Super-resolution Reconstruction for Diffusion-weighted Images using High Order SVD

Ying Fu¹, Yan Wang², Zhipeng Yang³, and Xi Wu¹

¹Department of Computer Science, Chengdu University of Information Technology, Chengdu, China
²Department of Computer Science, Sichuan University, China
³Department of Electronic Engineering, Chengdu University of Information Technology, Chengdu, China

"Regular Research Paper"

Abstract - The spatial resolution of diffusion-weighted imaging (DWI) is limited because of the loss of high-frequency information such as edges during data acquisition process. In this paper, method based on the patch-based super-resolution framework is proposed for single image super-resolution reconstruction of DWI dataset and the high order SVD was introduced to achieve more accurate image reconstruction and reduce the computational complexity. Experimental results demonstrate that the proposed method outperformed currently methods in both DWI reconstruction and its further applications.

Keywords: Diffusion weighted imaging; high-order SVD; super-resolution reconstruction

1 Introduction

Diffusion Weight Imaging (DWI) is a non-invasive magnetic resonance technology and can be used to infer features of the local tissue anatomy, composition and microstructure from water displacement measurements [1]. Water does not diffuse equally and this property has been applied widely for in vivo analysis of white matter architecture and neuronal diseases [2]. Despite the rapid development and interesting property, DWI is an inherent low sensitivity for the analysis of brain structure and clinical disease [3]. Moreover, the high resolution DWI could improve the estimation accuracy of diffusion tensor imaging, and was involved as one of the classical method for inverse problem such as denoising [12] and restoration [13]. Recently, high order single value decomposition (HOSVD) generalized SVD of matrix into high order matrix and offered a simple and elegant method for handling similar patches [14]. Besides this, HOSVD bases were adapted from image content and may achieve more sparse representation than the fixed bases. Have this in mind, we introduced the HOSVD into DWI super resolution. Based on the patch-based SR approaches implemented successfully in both MRI and DWI [5][9][10], HOSVD was involved to construct the regularization framework. The merit of the HOSVD SR method falls into the adaptive HOSVD bases which produce a more accurate reconstruct results. Besides this, HOSVD only implemented over a similar patches stack which effectively decreased the computation complexity. This is especially useful for DWI, since the involvement of joint information from adjacent directions DWI datasets causing extra computation burden dramatically [15].
2 Methods

Image SR leads to an ill-posed inverse problem which is related to LR image y and HR image x, the general model is:

$$\mathbf{y} = \mathbf{DHx} + \mathbf{n}$$  \hspace{1cm} (1)

where \(\mathbf{n}\) represents acquisition noise, \(\mathbf{D}\) is decimator operator and \(\mathbf{H}\) is degradation function [5][9][10].

Based on this model, the SR image can be estimated by minimizing a least-square cost function as follow:

$$\hat{x} = \arg \min_{x} \| \mathbf{y} - \mathbf{DHx} \|^2$$  \hspace{1cm} (2)

For such inverse problem, regularization term should be added to stabilize it [16], thus, the HR image \(\hat{x}\) can be estimated from LR observation \(\mathbf{y}\) by the following equation:

$$\hat{x} = \arg \min_{x} \{ |\mathbf{y} - \mathbf{DHx}| + \lambda R(\mathbf{x}) \}$$  \hspace{1cm} (3)

where \(R(\mathbf{x})\) is regulation term, \(|\mathbf{y} - \mathbf{DHx}|\) is fidelity term, and \(\lambda\) is the parameter to balance them. As shown in Coupe et al. [5], an efficient way to define the regulation term is to use non-local patches methods. Instead of nonlocal mean estimator, we proposed to implement high order SVD as the estimator in this work owing to its simple application and promising performance [14].

The HOSVD estimator clusters similar patches into a stack in a similar manner as other patche-based methods [5][9][10] and then perform HOSVD transform on it to obtain the HOSVD base and coefficients. After the truncation of the coefficient, the patches were then reconstructed by inverse HOSVD transform.

Have this in mind, the regulation term of the super resolution process in equation 3 can be defined as follow:

$$R(\mathbf{x}) = \sum_{i} \| \mathbf{x}(i) - \psi_{\text{HOSVD}}(\mathbf{x}(i)) \|$$  \hspace{1cm} (4)

where \(\psi_{\text{HOSVD}}\) is the HOSVD based estimator.

Given a \(n \times n\) patch \(\mathbf{P}_i\) centered in \(i\), define \(K\) such similar patches (including \(\mathbf{P}_i\)) as \{\(\mathbf{P}_n\)\}, where \(1 < n < K\), and the K-1 similar patches are obtained as follow:

Let us denote \{\(\mathbf{P}_n\)\} as stack \(\mathbf{Z} \in \mathbb{L}^{n \times n \times K}\), the HOSVD of the stack can be defined as [17] :

$$\mathbf{L} = \mathbf{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)}$$  \hspace{1cm} (5)

where \(\mathbf{S}\) is coefficient matrices of three order tensor with \(p \times p \times K\), \(x_j\) stands for the \(j\)th mode tensor product defined in [17] and \(\mathbf{U}^{(1)} \in \mathbb{L}^{p \times n}, \mathbf{U}^{(2)} \in \mathbb{L}^{n \times n}, \mathbf{U}^{(3)} \in \mathbb{L}^{n \times K}\) are orthonormal unitary matrix.

