Title: Denoising of 3D Magnetic Resonance Images by Using Higher-Order Singular Value Decomposition

Article Type: Research Paper

Keywords: MR, Volume data denoising, Nonlocal methods, Higher-order singular value decomposition, Sparseness

Corresponding Author: Dr. Yanqiu Feng,
Corresponding Author's Institution: Southern Medical University

First Author: Xinyuan Zhang

Order of Authors: Xinyuan Zhang; Zhongbiao Xu; Nan Jia; Wei Yang; Qianjin Feng; Wufan Chen; Yanqiu Feng

Abstract: The denoising of magnetic resonance (MR) images is important to improve the inspection quality and reliability of quantitative image analysis. Nonlocal filters by exploiting similarity and/or sparseness among patches or cubes achieve excellent performance in denoising MR images. Recently, higher-order singular value decomposition (HOSVD) has been demonstrated to be a simple and effective method for exploiting redundancy in the 3D stack of similar patches during denoising 2D natural image. This work aims to investigate the application and improvement of HOSVD to denoising MR volume data. The wiener-augmented HOSVD method achieves comparable performance to that of BM4D. For further improvement, we propose to augment the standard HOSVD stage by a second recursive stage, which is a repeated HOSVD filtering of the weighted summation of the residual and denoised image in the first stage. The appropriate weights have been investigated by experiments with different image types and noise levels. Experimental results over synthetic and real 3D MR data demonstrate that the proposed method outperforms current state-of-the-art denoising methods.
Denoising of 3D Magnetic Resonance Images by Using Higher-Order Singular Value Decomposition

Xinyuan Zhang, Zhongbiao Xu, Nan Jia, Wei Yang, Qianjin Feng, Wufan Chen, Yanqiu Feng

a. School of Biomedical Engineering, Southern Medical University, Guangzhou, China

Corresponding author:
Wufan Chen
Email: chenwf@fimmu.com
Yanqiu Feng
Email: foree@163.com

School of Biomedical Engineering,
Southern Medical University,
Guangzhou, China

Tel. +86 20 6164 8294
Fax +86 20 6164 8274

Grant Support: This study was supported by the National Basic Research Program of China (2010CB732502), National Natural Science Funds of China (81371539).
Abstract

The denoising of magnetic resonance (MR) images is important to improve the inspection quality and reliability of quantitative image analysis. Nonlocal filters by exploiting similarity and/or sparseness among patches or cubes achieve excellent performance in denoising MR images. Recently, higher-order singular value decomposition (HOSVD) has been demonstrated to be a simple and effective method for exploiting redundancy in the 3D stack of similar patches during denoising 2D natural image. This work aims to investigate the application and improvement of HOSVD to denoising MR volume data. The wiener-augmented HOSVD method achieves comparable performance to that of BM4D. For further improvement, we propose to augment the standard HOSVD stage by a second recursive stage, which is a repeated HOSVD filtering of the weighted summation of the residual and denoised image in the first stage. The appropriate weights have been investigated by experiments with different image types and noise levels. Experimental results over synthetic and real 3D MR data demonstrate that the proposed method outperforms current state-of-the-art denoising methods.

Keywords: MR, Volume data denoising, Nonlocal methods, Higher-order singular value decomposition, Sparseness
1. Introduction

Magnetic resonance imaging (MRI) is a non-invasive high-resolution multi-parameter imaging modality that is widely used in current clinical diagnosis and scientific research because of its ability to reveal the 3D internal structures of objects. However, magnetic resonance (MR) images may suffer from serious random noise, specifically when high-resolution and high-speed is required. The noise in MR images may affect the inspection quality in clinical diagnosis, and decrease the reliability of analysis tasks such as registration, segmentation and quantitative parameter mappings. Therefore, noise reduction is important for the interpretation and the computer-aided analysis of MR images.

Averaging repeated acquisitions can increase the signal-noise-ratio (SNR) in MR images at the expense of imaging time. By contrast, the denoising of MR images by post-processing techniques does not increase the acquisition time, thus is widely used in clinical practice and research. One simple filter is the low-pass Gaussian filter, which averages adjacent pixels by using a weight function of spatial distance between pixels (Lindenbaum M et al., 1994). This filter has been widely used in voxel-based morphometry (Ashburner and Friston, 2000) and has an inherent drawback of blurring edges, which may be clinically relevant.

Numerous edge-preserving denoising methods have been developed in the past few decades by exploiting different image and noise characteristics. Gradient-based filters, such as the anisotropic diffusion filter (Perona and Malik, 1990) and nonlinear total variation algorithm (Rudin et al., 1992), use gradient information to retain important image structures while removing noise. These filters have been extended to denoise MR images (Gerig et al., 1992; Krissian and Aja-Fernandez, 2009; Liang et al., 2011; Samsonov and Johnson, 2004). However, these filters usually erase small details and generate piece-wise constant images. A large number of transform-based filters have also been proposed to denoise images in a transform-threshold-inverse transform fashion by exploiting the uncorrelated properties of noise in the transform domain.
For example, various wavelet-based filters have been widely proposed to denoise MR images with different ways in manipulating transformed coefficients (Alexander et al., 2000; Nowak, 1999; Wirestam et al., 2006; Wood and Johnson, 1999). Transforms using adaptive bases such as the principal components analysis (Muresan and Parks, 2003) and using learned transform bases (Elad and Aharon, 2006) were also proposed to denoise images. Another popular denoising approach is based on the maximum likelihood (ML) estimation of the noise-free MR signal in the presence of Rician or non-central chi distributed noise (He and Greenshields, 2009; Rajan et al., 2011; Rajan et al., 2012; Sijbers and den Dekker, 2004).

Nonlocal denoising methods that exploit the similarity and sparsity among small patches have emerged as an effective way to preserve details while removing noise, and achieve state-of-the-art performances. A simple and classic method is the well-known nonlocal means (NLM) filter (Buades et al., 2005), which averages multiple pixels in a large search window and avoids the smoothing of edges by assigning only high weights to pixels with similar local patterns. The NLM filter has been adapted to denoise MR images and demonstrated better performance than the above mentioned gradient and wavelet-based filters (Coupe et al., 2008; Manjon et al., 2008; Wiest-Daessle et al., 2008). One drawback of the NLM filter is that exploiting self-similarities is impossible in the extreme case when a patch cannot find a similar patch in the image (Mairal et al., 2009). This situation can be addressed by exploiting the sparsity among overlapping patches. To take advantage of both self-similarity and sparsity, the discrete cosine transform (DCT) method (Guleryuz, 2007) is extended and combined with a rotationally invariant version of the NLM filter to denoise 3D MR data; the combined method demonstrates better results than both the oracle-based DCT and block-wise NLM filters (Manjon et al., 2012).

Another well-known nonlocal denoising method is the block-matching 3D (BM3D) filter proposed by Dabov et al. (2007). This method combines nonlocal and transform-domain approaches (Katkovnik et al., 2010) and presents an effective
denoising framework by grouping similar patches into a 3D array, then filtering the
3D array by using sparse representation in the transform domain, and finally
aggregating multiple estimates at each location. The framework of grouping similar
patches has the advantage of enhancing sparsity and is the reason behind the
outstanding performance of BM3D and also adopted in many recent papers
(Chatterjee and Milanfar, 2009; Dong et al., 2013; Mairal et al., 2009). The BM3D
filter has been extended to volumetric data and denominated BM4D, which
demonstrates state-of-the-art performance in denoising 3D MR images (Maggioni et
al., 2013). The BM3D filter assumes that the underlying truth patches can be well
represented using sparse coefficients with some fixed bases, which may be less
adaptive to varying image contents than learned bases (Dong et al., 2013; Elad and
Aharon, 2006).

Recently, the higher-order singular value decomposition (HOSVD) of grouped similar
patches offers a simple and elegant method for handling sparsity among similar
patches for denoising natural images (Rajwade et al., 2013). The HOSVD bases are
learned from image and thus more adaptable to the image content and may achieve a
more sparse representation than fixed bases such as wavelet and discrete cosine bases
used in the BM3D method. Although the learned HOSVD bases may be sensitive to
the noise in images, this disadvantage can be well addressed by learning bases from a
predenoised image. Compared other SVD-based nonlocal denoising methods (Dong et
al., 2013; Elad and Aharon, 2006), HOSVD does not require rearranging patches into
column vectors, thus may better preserve topological structures during exploiting
sparsity.

To the best of our knowledge, the HOSVD-based method has not been previously
considered for MR denoising, although it presents promising properties. The present
paper aims to investigate and improve the application of HOSVD to denoising MR
volume data. The Wiener filter-augmented HOSVD method proposed by Rajwade et
al (Rajwade et al., 2013) is firstly extended and applied to denoise 3D MR data. For
better denoising performance, we propose to augment the standard HOSVD stage by a second recursive HOSVD stage, where the "noise" predicted by the first HOSVD stage is partially added back to the denoised image and the sum is then filtered by another HOSVD procedure.

2. Materials and methods

2.1. HOSVD denoising for 2D images

HOSVD generalizes the SVD of matrix to higher-order matrixes (De Lathauwer et al., 2000; Rajwade et al., 2013). Similar to most contemporary techniques, the HOSVD denoising method is generally designed for the zero mean (i.i.d.) Gaussian noise, wherein the observed noisy image $Y$ from a noise free image $X$ can be modeled as follows:

$$Y = X + N$$  \hspace{1cm} (1)

where $N$ denotes a zero mean Gaussian noise with a known variance $\sigma^2$. In the case of a noisy image $Y$, denoising methods aim to find a good estimate $\hat{X}$ of $X$ from $Y$.

The HOSVD denoising method clusters similar patches into a stack in a similar manner as other patch-based methods (Dabov et al., 2007; Mairal et al., 2009) and then performs the HOSVD transform of such a stack to obtain its representation by the HOSVD bases and coefficients. One estimate of this stack is reconstructed by inverse HOSVD transform with truncated coefficients. This operation is repeated for each reference patch in a sliding window fashion, and multiple estimates at each pixel are averaged to obtain the final denoised image. The details of this standard HOSVD denoising process are described in the following.

Given a $p \times p$ reference patch in the noisy 2D image, $K$ similar patches (including the reference patch) are found and stacked into a 3D array $G$ of size $p \times p \times K$. The HOSVD of this stack can be formulated as follows (De Lathauwer et al., 2000; Rajwade et al., 2013):
\[ G = S \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)} \]  

where \( U^{(1)}, U^{(2)} \in \mathbb{R}^{p \times p} \) and \( U^{(3)} \in \mathbb{R}^{K \times K} \) are the orthogonal unitary matrices, \( S \in \mathbb{R}^{p \times p \times K} \) is a three-order tensor, and \( \times_n \) denotes the \( n \)-mode product of a tensor by a matrix \( U^{(n)} \). Eq. (2) shows that the HOSVD can exploit signal sparsity across each stack dimension.

