

A MULTISCALE APPROACH FOR TENSOR DENOISING

Alp Ozdemir¹, Mark A. Iwen^{1,2} and Selin Aviyente¹

¹Department of Electrical and Computer Engineering, Michigan State University

²Department of the Mathematics, Michigan State University

ABSTRACT

As higher-order datasets become more common, researchers are primarily focused on how to analyze and compress them. However, the most common challenge encountered in any type of data, including tensor data, is noise. Furthermore, the methods developed for denoising vector or matrix type datasets cannot be applied directly to higher-order datasets. This motivates the development of denoising methods for tensors. In this paper, we propose the use of a multiscale approach for denoising general higher-order datasets. The proposed approach works by decomposing the higher-order data into subtensors, and then denoises the subtensors by recursively exploiting filtered residuals. The method is validated on both hyperspectral image and brain functional connectivity network data.

Index Terms— Tensor algebra, higher order SVD, tensor denoising

1. INTRODUCTION

With recent advances in information technology it is now possible to simultaneously collect, store, and process data from multiple sources with different attributes. This type of data, also known as higher-order data or tensor data, is encountered frequently in multimedia, hyperspectral imaging (HSI), and seismic data analysis, as well as in multiple modality neuroimaging applications [1–3]. Furthermore, higher-order data encountered in such applications is often very noisy. Existing methods for data denoising primarily focus on either vector or matrix type data, and their extensions to high order data have been mostly limited to denoising each frame or slice of a given tensor dataset independently from all other frames/slices.

More recently, methods based on the higher-order singular value decomposition (HOSVD) and high-dimensional wavelet transforms have been proposed for denoising tensor type data [1, 3, 4]. These methods attempt to take advantage of signal correlations across tensor frames/slices in order to achieve better noise reduction. For example, Muti and Bourennane [2] presented multimode Wiener filtering for

noise removal in multidimensional seismic data. Chen and Qian [1] applied bivariate Wavelet shrinking combined with Principal Component Analysis to HSIs to do simultaneous denoising and dimensionality reduction. Similarly, Othman and Qian [5] proposed a hybrid method which performs both spatial and spectral wavelet shrinkage for noise reduction in HSIs. Yuan et al. [6] extended the total variation (TV) model to tensors by making it spectrally and spatially adaptive. Low-rank tensor approximation [7] and 3D wavelets [4] are also used to denoise multidimensional datasets. More recently, Zhang et al. [3] extended a HOSVD based image denoising algorithm [8] to 3-way tensors and showed that applying the algorithm recursively improves the denoising performance on MRI data. Moreover, Peng et al. [9] proposed a dictionary learning algorithm for tensors to denoise multispectral images. The suggested technique applies HOSVD to overlapping patches to eliminate spectral noise and takes the average of non-locally similar patches to reduce spacial noise. However, most of these approaches are limited to specific order of tensors such as 3-way tensors and specific data models or types. In contrast, we propose more general denoising approaches suitable for tensors of any order. Moreover, some of these approaches are direct extensions of image denoising algorithms to tensors, and so are specifically designed for denoising image-like datasets (e.g., hyperspectral images). In particular, wavelet based algorithms exploit both spatial and spectral properties of a given dataset, and their multiscale structure provides very effective denoising performance. However, network-like datasets (for example) do not generally exhibit the spatial and spectral characteristics on which wavelet based methods rely.

In this paper, we propose a recursive and multiscale denoising approach suitable for any type of tensor data. The proposed algorithm consists of two main steps: (i) Clustering the tensor into lower dimensional subtensors which are expected to have a lower rank and a more compact representation, thereby leading to a better discrimination between signal and noise, and then (ii) Denoising each subtensor. The performance of the proposed method is evaluated on both a 3-way tensor containing a hyperspectral image, and a 4-way tensor containing brain functional connectivity networks.

This work was in part supported by NSF CCF-1422262 and NSF DMS-1416752.

2. BACKGROUND

2.1. Tensor Algebra

A multidimensional array with N modes $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ is called a tensor, where x_{i_1, i_2, \dots, i_N} denotes the $(i_1, i_2, \dots, i_N)^{th}$ element of the tensor \mathcal{X} . Vectors obtained by fixing all indices of the tensor except the one that corresponds to n th mode are called mode- n fibers. Basic tensor operations are reviewed below [10].

