

A Low-rank Tensor Dictionary Learning Method for Multi-spectral Images Denoising

Xiao Gong and Wei Chen
Beijing Jiaotong University

Code is available at <https://www.dropbox.com/s/80ltjrxr2v9maeg/LTDL.zip?dl=0>

December 6, 2018

Abstract

As a 3-order tensor, a multi-spectral image (MSI) has dozens of spectral bands, which can deliver more information for real scenes. However, real MSIs are often corrupted by noises in the sensing process, which will further deteriorate the performance of higher-level classification and recognition tasks. In this paper, we propose a Low-rank Tensor Dictionary Learning (LTDL) method for MSI denoising. Firstly, we extract blocks from the MSI and cluster them into groups. Then instead of using the exactly low-rank model, we consider a nearly low-rank approximation, which is closer to the latent low-rank structure of the clean groups of real MSIs. In addition, we propose to learn an spatial dictionary and an spectral dictionary, which contain the spatial features and spectral features respectively of the whole MSI and are shared among different groups. Hence the LTDL method utilizes both the latent low-rank prior of each group and the correlation of different groups via the shared dictionaries. Experiments on synthetic data validate the effectiveness of dictionary learning by the LTDL. Experiments on real MSIs demonstrate the superior denoising performance of the proposed method in comparison to state-of-the-art methods.

1 Introduction

A multi-spectral image (MSI) has dozens of spectral bands, where the wavelengths may range from infrared to ultra-violet. Compared with a RGB image which only has three spectral bands, an MSI provides more information

which reveals features of the object hidden in the spectral domain. However, in many cases, MSIs suffer from corruptions or noises in the sensing process [2]. As a low level image processing technique, MSI denoising is key to many high-level computer vision tasks, such as segmentation and classification whose performance highly relies on the quality of the data.

As a model driven approach, dictionary learning methods have been used to find the basic atoms which comprise various signals of a training dataset. By using a learned dictionary of some signal ensembles, noises can be effectively removed via solving a sparse signal recovery problem for each patch of an image [13, 11]. For MSI denoising, applying the traditional dictionary learning methods, e.g., K-SVD [1], for each band leads to the poor performance, as it fails to exploit spectral information in MSIs [19].

Tensor dictionary learning, which keeps the multidimensional structure of tensors, has attracted growing interests of researchers to process images in the past years. Based on CANDECOMP/PARAFAC (CP) decomposition, Duan et al. extend the K-SVD method for tensors, where a higher order tensor dictionary is learned and each atom of the dictionary is a rank-one tensor [12]. By using the Tucker model of tensors, Zubair and Wang propose to learn multiple orthogonal dictionaries along different modes of tensors, where the core tensor have sparse non-zeros elements [28]. In [20], Qi et al. divide an MSI into small 3-order tensor blocks, and learn overcomplete dictionaries for each mode of the blocks via a two-phase block-coordinate-relaxation approach that includes sparse

coding and dictionary updating. However, they fails to further employ all the structural information embedded in images.

In [8, 18], non-local similar small patches in space are clustered into groups and processed together, which improves the image denoising performance. In addition, in view of the fact that blocks extracted from MSIs have both the spatial correlation and the spectral correlation, the low-rank model is employed in different dictionary learning methods for MSI denoising. For example, Peng et al. apply dictionary learning for each group (i.e., a tensor) separately and enforce a low-rank Tucker approximation [19]. A simultaneously sparse and low-rank structure is considered for each group of an MSI in [23]. In both two methods, each group of an MSI is processed separately and a learned dictionary only captures information of the group, which fails to exploit inter-group correlations. Learning a shared dictionary among all groups of an MSI would gain from the multitask learning concept and be more effective to capture the atoms that comprises the signal.

In this paper, we propose a Low-rank Tensor Dictionary Learning (LTDL) method for MSI denoising, which differs with existing methods in two aspects:

- Dictionaries of the spatial domain and the spectral domain are trained by all tensor groups in the proposed method, so the features in space and spectrum can be learned from the whole MSI, while the denoising methods in [19, 23] use different dictionaries for distinct groups.
- Instead of enforcing the denoised groups to be exactly low-rank, we consider a more flexible model that each group is decomposed into a low-rank component and a non-low-rank component, as clean MSI groups are nearly low-rank in practice.

A flowchart of the proposed MSI denoising method is shown in Figure 1. Full bands blocks of an MSI are extracted by window and then similar blocks are clustered into groups. Then we obtain 3-order tensor groups by unfolding each blocks at spectral domain. The three modes of a group corresponding to the spatial domain, the spectral domain and blocks. Shared overcomplete dictionaries of all groups in both the spatial mode and the spectral mode are learned via the proposed LTDL method, where

Method	Model	Data
NL-means [5]	Filter	Matrix
K-SVD [1]	Sparsity	Matrix
BM3D [8]	Filter	Matrix
CSR [11]	Sparsity	Matrix
WNNM [14]	Low Rank	Matrix
LRTA [21]	Low Rank	Tensor
NLM3D [7]	Filter	Tensor
PARAFAC [17]	Low Rank	Tensor
BM4D [18]	Filter	Tensor
Tdl [19]	Sparsity and Low Rank	Tensor
TenSR [20]	Sparsity	Tensor
KBReg [23]	Sparsity and Low Rank	Tensor
the proposed LTDL	Sparsity and Low Rank	Tensor

Table 1: A list of some recent denoising methods.

a nearly low-rank structure is enforced for the tensor approximation. An effective algorithm based on the alternating direction method of multipliers (ADMM) [4] is used to solve the optimization problem of the proposed LTDL method for MSI denoising. Experimental results demonstrate that the new method outperforms state-of-the-art methods in MSI denoising.

The rest of the paper is organized as follows: Section 2 reviews existing methods for image denoising based on the filter model, the sparsity model and the low-rank model. In section 3, we present the proposed LTDL method for MSI denoising. Experimental results on synthetic data and MSIs are provided in section 4. At last, conclusions are given in section 5.

2 Related Work

To decompose images with noises, one require the priors of the different signals. The “No Free Lunch” theory in machine learning suggests that all algorithms perform the same for the randomised data and we can achieve good performance only when the data has some structure and some appropriate model is used. In the past decade, various model driven methods have been proposed for image denoising, and some of them are summarised in Table 1.

