Low-Rank Plus Sparse Reconstruction Using Dictionary Learning For 3D-MRI

Wenxiong Zhong[1,2,3], Dongxiao Li[1,2,3], Lianghao Wang[1,2,3], Ming Zhang[1,2,3]
[1]College of Information Science and Electronic Engineering, Zhejiang University, P.R. China
[2]Zhejiang Provincial Key Laboratory of Information Processing, Communication and Networking, P.R. China
[3]State Key Lab. for Novel Software Technology, Nanjing University, P.R. China

***Abstract*—This work proposes a low-rank plus sparse model using dictionary learning for 3D-MRI reconstruction from down-sampling k-space data. The scheme decomposes the dynamic image signal into two parts: low-rank part L and sparse part S and then, constructing it as a constrained optimization problem. In the optimization process,a nonconvex penalty function is used to optimize the low rank part L. The sparse part S is expressed by a over-complete dictionary using blind compressed sensing and we formulate the sparsity of coffecient matrix using l_1 norm. To avoid the ill-posed of the problem, the Frobenius norm is used in dictionary. We adopt an alternate optimization algorithm to solve the problem, which cycles through the minimization of five subproblems. Finally, we prove the effectiveness of proposed method in two cardiac cine data sets. Experimental results were compared with exsiting L+S, L&S and BCS schemes, which demonstrate that the proposed method behaves better in removal of artifacts and maintaining the image details.**

***Keywords*—Blind compressed sensing(BCS); dynamic magnetic resonance imaging(dMRI); low-rank plus sparse; dictionary learning; undersampled reconstrution**

I. Introduction

Dynamic magnetic resonance imaging (dMRI) is mainly used to obtain the dynamic changes of organs and it is an effective assistance in medical diagnosis such as cardiac, liver and lungs imaging. One of challenging tasks of dMRI is to accomplish High spatial and temporal resolution imaging. It requires to reconstruct the images from undersampled k-space data due to the slow data acquisition of hardwaries. Since the undersampled data do not meet the requirements of Nyquist sampling rate, the recovery problem is ill-posed.

The compressed sensing (CS) approaches have been applied to dMRI to address this issue recently. Compressed sensing points out that one can achieve accurate recovery from undersampled data if the signal is compressible or sparse in a certain transform domain and the sampling basis is incoherent [1][2]. Related methods like k-t FOCUSS [3] impose sparsity constraints in the Fourier transform domain. In k-t SLR [5], the author imposes the sparsity constraint in the Karhunen Louve transform domain.

In above approaches, a sparse transform domain has to be found as a prerequisite. The commonly used sparse transform domain includes discrete cosine transform, fast Fourier transform, discrete wavelet transform, Curvelet transform, etc. However, these fixed sparsifying basis are not flexible enough to deal with diverse data adptively. To address this issue, approaches to learn the sparsifying basis from data itself were proposed [6]-[8], which provide more flexible representation. The blind compressed sensing [9] technique combines the advantages of both compressed sensing and dictionary learning, which do not require heuristic knowledge about sparsity basis for the signal recovery and learn a dictionary adptively.

Inspired by the L+S reconstruction model proposed by Ricardo et al. [4] and Blind Compressed Sensing(BCS) framework proposed by Sivan et al.[9], we propose a low-rank plus sparse scheme using dictionary learning for accelerated dMRI. The L+S decomposition represents the signal more efficiently in which the L part represents static background and S part represents dynamic foreground. Instead of using a nuclear norm constraint on L part , we use a generalized Huber function [10] as the rank penalty. Results shows that the nonconvex penalty outperforms the convex nuclear norm. For the S part, we replace the fixed basis with a dictionary estimated from the undersampled data using BCS scheme. Unlike usual dictionary learning algorithm, BCS framework doesn't require a large number of training data and learns the sparsifying dictionary during signal reconstruction. Then, we reinforce the sparsity constraint in S part with the dictionary, which results in less noises and more clear details in the reconstruction images.

