Signal Processing 174 (2020) 107645

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Hyperspectral image denoising by total variation-regularized bilinear factorization

Yongyong Chen [a], Jiaxue Li [a], Yicong Zhou [a,*]

Department of Computer and Information Science, University of Macau, Macau 999078, China

ARTICLE INFO

Article history:
Received 8 January 2020
Revised 6 April 2020
Accepted 29 April 2020
Available online 3 May 2020

Keywords:
Hyperspectral image
Denoising
Total variation
Bilinear factorization

ABSTRACT

Hyperspectral image (HSI) denoising is a prevalent research topic in the remote sensing area. In general, HSIs are inevitably impaired by different types of noise during the data acquisition. To fully characterize the underlying structures of clean HSI and remove mixed noises, we introduce a novel HSI denoising method named total variation-regularized bilinear factorization (BFTV) model. Specifically, we first utilize the bilinear factorization term to explore the globally low-rank structure of the clean HSI and suppress a certain degree of Gaussian noise, so as to make BFTV free to the singular value decomposition. Then the l_1-norm is applied to detect and separate the mixed sparse noise including impulse noise, deadlines, and stripes. Besides, the TV regularization is introduced to describe the locally piecewise smoothness property of the clean HSI both in spatial and spectral domains. To solve this optimization problem, we devise an effective algorithm based on the augmented Lagrange multiplier method. Numerical experiments on five different kinds of mixed noise scenarios and one real world data have tested and demonstrated the superior denoising power of the proposed BFTV model compared with three state-of-the-art low-rank-based approaches.

1. Introduction

Hyperspectral images (HSIs) are originally collected by remote sensors for electromagnetic spectrum at the identical locations. Due to the broad spectral information of objects, HSIs have been successfully applied in many different fields such as agriculture, astronomy, military, etc. Among them, HSI denoising [1] has been a prevalent research topic in the remote sensing area. Because in the real-world environment, HSIs are inevitably impaired by different kinds of noise such as Gaussian noise, impulse noise, deadlines, and stripes [2,3] due to various factors, e.g., photon shot effect, transmission error [4].

To date, there have been various typical HSI denoising techniques, such as non-local means algorithm (NL-Means) [5], block-matching and 3D filtering algorithm (BM3D) [6], K-SVD algorithm [7] and so on. To cope with the three dimensional hyperspectral data, these traditional methods are directly processed band by band or even pixel by pixel. However, these traditional denoising methods are not always well-pleasing since both spectral and spatial structures cannot be well preserved simultaneously for the HSI restoration. To better restore the HSI data cube, the underlying prior knowledge of the clean HSI should be taken into account. For example, the clean HSI data cube has the low rank property, because besides the spatial correlations among neighboring pixels in HSIs, there are also existing high correlations among different spectral bands. As a result, each spectral frame can be represented by a linear combination of several other irrelevant frames. In this way, the clean HSI structure can be characterized and the component of mixture sparse noise is separated as well. For example, Zhang et al. proposed a HSI denoising method based on the low-rank matrix recovery (LRMR) model [2]. Considering the superiority of the weighted nuclear norm [8], Wu et al.[9] exploited the weighted nuclear norm instead of the traditional nuclear norm to investigate the high spectral correlation. Flowing this, the study in [10] proposed a low-rank constraint-based HSI denoising methods. To further improve the approximation of the nuclear norm, Xie et al.[11] and Chen et al.[3] developed the weighted schatten p-norm and γ-norm, respectively. Due to the convex nature of the nuclear norm, it has been successfully used as a regularizer to solve the ill-posed HSI denoising problem. However, the nuclear norm-based and these nonconvex approximation-based HSI denoising methods may face one essential limitation. That is, they cannot ignore the high computation cost of the singular value decomposition. This is mainly because they need to perform the singular value decomposition with high complexity in each iteration.

Except for the globally low-rank structure, the locally piecewise smoothness is another important property of the HSIs. The total

* Corresponding author.
E-mail address: yicongzhou@um.edu.mo (Y. Zhou).

https://doi.org/10.1016/j.sigpro.2020.107645
0165-1684/

variation (TV) regularizer [12], as a plausible tool for smoothness preservation, has been widely applied in image denoising by minimizing difference values among the neighboring pixels. For an HSI cube, the smoothness property is not only limited in the spatial domain, but also appeared in the spectral domain since the distribution of difference values among the adjacent bands is very similar [13].