The first category of image denoising methods [5, 8, 7,

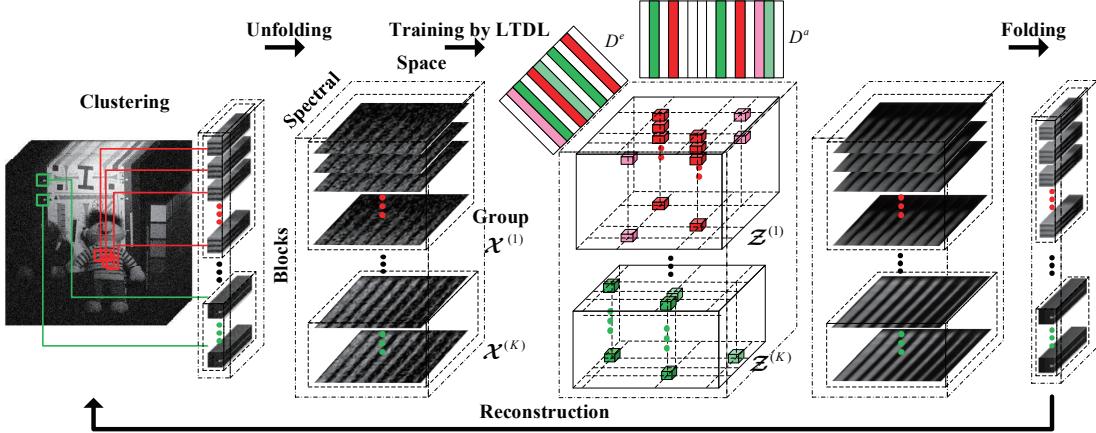


Figure 1: Flowchart of the proposed MSI denoising method.

[18] uses different filters, e.g., the mean-value filter and the Wiener filter, to exploit the local correlation between adjacent pixels and/or the non-local correlation between similar small patches/blocks of an image. Another category of widely used methods considers the property that natural images or clustered groups of their patches usually exhibit low-rank structures. Different low-rank approximation methods have been proposed for image denoising such as nuclear norm regularization [11], Tucker low-rank decomposition [21] and CP low-rank decomposition [17]. Based on the sparsity model, dictionary learning methods assume that image patches/blocks are linear compositions of very few atoms selected from a dictionary. Dictionary learning and sparse representation in [1, 11], is firstly applied for 2D image denoising, and then extended to higher-order image denoising such as TensSR [20]. In recent years, researchers consider to exploit both the sparsity model and low rank model to better utilize the prior of the groups extracted from an MSI [19, 23]. Unfortunately, in these recent methods, dictionaries (or called factors) are learned separately for each tensor group, which deviates from the principle of dictionary learning and significantly increases the total number of dictionary atoms. Our work differs with the existing methods for MSI denoising. Instead of using the exact low-rank model, we consider a nearly low-rank structure for each tensor group of an MSI, and learning dictionaries that are shared among all group-

s.

3 Low-rank tensor dictionary learning for MSI denoising

3.1 Notation

The following notations are used throughout this paper. The order of a tensor is the number of modes. Elements of an N -order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times \dots \times I_n \times \dots \times I_N}$ are denoted by $x_{i_1 \dots i_n \dots i_N}$, where i_n ($1 \leq i_n \leq I_n$) refers to the n th mode index. A mode- n vector of an N -order tensor \mathcal{X} is obtained from \mathcal{X} by varying index in the n th mode while keeping the indices of other modes fixed. The unfolding matrix of a tensor \mathcal{X} at the n th mode is denoted as $X_{(n)} \in \mathbb{R}^{I_n \times (I_1 \dots I_{n-1} I_{n+1} \dots I_N)}$, where the columns are all mode- n vectors of \mathcal{X} . The tensor \mathcal{X} can be obtained by folding the matrix $X_{(n)}$ at the n th mode. The n -mode product of the tensor \mathcal{X} and a matrix $U \in \mathbb{R}^{J \times I_n}$ is a tensor $\mathcal{X} \times_n U \in \mathbb{R}^{I_1 \times \dots \times I_{n-1} \times J \times I_{n+1} \times \dots \times I_N}$, whose elements are computed by $(\mathcal{X} \times_n U)_{i_1 \dots i_{n-1} j i_{n+1} \dots i_N} = \sum_{i_n=1}^{I_n} x_{i_1 \dots i_N} u_{j i_n}$. The inner product of two tensors $\mathcal{X}, \mathcal{Y} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ is the sum of the products of all entries, i.e., $\langle \mathcal{X}, \mathcal{Y} \rangle = \sum_{i_1=1}^{I_1} \dots \sum_{i_N=1}^{I_N} x_{i_1 \dots i_N} y_{i_1 \dots i_N}$. The ℓ_1 norm and the Frobenius norm of a tensor are defined as $\|\mathcal{X}\|_1 = \sum_{i_1=1}^{I_1} \dots \sum_{i_N=1}^{I_N} |x_{i_1 \dots i_N}|$ and

$\|\mathcal{X}\|_F = (\sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} x_{i_1 \dots i_N}^2)^{1/2}$, respectively. The nuclear norm of a matrix is denoted by $\|X\|_*$, which is the sum of all singular values of X . The symbol of \otimes denotes the Kronecker product of matrices.

3.2 The observation of the nearly low-rank tensor structure in real MSI groups

For an MSI $\mathcal{H} \in \mathbb{R}^{L \times W \times H}$ with $L \times W$ spatial size and H spectral bands, we extract S overlapping full-band blocks $\mathcal{H}^i \in \mathbb{R}^{d_L \times d_W \times H}$ by using a sampling window that traverses the whole MSI with step lengths p_L and p_W in the two spacial coordinates. Each block \mathcal{H}^i is unfolded in the spectral mode to be a matrix $H_{(3)}^i \in \mathbb{R}^{H \times d_L d_W}$, which has a spatial mode and a spectral mode. To exploit the non-local self-similarity of images, blocks of the MSI can be clustered into K groups, where each group forms a tensor (called a tensor group). The k th tensor group $\mathcal{X}^{(k)} \in \mathbb{R}^{d_L d_W \times H \times s^{(k)}}$ consists of $s^{(k)}$ similar blocks of the MSI and $\sum_{k=1}^K s^{(k)} = S$. Tensor groups are assumed with a low-rank structure in a variety of works [19, 24, 23], owing to the correlations across the spatial domain, the spectral domain and similar blocks.