Mode- n product The mode- n product of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times \dots \times I_n \times \dots \times I_N}$ and a matrix $\mathbf{U} \in \mathbb{R}^{J \times I_n}$ is denoted as $\mathcal{Y} = \mathcal{X} \times_n \mathbf{U}$, $(\mathcal{Y})_{i_1, i_2, \dots, i_{n-1}, j, i_{n+1}, \dots, i_N} = \sum_{i_n=1}^{I_n} x_{i_1, \dots, i_n, \dots, i_N} u_{j, i_n}$ and is of size $I_1 \times \dots \times I_{n-1} \times J \times I_{n+1} \times \dots \times I_N$.

Tensor matricization Process of reordering the elements of the tensor into a matrix is known as matricization or unfolding. The mode- n matricization of tensor \mathcal{Y} is denoted as $\mathbf{Y}_{(n)}$ and is obtained by arranging mode- n fibers to be the columns of the resulting matrix. Unfolding the tensor $\mathcal{Y} = \mathcal{X} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \dots \times_N \mathbf{U}^{(N)}$ along mode- n is equivalent to $\mathbf{Y}_{(n)} = \mathbf{U}^{(n)} \mathbf{X}_{(n)} (\mathbf{U}^{(N)} \otimes \dots \otimes \mathbf{U}^{(n+1)} \otimes \mathbf{U}^{(n-1)} \dots \otimes \mathbf{U}^{(1)})^\top$, where \otimes is the matrix Kronecker product.

Tensor Rank Unlike matrices, which have a unique definition of rank, there are multiple rank definitions for tensors including *tensor rank* and *tensor n -rank*. The *rank* of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times \dots \times I_n \times \dots \times I_N}$ is the smallest number of rank-one tensors that form \mathcal{X} as their sum. The *n -rank* of \mathcal{X} is the collection of ranks of mode matrices $\mathbf{X}_{(n)}$ and is denoted as:

$$n\text{-rank}(\mathcal{X}) = (\text{rank}(\mathbf{X}_{(1)}), \text{rank}(\mathbf{X}_{(2)}), \dots, \text{rank}(\mathbf{X}_{(N)})). \quad (1)$$

2.2. Higher Order Singular Value Decomposition

Any tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ can be decomposed as mode products of a core tensor $\mathcal{S} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ and N orthogonal projection matrices $\mathbf{U}^{(n)} \in \mathbb{R}^{I_n \times I_n}$ which are the left singular vectors of $\mathbf{X}_{(n)}$ [11]:

$$\mathcal{X} = \mathcal{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \dots \times_N \mathbf{U}^{(N)}, \quad (2)$$

where \mathcal{S} is computed as:

$$\mathcal{S} = \mathcal{X} \times_1 (\mathbf{U}^{(1)})^\top \times_2 (\mathbf{U}^{(2)})^\top \dots \times_N (\mathbf{U}^{(N)})^\top. \quad (3)$$

The noisy tensor \mathcal{X} can be filtered by thresholding the coefficients of the core tensor \mathcal{S} as follows. Let $\tilde{\mathcal{S}}$ be a denoised core tensor, $H_\tau(\cdot)$ be the hard thresholding operator, and $\tilde{\mathcal{S}}(i)$ be the $(i_1, i_2, \dots, i_N)^{th}$ element of $\tilde{\mathcal{S}}$, then $\tilde{\mathcal{S}}(i)$ is:

$$\tilde{\mathcal{S}}(i) = H_\tau(\mathcal{S}(i)) = \begin{cases} \mathcal{S}(i), & |\mathcal{S}(i)| \geq \tau \\ 0, & |\mathcal{S}(i)| < \tau \end{cases}, \quad (4)$$

where $\tau = \sqrt{2\sigma^2 \log_{10}(\prod_{n=1}^N I_n)}$ is the universal threshold [12] and σ^2 is the noise variance. The denoised tensor $\tilde{\mathcal{X}}$ is then obtained by:

$$\tilde{\mathcal{X}} = \tilde{\mathcal{S}} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \dots \times_N \mathbf{U}^{(N)}. \quad (5)$$

3. MULTISCALE DENOSING FOR TENSORS

3.1. Tensor Denoising by Local HOSVD

In this section, we propose a new procedure to remove noise from N th order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ named Local HOSVD Denoising (L-HOSVD). The proposed method applies the following two-step approach: (i) Reclustering the noisy tensor \mathcal{X} into subtensors, (ii) Denoising by HOSVD.