The noisy image can be filtered by nullifying the HOSVD transformation coefficients below a fixed threshold under the assumption that the original coefficients of the clean image have sparse distributions. The thresholding of the transform coefficients can be represented as follows:

\[ \hat{S} = H_\tau(S) \]

where \( H_\tau \) denotes the hard thresholding operator with threshold \( \tau \), which is determined as \( \tau = \sigma \sqrt{2 \log(p^2K)} \) for every stack of size \( p \times p \times K \) in 2D image, an optimal threshold from a statistical risk viewpoint according to the rule from (Donoho and Johnstone, 1994). It is noted that the coefficients in the tensor \( S \) are not necessarily positive in general (De Lathauwer et al., 2000). Therefore, hard thresholding is defined on the absolute value of the coefficient array \( S \) as follows:

\[ H_\tau(S) = \begin{cases} 
S_i & \text{if } \text{abs}(S_i) \geq \tau \\
0 & \text{if } \text{abs}(S_i) < \tau 
\end{cases} \]

where \( S_i \) represents the \( i \)th element of core tensor \( S \).

An estimate of the stack \( G \) then can be obtained by the inverse HOSVD transform with truncated coefficients:

\[ \hat{G} = \hat{S} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)} \]

The above operation is repeated for each reference patch in a sliding window fashion, thus producing multiple estimates at each pixel. The final estimate is obtained by aggregating these multiple estimates at each pixel. Weighted combination is usually
implemented to obtain the final denoised image with varying weighting strategies (Dabov et al., 2007; Manjon et al., 2012).

To improve further the denoising performance, the above HOSVD denoising process can be augmented by a Wiener filter step: Let \( \hat{G} \) denote the stack of similar patches regrouped from the filtered image and \( G_n \) denote the corresponding stack from the noisy image. Let the coefficients of \( G_n \) and \( \hat{G} \) in the HOSVD bases of \( \hat{G} \) be denoted as \( c_n \) and \( \hat{c} \), respectively. The Wiener filter shrinks the coefficients \( c_n \) in the presence of the coefficients \( \hat{c} \) such that \( \hat{c}_n = c_n \hat{c}_n^2 / (\hat{c}_n^2 + \sigma^2) \), where \( \hat{c}_n \) is the filtered coefficients of the stack \( G_n \). The Wiener filter-augmented HOSVD performs better than the standard HOSVD and works almost the same as the BM3D in denoising natural grayscale images (Rajwade et al., 2013). In this paper, the Wiener filter-augmented HOSVD is referred to as HOSVD-W.

2.2. Application of HOSVD-W to 3D MR image denoising

The above HOSVD-W denoising for 2D images can be straightforwardly extended to denoise 3D images by employing the four-order HOSVD transform of a stack of similar 3D cubes. In the truncation of HOSVD coefficients, the threshold \( \tau \) is picked to be \( \sigma \sqrt{2 \log(p^3K)} \) for the stack of size \( p \times p \times p \times K \) in a 3D image. The HOSVD-based method is designed for data that are contaminated by additive Gaussian noise independent of the signal and cannot be directly applied to denoise MR magnitude images, which are generally corrupted with Rician noise.

Owing to the advantage of optimal forward and inverse variance-stabilizing transformation (VST), the algorithms designed for additive white Gaussian noise achieve state-of-the-art performance in denoising MR images (Foi, 2011; Maggioni et al., 2013). HOSVD-based denoising algorithms can be integrated into Rician data by
using the following formula:

\[ \hat{Z} = \text{VST}^{-1}(\text{Denoising}_{\text{HOSVD}}(Z, \sigma_R), \sigma_{\text{VST}}), \sigma_R) \]  

(6)

where \( \text{VST}^{-1} \) denotes the inverse VST, \( Z \) denotes the MR data corrupted by Rician noise with standard deviation \( \sigma_R \), and \( \sigma_{\text{VST}} \) is the standard deviation of noise after VST and is generally stabilized to one.

2.3. Proposed recursive HOSVD denoising method

The Wiener filter requires accurate estimation of the SNR at each of transform bases.

To improve the performance of HOSVD denoising method, we propose to augment a standard HOSVD stage by using the recursive regularization in the second stage, which has the advantage of exploiting filtering residuals. The flow diagram of the proposed two-stage HOSVD algorithm is shown in Fig. 1. The second stage in this proposed method is the HOSVD denoising of the combined image from original and pre-denoised images; this stage can also be considered a two-stage recursive regularization by adding filtered noise back to the denoised image (Dong et al., 2013; Osher et al., 2005):

\[ Y_w = \hat{X}_{\text{HOSVD1}} + \xi(Y - \hat{X}_{\text{HOSVD1}}) = (1 - \xi)\hat{X}_{\text{HOSVD1}} + \xi Y \]  

(7)

where \( \xi \in [0, 1] \) is a relaxation parameter representing the part of original noisy image in the combination, \( \hat{X}_{\text{HOSVD1}} \) is the restored image after the first HOSVD stage.

The second stage also employs the hard-thresholding of the HOSVD transform coefficients with an updated threshold:

\[ \tau_w = \sigma_w \sqrt{2 \log(\rho K)} \]  

(8)

where \( \sigma_w \) is the updated noise variance after the feedback of filter noise and can be calculated as follows (Dong et al., 2013):

\[ \sigma_w = \gamma \sqrt{\sigma^2 - \|Y - Y_w\|^2} \]  

(9)
Fig. 1. Flow diagram of the proposed HOSVD-R algorithm.

where $\sigma$ is the standard deviation of noise in the original data $Y$, $\gamma \in [0, 1]$ is a scaling factor controlling the re-estimation of noise variance.

For clarity, the proposed HOSVD denoising augmented by the recursive regularization technique is referred to as HOSVD-R. HOSVD-R is combined with VST when applied to 3D MR data denoising. For the aggregation in the HOSVD-R method, the simple average of multiple estimates at each pixel is adopted.

2.4. Practical implementation to improve efficiency

The two HOSVD-based methods are applied to denoise MR volume data. Given that the most time-consuming parts of HOSVD-based algorithms are cube grouping and HOSVD transform. Cubes with Euclidean distances below a threshold are clustered into the stack of the reference cube in the HOSVD-W implementation (Rajwade et al., 2013), while the K-nearest-neighbors of the reference cube are found on the basis of the Euclidean distance between cubes and are then grouped into a stack in HOSVD-R implementation. In this section, we improve the computation efficiency without significantly affect the denoising efficacy by using similar techniques in the BM4D denoising of MR volume data (Maggioni et al., 2013).
Table 1. Parameter selections of the cube size $p$ and number $K$ of similar cubes under different noise levels.

<table>
<thead>
<tr>
<th>Noise level</th>
<th>1%</th>
<th>3%</th>
<th>5%</th>
<th>7%</th>
<th>9%</th>
<th>11%</th>
<th>13%</th>
<th>15%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P$</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>$K$</td>
<td>35</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
</tr>
</tbody>
</table>
Fig. 2. Influence of parameter $\zeta$ in the HOSVD-R algorithm on the PSNR and SSIM in denoising the T1w, T2w, and PDw images under different noise levels (3%, 7%, and 11%).
Fig. 3. Influence of parameter $\gamma$ in the HOSVD-R algorithm on the PSNR and SSIM in denoising the T1w, T2w, and PDw images under different noise levels (3%, 7%, and 11%).
Table 2. PSNR and SSIM comparisons of different algorithms on the T1w from the BrainWeb database.

<table>
<thead>
<tr>
<th>Noise level</th>
<th>1%</th>
<th>3%</th>
<th>5%</th>
<th>7%</th>
<th>9%</th>
<th>11%</th>
<th>13%</th>
<th>15%</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRI-NLM3D</td>
<td>43.97</td>
<td>38.19</td>
<td>35.34</td>
<td>33.37</td>
<td>31.94</td>
<td>30.74</td>
<td>30.14</td>
<td>28.88</td>
</tr>
<tr>
<td>BM4D</td>
<td>44.03</td>
<td>38.34</td>
<td>35.87</td>
<td>34.22</td>
<td>32.95</td>
<td>31.89</td>
<td>30.14</td>
<td>28.88</td>
</tr>
<tr>
<td>HOSVD-W</td>
<td>43.57</td>
<td>38.55</td>
<td>36.08</td>
<td>34.35</td>
<td>33.00</td>
<td>31.88</td>
<td>30.14</td>
<td>28.88</td>
</tr>
<tr>
<td>HOSVD-R</td>
<td>45.21</td>
<td>38.97</td>
<td>36.38</td>
<td>34.61</td>
<td>33.27</td>
<td>32.23</td>
<td>30.53</td>
<td>29.10</td>
</tr>
</tbody>
</table>

Two quantitative measures are adopted to evaluate the denoising performance in the presence of ground truth. The first measure is the peak SNR (PSNR), which is based on the root mean square error (RMSE) between the denoised data and ground truth:

$$\text{PSNR} = 20 \log_{10} \frac{255}{\text{RMSE}}$$  (10)

The second measure is the structural similarity index (SSIM), which is consistent with human eye perception (Manjon et al., 2012; Wang et al., 2004) and defined as follows:

$$\text{SSIM} = \frac{(2\mu_X \mu_{\hat{X}} + c_1)(2\sigma_{X\hat{X}} + c_2)}{(\mu_X^2 + \mu_{\hat{X}}^2 + c_1)(\sigma_X^2 + \sigma_{\hat{X}}^2 + c_1)}$$  (11)

where $c_1$ and $c_2$ are constant; $\mu_X$ and $\mu_{\hat{X}}$ are the mean of data $X$ and $\hat{X}$, respectively; $\sigma_X$ and $\sigma_{\hat{X}}$ are the standard noise variance of data $X$ and $\hat{X}$, respectively; $\sigma_{X\hat{X}}$ is the covariance of $X$ and $\hat{X}$, respectively. In this study, the SSIM is locally computed with a $3 \times 3 \times 3$ voxel Gaussian kernel, and the mean of all local SSIMs in the region-of-interest is estimated as a global SSIM. Both PSNR and SSIM are calculated on the anatomical region with the background removed.

The performance of the HOSVD-W and HOSVD-R algorithms are compared against PRI-NLM3D (Manjon et al., 2012) and BM4D (Maggioni et al., 2013) to provide relevant comparisons. Both the PRI-NLM3D and the BM4D algorithms are operated by using software from the homepage of PRI-NLM3D and BM4D with suggested filter parameters. The BM4D algorithm represents the current state-of-the-art method
Fig. 4. PSNR and SSIM comparisons of different denoising algorithms for the T1w, T2w, and PDw images under noise levels varying from 1% to 15% with an increase of 2%.

3.1. Filter parameters
Patch size is the only free parameter which influences the accuracy in the original HOSVD-W algorithm (Rajwade et al., 2013). Thus, patches with varying sizes are implemented and results with optimal PSNR and SSIM are obtained.

The performance of HOSVD-R denoising depends on the setting of two cube-related parameters (i.e., cube size $p$ and the number of similar cubes in a group $K$) and two noise-related parameters (i.e., relaxation parameter $\xi$ in Eq. (7) which controls the contribution of feed-back noise in the data $Y_w$, and scaling factor $\gamma$ in Eq. (9) which controls the re-estimation of noise variance).

Fig. 5. Denoised results of different algorithms on synthesized T1w brain image. Top row: the noise free T1w image and the image with Rician noise level of 15%. Second row: the denoised images with different algorithms. Third row: the corresponding error images (the absolute difference between the denoised and noise-free images).
In our denoising experiments with BrainWeb MR data by using HOSVD-R, the results show that a larger size of cube $p$ and a larger number of similar cubes $K$ yield higher denoising PSNR at high noise levels. Thus, in this study, $p$ and $K$ are tentatively chosen as values that increase with noise levels (Table 1). For real MR data, the HOSVD-R algorithm is implemented with experiential parameters of $p = 4$ and $K = 75$.