To enforce the low-rank structure of the recovered signal from a noisy tensor group $\mathcal{X}^{(k)}$ of an MSI, there are at least two approaches. One approach is using the low-rank regularization. For example, by employing the convex tensor nuclear norm [16], the tensor group denoising problem is casted into a convex optimization problem, which is given as

$$\min_{\hat{\mathcal{X}}^{(k)}} \frac{1}{2} \left\| \mathcal{X}^{(k)} - \hat{\mathcal{X}}^{(k)} \right\|_F^2 + \sum_{i=1}^3 \beta_i \left\| \hat{\mathcal{X}}_{(i)}^{(k)} \right\|_*, \quad (1)$$

where β_i ($i = 1, 2, 3$) are positive scalars that controls the low-rank penalties for different modes. The optimization problem in (1) can be efficiently solved by the singular value thresholding [6]. Another approach aims to find the rank- $(R_1^{(k)}, R_2^{(k)}, R_3^{(k)})$ approximation of the noisy tensor group $\mathcal{X}^{(k)}$ in an optimal least squares sense with $R_1^{(k)} < d_L d_W$, $R_2^{(k)} < H$ and $R_3^{(k)} < s^{(k)}$ in [19], which is given by

$$\min_{\mathcal{G}^{(k)}, \{U_i^{(k)}\}} \left\| \mathcal{X}^{(k)} - \mathcal{G}^{(k)} \times_1 U_1^{(k)} \times_2 U_2^{(k)} \times_3 U_3^{(k)} \right\|_F^2, \quad (2)$$

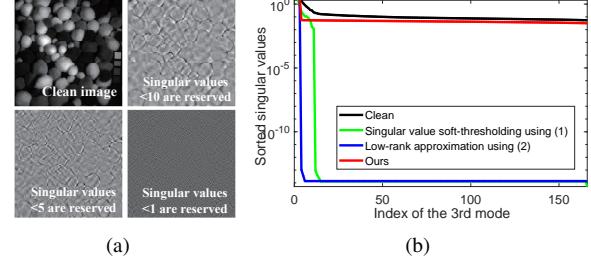


Figure 2: Investigation on the distribution of singular values of real MSIs and the effectiveness of different methods. (a) The clean “Pompoms” MSI and different “Pompoms” MSIs with large singular values removed. (b) The curves for the sorted singular values of a tensor group in the 3rd mode of different low-rank based methods.

where $\mathcal{G}^{(k)} \in \mathbb{R}_1^{R_1^{(k)}} \times \mathbb{R}_2^{R_2^{(k)}} \times \mathbb{R}_3^{R_3^{(k)}}$ is a smaller core tensor, and $U_i^{(k)}$ ($i = 1, 2, 3$) are factors. The denoised tensor group is $\hat{\mathcal{X}} = \mathcal{G}^{(k)} \times_1 U_1^{(k)} \times_2 U_2^{(k)} \times_3 U_3^{(k)}$. The optimization problem in (2) is nonconvex and different algorithms are proposed in literature such as higher-order singular value decomposition (HOSVD) [9] and higher-order orthogonal iteration (HOOI) [10].

However, we observe that usually tensor groups of clean real MSIs are not exactly low-rank. As shown in Fig. 2(b), the curve (in black) for the sorted singular values of a tensor group of the clean “pompoms” MSI in the 3rd mode has a very long tail, i.e., many singular values are close to 0 rather than exactly 0. As shown in 2(a), these small singular values contain the texture information of the MSI. Therefore, enforcing an exactly low-rank structure of a tensor group would lead to the lost of important information of the MSI. To avoid this drawback, we consider a nearly low-rank structure for each tensor group and pose the following optimization problem for denoising:

$$\begin{aligned} \min_{\hat{\mathcal{X}}^{(k)}, \mathcal{G}^{(k)}, \{U_i^{(k)}\}} & \left\| \mathcal{X}^{(k)} - \hat{\mathcal{X}}^{(k)} \right\|_F^2 \\ & + \lambda_r \left\| \hat{\mathcal{X}}^{(k)} - \mathcal{G}^{(k)} \times_1 U_1^{(k)} \times_2 U_2^{(k)} \times_3 U_3^{(k)} \right\|_F^2, \end{aligned} \quad (3)$$

where $\hat{\mathcal{X}}^{(k)}$ is the k th recovered tensor group, and λ_r is the weight to balance fidelity and the low-rank structure. Note that the small singular values of $\hat{\mathcal{X}}^{(k)}$ results

from $\hat{\mathcal{X}}^{(k)} = \mathcal{G}^{(k)} \times_1 U_1^{(k)} \times_2 U_2^{(k)} \times_3 U_3^{(k)}$, and are penalized via the Frobenius norm. The optimization problem in (3) can be solved by updating the variables in $\hat{\mathcal{X}}^{(k)}$ and $\mathcal{T}^{(k)} = \mathcal{G}^{(k)} \times_1 U_1^{(k)} \times_2 U_2^{(k)} \times_3 U_3^{(k)}$ alternately. As shown in Fig. 2(b), with appropriate choose of the weight λ_r , the curve for the sorted singular values of the proposed method is close to the curve of the clean MSI tensor group, while the methods corresponding to (1) and (2) lose information embedded in the small singular values.

3.3 The proposed LTDL method for MSI denoising

In this subsection, we introduce the proposed MSI denoising method that learns shared tensor overcomplete dictionaries and considers the nearly low-rank structure. Two dictionaries, i.e., D^a and D^e , are learned from all tensor groups of an MSI, where D^a corresponds to the spatial domain and D^e corresponds to the spectral domain. We define τ_a and τ_e as the redundancy ratios (i.e., the ratio of the number of columns to the number of rows, $\tau_a \geq 1$ and $\tau_e \geq 1$) corresponding to the spatial dictionary and the spectral dictionary, respectively. The proposed LTDL method for MSI denoising is formulated as follows:

$$\begin{aligned} & \min_{\substack{D^a, D^e \\ \{\mathcal{Z}^{(k)}, \mathcal{C}^{(k)}, \\ U_i^{(k)}\}}} \sum_{k=1}^K \left(\left\| \mathcal{X}^{(k)} - \mathcal{Z}^{(k)} \times_1 D^a \times_2 D^e \right\|_F^2 + \lambda_s \left\| \mathcal{Z}^{(k)} \right\|_1 \right. \\ & \quad \left. + \lambda_r \left\| \mathcal{Z}^{(k)} \times_1 D^a \times_2 D^e - \mathcal{G}^{(k)} \times_1 U_1^{(k)} \times_2 U_2^{(k)} \times_3 U_3^{(k)} \right\|_F^2 \right) \\ \text{s.t. } & \|D^a(:, r)\|_2^2 = 1 \text{ for } r = 1, \dots, \tau_a d_L d_W \\ & \|D^e(:, r)\|_2^2 = 1 \text{ for } r = 1, \dots, \tau_e H, \end{aligned} \quad (4)$$

where $\mathcal{Z}^{(k)} \in \mathbb{R}^{\tau_a d_L d_W \times \tau_e H \times s^{(k)}}$ is the representation of k th tensor group, λ_s and λ_r are the weights corresponding to regularization for sparsity and low-rank, respectively. For the last term of the object in (4), $\mathcal{T}^{(k)} = \mathcal{G}^{(k)} \times_1 U_1^{(k)} \times_2 U_2^{(k)} \times_3 U_3^{(k)}$ denotes the rank- $(R_1^{(k)}, R_2^{(k)}, R_3^{(k)})$ approximation of the reconstructed tensor group $\mathcal{Z}^{(k)} \times_1 D^a \times_2 D^e$.

