Accelerated Diffusion Spectrum Imaging
with Compressed Sensing Using Adaptive Dictionaries

Berkin Bilgic[1], Kawin Setsompop[2,3], Julien Cohen-Adad[2,3], Van Wedeen[2,3],
Lawrence L. Wald[2,3,4], and Elfar Adalsteinsson[1,4]

[1] Massachusetts Institute of Technology, MA, USA
[2] A.A. Martinos Center for Biomedical Imaging, MA, USA
[3] Harvard Medical School, MA, USA
[4] Harvard-MIT Division of Health Sciences and Technology, MA, USA

Abstract. Diffusion Spectrum Imaging (DSI) offers detailed information on complex distributions of intravoxel fiber orientations at the expense of extremely long imaging times (~1 hour). It is possible to accelerate DSI by sub-Nyquist sampling of the q-space followed by nonlinear reconstruction to estimate the diffusion probability density functions (pdfs). Recent work by Menzel et al. imposed sparsity constraints on the pdfs under wavelet and Total Variation (TV) transforms. As the performance of Compressed Sensing (CS) reconstruction depends strongly on the level of sparsity in the selected transform space, a dictionary specifically tailored for sparse representation of diffusion pdfs can yield higher fidelity results. To our knowledge, this work is the first application of adaptive dictionaries in DSI, whereby we reduce the scan time of whole brain DSI acquisition from 50 to 17 min while retaining high image quality. *In vivo* experiments were conducted with the novel 3T *Connectome* MRI, whose strong gradients are particularly suited for DSI. The RMSE from the proposed reconstruction is up to 2 times lower than that of Menzel et al.'s method, and is actually comparable to that of the fully-sampled 50 minute scan. Further, we demonstrate that a dictionary trained using pdfs from a single slice of a particular subject generalizes well to other slices from the same subject, as well as to slices from another subject.

1   Introduction

Diffusion weighted MR imaging is a widely used method to study the interconnectivity and structure of the brain. Diffusion Tensor Imaging (DTI) is an established diffusion weighted imaging method, which models the diffusion as a univariate Gaussian distribution (1). One limitation of this model arises in the presence of fiber crossings, and this can be addressed by using a more involved imaging method. Diffusion Spectrum Imaging (DSI) samples the full q-space and yields a complete description of the diffusion probability density function (pdf) (2). While DSI is capable of resolving complex distributions of intravoxel fiber orientations, full q-space coverage comes at the expense of substantially long scan times (~1 hour).

Compressed Sensing (CS) comprises algorithms that recover data from undersampled acquisitions by imposing sparsity or compressibility assumptions on the reconstructed

images (3). In the domain of DSI, acceleration with CS was successfully demonstrated by Menzel et al. (4) by imposing wavelet and Total Variation (TV) penalties in the pdf space. Up to an undersampling factor of 4 in q-space, it was reported that essential diffusion properties such as orientation distribution function (odf), diffusion coefficient, and kurtosis were preserved (4).

The performance of CS recovery depends strongly on the level of sparsity of the signal in the selected transform domain (3). While prespecified transformations such as wavelets and spatial gradients yield sparse signal representation, tailoring the sparsifying transform based on the characteristics of the particular signal type may offer even sparser results. K-SVD is an algorithm that designs a dictionary that achieves maximally sparse representation of the input training data (5). The benefit of using data-driven, adaptive dictionaries trained with K-SVD was also demonstrated in CS reconstruction of structural MR imaging (6).

