below to as SH1 and SH2, respectively. In this case also, it deserves mentioning that both SH1 and SH2 would have been equivalent to the proposed method, had we decided to ignore the spatial regularity of the diffusion data by setting $\mu = 0$ in (3). Finally, the proposed method will be referred below to as *spatially regularized spherical reconstruction* (SRSR, or simply SR^2 for short).

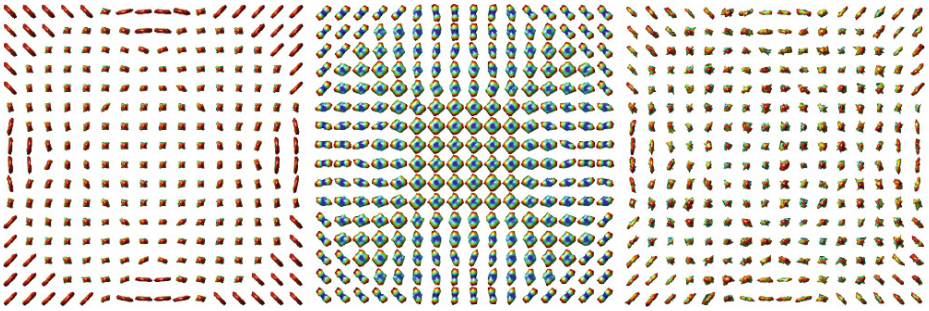

Fig. 1. (left) Phantom for in silico experiments (E), (center) ADC profile of the phantom on the left ($\overline{E}$), (right) Noisy (SNR=8) version of phantom on the left.

4.1 In Silico Experiments

To test the performance of the proposed and reference methods under controllable conditions, computer simulations have been carried out first. To this end, simulated HARDI data was generated in the form of a 16×16 array of spherical HARDI signals obeying a Gaussian mixture model (GMM) with a spatially dependent number of "crossing fibres" of various orientations [3]. The parameters of the GMMs were chosen so as to mimic a diffusion signal corresponding to a 16×16 region of interest supporting two "fibre bundles" crossing at a right angle with an additional "bundle" traversing the plane in a circular pattern. The b-value was set to 2500 s/mm^2, while the mean diffusivities and the fractional anisotropies of individual fibres comprising the Gaussian mixtures were set at random from the ranges $[0.6, 0.9]\cdot 10^{-3}$ and $[0.65, 0.85]$, respectively. All the simulated signals have been corrupted by different levels of Rician noise, giving rise to SNR values in the range [4, 20], where the SNR is given by the ratio between the maximum value of the normalized HARDI signal, E, and the scale (commonly denoted by σ) parameter of the Rician noise. The number of sampling points (diffusion-encoding directions) was set to be equal to $K = 51$ (a practical amount which suggests $n_{max} = 8$). The points have been positioned over the northern hemisphere in a quasi-uniform manner using the standard electrostatic repulsion algorithm. The original and sample noisy versions of the phantom, along with the ADC profile of the noise-free phantom and are shown in Fig. 1.

Table 1. NMSE produced by various compared methods. Each entry represents an ensemble average over 20 independent trials.

SNR	Raw	TV1	TV2	SH1	SH2	SR^2
4	0.520	0.326	0.272	0.333	0.282	**0.240**
8	0.261	0.189	0.158	0.159	0.150	**0.134**
12	0.175	0.145	0.122	0.110	0.106	**0.098**
16	0.131	0.119	0.100	0.086	0.083	**0.080**
20	0.105	0.101	0.085	0.072	0.069	**0.067**

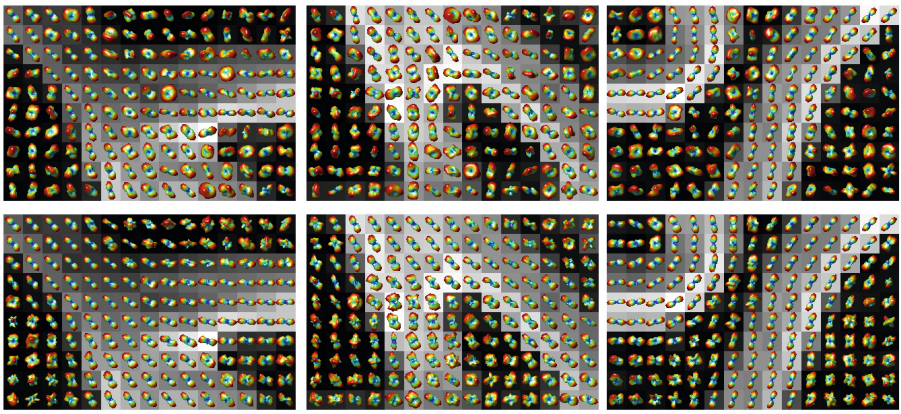

Fig. 2. Reconstructions by means of: SH2 (top row) and SR^2 (bottom row)

All the tested algorithms have been compared in terms of the normalized mean-squared error (NMSE). Moreover, the regularization parameters of the algorithms have been optimized so as to produce the smallest possible NMSE at every given level of SNR. Table 1 summarizes the results produced by different algorithms, at different levels of SNR. It clearly shows that the proposed method provides the most accurate reconstruction results.