This paper proposed a total variation-regularized blinear factorization-based (BFTV) HSI denoising model which removes the mixed noise from both spectral and spatial perspectives to recover the intrinsic structures of the clean component from noisy HSIs. For the clean HSI, BFTV factorizes the data matrix into two small factor matrices to explore the globally spectral correlations (the globally low-rank property). Unlike the nuclear norm-based HSI denoising methods, the bilinear factorization can update two small scale matrices to measure the low-rank property instead of calculating the singular value decomposition. We solve the proposed BFTV model based on the augmented Lagrange multiplier (ALM). Comprehensive experiments on simulated and real HSIs are conducted and validate the effectiveness of the proposed method over several state-of-the-art low-rank-based approaches. The main contributions are summarized as follows:

- A novel total variation-regularized bilinear factorization model (BFTV) is proposed for HSI denoising. BFTV not only exploits the low-rank characteristic but also the local piecewise smoothness property to better recover HSIs.
- Different from the existing HSI denoising methods which used the nuclear norm with coarse rank approximation, BFTV applies the bilinear matrix factorization to replace the original nuclear norm and make the singular value decomposition free. BFTV utilizes l_1-norm to detect the sparse noise including salt and pepper noise, deadlines, and stripes. To explore the local piecewise smoothness property in the spatial and spectral domains, BFTV integrates the tensor-based TV regularizer with the bilinear factorization as a unified model.
- An efficient algorithm is designed to solve the BFTV model based on the augmented Lagrange method. Extensive experiments on simulated and real-world HSI datasets demonstrate the superiority of the proposed BFTV over several state-of-the-art low-rank-based HSI denoising methods.

The rest of this paper is organized as follows. Section 2 reviews some related works including low-rank modeling and TV regularization. The proposed model and the iterative algorithm are introduced in Section 3. Extensive experiments and results analysis are reported in Section 4. Finally, we conclude this paper in Section 5.

2. Related works

Considering that HSIs have globally low-rank structure and locally piecewise smoothness property, we briefly introduce related works on low-rank modeling and TV regularization.

Low-rank modeling refers to learn a low dimensional matrix from the original high dimensional data. The existing approaches to the low-rank approximation can be basically divided into the rank minimization methods and matrix factorization ones. In the rank minimization methods, the nuclear norm [14] was usually applied to replace the rank function due to its convex advantage [15]. To remove the mixed noises, many HSI denoising methods [3,16,17] incorporated the nuclear norm into l_1-norm to adapt gross corruptions and outlier scenarios. Besides minimizing rank function, matrix factorization is another efficient way to exploit low-rank property. In matrix factorization approaches, the matrix $L \in R^{m \times n}$ is decomposed into two factor matrices $L = UV^T$, where $U \in R^{m \times r}$ and $v \in R^{n \times r}$. Then, we can obtain a low-rank matrix L by restricting r to a relatively small constant based on the mathematical knowledge that $rank(UV^T) \leq \min(rank(U), rank(V))$. To enhance the robustness of traditional factorization methods, the most typical way is to replace least square loss function with more robust loss functions, such as Geman-McClure function [18], which introduces a weighted parameter to measure the contribution of each element. Xu et al. [19] used a low-rank matrix factorization scheme combining with the logdet function to restore the clean HSIs. Following this idea, the study in [20] used the matrix factorization and the bi-nuclear quasi-norm to measure the low rank characteristic of HSI. However, it overlooks the local structures of HSI. Recently, the deep convolutional neural network-based HSI denoising methods [21–23] have been proposed. For example, Chang et al [21]. was the first time to use the fully convolutional neural network to denoise HSI. It incorporated the residual learning, dilated convolution, and multichannel filtering into the network to better model the HSIs. Yuan et al.[22] developed to learn a nonlinear end-to-end mapping between the noisy and clean HSIs with a combined spatialâspectral deep convolutional neural network. Although the aforementioned methods have achieved a great success in removing complex noise, they need a large number of HSI datasets to train the whole networks.

Except for the global low-rank structure, the local piecewise smoothness is another major property for HSIs. TV regularizer [12] was proposed as a powerful smoothness estimation tool for natural image denoising. Recently, many researches have focused on combining both the low-rank property and piecewise smoothness property together to recover clean HSIs. For instance, Golbabaee and Vandergheynst [24] firstly proposed to combine the nuclear norm and TV norm as one integrated convex function and achieved promising performance for HSIs restoration. Considering that the noise intensity in different bands is generally not the same, a spectral spatial adaptive hyperspectral TV denoising method was devised [25]. But this model considered only different Gaussian noise distribution and stripe noise, which is not fully qualified to deal with more complex noisy cases in HSIs. He et al. [17] developed a total variation-regularized low-rank matrix factorization (LRTV) model which integrated the nuclear norm and TV norm to remove the mixed noises. However, it may face two limitations: (1) LRTV may suffer from high computation cost since the nuclear norm-based problem needs inevitably to perform the singular value decomposition, especially for large-scale HSIs; (2) LRTV may not take full use of the spectral structure correlation since the total variation was performed band-by-band. Therefore, removing complex mixed noise is still a research hotspot in the remote sensing area.

3. HSI Denoising via TV-regularized bilinear factorization

Many works have stated that the HSI data are inevitably contaminated by different kinds of noise including sparse noise, strips, deadlines and Gaussian noise [3,16]. To well remove the mixed noise in HSIs, we proposed the total variation-regularized bilinear factorization (BFTV) model. Considering that the HSIs have two essential characteristics, *i.e.*, the globally low-rank structure and locally piecewise smoothness property, we exploited the bilinear factorization strategy [26,27] and the spatial-spectral total variation regularizer to explore the above two characteristics, respectively. Different from most of the existing HSI denoising methods which adopted the nuclear norm to depict the low-rank property, BFTV used the bilinear factorization to factorize a large-scale HSI data matrix into two smaller factor matrices, leading to the singular value decomposition-free nature.