In this work, we employ the K-SVD algorithm to design a sparsifying transform that captures the structure in diffusion pdfs and yields a signal representation with increased level of sparsity. Coupling this adaptive dictionary with the FOcal Underdetermined System Solver (FOCUSS) algorithm (7), we obtain a parameter-free CS algorithm. With 3-fold undersampling of q-space, we demonstrate *in vivo* up to 2-fold reduced pdf reconstruction errors relative to our implementation of the CS algorithm that uses wavelets and variational penalties by Menzel et al. (4). At higher acceleration factors of 5 and 9, we still demonstrate substantial improvement. For additional validation, the RMSE of the reconstructed 'missing' diffusion images were calculated by comparing them to a gold standard dataset obtained with 10 averages. In this case, dictionary-based reconstructions were seen to be comparable to the fully-sampled 1 average data. Further, we show that a dictionary trained on data from a particular subject generalizes well to reconstruction of another subject's data, still yielding significantly reduced reconstruction errors. Hence, application of the proposed method might reduce a typical 50-minute DSI scan to 17 minutes (upon 3× acceleration) while retaining high image quality. Additionally, we also investigate using a simple ℓ_1-norm penalty in the pdf space with the FOCUSS algorithm, and show that this approach gives comparable results to the more involved wavelet- and TV-based reconstruction by Menzel et al. (4), while being computationally more efficient.

2   Theory

CS Recovery with Prespecified Transforms

Letting $\boldsymbol{p} \in \mathbb{C}^N$ represent the 3-dimensional diffusion pdf at a particular voxel as a column vector, and $\boldsymbol{q} \in \mathbb{C}^M$ denote the corresponding undersampled q-space information, CS recovery with wavelet and TV penalties aim to solve the convex optimization problem at a single voxel,

$$min_{\boldsymbol{p}} \ \|\mathbf{F}_\Omega \boldsymbol{p} - \boldsymbol{q}\|_2^2 + \alpha \cdot \|\boldsymbol{\Psi} \boldsymbol{p}\|_1 + \beta \cdot \text{TV}(\boldsymbol{p}) \qquad\qquad \text{(Eq. 1)}$$

where $\mathbf{F}_\Omega$ is the undersampled Fourier transform operator, $\boldsymbol{\Psi}$ is a wavelet transform operator, $\text{TV}(.)$ is the Total Variation penalty, and α and β are regularization parameters that need to be determined.

Training an Adaptive Transform with K-SVD

Given an ensemble $\mathbf{P} \in \mathbb{C}^{N \times L}$ formed by concatenating L example pdfs $\{\boldsymbol{p}_i\}_{i=1}^{L}$ collected from a training dataset as column vectors, the K-SVD algorithm (5) aims to find the best possible dictionary for the sparse representation of this dataset by solving,

$$min_{\mathbf{P},\mathbf{D}} \sum_{i=1}^{L} \|\boldsymbol{x}_i\|_0 \quad \text{subject to} \quad \|\mathbf{P} - \mathbf{DX}\|_F^2 \leq \varepsilon \qquad \text{(Eq. 2)}$$

where $\mathbf{X}$ is the matrix that contains the transform coefficient vectors $\{\boldsymbol{x}_i\}_{i=1}^{L}$ as its columns, $\mathbf{D}$ is the adaptive dictionary formed by example pdfs, ε is a fixed constant that adjusts the data fidelity, and $\|.\|_F$ is the Frobenius norm. The K-SVD works iteratively, first by fixing $\mathbf{D}$ and finding an optimally sparse $\mathbf{X}$ using orthogonal matching pursuit, then by updating each column of $\mathbf{D}$ and the transform coefficients corresponding to this column to increase data consistency.

CS Recovery with an Adaptive Transform Using FOCUSS

The FOCUSS algorithm aims to find a sparse solution to the underdetermined linear system $\mathbf{F}_\Omega \mathbf{D} \boldsymbol{x} = \boldsymbol{q}$, where $\boldsymbol{x}$ is the vector of transform coefficients in the transform space defined by the dictionary $\mathbf{D}$ using the following iterations,

For iteration number $t = 1, \dots T$,

$$W_{j,j}^t = diag\left(\left|x_j^t\right|^{1/2}\right) \qquad \text{(Eq. 3)}$$

$$\boldsymbol{s}^t = \text{argmin}_{\boldsymbol{s}} \|\boldsymbol{s}\|_2^2 \quad \text{such that} \quad \mathbf{F}_\Omega \mathbf{D} \mathbf{W}^t \boldsymbol{s} = \boldsymbol{q} \qquad \text{(Eq. 4)}$$

$$\boldsymbol{x}^{t+1} = \mathbf{W}^t \boldsymbol{s}^t \qquad \text{(Eq. 5)}$$

Here, $\mathbf{W}^t$ is a diagonal weighting matrix whose jth diagonal entry is denoted as $W_{j,j}^t$, $\boldsymbol{x}^t$ is the estimate of transform coefficients at iteration t whose jth entry is x_j^t. The final reconstruction in diffusion pdf space is obtained via the mapping $\boldsymbol{p} = \mathbf{D}\boldsymbol{x}^{T+1}$.