4.2 Experiments with Real-Life Data

The real-data experiments have been performed using the data from the Fiber Cup diffusion phantom reported in [14](available at `www.lnao.fr/spip.php?article112`) with the isotropic spatial resolution of 3 mm^3, $b = 2000$ s/mm^2 and $K = 64$. No averaging over multiple acquisitions was applied.

In this section, as a reference method we use SH2, which has shown to be the "second best" according to the simulation results. Fig. 2 compares the performance of SH2 (top row) with that of the proposed SR^2 algorithm (bottom row). In both cases, the regularization parameters of the algorithms were tuned to result in the most consistent reconstructions with respect to the known geometry of the phantom. The portions of the phantom shown in Fig. 2 (with their related b_0-image in the background) correspond to two regions of "fibre branching" (left and center) and to a region with a "sharp turn" (right). It is seen that the underlying diffusion profiles are much more accurately resolved using the orientation distribution functions (ODFs) [3] computed based on the output of SR^2, as compared to the case of the reference method.

Finalizing the present article, we should note that the smoothness properties of the proposed method depend on the value of n_{max}, which requires a careful choice of K to avoid overly smoothing the reconstruction (with $K \geq 50$ being a reasonable minimum for $b \geq 2000$ s/mm^2).

References

[1] Basser, P.J., Mattiello, J., LeBihan, D.: Estimation of the effective self-diffusion tensor from the NMR spin echo. J. Magn. Reson. 103(3), 247–254 (1994)
[2] Tuch, D.S., Reese, T.G., Wiegell, M.R., Makris, N., Belliveau, J.W., Wedeen, V.J.: High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. Magn. Reson. Med. 48, 577–582 (2002)
[3] Tuch, D.S.: Q-ball imaging. Magn. Reson. Med. 52, 1358–1372 (2004)
[4] Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: Regularized, fast, and robust analytical Q-ball imaging. Magn. Res. Med. 58(3), 497–510 (2007)
[5] Michailovich, O., Rathi, Y.: On approximation of orientation distributions by means of spherical ridgelets. IEEE Trans. Image Proc. 19(3), 1–17 (2010)
[6] Aja-Fernández, S., Niethammer, M., Kubicki, M., Shenton, M., Westin, C.-F.: Restoration of DWI data using a Rician LMMSE estimator. IEEE Trans. Med. Imaging 27(10), 1389–1403 (2008)
[7] Kim, Y., Thompson, P.M., Toga, A.W., Vese, L., Zhan, L.: HARDI denoising: Variational regularization of the spherical apparent diffusion coefficient *sADC*. In: Prince, J.L., Pham, D.L., Myers, K.J. (eds.) IPMI 2009. LNCS, vol. 5636, pp. 515–527. Springer, Heidelberg (2009)
[8] Wiest-Daesslé, N., Prima, S., Coupé, P., Morrissey, S.P., Barillot, C.: Rician noise removal by Non-Local Means filtering for low signal-to-noise ratio MRI: Applications to DT-MRI. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part II. LNCS, vol. 5242, pp. 171–179. Springer, Heidelberg (2008)
[9] Michailovich, O., Rathi, Y., Dolui, S.: Spatially regularized compressed sensing for high angular resolution diffusion imaging. IEEE Trans. Med. Imaging 30(5), 1100–1115 (2011)
[10] Chen, Y., Guo, W., Zeng, Q., Yan, X., Huang, F., Zhang, H., Guojun, H., Vemuri, B.C., Liu, Y.: Estimation, smoothing, and characterization of apparent diffusion coefficient profiles from high angular resolution DWI. In: Proceed. of CVPR, pp. 588–593 (2004)
[11] Esser, E.: Applications of Lagrangian-based alternating direction methods and connections to split Bregman. Technical Report (2009)
[12] Chambolle, A.: An algorithm for total variation minimization and applications. J. Math. Imag. Vis. 20, 89–97 (2004)
[13] Bresson, X., Chan, T.F.: Fast dual minimization of the vectorial total variation norm and applications to color image processing. Inverse Problems and Imaging 2(4), 455–484 (2008)
[14] Poupon, C., Rieul, B., Kezel, I., Perrin, M., Poupon, F., Mangin, J.F.: New diffusion phantoms dedicated to the study and validation of high angular resolution diffusion imaging (HARDI) models. Magn. Res. Med. 60(6), 1276–1283 (2008)