II. Methods

Let $M \in C^{p\times q}$ denotes the unknown image matrix to be recovered from undersampled k-space data. Here, p represents the number of elements in a single frame, q represents the number of images in the data set. The L+S model points out that the matrix M is a decomposition of low-rank part L and sparse part S. The decomposition is unique as long as the condition satisfies L part and S part is incoherent. In general, the S part is not sparse enough itself and a sparsifying basis has to be found such that S has a sparse representation on it. As a fixed sparsifying basis is not flexible enough, a better idea is to learn the basis from the data itself adptively, which is known as dictionary learning.

A. L+S model using dictionary learning

In our proposed method, the dictionary learning method called BCS is employed to estimate the sparsifying dictionary of S component. In general, the dictionary learning process has two phases: train for the dictionary and then, the learned dictionary is employed for reconstruction. The BCS technique

978-1-5090-3710-0/16/$31.00 ©2016 IEEE

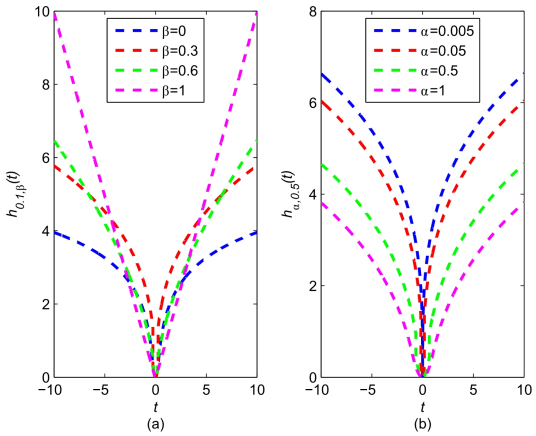

Fig. 1: (a) Huber function with α to be fixed as 0.1 (b) Huber function with β to be fixed as 0.5

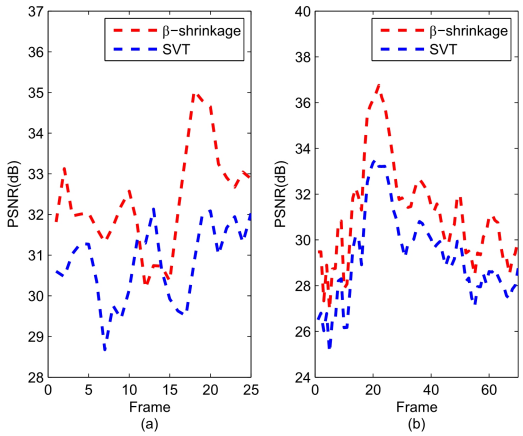

Fig. 2: Comparison of β-shrinkage and SVT of (a) data set 1 (b) data set 2

combines the two phases - it estimates the sparse matrix and dictionary simultaneously during the reconstruction.

The sparse part S of the unknown images is represented as the product of two matrices - sparse matrix $\Phi_{p\times N}$ and dictionary matrix $\Psi_{N\times q}$:

$$\begin{bmatrix} \phi_1(x_1) & \dots & \phi_N(x_1) \\ \dots & \dots & \dots \\ \phi_1(x_p) & \dots & \phi_N(x_p) \end{bmatrix} \begin{bmatrix} \psi_1(t_1) & \dots & \psi_1(t_q) \\ \dots & \dots & \dots \\ \psi_N(t_1) & \dots & \psi_N(t_q) \end{bmatrix} \quad (1)$$

Here, N is the number of temporal basis in the dictionary, ϕ_i is the ith spatial coefficient vector and ψ_i is the ith temporal basis vector. N is large enough to ensure the dictionary is overcomplete , which is not necessarily orthogonal. The matrix Φ is assumed to be a sparse matrix, which means it has very few nonzero elements. Let K denotes the data acquired in k-t space and E denotes the inverse Fourier transform. Specifically, for radial or spiral k-space data, E denotes the nonuniform Fourier transform [11]. The recovery of the signal matrix M is perfomed by tackling the constrained optimization problem as follow:

$$\min_{L,\Phi,\Psi} \|E(L+\Phi\Psi)-K\|_2^2 + \lambda_1 rank(L) + \lambda_2 \|\Phi\|_1$$
$$subject \quad to \|\Psi\|_2 \le \delta \quad (2)$$

where the first term denotes the data fidelity term to ensure the consistency between k-space measurement and reconstructed images, and λ_1, λ_2 are trade off parameters. The l_1 norm constraint on Φ is an approximate to l_0 norm, which ensures the sparsity of matrix Φ. The Frobenius norm constraint on Ψ is used to avoid scale ambiguity. Otherwise, the optimization will be ill-posed for the reason that matrix Φ can be arbitrarily small in magnitude.