3.1. BFTV for HSI denoising

A noisy HSI $\mathcal{Y}$ can be decomposed as three parts: the clean HSI $\mathcal{L}$, Gaussian noise $\mathcal{G}$, and sparse noise $\mathcal{S}$. The sparse noise usually includes deadlines, stripes, salt and pepper noise. Therefore, the degradation model of an HSI can be formulated as

$$\mathcal{Y} = \mathcal{L} + \mathcal{S} + \mathcal{G}, \tag{1}$$

where all of them are of size $M \times N \times B$. M and N are the spatial size while B is the number of the spectral band. By converting each band matrix of HSIs into a column vector, we constitute a Casorati matrix $\boldsymbol{Y} \in \mathbf{R}^{MN \times B}$. Then the matrix form of Eq. (1) is formulated as:

$$\boldsymbol{Y} = \boldsymbol{L} + \boldsymbol{S} + \boldsymbol{G}. \tag{2}$$

It is an ill-posed problem to recover the clean HSI $\boldsymbol{L}$ from its noisy observation $\boldsymbol{Y}$. To overcome this issue, some priors are imposed. In this paper, we consider two intrinsic priors of HSIs, i.e., the globally low-rank property and locally piecewise smoothness. Most of existing HSI denoising methods that utilized the nuclear norm to estimate the rank constraint of clean HSIs. However, when the spectral images are corrupted heavy Gaussian noise which effects data identically and independently, the rank estimated by the nuclear norm may not be a low value. Besides, the singular value decomposition is executed by the nuclear norm at each iteration, while the computational cost of SVD is high, which makes it impracticable to handle the real HSIs processing. The nuclear norm treated each singular value equally. However, as well known, the larger singular values usually contain the main information of the image while the smaller values may be controlled by noise. Therefore, the nuclear norm may not a close approximation [3,11]. To solve this issue, we exploit the bilinear factorization strategy to measure the low-rankness of clean HSIs and avoid this high computational trouble. Given any matrix $\boldsymbol{L} \in \mathbf{R}^{m \times n}$, the following holds:

$$\|\boldsymbol{L}\|_* = \min_{\boldsymbol{U},\boldsymbol{V},\boldsymbol{L}=\boldsymbol{U}\boldsymbol{V}} \frac{1}{2}(\|\boldsymbol{U}\|_F^2 + \|\boldsymbol{V}\|_F^2), \tag{3}$$

where two small scale matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ are of size $m \times r$ and $r \times n$, respectively. Thus, we propose the following bilinear factorization formulation to recover the clean HSI:

$$\begin{aligned} &\min_{\boldsymbol{L},\boldsymbol{U},\boldsymbol{V},\boldsymbol{S}} \frac{1}{2}(\|\boldsymbol{U}\|_F^2 + \|\boldsymbol{V}\|_F^2) + \lambda\|\boldsymbol{S}\|_1 \\ &s.t.\ \ \boldsymbol{Y} = \boldsymbol{L} + \boldsymbol{S},\ \ \boldsymbol{L} = \boldsymbol{U}\boldsymbol{V}. \end{aligned} \tag{4}$$

Although the bilinear factorization formulation characterizes the low-rankness of the clean HSI structure, it only separates a certain extent of Gaussian noise since the l_1-norm mainly detects mixture sparse noise. Another extreme noisy condition is assumed to be fixed sparse noise. For example, each spectral band is contaminated by the same deadlines, which is located at the identical spatial position as well. The aforementioned bilinear factorization model will treat these deadlines as a part of the clean HSI component and fail to remove the sparse noise. Thus, except for the globally low-rank structures, the locally piecewise smoothness property should be well explored. To preserve the edge information and smoothness details from spatial and spectral domains simultaneously, we introduce the following tensor-based TV regularization [28]:

$$\|\mathcal{L}\|_{SSTV} = \sum_{i,j,k} |\boldsymbol{l}_{i,j,k} - \boldsymbol{l}_{i,j,k-1}| + |\boldsymbol{l}_{i,j,k} - \boldsymbol{l}_{i,j-1,k}| + |\boldsymbol{l}_{i,j,k} - \boldsymbol{l}_{i-1,j,k}|, \tag{5}$$

where $\boldsymbol{l}_{i,j,k}$ denotes the entry of $\mathcal{L}$ located at the position of (i, j, k). This tensor-based TV norm can further remove Gaussian noise and restore image details from both spatial and spectral domains.