The tensor \mathcal{X} is first decomposed into subtensors as follows. Tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ is unfolded across each mode yielding $\mathbf{X}_{(n)} \in \mathbb{R}^{I_n \times \prod_{j \neq n} I_j}$ whose columns are the mode- n fibers of \mathcal{X} . For each mode, rows of $\mathbf{X}_{(n)}$ are partitioned into c_n non-overlapping clusters by a clustering algorithm. Cartesian products of the partitioning labels coming from different modes yields index sets of $K = \prod_{i=1}^N c_i$ subtensors $\mathcal{X}_{1,k}$ where $k \in \{1, 2, \dots, K\}$. In our experiments, Local Subspace Analysis (LSA) [13] is used to identify the clusters.

Let J_0^n and J_k^n be the index sets of the tensor \mathcal{X} and subtensor \mathcal{X}_k^n for the n th mode, respectively where $J_0^n = \{1, 2, \dots, I_n\}$ and $J_k^n \subset J_0^n$ with $n \in \{1, 2, \dots, N\}$. Index sets of subtensors satisfy $\cup_{k=1}^K J_k^n = J_0^n$ and $J_k^n \cap J_l^n = \emptyset$ when $k \neq l$ for all $k, l \in \{1, 2, \dots, K\}$. Therefore, index set of the first subtensor \mathcal{X}_1^n can be written as $J_1^1 \times J_1^2 \times \dots \times J_1^N$ and k th subtensor \mathcal{X}_k^n is obtained by

$$\mathcal{X}_k^n(i_1, i_2, \dots, i_N) = \mathcal{X}(J_k^1(i_1), J_k^2(i_2), \dots, J_k^N(i_N)), \quad (6)$$

$$\mathcal{X}_k^n = \mathcal{X}(J_k^1 \times J_k^2 \times \dots \times J_k^N),$$

where $i_n \in \{1, 2, \dots, J_k^n\}$. Each subtensor is then decomposed as:

$$\mathcal{X}_k^n = \mathcal{S}_k \times_1 \mathbf{U}_k^{(1)} \times_2 \mathbf{U}_k^{(2)} \dots \times_N \mathbf{U}_k^{(N)}, \quad (7)$$

where \mathcal{S}_k and $\mathbf{U}_k^{(n)}$ s correspond to the core tensor and projection matrices of \mathcal{X}_k^n , respectively. In order to denoise subtensors \mathcal{X}_k^n , the procedure in section 2.2 is applied to each subtensor \mathcal{X}_k^n individually. First, the coefficients of \mathcal{S}_k are thresholded with $\tau_k = \sqrt{2\sigma^2 \log_{10}(\prod_{n=1}^N |J_k^n|)}$ which yields $\tilde{\mathcal{S}}_k = H_{\tau_k}(\mathcal{S}_k)$. Then the denoised subtensors are obtained by $\tilde{\mathcal{X}}_k^n = \tilde{\mathcal{S}}_k \times_1 \mathbf{U}_{1,k}^{(1)} \times_2 \mathbf{U}_k^{(2)} \dots \times_N \mathbf{U}_k^{(N)}$. The denoised tensor $\tilde{\mathcal{X}}$ is formed by mapping all of the subtensors onto $\tilde{\mathcal{X}}_k^n$ as follows:

$$\tilde{\mathcal{X}}(J_k^1 \times J_k^2 \times \dots \times J_k^N) = \tilde{\mathcal{X}}_k^n. \quad (8)$$

A pseudo code of the algorithm is given in Algorithm 1.