To determine the appropriate values of the noise-related parameters $\xi$ and $\gamma$ for denoising MR images, the HOSVD-R algorithm was implemented on the BrainWeb data with varying $\xi$ and $\gamma$, and the resulting PSNR and SSIM values were adopted as the quantitative measures to evaluate the sensitivity to the setting of these two parameters. For T1w, T2w and PDw data at different noise levels (3%, 7% and 11%), Figs. 2 and 3 show the impact of the parameters $\xi$ and $\gamma$ on denoising performance, respectively. HOSVD-R can produce higher PSNR and SSIM in a large $\xi$ and $\gamma$ range than the BM4D and HOSVD-W methods. For the T2w image at 11% noise level, HOSVD-W produces a PSNR comparable to that of HOSVD-R but with the lower SSIM. Based on the observation from Figs. 2 and 3, $\xi$ and $\gamma$ are fixed at 0.38 and 0.65 in the following denoising experiments, respectively, which yield an improved PSNR and SSIM over the compared methods for data with varying noise levels and image types, although these two values are not strictly optimal for each dataset.

### 3.2. Denoising of BrainWeb MR images

The PSNR and SSIM of the PRI-NLM3D, BM4D, HOSVD-W, and HOSVD-R methods on T1w, T2w and PDw images are shown in Fig. 4. The HOSVD-W algorithm achieves comparable denoising performance to BM4D in terms of PSNR and SSIM. The HOSVD-R algorithm significantly outperforms the other three with regard to PSNR for all image types under low to moderate noise levels (from 1% to
5%). At noise levels above 7%, the HOSVD-R algorithm produces slightly higher
PSNRs than the BM4D and HOSVD-W algorithms, and all three algorithms produce
significantly higher PSNRs than the PRI-NLM3D. With regard to SSIM, the
HOSVD-R algorithm consistently yields SSIMs higher than the other three methods
for all noise levels and image types. Table 2 presents the discrete values of PSNRs
and SSIMs in denoising T1w images by using the four methods under investigation.
The PSNRs and SSIMs of PRI-NLM3D and BM4D are close to the values reported
by Maggioni et al (2013). The PSNRs and SSIMs of HOSVD-W are close to those of
BM4D. The proposed HOSVD-R algorithm outperforms BM4D with PSNR
improvements ranging from 0.3 dB to 1.2 dB and SSIM improvements from 0.002 to
0.009.

![Denoised results of different algorithms on synthesized T2w brain image. Top row: the noise free T2w image and the image with Rician noise level of 15%. Second row: the denoised images with different algorithms. Third row: the corresponding error images.](image)
Fig. 7. Denoised results of different algorithms on synthesized PDw brain image. Top row: the noise free PDw image and the image with Rician noise level of 15%. Second row: the denoised images with different algorithms. Third row: the corresponding error images.

Figs. 5 to 8 provide a visual evaluation of denoising results under T1w, T2w, and PDw data with 15% noise level. All compared algorithms show good performance in removing noise without significant detail loss. The PRI-NLM3D result has the sharpest edges but with the most prominent errors and most detail loss compared with the other denoising algorithms. In addition to PRI-NLM3D, HOSVD-R also produces results with fewer intensity oscillations in homogeneous areas than BM4D and HOSVD-W; this finding can be more clearly observed from the error images. To provide a better visual inspection, the enlarged detail is shown in Fig. 8. As seen in the region pointed by the arrows, PRI-NLM3D oversmooths low contrast details, and the results of the proposed HOSVD-R method are more visually pleasant than those
of other compared methods.

It should also be noted that PSNR and SSIM are calculated only in the anatomical region after removing the background. The error images show that the proposed approach removes background noise significantly better than the other denoising algorithms. If the background is included in the quantification, the PSNR of the proposed algorithm will be significantly higher than those of the other three algorithms.

3.3. Denoising of real MR data

This section is devoted to verifying the consistency of the proposed approach on real MR data. The experiments are conducted on two cross-sectional T1w MR brain datasets (OAS1_0112 and OAS1_0092) from the publicly available Open Access Series of Imaging Studies (OASIS) database (http://www.oasis-brains.org) (Marcus et al., 2007). The T1w real brain datasets are acquired by an MP-RAGE volumetric sequence on a Siemens 1.5 T Vision scanner. The acquisition parameters have a repetition time of 9.7 ms, echo time of 4 ms, flip angle of 10°, inversion time of 20 ms, duration time of 200 ms, matrix of 256×256×128, voxel resolution of 1×1×1.25 mm³.

The noise levels of the two selected datasets are approximately 3% and 4.5% of the maximum intensity according to the noise level estimation method of (Foi, 2011). The denoised results of HOSVD-R on these two datasets are presented in Fig. 9. The HOSVD-R method significantly reduces the noise, and no significant anatomical information can be observed in the residual image (the difference between the noisy and denoised images). As shown by the enlarged image in Fig. 10, PRI-NLM slightly oversmooths fine details, whereas BM4D and HOSVD-W preserves satisfactory details but with some intensity oscillations. The proposed HOSVD-R generates the most visually pleasant result. These observations are consistent with the results of synthetic data.
Fig. 8. Enlarged parts of the denoised images by different algorithms in Figs. 5 to 7. From top to bottom: T1w, T2w, and PDw images.

Fig. 9. Denoised results of the proposed HOSVD-R algorithm on two sets of real T1w brain data with different estimated Rician noise levels (a: 3%; b: 4.5%). From top to bottom are the original, denoised, and residual images. From left to right: sagittal, coronal and transverse frames. Note that the higher values in the residuals are due to
Rician noise-related bias on low intensity signals.

Fig. 10. Comparison of different algorithms on the enlarged views of the rectangular regions in Fig. 9. (a): enlarged parts of Fig. 9a; (b): enlarged parts of Fig. 9b.

4. Conclusion/Discussion

Two HOSVD-based methods that exploit the sparseness between similar cubes by using HOSVD transform have been presented to denoise 3D MR data. One method
is the simple extension of the original HOSVD method augmented by a Weiner filter to 3D MR data (HOSVD-W), and the other method is a novel HOSVD-R method that takes advantage of a recursive regularization to improve the denoising performance of HOSVD. Both HOSVD-based methods are applied and compared with current state-of-the-art methods on synthetic and real MR volumetric data. The experimental results demonstrate that the proposed HOSVD-R algorithm outperforms current state-of-the-art algorithms in 3D MR denoising.

The excellent denoising performance of the HOSVD-R method can be attributed to two main features. The first feature is that the HOSVD-R denoising method represents a stack of similar cubes by using learned orthogonal bases; the transform of such bases is more adaptable to different data types and can achieve a more efficient and sparse signal representation than a transform using fixed bases such as wavelet and DCT in the BM4D method. The second feature is that the feedback of filtered noise in the second stage maintains a two-step regularization effect to encourage the solution toward a better result; this feature distinguishes HOSVD-R from HOSVD-W. The proposed HOSVD-R also benefits from the manipulation of “better” HOSVD bases from combined images in the second stage, compared with bases from the original noisy data in the first stage.

Considering that both the BM4D and the HOSVD-based methods operate on stacks of similar cubes, comparing the time complexity of these transforms over a single stack clarifies the time costs. For a stack of size \( p \times p \times p \times K \), BM4D requires \( O(Kp^4) \) and \( O(K^2p^3) \) operations for 3D and 1D transforms, respectively, thus leading to a total complexity of \( O(Kp^4+K^2p^3) \). The time complexity of HOSVD is \( O(Kp^4+\min(K^2p^3, Kp^5)) \) (Rajwade et al., 2013). Thus, if the two types of denoising methods construct stacks with the same size, the computation complexity of the HOSVD-based method is comparable to that of BM4D. All denoising methods were performed in MATLAB 7.12.0 (R2011a) on a Windows 7 computer equipped with an Intel(R) Core(TM)2,
2.33 GHz CPU and 8 GB RAM. To denoise a typical 3D dataset of size $181 \times 217 \times 181$ pixels, the proposed HOSVD-R method took 24 min on average (multithreaded Matlab implementation on four cores), whereas HOSVD-W took 8 min (multithreaded Matlab implementation on four cores), BM4D 16 min (single-threaded Matlab/C implementation), PRI-NLM3D 1 min (multithreaded Matlab/C implementation on four cores). We believe that the implementation of HOSVD-based denoising methods using Matlab/C MEX techniques and parallel computations on graphic processing units may significantly further accelerate the filtering.

One limitation of the proposed HOSVD-R method is that its several filtering parameters (the number of similar cubes $K$, the cube size $p$, the relaxation parameter $\xi$ and the scaling factor $\gamma$) are currently manually determined by experience. The optimal parameter setting may change with noise levels and image characteristics. Although experiments in this study show that the proposed method can achieve state-of-the-art performance with experiential parameters, the automatic determination of filtering parameters with theoretical foundations is warranted in a future study.

Stein's unbiased risk estimate (Stein, 1981) or the no-reference metric $Q$ (Zhu and Milanfar, 2010) can serve as quantitative metrics for automatic parameter tuning when no "ground-truth" reference is available, but will require a number of iterations.

Finally, the proposed method is limited to processing spatially invariant Rician noise across the whole images, thus cannot be directly applied to denoising MR images reconstructed from multichannel data. For images reconstructed by the root-sum-squares operation from non-subsampled multichannel data, the noise is assumed to follow the noncentral Chi distribution (Constantinides et al., 1997; Koay and Basser, 2006). Thus, the noncentral Chi version of VST is required to be developed for the application of HOSVD-based methods. For images reconstructed from subsampled data by using parallel imaging techniques such as sensitivity encoding (Pruessmann et al., 1999) and generalized autocalibrating partially parallel
acquisitions (Griswold et al., 2002), the noise generally follows nonstationary Rician and noncentral Chi distributions, respectively. In such scenarios, the g-factor (Robson et al., 2008), which characterizes the varying noise field, can be employed to guide the denoising. When no g-factor is available, one possible approach to handle the spatially varying noise is to adapt the HOSVD-based algorithms according to locally estimated noise variances (Landman et al., 2009; Manjon et al., 2010; Maximov et al., 2012), which is warranted in a future study.
Acknowledgements:

We want to thank Jose V. Manjon (Manjon et al., 2012), Matteo Maggioni (Maggioni et al., 2013), Ajit Rajwade (Rajwade et al., 2013), and Weisheng Dong (Dong et al., 2013) for publishing online their implementation of PRI-NLM3D, BM4D, HOSVD and spatially adaptive iterative singular-value thresholding (SAIST), respectively. This study was supported by the National Basic Research Program of China (2010CB732502), National Natural Science Funds of China (81371539).
References


Process 16, 2080-2095.


Graphical Abstract
Highlights:

1. The Wiener filter-augmented HOSVD is extended to denoise 3D MR data.
2. The Wiener filter-augmented HOSVD achieves comparable performance to that of BM4D.
3. A novel recursive HOSVD method is proposed to exploit filtering residual for better performance.
4. The recursive HOSVD outperforms current state-of-the-art 3D denoising algorithms.
Point-to-Point Responses for Resubmission of Manuscript

MEDIA-D-14-00106

Title: Denoising of 3D Magnetic Resonance Images by Using Higher-Order Singular Value Decomposition

Dear Jim Duncan,

We are grateful for the opportunity to revise our manuscript after the careful and insightful reviews provided.