After HOSVD transform, the patches can be estimated by nullifying the coefficients under the assumption that the coefficients of the clean image have sparse distributions. As indicated in [14], the coefficients can be truncated using the hard thresholding as below:

$$\mathbf{S}' = H_t(\mathbf{S})$$  \hspace{1cm} (6)

where \(H_t\) denotes the hard threshold defined with \(\tau = \sigma \sqrt{2 \log(p^2 K)}\) for the stack have \(K\) patches with size of \(n \times n\). As pointed out in [17], the coefficients in tensor \(\mathbf{S}\) are not necessarily positive, and the hard thresholding is defined on the absolute value of the coefficient array:

$$H_t(\mathbf{S}) = \begin{cases} S_i & \text{if} \ abs(S_i) \geq \tau \\ 0 & \text{if} \ abs(S_i) < \tau \end{cases}$$  \hspace{1cm} (7)

Where \(S_i\) denotes the \(i\)th element of tensor \(\mathbf{S}\).

After truncation, the stack \(\mathbf{Z}\) is reconstructed with the inverting transform with truncated coefficients to obtain the final HOSVD estimator \(\psi_{\text{HOSVD}}\):

$$\psi_{\text{HOSVD}} = \mathbf{S}' \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)}$$  \hspace{1cm} (8)

Since the DWI datasets are three dimensions, the above methods should be extend to four-order HOSVD transform of the stack with 3D similarity patches. Besides this, the threshold should be modified as \(\tau = \sigma \sqrt{2 \log(p^3 K)}\) for a stack of size \(n \times n \times n \times K\).

In Manjon et al (2010), mean consistency correction was followed the estimator to ensure coherence with the physical model of acquisition. This was implemented on the fidelity term:

$$\mathbf{Y}(i) - \frac{1}{L} \sum_{l=1}^{L} \hat{X}(i) = 0, \forall p \in \mathbf{Y}$$  \hspace{1cm} (9)

This is done for all the location \(p\) in the LR dataset and imposes sub sampling consistency to the reconstructed patches. At last, the iteration process can be summed up with equation 8 and 9, and is applied until convergence:

$$\mathbf{x}^{t+1} = \psi_{\text{HOSVD}}(\mathbf{x}^t(i))$$  \hspace{1cm} (10)

$$\mathbf{x} = \mathbf{x}^{t+1} - NN(DHx^{t+1} - \mathbf{y})$$  \hspace{1cm} (11)
where NN is the nearest neighbor interpolation, and $t$ is the iteration number.

To improve further the SR performance, the proposed HOSVD SR method can be augmented using joint information from adjacent directions of DWI dataset [15]. For each patch $P_k$, the corresponding stack $Z$ was constructed with $K$ similar patches and the choice of these $K$ similar patches was implemented as follow: the distance threshold selected all patches that $\|P_i - P_k\| < \tau_d$ was chosen to $\tau_d = 3\sigma^2 n^2$ where $\sigma$ is the variance of the noise. This threshold balanced between the estimation accuracy and the computational speed as indicated in [14]. The joint information was introduced through enlarging the searching window into the $M$ adjacent DWI datasets, where $M$ was defined as $M = 2m + 1$ and $m$ denotes the $m$ directions before and after it. In this section, the HOSVD super-resolution method using joint information with multiply directions is referred to as HOSVD-M.

3 Experiments

To evaluate the quality of reconstruction, B-spline interpolation, which have been introduced for DWI resolution enhancement in literature [18], are used for comparison. Besides this, non-local approach for image SR [9] as an effective non-local patch-based SR method is also involved for comparison. In this section, both synthetic and in vivo dataset were implemented for evaluation.

The simulation dataset consists of the 3D structure field as presented at the 2012 HARDI Reconstruction Challenge [19] and has $16 \times 16 \times 5$ volume attempting to simulate a realistic 3-D configuration of tracts occurring. As shown in Fig. 1a, this dataset is comprised of five different fiber bundles, which gave rise to the nonplanar configurations of bending, crossing, and kissing tracts. All fiber tracts were characterized with a fractional anisotropy between 0.75 and 0.90. To better explore the proposed method, this synthetic dataset was also corrupted by Rician noise (SNR = 30) and demonstrated in Fig. 1e. Both the original dataset and the noisy one were down-sampled by a factor 2 using the nearest neighbor interpolation along each axis (i.e., [2 2 2]), which resulted in simulated LR images of $1.4 \times 1.4 \times 4$ mm$^3$.