3.2. Tensor Denoising by Multiscale HOSVD

In order to improve the performance of L-HOSVD, we propose to use a two-scale procedure named Multiscale HOSVD Denoising (MS-HOSVD) by making use of recursive regularization as in [3, 14]. In the first scale, the N th order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ is denoised using HOSVD as in section 2.2 which yields a tensor $\tilde{\mathcal{X}}^{(1)}$. In the second scale,

Algorithm 1 Denoising by Local HOSVD

- 1: Input: $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$: tensor, $\mathbf{C} = (c_1, c_2, \dots, c_N)$: the desired number of clusters for n th mode, σ^2 : the noise variance
 - 2: Output: $\tilde{\mathcal{X}}$: denoised tensor
 - 3: $K \leftarrow \prod_{i=1}^N c_i$
 - 4: $[\{J_1^1, J_1^2, \dots, J_1^N, \dots, J_K^1, J_K^2, \dots, J_K^N\}] = \text{cluster}(\mathcal{X}, \mathbf{C})$
 - 5: **for** $k = 1$ to K **do**
 - 6: $\mathcal{X}_k = \mathcal{X}(J_k^1 \times J_k^2 \times \dots \times J_k^N)$
 - 7: $\mathcal{S}_k, \{\mathbf{U}_k^1, \mathbf{U}_k^2, \dots, \mathbf{U}_k^N\} \leftarrow \text{HOSVD}(\mathcal{X}_k)$
 - 8: $\tau_k = \sqrt{\sigma^2 2 \log_{10}(\prod_{n=1}^N |J_k^n|)}$
 - 9: $\tilde{\mathcal{S}}_k \leftarrow H_{\tau_k}(\mathcal{S}_k)$
 - 10: $\tilde{\mathcal{X}}_k \leftarrow \tilde{\mathcal{S}}_k \times_1 \hat{\mathbf{U}}_k^1 \times_2 \hat{\mathbf{U}}_k^2 \dots \times_N \hat{\mathbf{U}}_k^N$
 - 11: $\tilde{\mathcal{X}}(J_k^1, J_k^2, \dots, J_k^N) = \tilde{\mathcal{X}}_k$
 - 12: **end for**
-

we adapted the idea of iterative regularization that adds filtered noise back to the denoised image [15]. Then L-HOSVD is applied to a combined tensor $\mathcal{X}^{(2)}$ which is the sum of the denoised tensor $\tilde{\mathcal{X}}^{(1)}$ and residual tensor $\mathcal{R}^{(1)} = \mathcal{X} - \tilde{\mathcal{X}}^{(1)}$ obtained from the first stage as:

$$\begin{aligned} \mathcal{X}^{(2)} &= \tilde{\mathcal{X}}^{(1)} + \alpha \mathcal{R}^{(1)} \\ &= (1 - \alpha) \tilde{\mathcal{X}}^{(1)} + \alpha \mathcal{X}, \end{aligned} \quad (9)$$

where $\alpha \in [0, 1]$ is the relaxation parameter. The noise variance σ_2^2 used in second stage is estimated as $\sigma_2 = \alpha \sqrt{\sigma^2 - \|\mathcal{X} - \tilde{\mathcal{X}}^{(2)}\|}$. Applying L-HOSVD to $\mathcal{X}^{(2)}$ yields denoised tensor $\tilde{\mathcal{X}}^{(2)}$ (Algorithm 2).

3.3. Computational Complexity

Computational complexity of denosing an N-way tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ where $I_1 = I_2 = \dots = I_N = I$ using HOSVD is $O(I^{(N+1)})$. By assuming that the clustering is performed using K-means with $c_i = c$ along each mode, then the complexity of denoising by L-HOSVD becomes the sum of the complexity of clustering along each mode $N \cdot O(I^N \cdot c \cdot i)$

Algorithm 2 Denoising by Multiscale HOSVD

- 1: Input: $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$: tensor, $\mathbf{C} = (c_1, c_2, \dots, c_N)$: the desired number of clusters for n th mode, σ^2 : the noise variance, α : the relaxation parameter
 - 2: Output: $\tilde{\mathcal{X}}_2$: denoised tensor
 - 3: $\mathcal{S}_1, \{\mathbf{U}^1, \mathbf{U}^2, \dots, \mathbf{U}^N\} \leftarrow \text{HOSVD}(\mathcal{X})$
 - 4: $\tau_k = \sqrt{\sigma^2 2 \log_{10}(\prod_{n=1}^N |I^n|)}$
 - 5: $\mathcal{S}(i) = H_{\tau}(S(i))$
 - 6: $\tilde{\mathcal{X}}^{(1)} = \mathcal{S} \times_1 \mathbf{U}^1 \times_2 \mathbf{U}^2 \dots \times_N \mathbf{U}^N$.
 - 7: $\mathcal{X}^{(2)} = (1 - \alpha) \tilde{\mathcal{X}}^{(1)} + \alpha \mathcal{X}$
 - 8: $\sigma_2 = \alpha \sqrt{\sigma^2 - \|\mathcal{X} - \tilde{\mathcal{X}}^{(1)}\|}$
 - 9: $\tilde{\mathcal{X}}^{(2)} = \text{L-HOSVD}(\mathcal{X}^{(2)}, \mathbf{C}, \sigma_2)$
-