By combine the bilinear factorization model (4) and the tensor-based TV regularization (5) together into one unifed model, the proposed BFTV can be finally formulated as follows:

$$\begin{aligned} &\min_{\boldsymbol{L},\boldsymbol{U},\boldsymbol{V},\boldsymbol{S}} \frac{1}{2}(\|\boldsymbol{U}\|_F^2 + \|\boldsymbol{V}\|_F^2) + \lambda_1\|\boldsymbol{S}\|_1 + \lambda_2\|\mathcal{L}\|_{SSTV} \\ &s.t.\ \ \boldsymbol{Y} = \boldsymbol{L} + \boldsymbol{S},\ \ \boldsymbol{L} = \boldsymbol{U}\boldsymbol{V}, \end{aligned} \tag{6}$$

where λ_1 and λ_2 are two non-negative parameters. The first term of Eq. (6) is the bilinear factorization which is to explore the low-rankness of clean HSIs. The second term of Eq. (6) is the l_1 norm to separate the sparse noise including deadlines, stripes and salt and pepper noise. The last term of Eq. (6) is the tensor-based TV, which is used to preserve the locally piecewise smoothness of clean HSIs from the spatial and spectral domains.

3.2. ALM algorithm for BFTV optimization

It is easy to see that the objective function of Eq. (6) is coupled with respect to $\boldsymbol{L}$ since we impose the low-rankness and piecewise smoothness on the clean HSI. Therefore, we solve Eq. (6) by using ALM method [29]. First, we rewrite the constrained optimization Eq. (6) by introducing an auxiliary variable $\mathcal{M}$:

$$\begin{aligned} &\min_{\boldsymbol{L},\boldsymbol{U},\boldsymbol{V},\boldsymbol{S},\mathcal{M}} \frac{1}{2}(\|\boldsymbol{U}\|_F^2 + \|\boldsymbol{V}\|_F^2) + \lambda_1\|\boldsymbol{S}\|_1 + \lambda_2\|\mathcal{M}\|_1 \\ &s.t.\ \ \boldsymbol{Y} = \boldsymbol{L} + \boldsymbol{S},\ \ \boldsymbol{L} = \boldsymbol{U}\boldsymbol{V},\ \ \mathcal{M} = \boldsymbol{D}(\mathcal{L}). \end{aligned} \tag{7}$$

where $\boldsymbol{D}(\cdot) = [\boldsymbol{D}_h(\cdot);\ \boldsymbol{D}_v(\cdot);\ \boldsymbol{D}_t(\cdot)]$ represents a three-dimensional operator, which is utilized to calculate differences between the neighboring pixels along three dimensions respectively. The corresponding augmented Lagrangian function of Eq. (7) is defined as follows:

$$\begin{aligned} \mathcal{F}(\boldsymbol{U},\boldsymbol{V},\boldsymbol{S},\boldsymbol{L},\mathcal{M};\ \Pi_1,\Pi_2,\Pi_3) &= \tfrac{1}{2}(\|\boldsymbol{U}\|_F^2 + \|\boldsymbol{V}\|_F^2) + \lambda_1\|\boldsymbol{S}\|_1 \\ &+\lambda_2\|\mathcal{M}\|_1 + \tfrac{\rho}{2}(\|\boldsymbol{Y} - \boldsymbol{L} - \boldsymbol{S} + \tfrac{\Pi_1}{\rho}\|_F^2 + \|\boldsymbol{L} - \boldsymbol{U}\boldsymbol{V} + \tfrac{\Pi_2}{\rho}\|_F^2 \\ &+\|\boldsymbol{D}(\mathcal{L}) - \mathcal{M} + \tfrac{\Pi_3}{\rho}\|_F^2) \end{aligned} \tag{8}$$

where $\rho > 0$ denotes the penalty parameter, Π_1, Π_2 and Π_3 are the Lagrange multipliers related to three different constraints $\boldsymbol{Y} = \boldsymbol{L} + \boldsymbol{S}$, $\boldsymbol{L} = \boldsymbol{U}\boldsymbol{V}$, and $\mathcal{M} = \boldsymbol{D}(\mathcal{L})$. Eq. (8) can be decomposed into the following sub-problems.

(1) $\boldsymbol{U}$ and $\boldsymbol{V}$: For the bilinear factor matrices minimization, we extract the terms involving $\boldsymbol{U}$ and $\boldsymbol{V}$ and obtain the following equations:

$$\begin{aligned} \boldsymbol{U}^{(k+1)} &= \arg\min_{\boldsymbol{U}} \mathcal{F}(\boldsymbol{U}, \boldsymbol{V}^{(k)}, \boldsymbol{S}^{(k)}, \mathcal{M}^{(k)}, \boldsymbol{L}^{(k)};\ \Pi_1^{(k)}, \Pi_2^{(k)}, \Pi_3^{(k)}) \\ &= \arg\min_{\boldsymbol{U}} \tfrac{1}{2}\|\boldsymbol{U}\|_F^2 + \tfrac{\rho^{(k)}}{2}\|\boldsymbol{U}\boldsymbol{V}^{(k)} - (\boldsymbol{L}^{(k)} + \tfrac{\Pi_2^{(k)}}{\rho^{(k)}})\|_F^2, \end{aligned} \tag{9}$$

$$\begin{aligned} \boldsymbol{V}^{(k+1)} &= \arg\min_{\boldsymbol{V}} \mathcal{F}(\boldsymbol{U}^{(k+1)}, \boldsymbol{V}, \boldsymbol{S}^{(k)}, \mathcal{M}^{(k)}, \boldsymbol{L}^{(k)}; \Pi_1^{(k)}, \Pi_2^{(k)}, \Pi_3^{(k)}) \\ &= \arg\min_{\boldsymbol{V}} \tfrac{1}{2}\|\boldsymbol{V}\|_F^2 + \tfrac{\rho^{(k)}}{2}\|\boldsymbol{U}^{(k+1)}\boldsymbol{V} - (\boldsymbol{L}^{(k)} + \tfrac{\Pi_2^{(k)}}{\rho^{(k)}})\|_F^2. \end{aligned} \tag{10}$$