B. The optimization Algorithm

To address the constrained optimization problem in (2), we start by introducing two proxy variables for the variables Γ, Θ respectively. Then, the rank penalty term and norm term were rewritten as follow [13]:

$$\min_{\Gamma} \rho_1 \|L-\Gamma\|_F^2 + rank(\Gamma) \quad (3)$$

$$\min_{\Theta} \rho_2 \|\Phi-\Theta\|_F^2 + \|\Theta\|_1 \quad (4)$$

where ρ_1, ρ_2 is a huge positive constant. Substituting (3)(4) in (2), and rewriting it as using regularization rather than strict constraints, we obtain the following modified Lagrange function :

$$\min_{L,\Gamma,\Phi,\Theta,\Psi} \|E(L+\Phi\Psi)-K\|_2^2 + \lambda_1(\rho_1 \|L-\Gamma\|_F^2 +$$
$$rank(\Gamma)) + \lambda_2(\rho_2 \|\Phi-\Theta\|_F^2 + \|\Theta\|_1) + \gamma(\|\Psi\|_2 - \delta) \quad (5)$$

We solve for the object function (5) by an alternating minimization algorithm. The problem is divided into five subproblems. Each subproblem is solved by supposing all the other variables are not change. We employ an iterative approach through all subproblems until all variables convergent. For the proxy variabes Γ, Θ, respective subproblems are shown in (3)(4). For the other variables, we express these subproblems as follow:

$$\min_{L} \|E(L+\Phi\Psi)-K\|_2^2 + \lambda_1\rho_1 \|L-\Gamma\|_F^2 \quad (6)$$

$$\min_{\Phi} \|E(L+\Phi\Psi)-K\|_2^2 + \lambda_2\rho_2 \|\Phi-\Theta\|_F^2 \quad (7)$$

$$\min_{\Psi} \|E(L+\Phi\Psi)-K\|_2^2 + \gamma(\|\Psi\|_2 - \delta) \quad (8)$$

Since the subproblems (6), (7), (8) are simple least squares problems, we use conjugate gradient algorithm to solve them efficiently. Subproblem (3) is an rank minimization problem. The rank operator is not convex for the rank penalty term, which implies that we need a way to relaxation. A popularly used relaxation scheme is a convex nuclear norm [14] penalty. Inspired by approach of [10], We use a nonconvex penalty using generalize Huber function instead, which outperforms a convex penalty:

$$h_{\alpha,\beta}(t) = \begin{cases} |t|^2/2\alpha, & if\ |t| < \alpha^{1/(2-\beta)} \\ |t|^\beta/\beta - \tau, & if\ |t| \ge \alpha^{1/(2-\beta)} \end{cases} \quad (9)$$

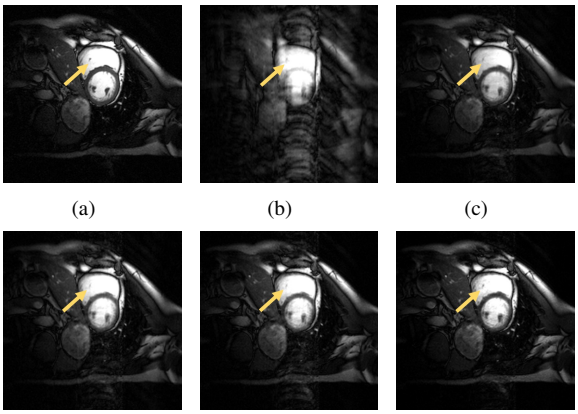

Fig. 3: data set 1 (a) ground truth (b) direct IFFT (c) L+S (d) L&S (e) BCS (f) proposed method