where i is the number of iterations and the complexity of HOSVD for each subtensor $c^N \cdot O((I/c)^{(N+1)})$. Similarly, denoising using MS-HOSVD requires $O(I^{(N+1)}) + N \cdot O(I^N \cdot c \cdot i) + c^N \cdot O((I/c)^{(N+1)})$ computations and order of the complexity is similar to HOSVD when $N \cdot c \cdot i$ is small compared to I .

4. RESULTS

4.1. Hyperspectral Image Denoising

In this experiment, we used a hyperspectral image from [16] to create a 3-mode tensor $\mathcal{X} \in \mathbb{R}^{401 \times 500 \times 148}$ where the modes are rows, columns and spectral bands respectively. Zero mean Gaussian noise with varying signal to noise ratio between 5 to 20 dB was added to \mathcal{X} , and HOSVD, L-HOSVD and MS-HOSVD were applied to the noisy tensor. Clustering parameters c_i s for both L-HOSVD and MS-HOSVD are set to 2 for each mode yielding 8 subtensors to denoise. To select the optimal relaxation parameter α for MS-HOSVD, different α values in the interval $[0, 1]$ are selected with a step size of 0.1. As seen in Fig. 1 $\alpha = 0.5$ maximizes the SNR for all of the noise levels. All of the experiments were repeated 20 times and the average SNR values of the denoised data were computed for each method (Table 1). As seen in Table 1, L-HOSVD provides better SNR than HOSVD while MS-HOSVD gives the highest SNR for all of the experiments. Slices from tensors denoised by HOSVD, L-HOSVD and MS-HOSVD are illustrated in Fig. 2, and it can be seen that the tensor denoised by HOSVD has more distortion than the ones denoised by L-HOSVD and MS-HOSVD.

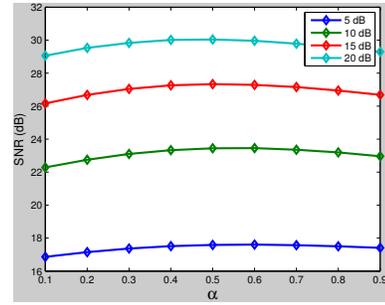


Fig. 1: Mean SNR values computed for MS-HOSVD for varying α values at 5, 10, 15 and 20 dB noise levels.

4.2. Functional Connectivity Network Denoising

The proposed denoising approach MS-HOSVD is applied to functional connectivity networks constructed from EEG data containing the error-related negativity (ERN). The ERN is a brain potential response that occurs following performance errors in a speeded reaction time task usually 25-75 ms after

Table 1: Average SNR for the denoised 3-way tensor $\mathcal{X} \in \mathbb{R}^{401 \times 500 \times 148}$ obtained by HOSVD, L-HOSVD and MS-HOSVD approaches at varying noise levels over 20 trials.

Noise Level	HOSVD	L-HOSVD	MS-HOSVD
5dB	16.49	17.19	17.60
10dB	21.70	22.62	23.45
15dB	25.54	26.41	27.33
20dB	28.47	29.12	30.04