We note that it is possible to impose sparsity-inducing ℓ_1 penalty directly on the pdf coefficients by taking $\mathbf{D}$ to be the identity matrix $\mathbf{I}$. A detailed description of application of FOCUSS algorithm to MRI can be found in (8), where it is shown that reweighted ℓ_2 norm solutions for the auxiliary variable $\boldsymbol{s}^t$ induce ℓ_1 penalty on $\boldsymbol{x}^t$.

3   Methods

Diffusion EPI acquisitions were obtained from three healthy volunteers (subjects A, B and C) using a novel 3T system (Magnetom Skyra Connectom, Siemens Healthcare, Erlangen, Germany) equipped with the AS302 "Connectom" gradient with $G_{max} = 300$ mT/m (here we used $G_{max} = 200$ mT/m) and Slew = 200 T/m/s. A custom-built 64-channel RF head array (9) was used for reception with imaging parameters of 2.3 mm isotropic voxel size, FOV = 220×220×130, matrix size = 96×96×57, $b_{max} = 8000$ s/mm^2, 514 directions full sphere q-space sampling organized in Cartesian grid with interspersed b=0 image every 20 TRs (for motion correction), in-plane acceleration = 2× (using GRAPPA algorithm), TR/TE = 5.4 s / 60 ms, total imaging time ~50 min. In addition, at

5 q-space points ($[1,1,0]$, $[0,2,-1]$, $[0,0,3]$, $[0,4,0]$, and $[5,0,0]$) residing on 5 different shells, 10 averages were collected for noise quantification. Eddy current related distortions were corrected using the reversed polarity method (10). Motion correction (using interspersed b=0) was performed using FLIRT (11) with sinc interpolation.

Variable-density undersampling (using a power-law density function (3)) with $R = 3$ acceleration was applied in q-space on a 12×12×12 grid. Three different adaptive dictionaries were trained with data from slice 30 of subjects A, B and C. Reconstruction experiments were applied on test slices that are different than the training slices. In particular, two reconstruction experiments were performed. First, voxels in slice 40 of subject A were retrospectively undersampled in q-space, and reconstructed using 5 different methods: wavelet+TV method of Menzel et al. (4), ℓ_1-regularized FOCUSS, and Dictionary-FOCUSS with the three dictionaries trained on three different subjects. Second, voxels in slice 25 of subject B were undersampled with the same $R = 3$ sampling pattern, and again reconstructed with wavelet+TV, ℓ_1-FOCUSS, and the three dictionaries trained on three different subjects. For Menzel et al.'s method, Haar wavelets in MATLAB's wavelet toolbox were used. The regularization parameters α and β in Eq.1 were chosen by parameter sweeping with values $\{10^{-4}, 3 \cdot 10^{-4}, 10^{-3}, 3 \cdot 10^{-3}\}$ to minimize the reconstruction error of 100 randomly selected voxels in slice 40 of subject A. The optimal regularization parameters were found to be $\alpha = 3 \cdot 10^{-4}$ for wavelet and $\beta = 10^{-4}$ for the TV term. By taking the fully-sampled data as ground-truth, the fidelity of the five methods were compared using root-mean-square error (RMSE) normalized by the ℓ_2-norm of ground-truth as the error metric both in pdf domain and q-space.

Since the fully-sampled data are corrupted by noise, computing RMSEs relative to them will include contributions from both reconstruction errors and additive noise. To address this, the additional 10 average data acquired at the selected 5 q-space points were used. As a single average full-brain DSI scan takes ~50 min, it was not practical to collect 10 averages for all of the undersampled q-space points. As such, we rely on both error metrics, namely: the RMSE relative to one average fully-sampled dataset and the RMSE relative to gold standard data for 5 q-space points.