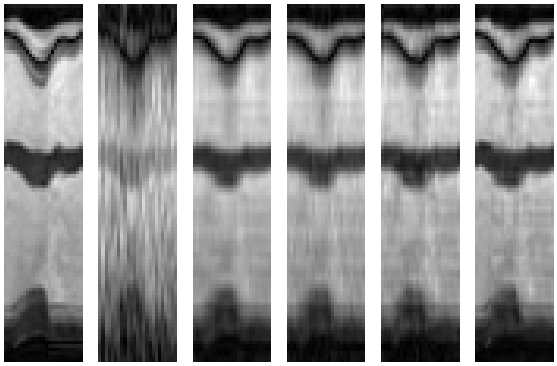

Fig. 4: Profiles of data set 1, from left to right, ground truth, direct IFFT, L+S L&S, BCS, proposed method

with $\tau = (1/\beta - 1/2)\alpha^{\beta/(2-\beta)}$. (Notice that when $\beta = 0, |t|^{\beta}/t$ becomes $log|t|$, with $\tau = (log\alpha - 1)/2$.) Examples of the β-Huber function is shown in Fig.1. It shows that the function is non-convex except for the region approaches the origin as $0 \leq \beta < 1$. As α increases, the convex region enlarges. Then, we define the rank penalty as following:

$$\|\Gamma\|_{h_{\alpha,\beta}} = \sum_{k=1}^{Rank(\Gamma)} h_{\alpha,\beta}(\sigma_k(\Gamma)), 0 \leq \beta \leq 1 \qquad (10)$$

Note that for the case $\beta = 1$, the rank penalty becomes the popular nuclear norm. Rick [10] shows that the solution of the above optimization problem is:

$$\Gamma_{n+1} = Ushrink_{\beta}(\Sigma, \alpha)V^H \qquad (11)$$

Here, $U\Sigma V^H$ represents the SVD (singular value decomposition) of $\Gamma_n, \alpha = 2/\rho_1$. The β-shrinkage operation $shrink_{\beta}(\Sigma, \alpha)$ represents the shrinkage operator[10]:

$$shrink_{\beta}(t,\alpha) = \begin{cases} 0, & if\ |t| < \alpha|t|^{(\beta-1)} \\ (|t| - \alpha|t|^{\beta-1})t/|t|, & if\ |t| \geq \alpha|t|^{(\beta-1)} \end{cases} \qquad (12)$$

Subproblem (4) is a typical l_1 norm proximal mapping problem and can be solved analytically, we solve it by performing

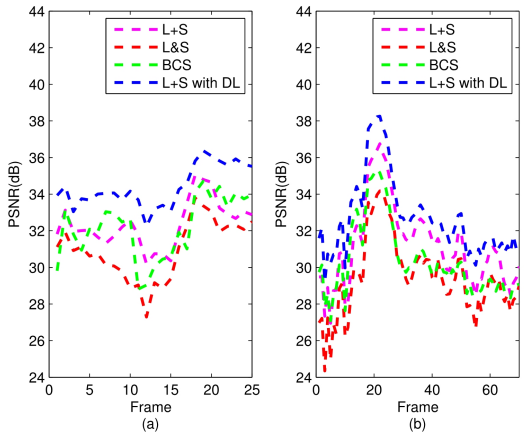

Fig. 5: PSNR of each scheme in (a)data set 1 (b)data set 2

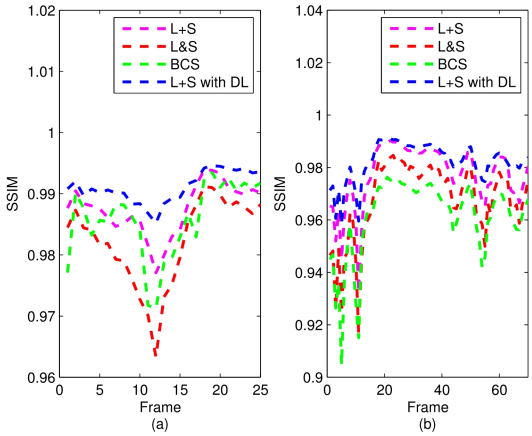

Fig. 6: SSIM of each scheme in (a)data set 1 (b)data set 2

the soft-thresholding in iterations:

$$\Theta_{n+1} = \Phi_n/|\Phi_n|/max\{|\Phi_n| - 2/\rho_2, 0\} \qquad (13)$$

We adopt a steepest descent algorithm to update the variable γ:

$$\gamma_{n+1} = max\{\gamma_n, \gamma_n + \|\Psi\|_F^2 - \delta\} \qquad (14)$$

III. Experimental Results

A. The scheme of experiments

We used two cardiac cine data sets to evaluate our proposed method. The first Cartesian data set from a 1.5T Philips scanner contains 25 frames of ground truth [3]. The size of each frame is 256×256, which implies 256 voxels along phase encoding direction and 256 voxels along frequency coding in the scanning. The second radial data set obtained from a Siemens 3T MRI by radial perfusion imaging [12]. For each frame, there are 256 samples per ray and 72 radial rays uniformly distribute around 360 degrees. The matrix size in image domain is 90×190×70. In our experiment, the undersampling factor is 6 for the first data set and 7.5 for the second data set.

We compared the performance of proposed algorithm against three compressed sensing approaches: L+S reconstruc-

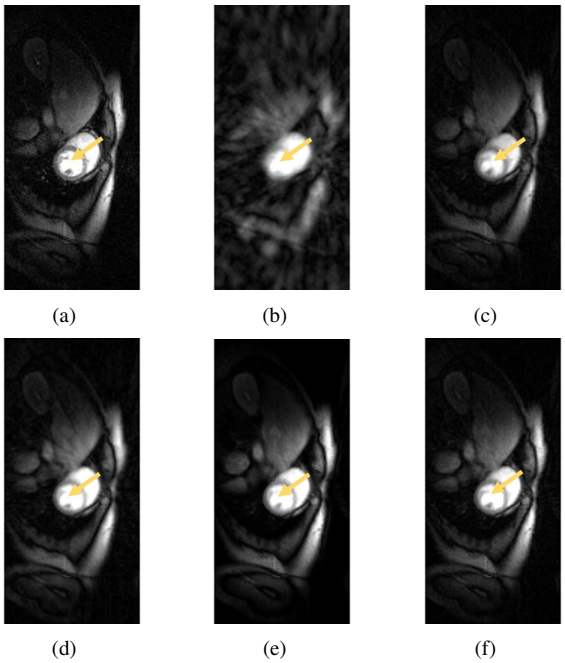

Fig. 7: data set 2 (a) ground truth (b) direct IFFT (c) L+S (d) L&S (e) BCS (f) proposed method

tion, L&S(low rank and sparse) reconstruction [15] and BCS reconstruction [12]. The implementation for L+S and BCS reconstruction were illustrated in section II and available from the authors website. The L&S approach is implemented by assuming the unknown image matrix M is low-rank and sparse simultaneously. We implemented the algorithm by ourselves.

B. Selection of parameters

The behavior of our proposed method is conditional on the selection of five parameters: smooth parameters λ_1 and λ_2, the parameters of β-shrinkage operator α, β, the number of temporal basis N in the dictionary matrix. The selection of δ is not very important as varying it results change of regularization parameters and we set δ as 500. The parameters ρ_1, ρ_2 should be large enough as they approach to $+\infty$ theoretically and we set them as 400, which results $\alpha = 2/\rho_1 = 0.05$.

As both data sets have ground truth data ,we choose the parameters λ_1 ,λ_2 and β such that errors between the reconstruction and ground truth is minimized. To tune the optimal parameter λ_1, we set the λ_2 to zero and varies λ_1 in log scale. To tune the optimal parameter λ_2, we fix λ_1 and find it by the same way.For the parameter β, we consider the case of $\beta \in \{1, 0.95, 0.9, ..., 0\}$.

The choice of dictionary size N is refer to BCS scheme [12]. We tune the parameter N ranging from 10 to 100, and plot the reconstruction errors and nonzero coefficients in sparse matrix Φ as a function of the N. BCS scheme indicates that the reconstruction is insensitive to N as long as it is greater than a reasonable size, which we can observe from the plot.

For the data set 1, the optimal parameters are $\lambda_1 = 0.06, \lambda_2 = 0.005, \beta = 0.55, N = 65$. For the data set 2, $\lambda_1 = 0.08, \lambda_2 = 0.02, \beta = 0.7, N = 50$.