The main concern raised by reviewers was that our method lacks the novelty. We understand reviewer’s concern. The main contribution of our work can be summarized as follows:

The HOSVD denoising method (Rajwade et al., 2013) provides a simple and elegant way to exploit sparseness among similar image patches or cubes, but its performance in denoising medical MR images has not been investigated up to date. The purpose of this work is to study and improve the application of the HOSVD method to denoise MR images. We first present the MR denoising result using the HOSVD-W method proposed by Rajwade et al. The result shows that the optimal performance of this HOSVD filter is comparable to that of the BM4D method (Maggioni et al., 2013). To further improve the denoising performance of HOSVD applied to MR data, we propose to augment the standard HOSVD stage by a second recursive HOSVD stage, where the "noise" predicted by the first HOSVD stage is partially added back to the denoised image and the sum is then filtered by another HOSVD procedure, instead of the wiener filter augmentation stage adopted by Rajwade et al. This idea is borrowed from the iterative regularization originally developed by Osher et al (Osher et al., 2005) and modified by Dong et al (Dong et al., 2013). Although this combination...
seems to be simple and lack of novel ideas, this combination appears for the first time and outperforms the wiener-augmented HOSVD and BM4D filters in denoising MR images, the latter of which demonstrated the state-of-the-art performance of denoising MR 3D data. It is well known that BM3D represents the state-of-the-art performance of natural image denoising, and only few of the methods developed after BM3D have comparable or slightly better performance to that of BM3D.

We have been trying to develop better HOSVD-based algorithm including the automatic determination of filtering parameters using Stein's unbiased risk estimate, but which requires numerous iterations and are very time-consuming. Thus, we are developing the GPU implementation of the HOSVD denoising algorithm to greatly shorten the filtering time, which will resolve the time cost issue so that better algorithm, which has better or more reliable denoising performance but may require several iterations, can be developed. We feel it is difficult to include such work in this manuscript due to the limit of scope but added some discussion on this issue.

We have responded to all comments of the reviewers and revised the manuscript accordingly. We feel that our revisions have improved the manuscript and hope that it is now acceptable for publication in your journal.
Comments to the Author:

**Reviewer #1**

This paper presents a "new" filter for MRI denoising based on HOSVD transform. The paper is clear and reads well (although the English should be improved in my opinion) and the results are appealing.

We thank reviewer 1 for the constructive criticism. We will respond to all issues addressed by the reviewer below.

However, my major concern about this work is the lack of novel ideas since the proposed method looks like a mixture of already proposed methods. Basically the proposed method is based on a 3D extension of the HOSVD transform (Rajwade et al, 2013) borrowing the two stage framework from BM3D (Davov et al, 2007). It also uses the aggregation parts from ODCT (Manjon et al, 2012) and the image mixing from SAIST (Dong et al., 2013). VST is also used to deal with Rician noise (Foi et al, 2011).

R1.C1: We thank the reviewer for the constructive criticism.

The HOSVD transform (Rajwade et al., 2013) provides a simple and elegant way to exploit sparseness among similar image patches or cubes, but its performance in denoising medical MR images has not been investigated up to date. The purpose of this work is to study and improve the application of the HOSVD method to denoise MR images. We first present the MR denoising result using the HOSVD-W method proposed by Rajwade et al. The result shows that the optimal performance of this HOSVD filter is comparable to that of the BM4D method (Maggioni et al., 2013). To further improve the denoising performance of HOSVD applied to MR data, we propose to augment the standard HOSVD stage by a second recursive HOSVD stage, where the "noise" predicted by the first HOSVD stage is partially added back to the
denoised image and the sum is then filtered by another HOSVD procedure, instead of the wiener filter augmentation stage adopted by Rajwade et al. This idea is borrowed from the iterative regularization originally developed by Osher et al. (Osher et al., 2005) and modified by Dong et al. (Dong et al., 2013). Although this combination seems to be simple and lack of novel ideas, this combination appears for the first time and outperforms the BM4D filter in denoising MR images, which is a simple extension of BM3D that represents the current state-of-art performance of image denoising.

With regard to the aggregation using the weighted average method from ODCT (Manjon et al., 2012), our recent experimental results show that aggregation using a simple average slightly outperform that from ODCT in terms of PSNR and SSIM for both HOSVD-based methods in the present work, as demonstrated in Table R1.C1. Thus, we deleted the description of the aggregation method proposed by Manjon et al, and updated the corresponding text in the revised manuscript.

Table R1.C1. PSNR and SSIM comparisons of using average (AVG) and weighted average (wAVG) for aggregation in the HOSVD-based methods on the BrainWeb T1w, T2w and PDw data.

<table>
<thead>
<tr>
<th>Noise level</th>
<th>1%</th>
<th>3%</th>
<th>5%</th>
<th>7%</th>
<th>9%</th>
<th>11%</th>
<th>13%</th>
<th>15%</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOSVD-W</td>
<td>AVG</td>
<td>43.56</td>
<td>38.50</td>
<td>36.04</td>
<td>34.30</td>
<td>32.95</td>
<td>31.82</td>
<td>30.87</td>
</tr>
<tr>
<td></td>
<td>wAVG</td>
<td>43.57</td>
<td>38.55</td>
<td>36.08</td>
<td>34.35</td>
<td>33.00</td>
<td>31.88</td>
<td>30.94</td>
</tr>
<tr>
<td>HOSVD-R</td>
<td>AVG</td>
<td>45.12</td>
<td>39.92</td>
<td>36.28</td>
<td>34.48</td>
<td>33.10</td>
<td>31.88</td>
<td>30.94</td>
</tr>
<tr>
<td></td>
<td>wAVG</td>
<td>45.21</td>
<td>39.97</td>
<td>36.38</td>
<td>34.61</td>
<td>33.23</td>
<td>31.93</td>
<td>30.98</td>
</tr>
<tr>
<td>T2w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOSVD-W</td>
<td>AVG</td>
<td>41.92</td>
<td>35.23</td>
<td>32.79</td>
<td>31.30</td>
<td>30.10</td>
<td>29.16</td>
<td>28.10</td>
</tr>
<tr>
<td></td>
<td>wAVG</td>
<td>41.93</td>
<td>35.25</td>
<td>32.83</td>
<td>31.36</td>
<td>30.16</td>
<td>29.10</td>
<td>28.15</td>
</tr>
<tr>
<td>HOSVD-R</td>
<td>AVG</td>
<td>43.40</td>
<td>36.43</td>
<td>33.55</td>
<td>31.69</td>
<td>30.04</td>
<td>28.99</td>
<td>28.11</td>
</tr>
<tr>
<td></td>
<td>wAVG</td>
<td>43.48</td>
<td>36.48</td>
<td>33.63</td>
<td>31.78</td>
<td>30.06</td>
<td>29.02</td>
<td>28.15</td>
</tr>
<tr>
<td>PDw</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOSVD-W</td>
<td>AVG</td>
<td>43.30</td>
<td>37.42</td>
<td>35.02</td>
<td>33.37</td>
<td>32.03</td>
<td>31.02</td>
<td>30.12</td>
</tr>
<tr>
<td></td>
<td>wAVG</td>
<td>43.31</td>
<td>37.46</td>
<td>35.08</td>
<td>33.42</td>
<td>32.12</td>
<td>31.07</td>
<td>30.17</td>
</tr>
<tr>
<td>HOSVD-R</td>
<td>AVG</td>
<td>44.69</td>
<td>38.32</td>
<td>35.49</td>
<td>33.65</td>
<td>32.21</td>
<td>31.16</td>
<td>30.28</td>
</tr>
<tr>
<td></td>
<td>wAVG</td>
<td>44.79</td>
<td>38.34</td>
<td>35.58</td>
<td>33.77</td>
<td>32.26</td>
<td>31.23</td>
<td>30.36</td>
</tr>
</tbody>
</table>
The VST, which had been used in the compared BM4D method (Foi, 2011), is only a necessary step to deal with the Rician noise and does not account for the gain of the proposed HOSVD-R method over BM4D.

In all, the main contribution of this work is the introduction of a MR image denoising algorithm which achieves the state-of-the-art performance by the combination of the HOSVD filter and a two-step recursive regulation, which is simple but effective and novel.

The only original contribution of the paper is the optimal parameter setting in page 13 which seems to me totally ad hoc.

\[ p = \log(100 \times \text{noiselevel} + 0.3) + 2 \]  

where this comes from? Is the noiselevel in percentage or in absolute value?

R1.C2: We thank the reviewer for the constructive criticism. We agree with the reviewer that the above optimal parameter setting is ad hoc and lacks theoretical foundation. Thus, we deleted the corresponding sentence and gave the experienced parameter settings as follows.

"For real MR data, the HOSVD-R algorithm is implemented with experiential parameters of \( p = 4 \) and \( K = 75 \)."

Time complexity is not enough; the authors must show the typical temporal cost in seconds, minutes, etc. so the possible users judge the feasibility of using the proposed method.

R1.C3: We thank the reviewer for the suggestion. We have added some text to describe the running time of compared methods as follows:

“All denoising methods were performed in MATLAB 7.12.0 (R2011a) on a Windows
To denoise a typical 3D dataset of size 181×217×181 pixels, the proposed HOSVD-R method took 24 min on average (multithreaded Matlab implementation on four cores), whereas HOSVD-W took 8 min (multithreaded Matlab implementation on four cores), BM4D 16 min (single-threaded Matlab/C implementation), PRI-NLM3D 1 min (multithreaded Matlab/C implementation on four cores).”

All in all, the experimental section shows interesting results. This part may justify the publication of the paper if the authors present their contributions more explicitly and rigorously.

R1.C4. We thank the reviewer for the suggestion and have modified the corresponding text accordingly for a more explicit and rigorous presentation of our contribution in the last paragraph of Introduction as follows:

“To the best of our knowledge, the HOSVD-based method has not been previously considered for MR denoising, although it presents promising properties. The present paper aims to investigate and improve the application of HOSVD to denoising MR volume data. The Wiener filter-augmented HOSVD method proposed by Rajwade et al (Rajwade et al., 2013) is firstly extended and applied to denoise 3D MR data. For better denoising performance, we proposed to augment the standard HOSVD stage by a second recursive HOSVD stage, where the "noise" predicted by the first HOSVD stage is partially added back to the denoised image and the sum is then filtered by another HOSVD procedure.”

Reviewer #2

The paper is dedicated to a very important problem such as a noise correction in MRI. Authors developed original algorithm for 3D MRI data denoising based on high-order singular value decomposition.
We thank reviewer 2 for the constructive criticism. We will respond to all issues addressed by the reviewer below.

However, in my opinion the paper suffers from a lack of details for practical implementation of developed approach. For example, authors did not mention how to choose the optimal parameters such as $\xi$ or $\gamma$. In turn, these parameters play a crucial role in resulting quality of denoising. Moreover, in the case of bad luck it might lead to worse results comparing to the state-of-the-art BM4D.