By setting the derivative of Eqs. (9) and (10) with respect to $\boldsymbol{U}$ and $\boldsymbol{V}$ to zero, respectively, we deduce their closed-form solutions as follows:

$$\boldsymbol{U}^{(k+1)} = (\rho^{(k)}\boldsymbol{L}^{(k)} + \Pi_2^{(k)})(\boldsymbol{V}^{(k)})^T(\rho^{(k)}\boldsymbol{V}^{(k)}(\boldsymbol{V}^{(k)})^T + \boldsymbol{I})^{-1}, \tag{11}$$

$$\boldsymbol{V}^{(k+1)} = (\boldsymbol{I} + \rho^{(k)}(\boldsymbol{U}^{(k+1)})^T\boldsymbol{U}^{(k+1)})^{-1}[(\boldsymbol{U}^{(k+1)})^T(\rho^{(k)}\boldsymbol{L}^{(k)} + \Pi_2^{(k)})]. \tag{12}$$

(2) $\boldsymbol{S}$: For the sparse noise removal, we extract the terms involving $\boldsymbol{S}$ and obtain the following equation:

$$\begin{aligned} \boldsymbol{S}^{(k+1)} &= \arg\min_{\boldsymbol{S}} \mathcal{F}(\boldsymbol{U}^{(k+1)}, \boldsymbol{V}^{(k+1)}, \boldsymbol{S}, \mathcal{M}^{(k)}, \boldsymbol{L}^{(k)};\ \Pi_1^{(k)}, \Pi_2^{(k)}, \Pi_3^{(k)}) \\ &= \arg\min_{\boldsymbol{S}} \tfrac{\lambda_1}{\rho^{(k)}}\|\boldsymbol{S}\|_1 + \tfrac{1}{2}\|\boldsymbol{S} - (\boldsymbol{Y} - \boldsymbol{L}^{(k)} + \tfrac{\Pi_1^{(k)}}{\rho^{(k)}})\|_F^2. \end{aligned} \tag{13}$$

Different from Eqs. (9) and (10) which are composed of two different Frobenius norms, the Eq. (13) consists of l_1-norm and Frobenius norm simultaneously. The solution of $\boldsymbol{S}$ [30] is:

$$\boldsymbol{S}^{(k+1)} = sign(\boldsymbol{B}). * \max\{|\boldsymbol{B}| - \lambda_1/\rho_k, 0\}. \tag{14}$$

(3) $\boldsymbol{\mathcal{M}}$: For the spatial-spectral smoothness reservation, we extract the terms involving $\boldsymbol{\mathcal{M}}$ from Eq. (8) and obtain the following equation:

$$\begin{aligned}\boldsymbol{\mathcal{M}}^{(k+1)} &= \arg\min_{\mathcal{M}} \mathcal{F}(\boldsymbol{U}^{(k+1)}, \boldsymbol{V}^{(k+1)}, \boldsymbol{S}^{(k+1)}, \boldsymbol{\mathcal{M}}, \boldsymbol{L}^{(k)};\ \Pi_1^{(k)}, \Pi_2^{(k)}, \Pi_3^{(k)}) \\ &= \arg\min_{\mathcal{M}} \tfrac{\lambda_2}{\rho^{(k)}} \|\boldsymbol{\mathcal{M}}\|_1 + \tfrac{1}{2}\|\boldsymbol{\mathcal{M}} - (\boldsymbol{D}(\mathcal{L}^{(k)}) + \tfrac{\Pi_3^{(k)}}{\rho^{(k)}})\|_F^2.\end{aligned} \tag{15}$$

Similar to the sparse noise removal, the solution can be obtained as follows:

$$\boldsymbol{\mathcal{M}}^{(k+1)} = sign(\mathcal{P}). * \max\{|\mathcal{P}| - \lambda_2/\rho_k, 0\}, \tag{16}$$

where $\mathcal{P} = \boldsymbol{D}(\mathcal{L}^{(k)}) + \frac{\Pi_3^{(k)}}{\rho^{(k)}}$.