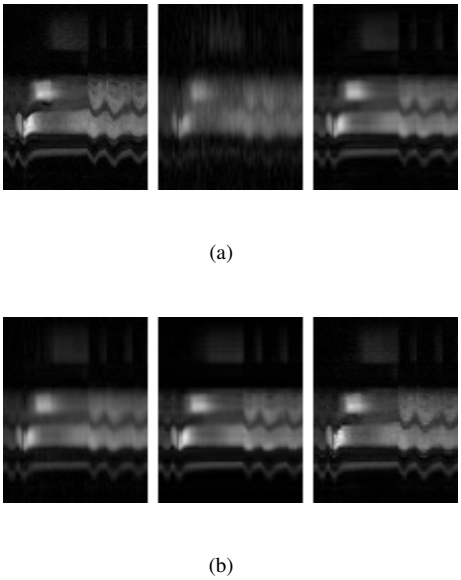

Fig. 8: From left to right,profiles of data set 2 (a) ground truth, direct IFFT, L+S (b) L&S, BCS, proposed method

C. Numerical results

We designed an experiment to validate that the non-convex rank penalty term using Huber function outperforms a convex nuclear norm penalty. We reconstructed the image sequence by L+S model [4] with β-shrinkage operation and singular-value threshold (SVT) operation respectively. The PSNR (Peak Signal-to-Noise Ratio) is popularly used in the assessment of images. Fig.2. shows that β-shrinkage operation achieves higher PSNR than SVT operation for both data sets. This proves the effectiveness of non-convex rank penalty.

Fig.3. shows the comparison of reconstruction of the fourth frame in first data set. Fig.3 (a) is ground truth as reference. The direct inverse FFT reconstruction of undersampled k-space data is badly blurred with artifacts. As arrows indicates, the L&S and BCS schemes lost some details and the L+S scheme caused shadows in local region, while The proposed method behaved better in both aspects.

data set 1	L+S	L&S	BCS	proposed
PSNR(dB)	32.3595	30.9444	32.1948	34.3995
SSIM	0.9875	0.9823	0.9858	0.9911

TABLE I: Average of NSME,PSNR,SSIM in data set 1

data set 2	L+S	L&S	BCS	proposed
PSNR(dB)	31.2019	29.4289	30.3719	32.6869
SSIM	0.9773	0.9679	0.9615	0.9823

TABLE II: Average of NSME,PSNR,SSIM in data set 2

To observe the cardiac motion in data set 1, Fig.4. shows the temporal profiles of a fixed phase encoding line along

temporal direction. We observe that the temporal variation of ground truth smoothly transits between frames. The zero-filled method is badly blurry and lost most of details of variation. Although L+S, L&S, BCS methods capture main motion feature, they lost some local contrasts and the temporal variations of them are not smooth enough. The proposed method preserves the smooth transition between frames and more key motion details.

It is diffcult to observe difference of reconstructed images' quality by naked eyes. Fig.5. shows the comparison of different scheme using PSNR in both data sets. As we can see, our proposed method achieves the highest PSNR among them. This proves that with the sparse basis learned from data itself, we have a better sparsity constraint for dMRI reconstruction.

The most commonly used image quality evaluation methods like NMSE and PSNR are based on statistical characteristics and have clear physical meaning. However, they are just pure mathematical absolute error between the pixels without considering the relevance and perceptual human vision properties. Z.Wang et al. proposed the SSIM [16] to measure the structural similarity between two images and it is more in line with the human vision properties than MSE and PSNR. Fig.6. shows the comparison of different schemes using SSIM in both data sets. It can be seen from the figures that our proposed method have better structural similarity than other methods compare with ground truth.

In comparison, we make quantiative analysis by calculating the average of PSNR, SSIM, NMSE and illustrates them in table I and table II. For data set 1, reconstruction errors get 0.5%-1.2% decrease and PSNR get 2.0-3.4 dB improvement and structural similarity get 0.4%-0.9% improvement by proposed method. For data set 2, reconstruction errors get 0.6%-1.8% decrease and PSNR get 1.5-3.2 dB improvement and structural similarity get 0.5%-1.0% improvement by proposed method.

Fig.7 shows the comparison of reconstruction of the forty-eighth frame for data set 2. The arrows indicate the detail that shows the difference. The L+S reconstruction caused ambiguity and L&S, BCS reconstruciton lost some details. Fig.8 shows the the temporal profiles for data set 2. These results also implies that proposed method obtains better reconstructed images than the other methods.