Tensor estimation of the in vivo DWI dataset was acquired using a 7T Philips Achieva whole body scanner (Philips Healthcare, Cleveland, OH) equipped with a volume head coil for transmission and 32-channels. A DW dual spin-echo, SENSE accelerated msh-EPI was used to acquire the DWI data (b-value: 700 s/mm$^2$; 15 diffusion directions); FOV = $210 \times 30 \times 21$ mm$^3$; matrix size = $300 \times 300$ with 15 slices and a spatial resolution of $0.7 \times 0.7 \times 2$ mm$^3$. In order to validate the proposed approach quantitatively and qualitatively, a gold standard image was constructed based on this in vivo HR DWI dataset. To do that, we averaged 10 acquisitions of high-resolution DW images in the image space ($0.7 \times 0.7 \times 2$ mm$^3$). Then the LR images used for the experiment were simulated by down-sampling our gold standard by a factor of 2 using the nearest neighbor interpolation along each axis (i.e., [2 2 2]), which resulted in simulated LR images of $1.4 \times 1.4 \times 4$ mm$^3$.

where the unitary vector $d_{\text{true}}$ and $d_{\text{estimated}}$ are a true fiber population in the voxel and the closest of the estimated directions.

4 Results

Fig. 1 demonstrates the principle eigenvector of the tensor model in the synthetic phantom and reconstructed results using the B-spline interpolation, non-local upsampling, proposed HOSVD and proposed HOSVD-M. It can be observed that all the results of super-resolved methods were outperform the interpolated results dramatically. The proposed method achieved best results in both noisy and no-noise situations. This may probably due to the adaptive HOSVD bases derived from the stacked patches which is more suitable for the reconstruction.
The reconstructed results of in vivo DWI data were demonstrated in Fig. 2 quantitatively and qualitatively. Fig. 2 shows the visual comparison of the reconstructed DWI images. Interpolated results (Fig. 2b) achieved blurriest results. The proposed method reconstructed the most similar results with the original images. The enlarged region (Fig. 2j) demonstrated that the proposed HO-SVD reconstructed the clear structure of the crack area compared with the same area constructed by other methods were blur and hard to distinguish the edges.

Fig. 3 and Fig. 4 demonstrate the tensor estimation results using the super-resolved DWI datasets. Fig. 3 shows the FA map of the estimated DTI datasets. The proposed HOSVD-M method achieved the best results in the enlarged areas and remained most of the structure and tissues in the original images. The fiber direction indicated with the FA...
 colormap was demonstrated in Fig. 4. It can be seen that, in Fig. 4e, the proposed HOSVD-M obtained the robust direction reconstruction results. For example, in the bundle of corpus callosum, the color of most of the voxels remained the same.

![Fig. 3](image)

**Fig. 3.** FA maps estimated on the gold standard and several methods. (a) FA maps estimated on the gold standard; (b-e) FA maps obtained on the reconstructed dataset using B-spline, the non-local method, proposed HOSVD, proposed HOSVD-M. (f-j) The enlarged details of the B-spline reconstruction, the non-local method, proposed HOSVD, proposed HOSVD-M. The red ROIs indicate the detailed reconstruction. Visually, the FA map obtained using the proposed method is closer to the FA of the gold standard.

![Fig. 4](image)

**Fig. 4** (a) FA colormap on the gold standard; (b-e) FA colormap on reconstructed dataset using B-spline, the non-local method, proposed HOSVD, proposed HOSVD-M

5 Conclusions

In this paper, we proposed a patch-base single image super-resolution method which involved high order SVD for DWI dataset. The adaptive HOSVD bases learned from image ensured a more accurate image reconstruction and manipulation on similar patches stack leaded to a reduction on computational complexity. Quantitative and qualitative comparisons of the traditional interpolation method and non-local patch-based method demonstrate the competitive results in both DWI reconstruction and DTI estimation.

Compared with the currently used interpolation method and patch-based SR method, the improvement of the proposed HOSVD based method can be contributed to two features. The first one is the adapted HOSVD bases learned from a stack of similar patches. This method obtained the bases adaptively according to the image content and achieved more effective reconstruction results. The second feather is the introduction of the joint information from the adjacent direction of the DWI datasets. As pointed out in (Tristán-Vega and Aja-Fernández, 2010)[16], the adjacent directions contained a great many of image redundancy, the
encapsulation of both processed direction and its adjacent ones benefits the reconstruction effectively.

Computational complexity is another important issue for patch-based method as well as DWI processing. All experiments were performed on the Windows 7 computer equipment with an Intel(R) core i7-4600U AND 8g RAM, MATLAB R2013b. For a commonly used DWI datasets with 128×128 matrix size, 60 slices and 32 directions, the running time for single direction of non-local upsampling, proposed HOSVD and the proposed HOSVD-M were around 8 minutes, 3 minutes and 5 minutes respectively. This speed-up is probably due to the inherent dimension decreasing property of the SVD approaches and is specifically useful to HOSVD-M method which introduce many times of extra computational burden.

6 Acknowledgements

This work is supported by National Basic Research Program of China (973 Program, No. 2014CB360506), National Science Foundation of China (No. 61303126), Ministry of Education, China: Chunhui Project (No. Z2015108 ), Sichuan Province Science and Technology Support Program (No. 2016RZ0 051), the Open Research Fund from Province Key Laboratory of Xihua University (No. szj2013-022), Sichuan Province Science and Technology Support Program (No.2015RZ20008) and Educational Commission of Sichuan Province of China (NO. 15ZA0186).

7 References