(4) $\boldsymbol{L}$**:** For our desired HSI, we extract the terms involving $\boldsymbol{L}$ as follows:

$$\begin{aligned}\boldsymbol{L}^{(k+1)} &= \arg\min_{\boldsymbol{L}} \mathcal{F}(\boldsymbol{U}^{(k+1)}, \boldsymbol{V}^{(k+1)}, \boldsymbol{S}^{(k+1)}, \boldsymbol{\mathcal{M}}^{(k+1)}, \boldsymbol{L};\ \Pi_1^{(k)}, \Pi_2^{(k)}, \Pi_3^{(k)}) \\ &= \arg\min_{\boldsymbol{L}} \|\boldsymbol{L} - (\boldsymbol{Y} - \boldsymbol{S}^{(k+1)} + \tfrac{\Pi_1^{(k)}}{\rho^{(k)}})\|_F^2 + \|\boldsymbol{L} \\ &\quad -(\boldsymbol{U}^{(k+1)}\boldsymbol{V}^{(k+1)} - \tfrac{\Pi_2^{(k)}}{\rho^{(k)}})\|_F^2 + \|\boldsymbol{D}(\mathcal{L}) - (\boldsymbol{\mathcal{M}}^{(k+1)} - \tfrac{\Pi_3^{(k)}}{\rho^{(k)}})\|_F^2.\end{aligned} \tag{17}$$

One can see that these three terms of Eq. (17) are convex with respect to variable L. Therefore, we set the derivative of Eq. (17) with respect to L to zero and yield:

$$\begin{aligned}&\boldsymbol{L} - (\boldsymbol{Y} - \boldsymbol{S}^{(k+1)} + \tfrac{\Pi_1^{(k)}}{\rho^{(k)}}) + \boldsymbol{L} - (\boldsymbol{U}^{(k+1)}\boldsymbol{V}^{(k+1)} - \tfrac{\Pi_2^{(k)}}{\rho^{(k)}}) \\ &+\boldsymbol{D}^*(\boldsymbol{D}(\mathcal{L}) - (\boldsymbol{\mathcal{M}}^{(k+1)} - \tfrac{\Pi_3^{(k)}}{\rho^{(k)}})) = 0.\end{aligned} \tag{18}$$

Finally, we obtain the closed-form solution as follows:

$$\boldsymbol{L}^{(k+1)} = (\boldsymbol{E} + \boldsymbol{D}^*(\boldsymbol{\mathcal{M}}^{(k+1)} - \Pi_3^{(k)}/\rho^{(k)}))(2\boldsymbol{I} + \boldsymbol{D}^*\boldsymbol{D})^{-1}. \tag{19}$$

where $E = \boldsymbol{Y} - \boldsymbol{S}^{(k+1)} + \frac{\Pi_1^{(k)} - \Pi_2^{(k)}}{\rho^{(k)}} + \boldsymbol{U}^{(k+1)}\boldsymbol{V}^{(k+1)}$.

(4) Π_1, Π_2 and Π_3: Three Lagrange multipliers are updated as

$$\begin{aligned}\Pi_1^{(k+1)} &= \Pi_1^{(k)} + \rho^{(k)}\left(\boldsymbol{Y} - \boldsymbol{L}^{(k+1)} - \boldsymbol{S}^{(k+1)}\right); \\ \Pi_2^{(k+1)} &= \Pi_2^{(k)} + \rho^{(k)}\left(\boldsymbol{L}^{(k+1)} - \boldsymbol{U}^{(k+1)}\boldsymbol{V}^{(k+1)}\right); \\ \Pi_3^{(k+1)} &= \Pi_3^{(k)} + \rho^{(k)}\left(\boldsymbol{\mathcal{M}}^{(k+1)} - \boldsymbol{D}(\mathcal{L}^{(k+1)})\right).\end{aligned} \tag{20}$$

Specifically, Algorithm 1 presents the BFTV solver. In Algorithm 1, the inputs contain the noisy HSI $\boldsymbol{Y} \in \mathbf{R}^{MN \times B}$, desired rank r, the regularization parameters λ_1 and λ_2, and the stopping criteria ϵ. The output is the restored clean image $\boldsymbol{L}$, which is the same size as $\boldsymbol{Y}$. The ALM algorithm has been proven [29] that it will converge to the exact optimal solution iteratively and has been successfully applied to various applications.

Algorithm 1 BFTV.

Input: The noisy HSI $\boldsymbol{Y}$, desired rank r, the regularization parameters λ_1 and λ_2, and the stopping criteria ϵ

Output: Desired clean image $\boldsymbol{L}$

1: Initialize $\boldsymbol{U}^0 = \boldsymbol{V}^0 = \boldsymbol{S}^0 = \boldsymbol{\mathcal{M}}^0 = \boldsymbol{L}^0 = \boldsymbol{0}$;
$\Pi_1^0 = \Pi_2^0 = \Pi_3^0 = \boldsymbol{0}$; $\rho^0 = 10^{-3}$, $\rho_{\max} = 10^6$, $\beta = 1.1$ and $k = 0$

2: Repeat until convergence
Update $\boldsymbol{U}^{(k+1)}, \boldsymbol{V}^{(k+1)}, \boldsymbol{S}^{(k+1)}, \boldsymbol{\mathcal{M}}^{(k+1)}, \boldsymbol{L}^{(k+1)}; \Pi_1^{(k+1)}, \Pi_2^{(k+1)}, \Pi_3^{(k+1)}$
via (11), (12), (14), (16), (19), and (20) iteratively
Update the parameter $\rho := \min\{\beta\rho, \rho_{\max}\}$
Check the convergence condition

$$\max\left\{\begin{array}{l}\|\boldsymbol{Y} - \boldsymbol{L}^{(k+1)} - \boldsymbol{S}^{(k+1)}\|_\infty \\ \|\boldsymbol{L}^{(k+1)} - \boldsymbol{U}^{(k+1)}\boldsymbol{V}^{(k+1)}\|_\infty \\ \|\boldsymbol{\mathcal{M}}^{(k+1)} - \boldsymbol{D}(\mathcal{L}^{(k+1)}\|_\infty\end{array}\right\} \le \epsilon, \tag{21}$$