4   Results

Fig.1 depicts the error of the five reconstruction methods in the pdf domain for each voxel in slice 40 of subject A. At $R = 3$ acceleration, reconstruction error of Menzel et al.'s method averaged over brain voxels in the slice was 15.8%, while the error was 15.0% for ℓ_1-regularized FOCUSS. Adaptive dictionary trained on subject A yielded 7.8% error. Similarly, reconstruction with dictionaries trained on pdfs of the other subjects B and C returned 7.8% and 8.2% RMSE, respectively. At $R = 5$ and 9, Dictionary-FOCUSS with training on subjects C and B returned 9.3% and 10.0% RMSE, respectively.

In Fig.2, reconstruction errors at $R = 3$ on slice 25 of subject B are presented. In this case, Menzel et al.'s method yielded 17.5% average RMSE, and ℓ_1-FOCUSS had 17.3% error. Dictionary trained on slice 40 of subject B returned 11.4% RMSE, while adaptive transforms trained on subjects A and C had 11.4% and 11.8% error, respectively. At higher acceleration factors of $R = 5$ and 9, Dictionary-FOCUSS with training on subjects C and A returned 13.5% and 14.2% RMSE, respectively.
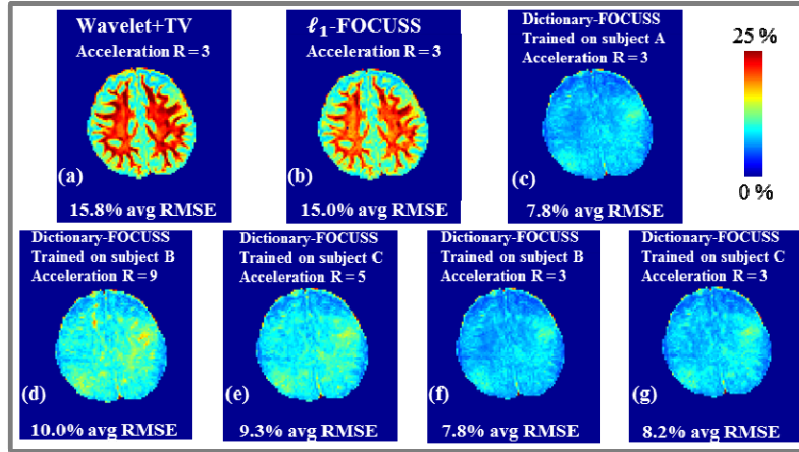
Fig. 1. RMSE at each voxel in slice 40 of subject A upon $R=3$ acceleration and reconstruction with Menzel et al.'s method (a), ℓ_1-FOCUSS (b), Dictionary-FOCUSS trained on subjects A (c), B (f), and C (g). Dictionary-FOCUSS errors in (d) and (e) are obtained at higher acceleration factors of $R=9$ and 5.
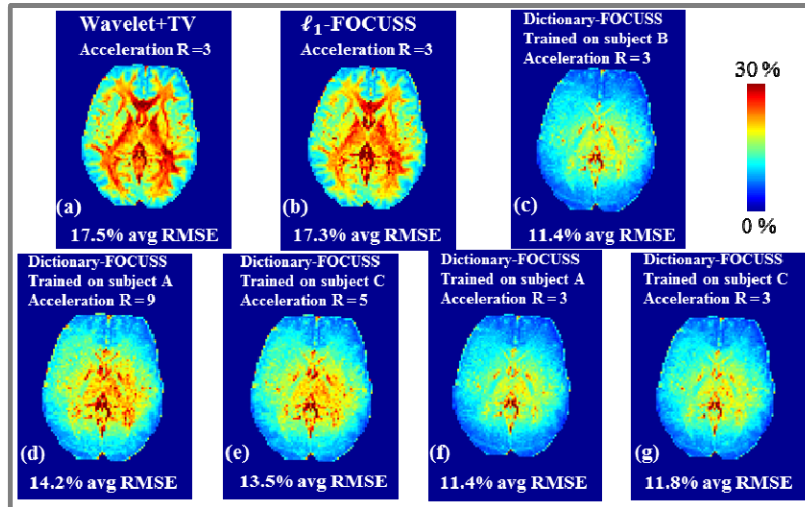