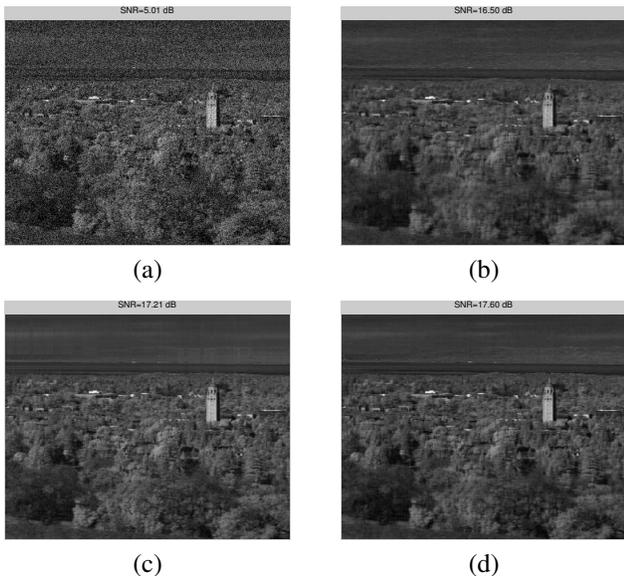


Fig. 2: Sample tensor slices corresponding to 100th band of the (a) 5dB Noisy image, (b) Denoised image by HOSVD (b) Denoised image by L-HOSVD, (d) Denoised image by MS-HOSVD.

the response [17]. Previous work [18] indicates that there is increased coordination between the lateral prefrontal cortex (IPFC) and medial prefrontal cortex (mPFC) within the theta frequency band (4-8 Hz) and ERN time window. EEG data from 63-channels was collected in accordance with the 10/20 system on a Neuroscan Synamps2 system (Neuroscan, Inc.) sampled at 128 Hz from 91 subjects. A speeded-response flanker task was employed, and response-locked averages were computed for each subject. All EEG epochs were converted to current source density (CSD) using published methods [19]. For each subject during error (ERN) response, the pairwise phase locking value in the theta frequency band was computed as described in [20].

In this section, we constructed 4-way tensors $\mathcal{X} \in \mathbb{R}^{63 \times 63 \times 14 \times 91}$ for ERN data where the first and second mode represent the adjacency matrix of the connectivity graphs while the third and fourth mode corresponds to time points and the subjects respectively. We only considered the 25-75 ms time interval which corresponds to the ERN response. We applied MS-HOSVD to denoise the constructed tensor where

clustering parameters c_i s are set to 2 for each mode yielding 16 subtensors and the relaxation parameter α is selected as 0.5 empirically. The noise variance is estimated as 10 percent of the data variance. In order to better interpret the denoising performance of MS-HOSVD, the denoised networks were clustered using a multiple graph clustering algorithm FCCA (see [21] for details). As seen in Fig. 3, denoising the data with MS-HOSVD yields more localized and more stable clusters with each cluster having at least 4 nodes. Moreover, the obtained clusters are in line with previous results indicating that separable clusters are apparent relative to left and right motor areas, and left and right lateral-PFC regions during ERN [21].

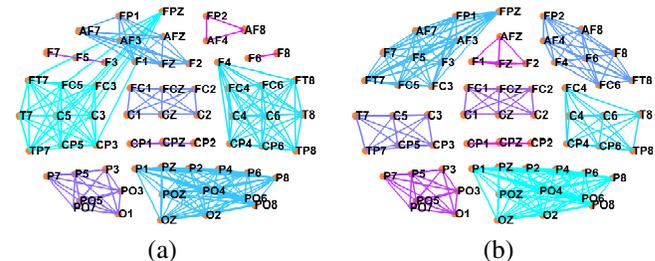


Fig. 3: Identified clusters from functional connectivity networks a) without applying noise removal, b) by applying MS-HOSVD

5. CONCLUSIONS

In this study we introduced a new multiscale tensor denoising technique. The method works by reclustering the tensor data together into subtensors in a way which is expected to increase the similarity of the signal data within each subtensor, thereby improving discrimination between signal and noise. An improvement of the proposed approach then makes use of recursive regularization when denoising subtensors in order to improve performance. The proposed approach is evaluated by applying it to both 3-way and 4-way tensors constructed from a hyperspectral image, and functional connectivity networks, respectively. Future work will consider automatic selection of parameters in the algorithm, e.g., the number of clusters along each mode, and the relaxation parameter α . The proposed method is currently limited to 2-scale analysis but can be easily extended to multiple scales to allow finer scale denoising.

6. REFERENCES

- [1] Guangyi Chen and Shen-En Qian, “Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage,” *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 49, no. 3, pp. 973–980, 2011.