4. Experimental results

To illustrate the effectiveness of the proposed BFTV model for HSIs denoising, we conducted extensive experiments on one simulated and one real-world HSIs. For a comprehensive comparison, the proposed BFTV is compared with three state-of-the-art HSI denoising methods *i.e.*, the noise-adjusted iterative low-rank matrix approximation (NAILRMA) model [16], the low-rank matrix recovery (LRMR) model [2] with Go Decomposition algorithm, and total-variation-regularized low-rank matrix factorization (LRTV) [17]. The differences and connections between the proposed BFTV and all competitors are described as follows: (1) To establish the low-rank property, LRMR, NAILRMA, and LRTV utilized the bilateral random projection-based low-rank approximation, the randomized singular value decomposition, and the nuclear norm, respectively. While the proposed BFTV used the bilinear factorization to make the singular value decomposition free. (2) LRMR and NAILRMA considered only the globally low-rank property, resulting in the loss of the locally piecewise smoothness of clean HSIs. Although LRTV incorporated the hyper total variation into the low-rank matrix approximation model, it ignored the piecewise smoothness among the spectral dimension of HSIs since it performed the hyper total variation in a band-by-band way. The proposed BFTV not only considers the locally piecewise smoothness among the spatial dimension, but also among the spectral dimension.

4.1. Simulated data experiments

Following the experimental setting in LRTV, we used the synthetic data collected by using the Indian Pines dataset as the test data. This synthetic HSI dataset comprises 224 spectral bands with the wavelength ranging from 0.4 μm to 2.5 μm and the size of each band is 145 × 145 pixels. The reflectance values of all the voxels in the synthetic HSI were linearly mapped to [0, 1]. To simulate the real situation, different types of noise were added into the synthetic HSI. These specific details for five scenarios are listed as follows:

- **Scenario 1)** In this scenario, there were two types of noise including Gaussian noise and deadlines. We added Gaussian noise with zero mean and variance 0.1 to each spectral band. Besides, we also added deadlines to several selected bands 91–130. The number of deadlines was randomly changing from 3 to 10. Meanwhile, the width of each deadline was randomly fixed from 1 to 3 pixels.
- **Scenario 2)** In this scenario, the mixed noise including Gaussian noise with zero mean and variance 0.075 and impulse noise were added into different bands. The ratio of impulse noise is equal to 0.15.
- **Scenario 3)** In this scenario, there were three types of noise involving Gaussian noise, deadlines and impulse noise. We combined **Scenario 1)** and **Scenario 2)** together to generate the mixed noise. To be specific, the variance of Gaussian noise was equal to 0.1, the ratio of impulse noise was set to 0.15, and the deadlines was added to the same bands as **Scenario 1)**.
- **Scenario 4)** In this scenario, there were three types of noise including Gaussian noise, deadlines and impulse noise. How-

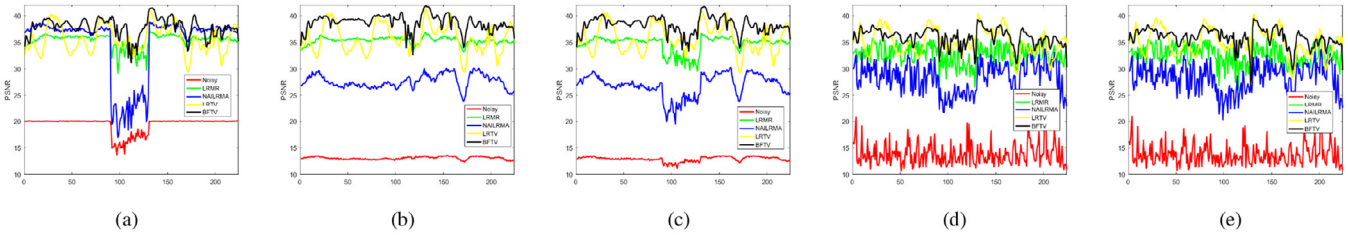

Fig. 1. PSNR values of different denoising methods in each band: (a) **Scenario 1)** (b) **Scenario 2)** (c) **Scenario 3)** (d) **Scenario 4)** (e) **Scenario 5)**.