3.4 Algorithm development

The proposed optimization problem in (4) is nonconvex. Here, we employ the alternating direction method of multipliers(ADMM) [4], which leads to solve several subproblems.

We first introduce K auxiliary tensors $\mathcal{C}^{(k)} (k = 1, \dots, K)$ and the optimization problem in (4) is equivalent to

$$\begin{aligned} & \min_{\substack{D^a, D^e \\ \{\mathcal{Z}^{(k)}, \mathcal{C}^{(k)}, \\ \mathcal{T}^{(k)}\}}} \sum_{k=1}^K \left(\left\| \mathcal{X}^{(k)} - \mathcal{Z}^{(k)} \times_1 D^a \times_2 D^e \right\|_F^2 + \lambda_s \left\| \mathcal{C}^{(k)} \right\|_1 \right. \\ & \quad \left. + \lambda_r \left\| \mathcal{Z}^{(k)} \times_1 D^a \times_2 D^e - \mathcal{T}^{(k)} \right\|_F^2 \right) \\ \text{s.t. } & \|D^a(:, r)\|_2^2 = 1 \text{ for } r = 1, \dots, \tau_a d_L d_W \\ & \|D^e(:, r)\|_2^2 = 1 \text{ for } r = 1, \dots, \tau_e H \\ & \mathcal{C}^{(k)} = \mathcal{Z}^{(k)} \text{ for } k = 1, \dots, K, \end{aligned} \quad (5)$$

The augmented Lagrangian function for the above problem can be given as:

$$\begin{aligned} & \mathcal{L}_\rho(D^a, D^e, \{\mathcal{Z}^{(k)}, \mathcal{C}^{(k)}, \mathcal{T}^{(k)}, \mathcal{Y}^{(k)}\}) \\ & = \sum_{k=1}^K \left(\left\| \mathcal{X}^{(k)} - \mathcal{Z}^{(k)} \times_1 D^a \times_2 D^e \right\|_F^2 \right. \\ & \quad \left. + \lambda_s \left\| \mathcal{C}^{(k)} \right\|_1 + \langle \mathcal{C}^{(k)} - \mathcal{Z}^{(k)}, \mathcal{Y}^{(k)} \rangle + \frac{\rho}{2} \left\| \mathcal{C}^{(k)} - \mathcal{Z}^{(k)} \right\|_F^2 \right. \\ & \quad \left. + \lambda_r \left\| \mathcal{Z}^{(k)} \times_1 D^a \times_2 D^e - \mathcal{T}^{(k)} \right\|_F^2 \right), \end{aligned} \quad (6)$$

where $\mathcal{Y}^{(k)} (k = 1, \dots, K)$ are the Lagrange multipliers, ρ is a positive scalar, and columns of the dictionaries D^a and D^e are constrained by a unit power.

Now we introduce the strategy for solving (5) based on the ADMM. To minimize the augmented Lagrange function (6), we fix the dictionaries D^a and D^e , and all multipliers $\mathcal{Y}^{(k)}$. Then (6) can be split into K separate opti-

mization problems with objects as:

$$\begin{aligned} \mathcal{L}_\rho^{(k)}(\mathcal{Z}^{(k)}, \mathcal{C}^{(k)}, \mathcal{T}^{(k)}) &= \left\| \mathcal{X}^{(k)} - \mathcal{Z}^{(k)} \times_1 D^a \times_2 D^e \right\|_F^2 \\ &+ \lambda_s \left\| \mathcal{C}^{(k)} \right\|_1 + \langle \mathcal{C}^{(k)} - \mathcal{Z}^{(k)}, \mathcal{Y}^{(k)} \rangle + \frac{\rho}{2} \left\| \mathcal{C}^{(k)} - \mathcal{Z}^{(k)} \right\|_F^2 \\ &+ \lambda_r \left\| \mathcal{Z}^{(k)} \times_1 D^a \times_2 D^e - \mathcal{T}^{(k)} \right\|_F^2. \end{aligned} \quad (7)$$

$\mathcal{T}^{(k)}$, $\mathcal{Z}^{(k)}$ and $\mathcal{C}^{(k)}$ can be updated alternatively by solving the optimization problem with all the other variables fixed. In specific, to update $\mathcal{T}^{(k)}$, one needs to solve the tensor low multilinear rank approximation problem given by

$$\min_{\mathcal{T}^{(k)}} \left\| \mathcal{Z}^{(k)} \times_1 D^a \times_2 D^e - \mathcal{T}^{(k)} \right\|_F^2, \quad (8)$$

which can be handled by HOOI [10] and the solution is denoted as

$$\mathcal{T}^{(k)} = \text{HOOI} \left(\mathcal{Z}^{(k)} \times_1 D^a \times_2 D^e, R_1^{(k)}, R_2^{(k)}, R_3^{(k)} \right). \quad (9)$$

To update $\mathcal{Z}^{(k)}$, with irrelevant terms removed, the problem of $\min_{\mathcal{Z}^{(k)}} \mathcal{L}_\rho^{(k)}$ becomes:

$$\begin{aligned} \min_{\mathcal{Z}^{(k)}} \left\| \mathcal{X}^{(k)} - \mathcal{Z}^{(k)} \times_1 D^a \times_2 D^e \right\|_F^2 + \langle \mathcal{C}^{(k)} - \mathcal{Z}^{(k)}, \mathcal{Y}^{(k)} \rangle \\ + \frac{\rho}{2} \left\| \mathcal{C}^{(k)} - \mathcal{Z}^{(k)} \right\|_F^2 + \lambda_r \left\| \mathcal{Z}^{(k)} \times_1 D^a \times_2 D^e - \mathcal{T}^{(k)} \right\|_F^2, \end{aligned} \quad (10)$$

which has a closed-form solution

$$\begin{aligned} Z_{(3)}^{(k)} &= \left(\left(2X_{(3)}^{(k)} + 2\lambda_r T_{(3)}^{(k)} \right) D + \rho C_{(3)}^{(k)} + Y_{(3)}^{(k)} \right) \\ &\quad \left((2 + 2\lambda_r) D^T D + \rho I \right)^{-1}, \end{aligned} \quad (11)$$

where $D = D^e \otimes D^a$ and I is an identity matrix. The tensor $\mathcal{Z}^{(k)}$ can be obtained by folding $Z_{(3)}^{(k)}$ at the 3rd mode. Updating $\mathcal{C}^{(k)}$ requires to solve the following optimization problem:

$$\min_{\mathcal{C}^{(k)}} \lambda_s \left\| \mathcal{C}^{(k)} \right\|_1 + \frac{\rho}{2} \left\| \mathcal{C}^{(k)} - \left(\mathcal{Z}^{(k)} - \frac{1}{\rho} \mathcal{Y}^{(k)} \right) \right\|_F^2, \quad (12)$$

which leads to the solution

$$\mathcal{C}^{(k)} = \text{soft}_{\frac{\lambda_s}{\rho}}(\mathcal{M}^{(k)}), \quad (13)$$

where $\text{soft}(\cdot)$ denotes the soft-thresholding operator and $\mathcal{M}^{(k)} = \mathcal{Z}^{(k)} - \mathcal{Y}^{(k)}/\rho$. The elements of $\mathcal{C}^{(k)}$ is $c_{ijk}^{(k)} = \text{sign}(m_{ijk}^{(k)}) \max\{0, |m_{ijk}^{(k)}| - \frac{\lambda_s}{\rho}\}$, where $\text{sign}(\cdot)$ denotes the sign function. Note that variables for different tensor groups ($k = 1, \dots, K$) can be updated in parallel.

Next we consider to update the dictionaries D^a and D^e , which are shared by all groups. By letting $\mathcal{O}^{(k)} = \frac{\mathcal{X}^{(k)} + \lambda_r \mathcal{T}^{(k)}}{1 + \lambda_r}$, the optimization problem turns to be

$$\begin{aligned} \min_{D^a, D^e} & \left\| \mathcal{O} - \mathcal{Z} \times_1 D^a \times_2 D^e \right\|_F^2, \\ \text{s.t. } & \|D^a(:, r)\|_2^2 = 1 \text{ for } r = 1, \dots, \tau_a d_L d_W \\ & \|D^e(:, r)\|_2^2 = 1 \text{ for } r = 1, \dots, \tau_e H \end{aligned} \quad (14)$$

where $\mathcal{O} = (\mathcal{O}^{(1)}, \dots, \mathcal{O}^{(K)})$ and $\mathcal{Z} = (\mathcal{Z}^{(1)}, \dots, \mathcal{Z}^{(K)})$ are obtained by stacking all groups of $\mathcal{O}^{(k)}$ and $\mathcal{Z}^{(k)}$ at the 3rd mode, respectively. Define $\mathcal{A} = \mathcal{Z} \times_2 D^e$. Then D^a can be updated by

$$\begin{aligned} \min_{D^a} & \left\| O_{(1)} - D^a A_{(1)} \right\|_F^2 \\ \text{s.t. } & \|D^a(:, r)\|_2^2 = 1 \text{ for } r = 1, \dots, \tau_a d_L d_W, \end{aligned} \quad (15)$$

where $O_{(1)} \in \mathbb{R}^{d_L d_W \times HS}$ and $A_{(1)} \in \mathbb{R}^{\tau_a d_L d_W \times HS}$. The optimization problem in (15) is a quadratically constrained quadratic programming problem and can be solved using a Lagrange dual [15]. Thus, the spatial dictionary D^a can be updated by

$$D^a = \left(O_{(1)} A_{(1)}^T \right) \left(A_{(1)} A_{(1)}^T + \Gamma \right)^{-1}, \quad (16)$$

where $\Gamma = \text{diag}([\gamma_1, \dots, \gamma_{\tau_a d_L d_W}])$ and γ_r ($r = 1, \dots, \tau_a d_L d_W$) are dual variables whose values are obtained by solving the dual problem. Similarly, letting $\mathcal{E} = \mathcal{Z} \times_1 D^e$, D^e can be obtained by solving the following problem:

$$\begin{aligned} \min_{D^e} & \left\| O_{(2)} - D^e E_{(2)} \right\|_F^2 \\ \text{s.t. } & \|D^e(:, r)\|_2^2 = 1 \text{ for } r = 1, \dots, \tau_e H. \end{aligned} \quad (17)$$

Then the spectral dictionary D^e is updated by

$$D^e = \left(O_{(2)} E_{(2)}^T \right) \left(E_{(2)} E_{(2)}^T + \Delta \right)^{-1}, \quad (18)$$

where $\Delta = \text{diag}([\delta_1, \dots, \delta_{\tau_e H}])$ and δ_r ($r = 1, \dots, \tau_e H$) are optimized dual variables of (17).

For the last step, we update the K Lagrange multipliers in turn by

$$\mathcal{Y}^{(k)} := \mathcal{Y}^{(k)} + \rho \left(\mathcal{C}^{(k)} - \mathcal{Z}^{(k)} \right). \quad (19)$$

The proposed LTDL method for MSI denoising is summarized in Algorithm 1.

Algorithm 1 Algorithm for MSI denoising

Input: the MSI of $\mathcal{H} \in \mathbb{R}^{L \times W \times H}$

Output: denoised MSI \mathcal{H}_{de} , spatial dictionary D^a and spectral dictionary D^e

- 1: Construct groups $\mathcal{X}^{(k)}$ ($k = 1, \dots, K$) by the extracting, clustering and unfolding process of the \mathcal{H}
 - 2: Initialize $D^a, D^e, \{\mathcal{C}^{(k)}, \mathcal{Z}^{(k)}, \mathcal{Y}^{(k)}\}_{k=1}^K$
 - 3: **while** not converged **do**
 - 4: **for** $k = 1 : K$ **do**
 - 5: update $\mathcal{T}^{(k)}$ by (9)
 - 6: update $\mathcal{Z}^{(k)}$ via folding $Z_{(3)}^{(k)}$ by (11)
 - 7: update $\mathcal{C}^{(k)}$ by (13)
 - 8: **end for**
 - 9: update D^a by (16)
 - 10: update D^e by (18)
 - 11: **for** $k = 1 : K$ **do**
 - 12: update $\mathcal{Y}^{(k)}$ by (19)
 - 13: **end for**
 - 14: $\rho := \mu\rho$
 - 15: **end while**
 - 16: Reconstruct groups $\hat{\mathcal{X}}^{(k)} = \mathcal{Z}^{(k)} \times_1 D^a \times_2 D^e$ ($k = 1, \dots, K$)
 - 17: Aggregate $\hat{\mathcal{X}}^{(k)}$ to form the denoised MSI \mathcal{H}_{de}
-

In the following theorem, we provide a weak convergence condition for the proposed algorithm, where we assume the dictionaries are fixed. By alternatively updating the dictionaries and the other variables, it can be guaranteed that the value of the cost function in (4) would not increase. In practice, to speed up the algorithm, one can update the dictionaries in each iteration without waiting for the other variables achieving convergence. Although we cannot provide formal convergence guarantees in this case, we do not have convergence problems in our experiments.