Fig. 2. RMSE at each voxel in slice 25 of subject B upon $R=3$ acceleration and reconstruction with Menzel et al.'s method (a), ℓ_1-FOCUSS (b), Dictionary-FOCUSS trained on subjects B (c), A (f), and C (g). Dictionary-FOCUSS errors in (d) and (e) are obtained at higher acceleration factors of $R=9$ and 5.

Reconstruction errors in q-space images of subject A obtained with Wavelet+TV, ℓ_1-FOCUSS and Dictionary-FOCUSS for the undersampled q-space directions are plotted in Fig.3. For two particular diffusion directions, q-space reconstructions obtained with the three methods are also presented.

To allow inter-subject comparison, slices that correspond to approximately the same anatomical region in subjects A, B and C were also reconstructed. As shown in Fig.1f, slice 40 of subject A when reconstructed using the dictionary trained on subject B gave 7.8% average error. Slices 38 of subjects B and C yielded 7.3% and 9.3% RMSE when reconstructed with the dictionary trained on subject A upon 3-fold undersampling (not shown).

In an attempt to quantify the noise in q-space and separate it from CS reconstruction error, we take the 10 average data acquired at 5 q-space directions as ground truth and compute RMSEs relative to them. Fig.4 shows the error plots for the 1 average fully sampled data, Wavelet+TV, ℓ_1-FOCUSS, and Dictionary-FOCUSS reconstructions relative to the 10 average data for slices from subjects A and B.
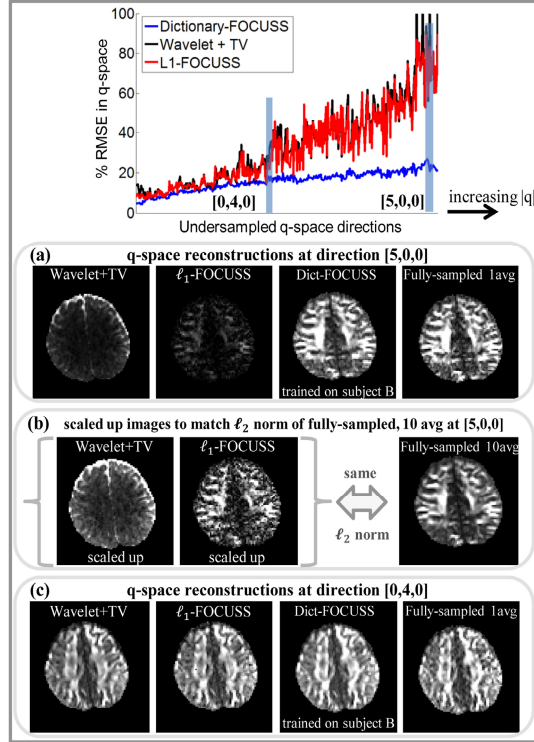

Fig. 3. Top: RMSEs in 'missing' q-space directions estimated with Wavelet+TV, ℓ_1 - and Dictionary-FOCUSS at R=3. q-space images at directions [5,0,0] (a) and [0,4,0] (c) are also depicted. In (b), Wavelet+TV and ℓ_1 recons of dir. [5,0,0] are scaled up to have same ℓ_2 norm as fully-sampled 10 avg .
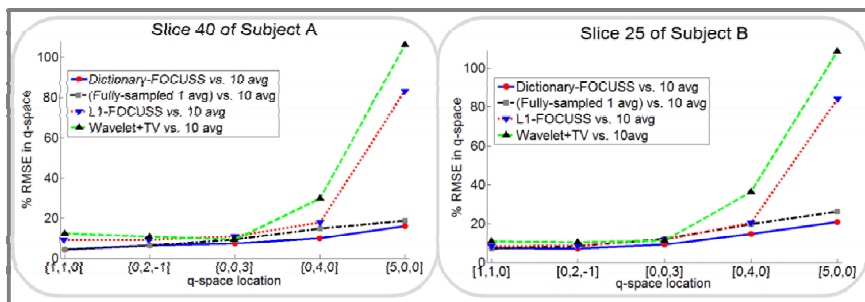
Fig. 4. Panel on the left depicts RMSEs of Wavelet+TV, ℓ_1-FOCUSS and Dictionary-FOCUSS at R=3 and fully-sampled 1 average data computed in 5 q-space locations relative to the 10 average data for subject A. Panel on the right shows the same comparison for the slice belonging to subject B.