IV. Conclusion

This work proposed a low-rank plus sparse scheme using dictionary learning for dynamic MRI reconstruction. It combines the avantages of L+S scheme and BCS scheme. A generalized Huber function is used as a nonconvex rank penalty for L part, which outperforms convex nuclear norm penalty. BCS scheme learns the dictionary by undersampled data itself and represents the S part in a sparser way. The experimental results imply that proposed method behaves better in removal of artifacts and reconstructing clear images compared with L+S scheme, L&S scheme and BCS reconstruction.

In the future work, the proposed method can be extended in several directions. For example, we can applicate it on 4-D dMRI reconstruction. The extension can be easily done by rearranging every 2-D slice in a 3-D frame to a columns and concatenating these column into a 2-D matrix. Otherwise, to deal with large deformable motion between organs, motion compensation can be added to reconstruction scheme. The motion can be continually removed during the iterations with a reliable motion estimation scheme. We are intended to investigate and realize these extensions in the next steps.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China (Grant No. 61271338, 61401390), the Zhejiang Provincial Natural Science Foundation of China (Grant No. LQ14F010005).

References

[1] D. L. Donoho, "Compressed sensing," IEEE Trans. Inf. Theory, vol. 52, no. 4, pp. 1289-1306, Apr. 2006.

[2] E. J. Candes, J. Romberg, and T. Tao,"Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," Information Theory, IEEE Transactions on, vol. 52, no. 2, pp. 489-509, 2006.

[3] H. Jung, K. Sung, K. Nayak, E. Kim, and J. Ye,"k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI," Magn. Reson. Med., vol. 61, no. 1, pp. 103-116, 2009.

[4] R. Otazo, E. Candes, and D. K. Sodickson,"Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components," Magnetic Resonance in Medicine, vol. 73, no. 3, pp. 1125-1136, Apr. 2014.

[5] S. G. Lingala, Y. Hu, E. DiBella, and M. Jacob,"Accelerated dynamic MRI exploiting sparsity and low-rank structure: k-t SLR," IEEE Trans.Med. Imag., vol. 30, no. 5, pp. 1042-1054, May 2011.

[6] Tosic, I., Pascal, F.," Dictionary learning," J. Signal Process. Mag., vol. 28, no. 2, pp. 27-38 , 2011.

[7] Ravishankar, S., Bresler, Y.," MR image reconstruction from highly undersampled k-space data by dictionary learning," IEEE Trans. Med. Imaging, vol. 30, no. 5, pp. 1028-1041, 2011.

[8] Aharon, M., Elad, M., Bruckstein, A."K-SVD: an algorithm for designing overcomplete dictionaries for sparse representatio," IEEE Trans. Signal Process., vol. 54, no. 11, pp. 4311-4322, 2006.

[9] S. Gleichman and Y. Eldar,"Blind compressed sensing" IEEE Trans Inf. Theory, vol. 57, no. 10, pp. 6958-6975, Oct. 2011.

[10] R. Chartrand,"Nonconvex splitting for regularized low-rank + sparse decomposition," IEEE Trans. Signal Process., vol. 60, no. 11, pp. 5810-5819, Nov. 2012.

[11] Fessler JA, Sutton BP.,"Nonuniform fast Fourier transforms using min-max interpolation," IEEE Trans Signal Process, vol.51, no. 2, pp. 560-574, 2003.

[12] Lingala, S.G., Jacob, M.,"Blind compressive sensing dynamic MRI," IEEE Trans. Med. Imaging, vol. 32 , no. 6, pp. 1132-1145 , 2013.

[13] Y. Hu, S. Lingala, and M. Jacob,"A fast majorize-minimize algorithm for the recovery of sparse and low rank matrices," IEEE Trans. Image Process., vol. 21, no. 2, pp. 742-753, Feb. 2012.

[14] E. Candes and B. Recht, "Exact matrix completion via convex optimization," Foundat. Computat.Math., vol. 9, no. 6, pp. 717-772, 2009.

[15] Otazo R, Candes E, Sodickson DK,"Low-rank and sparse matrix decomposition for acclerated DCE-MRI with background and contrast separation," Data Sampling and Image Reconstruction, Sedona, Arizona, USA, 2013.

[16] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, Apr. 2004.