R2.C1: We thank the reviewer for pointing it out. We agree with the reviewer that the performance of the HOSVD-R depends on the setting of parameters: $\xi$ and $\gamma$. In the revised manuscript, we show the impact of these two parameters on denoising performance applied to the BrainWeb MR images at three typical noise levels (3%, 7% and 11%) in Figs. 2 and 3. It can be observed that HOSVD-R outperforms BM4D in quite a large range of $\xi$ and $\gamma$. Based on this observation, the parameter $\xi$ and $\gamma$ was respectively set 0.38 and 0.65 in all the following denoising experiments of the present work. The HOSVD-R with the above experiential-determined values of the parameters $\xi$ and $\gamma$, although are not optimal, are demonstrated to outperform BM4D in terms of PSNR and SSIM for varying image types and noise levels. We agree with the reviewer that it might lead to worse results comparing to the state-of-the-art BM4D in the case of bad luck, but it should be also noted that there are quite a few experientially-determined parameters in the BM4D filter. We added some text to discuss the feasibility of automatic parameter tuning using no-reference measures such as Stein's unbiased risk estimate (SURE) as follows:

"One limitation of the proposed HOSVD-R method is that its several filtering parameters (the number of similar cubes $K$, the cube size $p$, the relaxation parameter $\xi$ and the scaling factor $\gamma$) are currently manually determined by experience. The optimal parameter setting may change with noise levels and image characteristics.
Although experiments in this study show that the proposed method can achieve state-of-the-art performance with experiential parameters, the automatic determination of filtering parameters with theoretical foundations is warranted in a future study. Stein's unbiased risk estimate (Stein, 1981) or the no-reference metric Q (Zhu and Milanfar, 2010) can serve as quantitative metrics for automatic parameter tuning when no "ground-truth" reference is available, but will require a number of iterations."

It should also be noted that the optimal quantitative PSNR or SURE measure does not guarantee optimal filtering in the visual perception sense, thus it is difficult to find a strictly optimal filtering parameter in practice and the filtering parameters are usually manually and carefully tuned according to different image contents. From Figs. 2 and 3, it can be observed that the proposed method is not very sensitive to the parameter $\xi$ and $\gamma$, which may facilitate the manual adjusting of filtering parameters in practice.

Excepting the lack of details, in my opinion, authors have to concentrate their efforts to better algorithm based on HOSVD class instead of proposing a few variants and implementations such as the HOSVD-W, R.

R2.C2: We thank for the reviewer for the constructive criticism. Although the combination of HOSVD and recursive regulation seems to be simple, this combination appears for the first time and outperforms the BM4D filter in denoising MR images. The simple nature and outstanding performance of the proposed HOSVD-R method will facilitate its application in practice. Please see our response to R1.C1 for more detailed interpretation.

We agree with the reviewer that developing better algorithm based on HOSVD class is a more important issue. We have been trying to develop better HOSVD-based algorithm including the automatic determination of filtering parameters using Stein's unbiased risk estimate, but which requires numerous iterations and are very
time-consuming. Thus, we are developing the GPU implementation of the HOSVD denoising algorithm which may greatly shorten the filtering time. We feel it is difficult to include such work in this manuscript due to the limit of scope but added some discussion on this issue.

Minor remarks:

1. It would be very helpful if authors provide page numbers.

R2.C3: We thank the reviewer for pointing it out and have added page numbers.

2. In introduction authors did not explain very well pros and cons of the HOSVD class algorithms.

R2.C4: We thank the reviewer for pointing out this question and have added some text in introduction to explain the pros and cons of the HOSVD class algorithms.

“The HOSVD bases are learned from image and thus more adaptable to the image content and may achieve a more sparse representation than fixed bases such as wavelet and discrete cosine bases used in the BM3D method. Although the learned HOSVD bases may be sensitive to the noise in images, this disadvantage can be well addressed by learning bases from a predenoised image. Compared other SVD-based nonlocal denoising methods (Dong et al., 2013; Elad and Aharon, 2006), HOSVD does not require rearranging patches into column vectors, thus may better preserve topological structures during exploiting sparsity.”

3. It has no sense to mention b-values and related approaches such as diffusion-weighted imaging/DTI if authors anyway are not going to use it in their work. Otherwise, comprehensive definition should be performed.

R2.C5: We thank the reviewer for this advice and have deleted corresponding text in
introduction accordingly.

4. Please, check the sign > in Eq. (4)

R2.C6: We apologise for the typo and have corrected this sign in Eq. (4).

5. Figs. 4-6 can be much more informative with a scale bar for error images.

R2.C7: We thank the reviewer for this advice and have added the scale bar for error images.

6. Fig. 8 - definition of used frames is missed in figure description.

R2.C8: We thank the review for pointing it out and have added text to indicate the used frames in the legend of Fig. 9 (i.e., Fig. 8 in the original manuscript) as follows:

"From top to bottom are the original, denoised, and residual images. From left to right: sagittal, coronal and transverse frames."

7. In conclusion authors might pay more attention to the limitations of the developed approach, in particular, in the case of parallel imaging.

R2.C9: We thank the reviewer for this advice and have added some text to discuss the limitations of our proposed method including in the case of paralleling imaging in Conclusion/Discussion section.

“Finally, the proposed method is limited to processing spatially invariant Rician noise across the whole images, thus cannot be directly applied to denoising MR images reconstructed from multichannel data. For images reconstructed by the root-sum-squares operation from non-subsampled multichannel data, the noise is
assumed to follow the noncentral Chi distribution (Constantinides et al., 1997; Koay and Basser, 2006). Thus, the noncentral Chi version of VST is required to be developed for the application of HOSVD-based methods. For images reconstructed from subsampled data by using parallel imaging techniques such as sensitivity encoding (Pruessmann et al., 1999) and generalized autocalibrating partially parallel acquisitions (Griswold et al., 2002), the noise generally follows nonstationary Rician and noncentral Chi distributions, respectively. In such scenarios, the g-factor (Robson et al., 2008), which characterizes the varying noise field, can be employed to guide the denoising. When no g-factor is available, one possible approach to handle the spatially varying noise is to adapt the HOSVD-based algorithms according to locally estimated noise variances (Landman et al., 2009; Manjon et al., 2010; Maximov et al., 2012), which is warranted in a future study.”

References:


Denoising of 3D Magnetic Resonance Images by Using Higher-Order Singular Value Decomposition

Xinyuan Zhang, Zhongbiao Xu, Nan Jia, Wei Yang, Qianjin Feng, Wufan Chen, Yanqiu Feng

a. School of Biomedical Engineering, Southern Medical University, Guangzhou, China

Corresponding author:
Wufan Chen
Email: chenwf@fimmu.com
Yanqiu Feng
Email: foree@163.com

Grant Support: This study was supported by the National Basic Research Program of China (2010CB732502), National Natural Science Funds of China (81371539).
The denoising of magnetic resonance (MR) images is important to improve the inspection quality and reliability of quantitative image analysis. Several nonlocal filters, by which exploiting similarity and/or sparseness among patches or cubes achieve excellent performance, have been proposed and in denoising applied to MR images and have achieved excellent denoising performance. Recently, higher-order singular value decomposition (HOSVD) has been demonstrated to be a simple and effective method for exploiting redundancy in the 3D stack of similar patches proposed recently as a transform inducing denoising 2D natural image. HOSVD has been demonstrated to be a simple, elegant, and effective method for exploiting redundancy in the 3D stack of similar patches. In this work, we aim to investigate the application and improvement of HOSVD to denoising MR volume data. The wiener-augmented HOSVD method achieves comparable performance to that of BM4D. For further improvement, we propose to augment the standard HOSVD stage by a second recursive stage, which is a repeated HOSVD filtering of the weighted summation of the residual and denoised image in the first stage. Wiener filter-augmented HOSVD is applied to denoising 3D MR data. This study proposes a novel recursive regularization-augmented HOSVD method wherein the second denoising stage involves the HOSVD filtering of the weighted average of the original image and the pre-denoised image in the first stage. The appropriate weights have been investigated by experiments with different image types and noise levels. Experimental results over synthetic and real 3D MR data demonstrate that the proposed approach outperforms current state-of-the-art denoising methods.

Keywords: MR, Volume data Denoising, Nonlocal methods, Higher-order singular value decomposition, Sparseness
1. Introduction

Magnetic resonance imaging (MRI) is a non-invasive high-resolution multi-parameter imaging modality that is widely used in current clinical diagnosis and scientific research because of its ability to reveal the 3D internal structures of objects. However, magnetic resonance (MR) images may suffer from serious random noise, specifically when high-resolution and high-speed or high-b-value acquisition is required. The noise in MR images may affect the inspection quality in clinical diagnosis, and decrease the reliability of analysis tasks such as registration, segmentation and quantitative parameter mappings, and diffusion tensor imaging and the quantification of tissue-related parameters, such as perfusion-related and relaxation parameters (T1, T2 and T2*). Therefore, noise reduction is important for the interpretation and the computer-aided quantitative analysis of MR images.

Averaging repeated acquisitions can increase the signal-noise-ratio (SNR) in MR images at the expense of imaging time. By contrast, the denoising of MR images by post-processing filtering techniques does not increase the acquisition time; thus, post-processing filtering techniques is widely used in clinical practice and research. One simple filter is the low-pass Gaussian filter, which averages adjacent pixels by using a weight function of spatial distance between pixels (Lindenbaum M et al., 1994). This filter has been widely used in voxel-based morphometry (Ashburner and Friston, 2000) and has an inherent drawback of blurring edges, which may be clinically relevant.

Numerous edge-preserving denoising methods have been developed in the past few decades by exploiting different images and noise characteristics. Gradient-based filters, such as the anisotropic diffusion filter (Perona and Malik, 1990) and nonlinear total variation algorithm (Rudin et al., 1992), use gradient information to retain important image structures while removing noise. These filters have been extended to denoise MR images (Gerig et al., 1992; Krissian and Aja-Fernandez, 2009; Liang et al., 2011; Samsonov and Johnson, 2004). However, these filters usually erase small
details and generate piece-wise constant images. A large number of transform-based filters have also been proposed to denoise images in a transform-threshold-inverse transform fashion by exploiting the uncorrelated properties of noise in the transform domain. For example, various wavelet-based filters have been widely proposed to denoise MR images with different ways in manipulating transformed coefficients (Alexander et al., 2000; Nowak, 1999; Wirestam et al., 2006; Wood and Johnson, 1999). Transforms using adaptive bases such as the principal components analysis (Muresan and Parks, 2003) and using learned transform bases (Elad and Aharon, 2006) were also proposed to denoise images. Another popular denoising approach is based on the maximum likelihood (ML) estimation of the noise-free MR signal in the presence of Rician or non-central chi distributed noise (He and Greenshields, 2009; Rajan et al., 2011; Rajan et al., 2012; Sijbers and den Dekker, 2004).

Nonlocal denoising methods that exploit the similarity and sparsity among sparseness of small patches have emerged as an effective way to preserve satisfactory details while removing noise, and achieve state-of-the-art performances. A simple and classic method is the well-known nonlocal means (NLM) filter (Buades et al., 2005), which averages multiple pixels in a large search window and avoids the smoothing of edges by assigning only high weights to pixels with similar local patterns. The NLM filter has been adapted to denoise MR images and received considerable attention and has been improved and adapted to denoise MR images. NLM has also demonstrated a better performance than the above mentioned gradient and wavelet-based filters (Coupe et al., 2008; Manjon et al., 2008; Wiest-Daessle et al., 2008). One drawback of the NLM filter is that exploiting self-similarities is impossible in the extreme case when a patch cannot find a similar patch in the image (Mairal et al., 2009). This situation can be addressed by exploiting the sparsity among redundancy between overlapping patches. To take advantage of both self-similarity and sparsity among sparseness, the discrete cosine transform (DCT) method (Guleryuz, 2007) is extended and combined with a rotationally invariant version of the NLM filter to denoise 3D MR data; the combined method demonstrates better
results than both the oracle-based DCT and block-wise NLM filters (Manjon et al., 2012).