Fig.5 presents Orientation Distribution Function (ODF) visualization of reconstructions obtained at 3-fold acceleration using Wavelet+TV and Dictionary-FOCUSS, and compares the tractography solutions obtained with adaptive reconstruction and fully-sampled data.
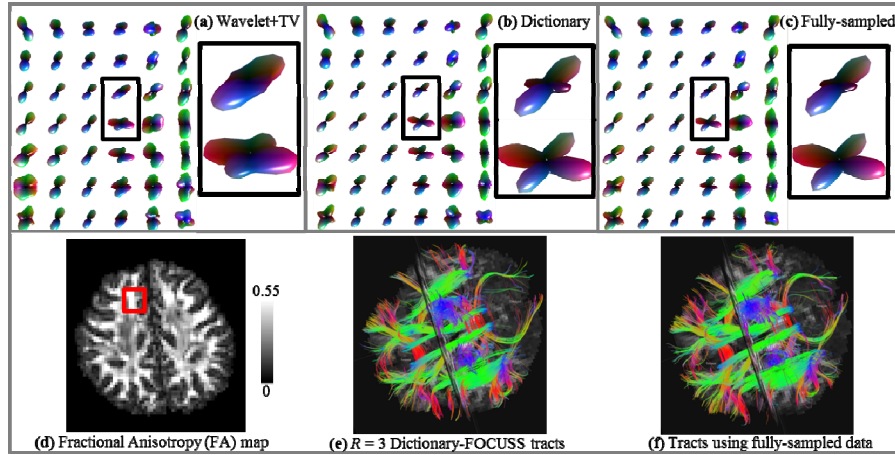
Fig. 5. ODFs for subject A using Wavelet+TV (a), dictionary (b), and fully-sampled data (c) within the ROI in the FA map in (d). Tracts with R=3 dictionary recon (e) and fully-sampled data (f) are also presented.

5 Discussion

This work presented the first application of adaptive transforms to voxel-by-voxel CS reconstruction of undersampled q-space data. Relative to reconstruction with prespecified transforms, the proposed algorithm has up to 2 times reduced error in the pdf domain at the same acceleration factor ($R = 3$), while requiring no regularization parameter tuning. When the undersampling ratio was increased to $R = 5$ and even up to $R = 9$, the proposed method still demonstrated substantial improvement relative to using prespecified transforms at $R = 3$ (Figs.1 and 2). This will render DSI clinically feasible, by cutting a 50 min scan to 5.5 min upon 9-fold acceleration. As demonstrated, a dictionary trained with pdfs from a single slice of a particular subject generalizes to other slices of the same subject, as well as to different subjects. However, further tests are needed to see if dictionaries can generalize across healthy and patient populations, across age groups, or fundamentally different anatomical locations.

Since the acquired 1 average DSI data is corrupted by noise (especially in the outer shells), it is desired to obtain noise-free data for more reliable computation of CS reconstruction errors. Because even the 1 average full-shell acquisition takes ~50 min, it is practically not possible to collect multiple-average data at all q-space points. To address this, one representative q-space sample at each shell was collected with 10 averages to serve as "(approximately) noise-free" data. When the noise-free data were

taken to be ground-truth, the dictionary reconstruction with 3-fold undersampling was comparable to the fully-sampled 1 average data for both subjects (Fig.4).