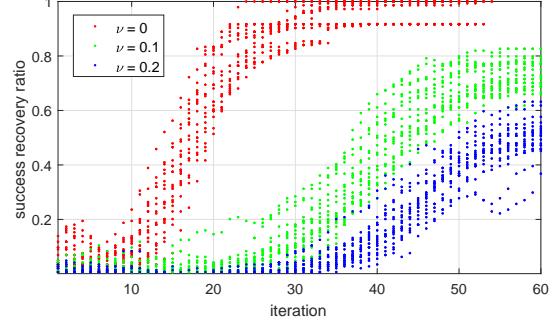


Figure 3: Success recovery ratio of the ground truth dictionary.

Theorem 1. When D^a and D^e are fixed, for l th iteration, the sequences $\{\mathcal{Z}_l^{(k)}\}$, $\{\mathcal{C}_l^{(k)}\}$ and $\{\hat{\mathcal{X}}_l^{(k)}\}$ ($k = 1, \dots, K$) satisfy:

$$\begin{aligned} \|\mathcal{Z}_l^{(k)} - \mathcal{C}_l^{(k)}\|_F &\leq O\left(\frac{1}{\mu^{l-1}\rho_0}\right) \\ \|\mathcal{Z}_{l+1}^{(k)} - \mathcal{Z}_l^{(k)}\|_F &\leq O\left(\frac{1}{\mu^{l-1}\rho_0}\right) \\ \|\hat{\mathcal{X}}_{l+1}^{(k)} - \hat{\mathcal{X}}_l^{(k)}\|_F &\leq O\left(\frac{1}{\mu^{l-1}\rho_0}\right). \end{aligned} \quad (20)$$

where ρ_0 is initial value. Proofs are provided in the supplementary document.

4 Experimental results

In this section, we evaluate the effectiveness of the proposed LTDL using both synthetic data and real MSIs.

4.1 Dictionary Learning Performance with Synthetic Data

We first evaluate the dictionary learning performance of the proposed method on synthetic data. Both the spatial dictionary $D^a \in \mathbb{R}^{10 \times 12}$ and the spectral dictionary $D^e \in \mathbb{R}^{10 \times 12}$ are generated randomly with normalized columns. For each tensor group, its tensor representation $\mathcal{Z}^{(k)} \in \mathbb{R}^{12 \times 12 \times 12}$ is generated randomly with the sparsity level 6, i.e., no more than 6 dictionary atoms are used to

constitute the tensor group. The nonzero components in $\mathcal{Z}^{(k)}$ can be rewritten as a 12×6 matrix, which is generated as the production of two random matrices $G_k \in \mathbb{R}^{12 \times 2}$ and $F_k \in \mathbb{R}^{2 \times 5}$. Therefore, all generated tensor groups have ranks no higher than $(6, 6, 2)$. We generate 200 different tensor groups in this way, and then add Gaussian noises with standard deviation ν . The learned dictionaries are compared against all atoms of the generated dictionaries and we find the most close pair via $1 - |d_i^T \hat{d}_i|$, where d_i is a generated dictionary atom and \hat{d}_i is a recovered dictionary atom¹. The learning is seemed as success if the distance is less than 0.1. The convergence behaviour of the proposed algorithm, i.e., success recovery ratio of the ground truth dictionary versus the number of iterations is shown in Figure 3, where we provide the results of 20 trials for each noise level. With the decrease of the noise variance, the performance of success recovery ratio is improved.

4.2 Denoising Performance for Real MSIs

MSI datasets: Two real MSI datasets, i.e., the Columbia dataset [25]² and the urban area HYDICE MSI³, are used to evaluate the denoising performance of the proposed algorithm. The Columbia dataset includes 32 scenes that are separated into 5 sections. Each MSI has the size of 512×512 in space and includes full spectral resolution reflectance data from $400nm$ to $700nm$ at $10nm$ steps, which leads to 31 bands. These MSIs are adopted as clean data, and we generate additive noises to evaluate the denoising performance of the proposed algorithm. For the natural urban area HYDICE MSI, we select the bands from 178 to 207, which are severely damaged (i.e., the noise is not generated).

Experimental settings: In the proposed LTDL method, the spatial window size is set as $d_L = d_W = 7$ and the step size is set as $p_L = p_W = 3$. Groups of an MSI are clustered by using the k-means++ [3], and each cluster forms a tensor group. We set $\rho = 0.01$ and $\mu = 1.3$ as the parameters in the ADMM of our algorithm. The redundancy ratio of dictionaries are set as $\tau_a = \tau_e = 1.5$. We add Gaussian noise with the mean

¹We consider the equivalent dictionary $D^e \otimes D^a$ in the experiments.

²<http://www.cs.columbia.edu/CAVE/databases/multispectral/>

³<https://erdc-library.erdc.dren.mil/xmlui/handle/11681/2925>

0 and the standard deviation ν on the whole MSIs in the Columbia dataset. We set the sparsity weight $\lambda_s = 0.1\nu$ and the low-rank weight $\lambda_r = 500\nu$. The rank parameters $\{R_1^{(k)}, R_2^{(k)}, R_3^{(k)}\}^K$ in (9) are estimated by using the tensor rank estimation method proposed in [26]. For unknown noises, we suggest that the $(R_1^{(k)}, R_2^{(k)}, R_3^{(k)})$ is smaller than the size of k th tensor group.

The proposed algorithm is compared with both denoising methods for 2D images including K-SVD [1] and BM3D [8] and denoising methods for 3D images including NLM3D [7], LRTA [21], PARAFAC [17], BM4D [18], Tdl [19], and KBRreg [23]. For denoising methods for 2D images, each MSI is processed band by band as multiple 2D images.

MSI denoising with generated noises: We generate Gaussian noises for MSIs in the Columbia dataset, and employ four different performance indicators including peak signal-to-noise ratio (PSNR), structural similarity (SSIM), spectral angle mapper (SAM) [27] and dimensionless global relative error of synthesis (ERGAS) [22], which are widely used to evaluate recovery quality for MSIs. Recovered MSIs with higher PSNR and SSIM or lower SAM and ERGAS are considered with better quality. PSNR and SSIM are classical spatial-based quality indices, while ERGAS and SAM are spectral-based quality indices. We report the averaged denoising performance and the standard deviation of all the compared methods in Table 2. It can be observed that the proposed method outperforms all the competing methods under all of the four different quality indices.

To visualize the denoising performance, we display the denoised 590nm band image of the “flowers” MSI in Figure 4. By zooming in the petal part of the image, it can be observed that the proposed LTDL retains the gynoecium details of the flower, while most other methods blur the texture.