Another well-known nonlocal denoising method is the block-matching 3D filtering (BM3D) filter, which was proposed by Dabov et al. (2007). This method combines nonlocal and transform-domain approaches (Katkovnik et al., 2010) and presents an novel and effective denoising framework by grouping similar patches into a 3D array, then filtering the 3D array by using sparse representation in the transform domain, and finally combining aggregating multiple estimates at each location the filtered patches by using aggregation. This framework of grouping similar patches has the advantage of enhancing sparsity and is the reason behind the outstanding performance of BM3D and also adopted in many recent papers (Chatterjee and Milanfar, 2009; Dong et al., 2013; Mairal et al., 2009). The BM3D filter has been extended to volumetric data and denominated BM4D, which demonstrates state-of-the-art performance in denoising 3D MR images (Maggioni et al., 2013). The BM3D filter assumes that the underlying truth patches can be well represented using sparse coefficients with some fixed bases, which may be less adaptive to varying image contents than learned bases (Dong et al., 2013; Elad and Aharon, 2006).

The higher-order singular value decomposition (HOSVD) of patches has recently been proposed for the denoising of natural images (Rajwade et al., 2013). Although studies have shown that HOSVD lags slightly behind BM3D in denoising grayscale images, HOSVD method offers a simple and elegant method for handling sparseness among similar patches. Recently, the higher-order singular value decomposition (HOSVD) of grouped similar patches offers a simple and elegant method for handling sparsity among similar patches for denoising natural images (Rajwade et al., 2013). The HOSVD bases are learned from image and thus more adaptable to the image content and may achieve a more sparse representation than fixed bases such as wavelet and discrete cosine bases used in the BM3D method. Although the learned HOSVD bases may be sensitive to the noise in images, this disadvantage can be well addressed by learning bases from a
predenoised image. Compared other SVD-based nonlocal denoising methods (Dong et al., 2013; Elad and Aharon, 2006), HOSVD does not require rearranging patches into column vectors, thus may better preserve topological structures during exploiting sparsity.

To the best of our knowledge, the HOSVD-based method has not been previously considered for MR denoising, although it presents promising properties. In the present paper, aims to investigate and improve the application of HOSVD to denoising MR volume data. The Wiener filter-augmented HOSVD method proposed by Rajwade et al. (Rajwade et al., 2013) is firstly extended and applied to denoise 3D MR data. For better denoising performance, we propose to augment the standard HOSVD stage by a second recursive HOSVD stage, where the "noise" predicted by the first HOSVD stage is partially added back to the denoised image and the sum is then filtered by another HOSVD procedure. A novel HOSVD denoising method wherein the first standard HOSVD stage is augmented by a recursive regularization stage is also proposed.

2. Materials and methods

2.1. HOSVD denoising for 2D images

HOSVD generalizes the SVD of matrix to higher-order matrixes (De Lathauwer et al., 2000; Rajwade et al., 2013). Similar to most contemporary techniques, the HOSVD denoising method is generally designed for the zero mean (i.i.d.) Gaussian noise, wherein the observed noisy image $Y$ from a noise free image $X$ can be modeled as follows:

$$ Y = X + N $$

where $N$ denotes a zero mean Gaussian noise with a known variance $\sigma^2$. In the case of a noisy image $Y$, denoising methods aim to find a good estimate $\hat{X}$ of $X$ from $Y$. The HOSVD denoising method clusters similar patches into a stack in a similar
manner as other patch-based methods (Dabov et al., 2007; Mairal et al., 2009) and then performs the HOSVD transform of such a stack to obtain its representation by the HOSVD bases and coefficients. One estimate of this stack is reconstructed by inverse HOSVD transform with truncated coefficients. This operation is repeated for each reference patch in a sliding window fashion, and multiple estimates at each pixel are averaged to obtain the final denoised image. The details of this standard HOSVD denoising process are described in the following.

Given a $p \times p$ reference patch in the noisy 2D image, $K$ similar patches (including the reference patch) are found and stacked into a 3D array $G$ of size $p \times p \times K$. The HOSVD of this stack can be formulated as follows (De Lathauwer et al., 2000; Rajwade et al., 2013):

$$G = S \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)}$$

(2)

where $U^{(1)}, U^{(2)} \in R^{p \times p}$ and $U^{(3)} \in R^{K \times K}$ are the orthogonal unitary matrixes, $S \in R^{p \times p \times K}$ is a three-order tensor, and $\times_n$ denotes the $n$-mode product of a tensor by a matrix $U^{(n)}$. Eq. (2) shows that the HOSVD can exploit signal sparsity across each stack dimension.

The noisy image can be filtered by nullifying the HOSVD transformation coefficients below a fixed threshold under the assumption that the original coefficients of the clean image have sparse distributions. The thresholding of the transform coefficients can be represented as follows:

$$\hat{S} = H_\tau(S)$$

(3)

where $H_\tau$ denotes the hard thresholding operator with threshold $\tau$, which is determined as $\tau = \sigma \sqrt{2 \log(p^2 K)}$ for every stack of size $p \times p \times K$ in 2D image, an optimal threshold from a statistical risk viewpoint according to the rule from (Donoho and Johnstone, 1994). It is noted that the coefficients in the tensor $S$ are not...
necessarily positive in general (De Lathauwer et al., 2000). Therefore, hard thresholding is defined on the absolute value of the coefficient array $S$ as follows:

$$H_T(S) = \begin{cases} S_i & \text{if } \text{abs}(S_i) \geq \tau \\ 0 & \text{if } \text{abs}(S_i) < \tau \end{cases}$$ (4)

where $S_i$ represents the $i^{th}$ element of core tensor $S$.

An estimate of the stack $G$ then can be obtained by the inverse HOSVD transform with truncated coefficients:

$$\hat{G} = \hat{\mathcal{S}} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)}$$ (5)

The above operation is repeated for each reference patch in a sliding window fashion, thus producing multiple estimates at each pixel. The final estimate is obtained by aggregating these multiple estimates at each pixel. Weighted combination is usually implemented to obtain the final denoised image with varying weighting strategies (Dabov et al., 2007; Manjon et al., 2012), are averaged to obtain the final denoised image by using the following weighted average rule (Manjon et al., 2012):

$$\hat{X}(i) = \frac{\sum_{j=1}^{V} w_j \hat{X}_j(i)}{\sum_{j=1}^{V} w_j}, \quad w_j = \frac{1}{1 + \left\| \hat{S}_j \right\|_0}$$ (6)

where $V$ is the number of patches contributing to pixel $i$, $w_j$ is the weight of each patch $j$, $\hat{X}_j(i)$ is the estimate of pixel $i$ from patch $j$, $\left\| \hat{S}_j \right\|_0$ is the number of nonzero coefficients in $\hat{S}_j$ after hard threshold operation.

To improve further the denoising performance, the above HOSVD denoising process can be augmented by a Wiener filter step: Let $\hat{G}$ denote the stack of similar patches regrouped from the filtered image and $G_n$ denote the corresponding stack from the noisy image. Let the coefficients of $G_n$ and $\hat{G}$ in the HOSVD bases of $\hat{G}$ be
denoted as $c_n$ and $\hat{c}_n$, respectively. The Wiener filter shrinks the coefficients $c_n$ in the presence of the coefficients $\hat{c}_n$ such that $\hat{c}_n = c_n \hat{c}^2 / (\hat{c}^2 + \sigma^2)$, where $\hat{c}_n$ is the filtered coefficients of the stack $G_n$. The Wiener filter-augmented HOSVD performs better than the standard HOSVD and works almost the same as the BM3D in denoising natural grayscale images (Rajwade et al., 2013). In this paper, the Wiener filter-augmented HOSVD is referred to as HOSVD-W.

2.2. Application of HOSVD-W to 3D MR image denoising

The above HOSVD-W denoising for 2D images can be straightforwardly extended to denoise 3D images by employing the four-order HOSVD transform of a stack of similar 3D cubes. In the truncation of HOSVD coefficients, the threshold $\tau$ is picked to be $\sigma \sqrt{2 \log(p^2 K)}$ for the stack of size $p \times p \times p \times K$ in a 3D image. The HOSVD-based method is designed for data that are contaminated by additive Gaussian noise independent of the signal and cannot be directly applied to denoise MR magnitude images, which are generally corrupted with Rician noise.

Owing to the advantage of optimal forward and inverse variance-stabilizing transformation (VST), the algorithms designed for additive white Gaussian noise achieve state-of-the-art performance in denoising MR images (Foi, 2011; Maggioni et al., 2013). HOSVD-based denoising algorithms can be integrated into Rician data by using the following formula:

$$\hat{Z} = \text{VST}^{-1}(\text{Denoising}_{\text{HOSVD}}(\text{VST}(Z, \sigma_R, \sigma_{\text{VST}}), \sigma_R))$$

(26)

where $\text{VST}^{-1}$ denotes the inverse VST, $Z$ denotes the MR data corrupted by Rician noise with standard deviation $\sigma_R$, and $\sigma_{\text{VST}}$ is the standard deviation of noise after VST and is generally stabilized to one.

2.3. Proposed regularization-Augmented recursive HOSVD denoising method

The Wiener filter requires accurate estimation of the SNR at each of transform bases.
To improve the performance of HOSVD denoising method, we propose to augment the standard HOSVD staged denoising by using the recursive regularization in the second stage, which has the advantage of exploiting filtering residuals. The flow diagram of the proposed two-stage HOSVD algorithm is shown in Fig. 1. The second stage in this proposed method is the HOSVD denoising of the combined image from original and pre-denoised images; this stage can also be considered a two-stage recursive regulariization by adding filtered noise back to the denoised image (Dong et al., 2013; Osher et al., 2005):

\[ Y' = \hat{X}_{\text{HOSVD}1} + \xi(Y - \hat{X}_{\text{HOSVD}1}) = (1 - \xi)\hat{X}_{\text{HOSVD}1} + \xi Y \] (87)

where \( \xi \in [0,1] \) is a relaxation parameter representing the part of original noisy image in the combination ranging from zero to one. \( \hat{X}_{\text{HOSVD}1} \) is the restored image after the first HOSVD denoising stage.

The second stage also employs the hard-thresholding of the HOSVD transform coefficients with an updated threshold:

\[ \tau_w = \sigma_w \sqrt{2\log(pK)} \] (98)

where \( \sigma_w \) is the updated noise variance after the feedback of filter noise and can be updated calculated as follows (Dong et al., 2013):

\[ \sigma_w = \gamma \sqrt{\sigma^2 - \|Y' - Y\|} \] (499)
where $\sigma$ is the standard deviation of noise in the original data $Y$. $\gamma \in [0,1]$ is a scaling factor controlling the re-estimation of noise variance.

For clarity, the proposed HOSVD denoising augmented by the recursive regularization technique is referred to as HOSVD-R. HOSVD-R is combined with VST when applied to 3D MR data denoising. For the aggregation in the HOSVD-R method, the simple average of multiple estimates at each pixel is adopted.