RMSE in Fig.2 was overall higher than in Fig.1. A possible explanation is the inherently lower signal-noise-ratio (SNR) in the lower axial slice, particularly in the center area of the brain which is further away from the receive coils. In particular, the error is higher in the central region of the image where the SNR is expected to be lowest. Future work includes a detailed analysis of how SNR level affects the reconstruction performance of the proposed CS algorithm and the dictionary learning step.

As seen in Fig.3, wavelet and TV penalized reconstruction and ℓ_1-FOCUSS yield especially poor quality results in estimating the high q-space samples. In particular, as depicted in Fig.3a, these CS methods tend to underestimate the high q-space content. However, this is not a simple scaling problem, even when q-space images are scaled to have the same ℓ_2-norm as the fully-sampled 10 average data, they yield either flat (Wavelet+TV) or grainy (ℓ_1-FOCUSS) results (Fig.3b). ODF visualization and tractography solutions in Fig.5 show good agreement between the adaptive reconstruction and the fully-sampled dataset. Average FA and tract volume metrics obtained from 18 major white-matter pathways were seen to support this good agreement (data not shown due to space limitation). As the tract results show that main fiber bundles are not corrupted, adaptive reconstruction causes no regional bias in the reconstruction.

In our implementation, *per voxel* processing time of ℓ_1-FOCUSS was 0.6 seconds, while this was 12 seconds for Dictionary-FOCUSS and 27 seconds for Wavelet+TV method on a workstation with 12GB memory and 6 processors. Hence, full-brain reconstruction using the Dictionary-FOCUSS algorithm would still take *several days*. Because each voxel can be processed independently, parallel implementation is likely to be a significant source of performance gain.

The proposed CS acquisition/reconstruction can be combined with other techniques to further reduce the acquisition time and/or improve reconstruction quality. In particular, combining the proposed method with the Blipped-CAIPI Simultaneous MultiSlice (SMS) acquisition (12) could reduce a 50 minute DSI scan to mere 5.5 minutes upon 9-fold acceleration (3×3 CS-SMS) while retaining high image quality.

References

1. Basser, P.J., Mattiello, J., LeBihan, D.: MR diffusion tensor spectroscopy and imaging. Biophys. J. 66(1), 259–267 (1994)
2. Wedeen, V.J., et al.: Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. Magn. Reson. Med. 54(6), 1377–1386 (2005)
3. Lustig, M., Donoho, D., Pauly, J.M.: Sparse MRI: The application of compressed sensing for rapid MR imaging. Magn. Reson. Med. 58(6), 1182–1195 (2007)
4. Menzel, M.I., et al.: Accelerated diffusion spectrum imaging in the human brain using compressed sensing. Magn. Reson. Med. 66(5), 1226–1233 (2011)
5. Aharon, M., Elad, M., Bruckstein, A.: K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. IEEE T. Signal Proces. 54(11), 4311–4322 (2006)

6. Ravishankar, S., Bresler, Y.: MR image reconstruction from highly undersampled k-space data by dictionary learning. IEEE Trans. Med. Imaging 30(5), 1028–1041 (2011)
7. Gorodnitsky, I.F., Rao, B.D.: Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm. IEEE T. Signal Proces. 45(3), 600–616 (1997)
8. Ye, J.C., Tak, S., Han, Y., Park, H.W.: Projection reconstruction MR imaging using FOCUSS. Magn. Reson. Med. 57(4), 764–775 (2007)
9. Keil, B., et al.: A 64-channel brain array coil for 3T imaging. In: 20th Annual ISMRM Scientific Meeting and Exhibition (2012)
10. Bodammer, N., et al.: Eddy current correction in diffusion-weighted imaging using pairs of images acquired with opposite diffusion gradient polarity. Magn. Reson. Med. 51(1), 188–193 (2004)
11. Jenkinson, M., et al.: Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17(2), 825–841 (2002)
12. Setsompop, K., et al.: Blipped-controlled aliasing in parallel imaging for simultaneous multislice Echo Planar Imaging with reduced g-factor penalty. Magn. Reson. Med. (2011)