Real MSI denoising: In reality, an MSI may suffer from different kind of noises including not only the Gaussian noise but also the non-Gaussian noise, such as the stripe noise. Now we investigate the effectiveness of the proposed LTDL for recovering a real corrupted MSI. The denoising results are compared in Figure 5, where the 197th band of the HYDICE MSI has light noise and the 207th band has severe noise. The results of the LRTA, the PARAFAC and the Tdl are much poor than the others,

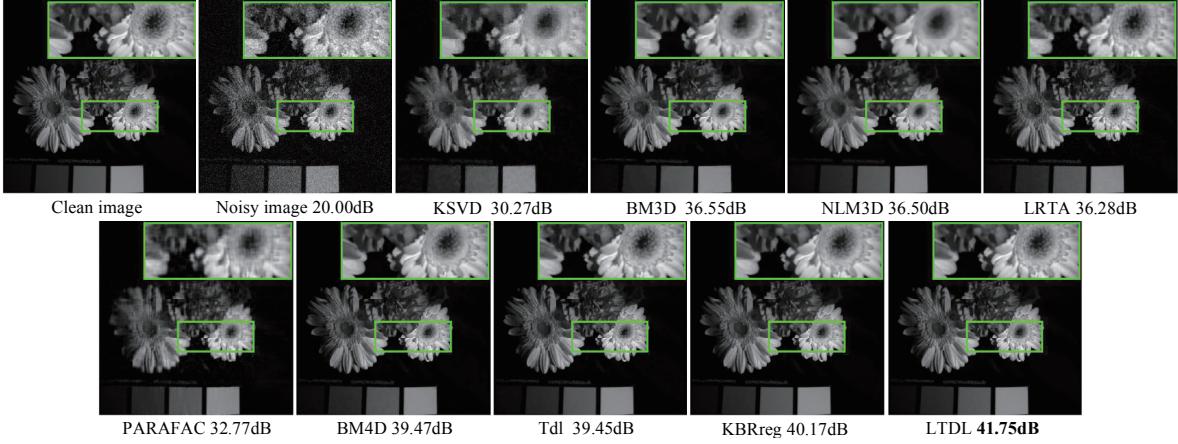


Figure 4: Denoising results of the 590nm band image of the “flowers” MSI ($\nu = 0.1$).

which are not provided here owing to the space limitation. The BM4D is more appropriate for MSIs than the BM3D, and thus we also omit the results of BM3D to save the space. We zoom in a part of the restored image which involves stripe noise. It can be observed that although most methods achieve similar performance in the 197th band image that has light noise, the proposed LTDL is able to restore the 207th band which suffers from more complex and severe noise.

The learned dictionaries of the proposed method are shown in Fig. 6. Atoms in the spatial dictionary D^a represent the spatial features of the HYDICE MSI. To enhance visualization, we reorganize each column into a patch of the size 7×7 . Atoms of the learned spectral dictionary D^e correspond to various spectral features of different bands.

5 Conclusion

This paper presents an effective tensor dictionary learning method for restore high dimensional MSIs. The proposed LTDL method exploits the nearly low-rank structure in a group of similar blocks in the natural MSI and also exploits shared dictionaries among different groups, which makes the proposed method distinct to existing methods. Experimental results show the superior performance of the proposed method for denoising MSIs with both simulated corruptions and real corruptions.

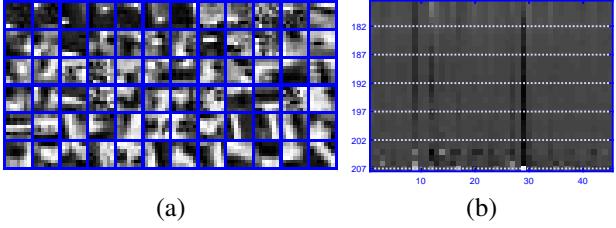


Figure 6: Visualization of the learned dictionaries of the HYDICE MSI. (a) The learned spatial dictionary D^a ; (b) The learned spectral dictionary D^e .

References

- [1] M. Aharon, M. Elad, A. Bruckstein, et al. K-svd: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Signal Processing*, 54(11):4311, 2006. [1](#), [2](#), [3](#), [8](#)
- [2] B. Aiazzi, L. Alparone, A. Barducci, S. Baronti, and I. Pipipi. Information-theoretic assessment of sampled hyperspectral imagers. *IEEE Trans. Geoscience and Remote Sensing*, 39(7):1447–1458, 2001. [1](#)
- [3] D. Arthur and S. Vassilvitskii. K-means++: the advantages of careful seeding. pages 1027–1035. Society for Industrial and Applied Mathematics, 2007. [8](#)
- [4] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, et al. Distributed optimization and statistical learning via the al-

Table 2: MSI denoising performance on the Columbia dataset.

Method	$\nu = 0.1$				$\nu = 0.2$			
	PSNR	SSIM	SAM	ERGAS	PSNR	SSIM	SAM	ERGAS
Noisy image	20.00±0.00	0.14±0.07	0.94±0.24	552.89±159.26	13.98±0.00	0.05±0.03	1.13±0.20	1105.74±318.50
K-SVD	30.07±1.30	0.61±0.04	0.53±0.22	170.02±38.26	27.30±1.44	0.46±0.05	0.61±0.23	233.87±51.99
BM3D	36.92±2.88	0.92±0.03	0.21±0.09	78.43±23.40	33.39±2.93	0.86±0.06	0.28±0.11	117.97±36.64
NLM3D	36.47±2.93	0.93±0.04	0.24±0.10	85.64±21.54	33.44±2.83	0.87±0.05	0.33±0.15	119.53±30.15
LRTA	36.58±2.78	0.89±0.05	0.22±0.11	82.09±26.02	33.08±2.68	0.82±0.08	0.28±0.12	122.36±37.41
PARAFAC	33.16±4.17	0.84±0.11	0.28±0.14	127.31±65.73	31.10±3.00	0.73±0.08	0.42±0.20	155.92±61.05
BM4D	39.49±2.29	0.94±0.02	0.22±0.12	57.81±12.63	35.53±2.10	0.87±0.03	0.34±0.17	91.37±19.21
Tdl	39.22±2.37	0.94±0.02	0.17±0.10	59.67±14.28	35.12±2.05	0.87±0.03	0.27±0.15	95.33±22.70
KBRreg	40.63±2.15	0.95±0.03	0.24±0.22	51.80±13.47	37.57±2.51	0.92±0.05	0.25±0.24	72.98±20.00
Ours	41.50±2.82	0.97±0.01	0.09±0.03	46.18±12.84	38.13±2.73	0.95±0.02	0.13±0.07	67.65±18.40