### 2.4. Practical implementation to improve efficiency

The two HOSVD-based methods are applied to denoise MR volume data. Given that the most time-consuming parts of HOSVD-based algorithms are cube grouping and HOSVD transform. Cubes with Euclidean distances below a threshold are clustered into the stack of the reference cube in the HOSVD-W implementation (Rajwade et al., 2013), while the K-nearest-neighbors of the reference cube are found on the basis of the Euclidean distance between cubes and are then grouped into a stack in HOSVD-R implementation. In this section, we improve the computation efficiency without significantly affect the denoising efficacy by using similar techniques in the BM4D denoising of MR volume data (Maggioni et al., 2013).
Table 1. Parameter selections of the cube size $p$ and number $K$ of similar cubes under different noise levels.

<table>
<thead>
<tr>
<th>Noise level</th>
<th>1%</th>
<th>3%</th>
<th>5%</th>
<th>7%</th>
<th>9%</th>
<th>11%</th>
<th>13%</th>
<th>15%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P$</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>$K$</td>
<td>35</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>1</td>
<td>1</td>
<td>20</td>
</tr>
</tbody>
</table>

Restricted searching space: Maggioni et al. (Maggioni et al., 2013) demonstrated that bigger search windows do not always have better results. Thus, the similarity cube is searched only in a restricted window of size $N \times N \times N$, which is significantly smaller than the data dimensions, to reduce time complexity. The size of search window in the two HOSVD-based methods is set at 13 similar to that in the BM4D method (Maggioni et al., 2013); further increasing the search window size may reduce denoising efficacy and efficiency.

Reference cube selecting: The HOSVD transform time is linearly correlated with the number of reference cubes selected in the implementation. To improve the computational efficiency, the reference cube with a sliding step of $p-1$ is moved along each data dimension to achieve an acceleration of $(p-1)^3$ compared with moving the reference cube voxel-by-voxel.

3. Experiments and results

To evaluate and compare the performance of denoising methods quantitatively, volume data from BrainWeb database (Collins et al., 1998; Kwan et al., 1999) are used. Noise-free T1 weighted data (T1w), T2 weighted data (T2w), and proton density weighted data (PDw) corresponding to the normal brain database are downloaded. The data have a size of $181 \times 217 \times 181$ with $1\times 1 \times 1\text{mm}^3$ resolution. To evaluate the performance under different noise levels, the downloaded data are added with varying levels of Rician noise (from 1% to 15% of the maximum intensity with an increase of 2%).
Fig. 2. Influence of parameter $\xi$ in the HOSVD-R algorithm on the PSNR and SSIM in denoising the T1w, T2w, and PDw images under different noise levels (3%, 7%, and 11%).
Fig. 3. Influence of parameter $\gamma$ in the HOSVD-R algorithm on the PSNR and SSIM in denoising the T1w, T2w, and PDw images under different noise levels (3%, 7%, and 11%).
Table 2. PSNR and SSIM comparisons of different algorithms on the T1w from the BrainWeb database.

<table>
<thead>
<tr>
<th>Noise level</th>
<th>1%</th>
<th>3%</th>
<th>5%</th>
<th>7%</th>
<th>9%</th>
<th>11%</th>
<th>13%</th>
<th>15%</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRI-NLM3D</td>
<td>43.97</td>
<td>38.19</td>
<td>35.34</td>
<td>33.79</td>
<td>31.94</td>
<td>30.74</td>
<td>30.19</td>
<td>29.88</td>
</tr>
<tr>
<td>BM4D</td>
<td>44.03</td>
<td>38.34</td>
<td>35.87</td>
<td>34.22</td>
<td>32.95</td>
<td>31.89</td>
<td>31.21</td>
<td>30.88</td>
</tr>
<tr>
<td>HOSVD-W</td>
<td>43.57</td>
<td>38.55</td>
<td>36.08</td>
<td>34.58</td>
<td>33.00</td>
<td>31.88</td>
<td>31.33</td>
<td>30.53</td>
</tr>
<tr>
<td>HOSVD-R</td>
<td>45.21</td>
<td>40.97</td>
<td>36.90</td>
<td>34.61</td>
<td>33.27</td>
<td>32.23</td>
<td>31.33</td>
<td>30.53</td>
</tr>
</tbody>
</table>

Two quantitative measures are adopted to evaluate the denoising performance in the presence of ground truth. The first measure is the peak SNR (PSNR), which is based on the root mean square error (RMSE) between the denoised data and ground truth:

$$\text{PSNR} = 20 \log_{10} \frac{255}{\text{RMSE}}$$  \hspace{1cm} (11)

The second measure is the structural similarity index (SSIM), which is consistent with human eye perception (Manjon et al., 2012; Wang et al., 2004) and defined as follows:

$$\text{SSIM} = \frac{(2 \mu_X \mu_{\hat{X}} + c_1)(2 \sigma_{X\hat{X}} + c_2)}{(\mu_X^2 + \mu_{\hat{X}}^2 + c_1)(\sigma_X^2 + \sigma_{\hat{X}}^2 + c_2)}$$  \hspace{1cm} (12)

where $c_1$ and $c_2$ are constant; $\mu_X$ and $\mu_{\hat{X}}$ are the mean of data $X$ and $\hat{X}$, respectively; $\sigma_X$ and $\sigma_{\hat{X}}$ are the standard noise variance of data $X$ and $\hat{X}$, respectively; $\sigma_{X\hat{X}}$ is the covariance of $X$ and $\hat{X}$, respectively. In this study, the SSIM is locally computed with a $3 \times 3 \times 3$ voxel Gaussian kernel, and the mean of all local SSIMs in the region-of-interest is estimated as a global SSIM. Both PSNR and SSIM are calculated on the anatomical region by removing the background removed; in such a calculation, higher values are always better.

The performance of the HOSVD-W and HOSVD-R algorithms are compared against PRI-NLM3D (Manjon et al., 2012) and BM4D (Maggioni et al., 2013) to provide
relevant comparisons. Both the PRI-NLM3D and the BM4D algorithms are operated by using software from the homepage of PRI-NLM3D and BM4D with suggested filter parameters. The BM4D algorithm represents the current state-of-the-art method in denoising 3D MR images.

Fig. 3. PSNR and SSIM comparisons of different denoising algorithms for the T1w, T2w, and PDw images under noise levels varying from 1% to 15% with an increase of 2%.
3.1. Filter parameters

Patch size is the only free parameter which influences the accuracy in the original HOSVD-W algorithm (Rajwade et al., 2013). Thus, patches with varying sizes are implemented and results with optimal PSNR and SSIM are obtained.

The performance of HOSVD-R denoising depends on the setting of two three cube-related
Fig. 2. Influence of parameter $\xi$ in the HOSVD-R algorithm on the PSNR and SSIM in denoising the T1w, T2w, and PDw images under different noise levels (3%, 7%, and 11%).
Fig. 3. PSNR and SSIM comparisons of different denoising algorithms for the T1w, T2w, and PDw images under noise levels varying from 1\% to 15\% with an increase of 2\%.

parameters (i.e., cube size $p$, search window size $N$, and the number of similar cubes in a group $K$ in groups) and two noise-related parameters (i.e., relaxation parameter $\lambda$ in (Eq. (7))) which controls the contribution of feed-back noise in the data $Y_{\omega}$ and
scaling factor $\gamma$ in Eq. (109), which controls the re-estimation of noise variance; relaxation parameter $\xi$ in Eq. (8), which controls the contribution of feedback noise in the data $Y_{w}$).

The search window size in HOSVD-R denoising is set at 13 similar to the search window size in the BM4D method (Maggioni et al., 2013). Further increasing the search window size may reduce denoising efficacy and efficiency. In our denoising experiments with BrainWeb MR data by using HOSVD-R, the results show that a larger size of cube $p$ and a larger number of similar cubes $K$ yield higher denoising PSNR at high noise levels. Thus, in this study, $p$ and $K$ are tentatively chosen as
values that increase with noise levels (Table 1). For real MR data, the HOSVD-R algorithm is implemented with experiential parameters of $p = 4$ and $K = 75$, we use

$$p = \lceil \log(100 \cdot \text{noise level} + 0.3) \rceil + 2,$$

where $\lceil \cdot \rceil$ maps a real number to the smallest following integer, that is, from a relation of $p$ to noise levels in Table 1, and determine the corresponding $K$ from the Table 1 with a given $p$.

To determine the appropriate values of the noise-related parameters $\xi$ and $\gamma$ for denoising MR images, the HOSVD-R algorithm was implemented on the BrainWeb data with varying $\xi$ and $\gamma$, and the resulting PSNR and SSIM values were adopted as the quantitative measures to evaluate the sensitivity to the setting of these two parameters. For T1w, T2w and PDw data at different noise levels (3%, 7% and 11%), Figs. 2 and 3 show the impact of the parameters $\xi$ and $\gamma$ on denoising performance, respectively. For noise-related parameters, the scaling factor $\gamma$ is set at 0.65, which is similar to that in the study of Dong et al. (Dong et al., 2013) and demonstrates good performance in our experiment. An exhaustive searching uses both the PSNR and SSIM criteria to determine an optimal value for parameter $\xi$. The results for T1w, T2w and PDw data at different noise levels (3%, 7% and 11%) are shown in Fig.2. HOSVD-R can produce higher PSNR and SSIM in a large range than the BM4D and HOSVD-W methods. For the T2w image at 11% noise level, HOSVD-W produces a PSNR comparable to that of HOSVD-R for the T2w image under 11% noise level but with the lower lowest SSIM. Based on the observation from Figs. 2 and 3 in the denoising experiments, $\xi$ and $\gamma$ are roughly fixed at 0.38 and 0.65 in the following denoising experiments, respectively, which yields an improved PSNR and SSIM over the compared methods for data with varying noise levels and image types, although these two values are not strictly optimal for each dataset.
3.2. Denoising of BrainWeb MR images

The PSNR and SSIM of the PRI-NLM3D, BM4D, HOSVD-W, and HOSVD-R methods on T1w, T2w and PDw images are shown in Fig. 3. The HOSVD-W algorithm achieves comparable denoising performance to BM4D in terms of PSNR and SSIM. The HOSVD-R algorithm significantly outperforms the other three with regard to PSNR for all image types under low to moderate noise levels (from 1% to 5%). At noise levels above 7%, the HOSVD-R algorithm produces slightly higher PSNRs than the BM4D and HOSVD-W algorithms, but all three algorithms produce significantly higher PSNRs than the PRI-NLM3D. With regard to SSIM, the HOSVD-R algorithm consistently yields SSIMs higher than the other three methods for all noise levels and image types. Table 2 presents the discrete values of PSNRs and SSIMs in denoising T1w images by using the four methods under investigation. The PSNRs and SSIMs of PRI-NLM3D and BM4D are close to the values reported by Maggioni et al (2013). The PSNRs and SSIMs of HOSVD-W are close to those of BM4D. The proposed HOSVD-R algorithm outperforms BM4D with PSNR improvements ranging from 0.2-3 dB to 1.2 dB and SSIM improvements from 0.002 to 0.0109.
Fig. 5. Denoised results of different algorithms on synthesized T2w brain image. Top row: the noise free T2w image and the image with Rician noise level of 15%. Second row: the denoised images with different algorithms. Third row: the corresponding error images.
Fig. 6. Denoised results of different algorithms on synthesized PDw brain image. Top row: the noise free PDw image and the image with Rician noise level of 15%. Second row: the denoised images with different algorithms. Third row: the corresponding error images.