- [2] Damien Muti and Salah Bourennane, "Multidimensional filtering based on a tensor approach," *Signal Processing*, vol. 85, no. 12, pp. 2338–2353, 2005.
- [3] Xinyuan Zhang, Zhongbiao Xu, Nan Jia, Wei Yang, Qianjin Feng, Wufan Chen, and Yanqiu Feng, "Denoising of 3d magnetic resonance images by using higher-order singular value decomposition," *Medical image analysis*, vol. 19, no. 1, pp. 75–86, 2015.
- [4] Behnood Rasti, Johannes R Sveinsson, Magnus O Ulfarsson, and Jon Atli Benediktsson, "Hyperspectral image denoising using 3d wavelets," in *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*. IEEE, 2012, pp. 1349–1352.
- [5] Hisham Othman and Shen-En Qian, "Noise reduction of hyperspectral imagery using hybrid spatial-spectral derivative-domain wavelet shrinkage," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 44, no. 2, pp. 397–408, 2006.
- [6] Qiangqiang Yuan, Liangpei Zhang, and Huanfeng Shen, "Hyperspectral image denoising employing a spectral-spatial adaptive total variation model," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 50, no. 10, pp. 3660–3677, 2012.
- [7] Damien Muti and Salah Bourennane, "Multidimensional signal processing using lower-rank tensor approximation," in *Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). 2003 IEEE International Conference on*. IEEE, 2003, vol. 3, pp. III–457.
- [8] Ajit Rajwade, Anand Rangarajan, and Adrish Banerjee, "Image denoising using the higher order singular value decomposition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, no. 4, pp. 849–862, 2013.
- [9] Yi Peng, Deyu Meng, Zongben Xu, Chenqiang Gao, Yi Yang, and Biao Zhang, "Decomposable nonlocal tensor dictionary learning for multispectral image denoising," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 2949–2956.
- [10] Tamara G Kolda and Brett W Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.
- [11] Lieven De Lathauwer, Bart De Moor, and Joos Vandewalle, "A multilinear singular value decomposition," *SIAM journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253–1278, 2000.
- [12] David L Donoho and Jain M Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.
- [13] Jingyu Yan and Marc Pollefeys, "A general framework for motion segmentation: Independent, articulated, rigid, non-rigid, degenerate and non-degenerate," in *Computer Vision—ECCV 2006*, pp. 94–106. Springer, 2006.
- [14] Weisheng Dong, Guangming Shi, and Xin Li, "Nonlocal image restoration with bilateral variance estimation: a low-rank approach," *Image Processing, IEEE Transactions on*, vol. 22, no. 2, pp. 700–711, 2013.
- [15] Stanley Osher, Martin Burger, Donald Goldfarb, Jinjun Xu, and Wotao Yin, "An iterative regularization method for total variation-based image restoration," *Multiscale Modeling & Simulation*, vol. 4, no. 2, pp. 460–489, 2005.
- [16] Torbjørn Skauli and Joyce Farrell, "A collection of hyperspectral images for imaging systems research," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2013, pp. 86600C–86600C.
- [17] Jason R Hall, Edward M Bernat, and Christopher J Patrick, "Externalizing psychopathology and the error-related negativity," *Psychological Science*, vol. 18, no. 4, pp. 326–333, 2007.
- [18] James F Cavanagh, Michael X Cohen, and John JB Allen, "Prelude to and resolution of an error: Eeg phase synchrony reveals cognitive control dynamics during action monitoring," *The Journal of Neuroscience*, vol. 29, no. 1, pp. 98–105, 2009.
- [19] Craig E Tenke and Jürgen Kayser, "Generator localization by current source density (csd): implications of volume conduction and field closure at intracranial and scalp resolutions," *Clinical neurophysiology*, vol. 123, no. 12, pp. 2328–2345, 2012.
- [20] Selin Aviyente, Edward M Bernat, Westley S Evans, and Scott R Sponheim, "A phase synchrony measure for quantifying dynamic functional integration in the brain," *Human brain mapping*, vol. 32, no. 1, pp. 80–93, 2011.
- [21] Alp Ozdemir, Marcos Bolanos, Edward Bernat, and Selin Aviyente, "Hierarchical spectral consensus clustering for group analysis of functional brain networks," *Biomedical Engineering, IEEE Transactions on*, vol. 62, no. 9, pp. 2158–2169, 2015.