- ternating direction method of multipliers. *Foundations and Trends in Machine learning*, 3(1):1–122, 2011. [2](#), [5](#)
- [5] A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In *CVPR*, volume 2, pages 60–65. IEEE, 2005. [2](#), [3](#)
- [6] J.-F. Cai, E. J. Candès, and Z. Shen. A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*, 20(4):1956–1982, 2010. [4](#)
- [7] P. Coupé, P. Yger, S. Prima, P. Hellier, C. Kervrann, and C. Barillot. An optimized blockwise nonlocal means denoising filter for 3-d magnetic resonance images. *IEEE Trans. Medical Imaging*, 27(4):425–441, 2008. [2](#), [3](#), [8](#)
- [8] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Trans. Image Processing*, 16(8):2080–2095, 2007. [2](#), [3](#), [8](#)
- [9] L. De Lathauwer, B. De Moor, and J. Vandewalle. A multilinear singular value decomposition. *SIAM journal on Matrix Analysis and Applications*, 21(4):1253–1278, 2000. [4](#)
- [10] L. De Lathauwer, B. De Moor, and J. Vandewalle. On the best rank-1 and rank-(r 1, r 2,..., rn) approximation of higher-order tensors. *SIAM journal on Matrix Analysis and Applications*, 21(4):1324–1342, 2000. [4](#), [6](#)
- [11] W. Dong, X. Li, L. Zhang, and G. Shi. Sparsity-based image denoising via dictionary learning and structural clustering. In *CVPR*, pages 457–464. IEEE, 2011. [1](#), [2](#), [3](#)
- [12] G. Duan, H. Wang, Z. Liu, J. Deng, and Y.-W. Chen. K-cpd: Learning of overcomplete dictionaries for tensor sparse coding. In *ICPR*, pages 493–496. IEEE, 2012. [1](#)
- [13] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans. Image Processing*, 15(12):3736–3745, 2006. [1](#)
- [14] S. Gu, L. Zhang, W. Zuo, and X. Feng. Weighted nuclear norm minimization with application to image denoising. In *CVPR*, pages 2862–2869, 2014. [2](#)
- [15] H. Lee, A. Battle, R. Raina, and A. Y. Ng. Efficient sparse coding algorithms. In *Advances in neural information processing systems*, pages 801–808, 2007. [6](#)
- [16] J. Liu, P. Musialski, P. Wonka, and J. Ye. Tensor completion for estimating missing values in visual data. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 35(1):208–220, 2013. [4](#)
- [17] X. Liu, S. Bourennane, and C. Fossati. Denoising of hyperspectral images using the parafac model and statistical performance analysis. *IEEE Trans. Geoscience and Remote Sensing*, 50(10):3717–3724, 2012. [2](#), [3](#), [8](#)
- [18] M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi. Nonlocal transform-domain filter for volumetric data denoising and reconstruction. *IEEE Trans. Image Processing*, 22(1):119–133, 2013. [2](#), [3](#), [8](#)
- [19] Y. Peng, D. Meng, Z. Xu, C. Gao, Y. Yang, and B. Zhang. Decomposable nonlocal tensor dictionary learning for multispectral image denoising. In *CVPR*, pages 2949–2956, 2014. [1](#), [2](#), [3](#), [4](#), [8](#)
- [20] N. Qi, Y. Shi, X. Sun, and B. Yin. Tensr: Multi-dimensional tensor sparse representation. In *CVPR*, pages 5916–5925, 2016. [1](#), [2](#), [3](#)
- [21] N. Renard, S. Bourennane, and J. Blanc-Talon. Denoising and dimensionality reduction using multilinear tools for hyperspectral images. *IEEE Geoscience and Remote Sensing Letters*, 5(2):138–142, 2008. [2](#), [3](#), [8](#)
- [22] L. Wald. *Data fusion: definitions and architectures: fusion of images of different spatial resolutions*. Presses des MINES, 2002. [8](#)

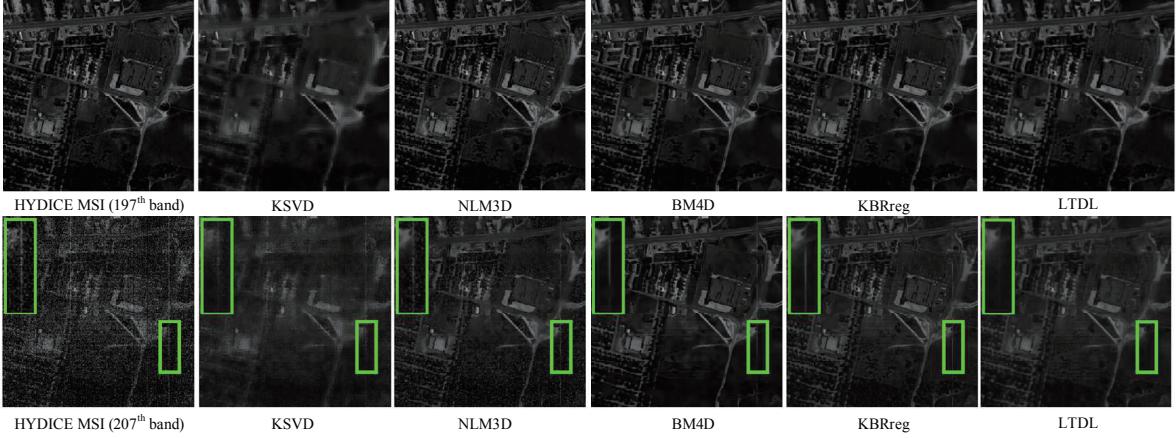


Figure 5: Denoising results for the HYDICE MSI (the 197th band on the top and the 207th band on the bottom).

- [23] Q. Xie, Q. Zhao, D. Meng, and Z. Xu. Kronecker-basis-representation based tensor sparsity and its applications to tensor recovery. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2017. 2, 3, 4, 8
- [24] Q. Xie, Q. Zhao, D. Meng, Z. Xu, S. Gu, W. Zuo, and L. Zhang. Multispectral images denoising by intrinsic tensor sparsity regularization. In *CVPR*, pages 1692–1700, 2016. 4
- [25] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar. Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum. *IEEE Trans. Image Processing*, 19(9):2241–2253, 2010. 8
- [26] T. Yokota, N. Lee, and A. Cichocki. Robust multilinear tensor rank estimation using higher order singular value decomposition and information criteria. *IEEE Trans. Signal Processing*, 65(5):1196–1206, 2017. 8
- [27] R. H. Yuhas, J. W. Boardman, and A. F. Goetz. Determination of semi-arid landscape endmembers and seasonal trends using convex geometry spectral unmixing techniques. 1993. 8
- [28] S. Zubair and W. Wang. Tensor dictionary learning with sparse tucker decomposition. In *DSP*, pages 1–6. IEEE, 2013. 1