Figs. 4-5 to 7-8 provide a visual evaluation of denoising results under T1w, T2w, and PDw data with 15% noise level. All compared algorithms show good performance in removing noise without significant detail loss. The PRI-NLM3D result has the sharpest edges but with the most prominent errors and most detail loss compared with the other denoising algorithms. In addition to PRI-NLM3D, HOSVD-R also produces results with fewer intensity oscillations in homogeneous areas than BM4D and HOSVD-W; this finding can be more clearly observed from the error images. To provide a better visual inspection, the enlarged detail is shown in Fig. 7b. As seen in the region pointed by the arrows, PRI-NLM3D oversmooths low contrast details, and the results of the proposed HOSVD-R method are more visually pleasant than those
of other compared methods.

It should also be noted that PSNR and SSIM are calculated only in the anatomical region after removing the background. The error images show that the proposed approach removes background noise significantly better than the other denoising algorithms. If the background is included in the quantification, the PSNR of the proposed algorithm will be significantly higher than those of the other three algorithms.

3.3. Denoising of real MR data

This section is devoted to verifying the consistency of the proposed approach on real MR data. The experiments are conducted on two cross-sectional T1w MR brain datasets (OAS1_0112 and OAS1_0092) from the publicly available Open Access Series of Imaging Studies (OASIS) database (http://www.oasis-brains.org) (Marcus et al., 2007). The T1w real brain datasets are acquired by an MP-RAGE volumetric sequence on a Siemens 1.5 T Vision scanner. The acquisition parameters have a repetition time of 9.7 ms, echo time of 4 ms, flip angle of 10°, inversion time of 20 ms, duration time of 200 ms, matrix of 256×256×128, voxel resolution of 1×1×1.25 mm³.

The noise levels of the two selected datasets are approximately 3% and 4.5% of the maximum intensity according to the noise level estimation method of (Foi, 2011). The denoised results of HOSVD-R on these two datasets are presented in Fig. 8. The HOSVD-R method significantly reduces the noise, and no significant anatomical information can be observed in the residual image (the difference between the noisy and denoised images). As shown by the enlarged image in Fig. 9, PRI-NLM slightly oversmooths fine details, whereas BM4D and HOSVD-W preserves satisfactory details but with some intensity oscillations. The proposed HOSVD-R generates the most visually pleasant result. These observations are consistent with the results of synthetic data.
Fig. 7. Enlarged parts of the denoised images by different algorithms in Figs. 5 to 7. From top to bottom: T1w, T2w, and PDw images.

Fig. 8. Denoised results of the proposed HOSVD-R algorithm on two sets of real T1w brain data with different estimated Rician noise levels (Left: 3%; Right: 4.5%). From top to bottom are the original, denoised, and
residual images. From left to right: sagittal, coronal and transverse frames. Note that the higher values in the residuals are due to Rician noise-related bias on low intensity signals.

<table>
<thead>
<tr>
<th>Noisy Image</th>
<th>PRI-NLM3D</th>
<th>BM4D</th>
<th>HOSVD-W</th>
<th>HOSVD-R</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image6.png" alt="Image" /></td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
<td><img src="image15.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image16.png" alt="Image" /></td>
<td><img src="image17.png" alt="Image" /></td>
<td><img src="image18.png" alt="Image" /></td>
<td><img src="image19.png" alt="Image" /></td>
<td><img src="image20.png" alt="Image" /></td>
</tr>
</tbody>
</table>

**Fig. 9**. Comparison of different algorithms on the enlarged views of the rectangular regions in Fig. 9. (a): enlarged parts of Fig. 8a; (b): enlarged parts of Fig. 8b.

4. Conclusion/Discussion
Two HOSVD-based methods that exploit the sparseness between similar cubes by using HOSVD transform have been presented to denoise 3D MR data. One method is the simple extension of the original HOSVD method augmented by a Weiner filter to 3D MR data (HOSVD-W), and the other method is a novel HOSVD-R method that takes advantage of a recursive regularization to improve the denoising performance of HOSVD. Both HOSVD-based methods are applied and compared with current state-of-the-art methods on synthetic and real MR volumetric data. The experimental results demonstrate that the proposed HOSVD-R algorithm outperforms current state-of-the-art algorithms in 3D MR denoising.

The excellent denoising performance of the HOSVD-R method can be attributed to two main features. The first feature is that the HOSVD-R denoising method represents a stack of similar cubes by using learned orthogonal bases; the transform of such bases is more adaptable to different data types and can achieve a more efficient and sparse signal representation than a transform using fixed bases such as wavelet and DCT in the BM4D method. The second feature is that the feedback of filtered noise in the second stage maintains a two-step regularization effect to encourage the solution toward an optimal result; this feature distinguishes HOSVD-R from HOSVD-W. The proposed HOSVD-R also benefits from the manipulation of “better” HOSVD bases from combined images in the second stage, compared with bases from the original noisy data in the first stage.

Considering that both the BM4D and the HOSVD-based methods operate on stacks of similar cubes, comparing the time complexity of these transforms over a single stack clarifies the time costs. For a stack of size $p \times p \times p \times K$, BM4D requires $O(Kp^4)$ and $O(K^2p^3)$ operations for 3D and 1D transforms, respectively, thus leading to a total complexity of $O(Kp^4 + K^2p^3)$. The time complexity of HOSVD is $O(Kp^4 + \min(K^2p^3, Kp^5))$ (Rajwade et al., 2013). Thus, if the two types of denoising methods construct stacks with the same size, the computation complexity of the HOSVD-based method
is comparable to that of BM4D. All denoising methods were performed in MATLAB 7.12.0 (R2011a) on a Windows 7 computer equipped with an Intel(R) Core(TM)2 2.33 GHz CPU and 8 GB RAM. To denoise a typical 3D dataset of size $181 \times 217 \times 181$ pixels, the proposed HOSVD-R method took 24 min on average (multithreaded Matlab implementation on four cores), whereas HOSVD-W took 8 min (multithreaded Matlab implementation on four cores), BM4D 16 min (single-threaded Matlab/C implementation), PRI-NLM3D 1 min (multithreaded Matlab/C implementation on four cores). We believe that the implementation of HOSVD-based denoising methods using Matlab/C MEX techniques and parallel computations on Developing a graphic processing unit implementation of HOSVD transform may significantly further accelerate the filtering and enable the real-time denoising of 3D data.

One limitation of the proposed HOSVD-R method is that its several filtering parameters (the number of similar cubes $K$, the cube size $p$, the relaxation parameter $\xi$, and the scaling factor $\gamma$) are currently manually determined by experience. The optimal parameter setting may change with noise levels and image characteristics. Although experiments in this study show that the proposed method can achieve state-of-the-art performance with experiential parameters, the automatic determination of filtering parameters with theoretical foundations is warranted in a future study. Stein's unbiased risk estimate (Stein, 1981) or the no-reference metric $Q$ (Zhu and Milanfar, 2010) can serve as quantitative metrics for automatic parameter tuning when no "ground-truth" reference is available, but will require a number of iterations.

Finally, the proposed method is limited to processing spatially invariant Rician noise across the whole images, thus cannot be directly applied to denoising MR images reconstructed from multichannel data. For images reconstructed by the root-sum-squares operation from non-subsampled multichannel data, the noise is assumed to follow the noncentral Chi distribution (Constantinides et al., 1997; Koay,
and Basser, 2006). Thus, the noncentral Chi version of VST is required to be
developed for the application of HOSVD-based methods. For images reconstructed
from subsampled data by using parallel imaging techniques such as sensitivity
encoding (Pruessmann et al., 1999) and generalized autocalibrating partially parallel
acquisitions (Griswold et al., 2002), the noise generally follows nonstationary Rician
and noncentral Chi distributions, respectively. In such scenarios, the g-factor (Robson
et al., 2008), which characterizes the varying noise field, can be employed to guide
the denoising. When no g-factor is available, one possible approach to handle the
spatially varying noise is to adapt the HOSVD-based algorithms according to locally
estimated noise variances (Landman et al., 2009; Manjon et al., 2010; Maximov et al.,
2012), which is warranted in a future study. Finally, the proposed method assumes
spatially invariant Rician noise across the whole images. However, the noise field can
be spatially variable across the image due to the parallel imaging using multiple
coils, the local fielded inhomogeneities, and imaging artifacts—(Landman et al., 2009;
Manjon et al., 2010; Maximov et al., 2012). Future studies are warranted to extend the
HOSVD-based method to consider spatially varying noise for images reconstructed
from subsampled data by using parallel imaging techniques.
Acknowledgements:

We want to thank Jose V. Manjon (Manjon et al., 2012), Matteo Maggioni (Maggioni et al., 2013), Ajit Rajwade (Rajwade et al., 2013), and Weisheng Dong (Dong et al., 2013) for publishing online their implementation of PRI-NLM3D, BM4D, HOSVD and spatially adaptive iterative singular-value thresholding (SAIST), respectively. This study was supported by the National Basic Research Program of China (2010CB732502), National Natural Science Funds of China (81371539).
References


1 Trans Image Process 22, 119-133.


Table 1. Parameter selections of the cube size $p$ and number $K$ of similar cubes under different noise levels.

<table>
<thead>
<tr>
<th>Noise level</th>
<th>1%</th>
<th>3%</th>
<th>5%</th>
<th>7%</th>
<th>9%</th>
<th>11%</th>
<th>13%</th>
<th>15%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p$</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>$K$</td>
<td>35</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
</tr>
<tr>
<td>Noise level</td>
<td>1%</td>
<td>3%</td>
<td>5%</td>
<td>7%</td>
<td>9%</td>
<td>11%</td>
<td>13%</td>
<td>15%</td>
</tr>
<tr>
<td>-------------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>PRI-NLM3D</td>
<td>43.97</td>
<td>38.19</td>
<td>35.34</td>
<td>33.37</td>
<td>31.94</td>
<td>30.74</td>
<td>29.75</td>
<td>28.88</td>
</tr>
<tr>
<td></td>
<td>0.992</td>
<td>0.975</td>
<td>0.956</td>
<td>0.935</td>
<td>0.914</td>
<td>0.891</td>
<td>0.868</td>
<td>0.847</td>
</tr>
<tr>
<td>BM4D</td>
<td>44.03</td>
<td>38.34</td>
<td>35.87</td>
<td>34.22</td>
<td>32.95</td>
<td>31.89</td>
<td>30.97</td>
<td>30.14</td>
</tr>
<tr>
<td></td>
<td>0.992</td>
<td>0.975</td>
<td>0.960</td>
<td>0.944</td>
<td>0.928</td>
<td>0.912</td>
<td>0.895</td>
<td>0.878</td>
</tr>
<tr>
<td>HOSVD-W</td>
<td>43.57</td>
<td>38.55</td>
<td>36.08</td>
<td>34.35</td>
<td>33.00</td>
<td>31.88</td>
<td>30.94</td>
<td>30.10</td>
</tr>
<tr>
<td></td>
<td>0.992</td>
<td>0.977</td>
<td>0.961</td>
<td>0.945</td>
<td>0.928</td>
<td>0.910</td>
<td>0.893</td>
<td>0.875</td>
</tr>
<tr>
<td>HOSVD-R</td>
<td>45.21</td>
<td>38.97</td>
<td>36.38</td>
<td>34.61</td>
<td>33.27</td>
<td>32.23</td>
<td>31.33</td>
<td>30.53</td>
</tr>
<tr>
<td></td>
<td>0.994</td>
<td>0.979</td>
<td>0.965</td>
<td>0.949</td>
<td>0.933</td>
<td>0.918</td>
<td>0.903</td>
<td>0.887</td>
</tr>
</tbody>
</table>