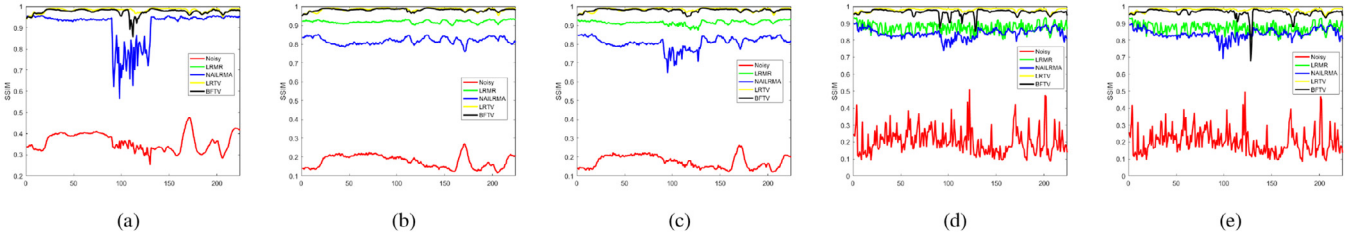

Fig. 2. SSIM values of different denoising methods in each band: **Scenario 1)** (b) **Scenario 2)** (c) **Scenario 3)** (d) **Scenario 4)** (e) **Scenario 5)**.

ever, compared with **Scenario 3)**, this scenario was more complicated because the noise intensity in different bands was randomly changed. The variance of Gaussian noise was generated from 0 to 0.2. Besides, the ratio of impulse noise was assigned from 0 to 0.2 randomly.

- **Scenario 5)** In this scenario, there were four kinds of noise containing Gaussian noise, deadlines, impulse noise and stripes. Except for noise as in **Scenario 4)**, 20–40 stripes were added into bands 161–190.

1) Quantitative Criteria: To evaluate the experimental results more precisely, three quantitative indexes including peak signal-to-noise ratio (PSNR), structural similarity index (SSIM) and spectral angle measure (SAM) [31] are selected as the performance criteria to analyze the results of five scenarios. PSNR and SSIM measure the difference between the groundtruth and the denoised HSI, while SAM aims to measure the spectral distortion. PSNR and SSIM are defined as follows:

$$PSNR_i = 10 * log_{10} \frac{MN}{\sum_{x=1}^{M}\sum_{y=1}^{N}[\hat{\mu}_i(x,y) - \mu_i(x,y)]^2} \tag{22}$$

$$SSIM_i = \frac{(2\mu_{\mu_i}\mu_{\hat{\mu}_i} + C_1)(2\sigma_{\mu_i}\mu_{\hat{\mu}_i} + C_2)(\sigma_{\mu_i\hat{\mu}_i} + C_3)}{(\mu_{\mu_i}{}^2 + \mu_{\hat{\mu}_i}{}^2 + C_1)(\sigma_{\mu_i}{}^2 + \sigma_{\hat{\mu}_i}{}^2 + C_2)(\sigma_{\mu_i}\sigma_{\hat{\mu}_i} + C_3)} \tag{23}$$

where M and N indicate the height and width of each spectral band image, while μ_i and $\hat{\mu}_i$ denote the ith bands of the groundtruth and the denoised image. μ_{μ_i} and $\mu_{\hat{\mu}_i}$ are the means of μ_i and $\hat{\mu}_i$ respectively. Besides, σ_{μ_i} and $\sigma_{\hat{\mu}_i}$ are their corresponding variances, and $\sigma_{\mu_i\hat{\mu}_i}$ is the covariance. C_1, C_2 and C_3 in the Eq (23) are constants utilized to maintain the stability of different terms. B is the number of total spectral bands. There are 224 spectral bands of the synthetic data, thus we computed the average values of PSNR, SSIM, and SAM denoted as MPSNR, MSSIM, and MSAM respectively. The higher values of PSNR and SSIM we get, and the smaller SAM is, the better performance of HSI denoising method is.

First, we report the PSNR and SSIM values of each spectral band to evaluate the performance of different denoising models in Figs. 1 and 2. One can see that all methods have achieved promising denoising performance for all scenarios. Generally speaking, the PSNR and SSIM values of BFTV are higher than those of all competing methods in most of all bands. MPSNR, MSSIM, and MSAM values of all methods are shown in Table 1 in the five scenarios. In each row of this table, the best results are tagged in bold. It can be seen that the proposed BFTV model performs best over another three popular algorithms in all scenarios. Specifically, BFTV outperforms the best peer algorithm LRTV by the improvement of 0.86 dB in **Scenario 5)** with respect to MPSNR. In addition, BFTV achieves the smallest MSAM value in all scenarios. This demonstrates that the results of BFTV are the closest to the groundtruth. LRTV and the proposed BFTV methods have achieved better denoising performance than LRMR and NAILRMA. The reason is that

Table 1
Quantitative evaluation of different HSI denoising approaches in different simulated scenarios.

Simulated scenario	Evaluation criteria	Noise	LRMR	NAILRMA	LRTV	BFTV
Scenario1)	MPSNR (dB)	19.34	35.89	35.30	35.56	***38.07***
	MSSIM	0.37	0.91	0.92	0.98	***0.98***
	MSAM	0.2152	0.0315	0.0566	0.0301	0.0214
Scenario2)	MPSNR (dB)	13.02	35.38	27.66	35.73	***38.90***
	MSSIM	0.19	0.90	0.82	0.98	***0.98***
	MSAM	0.4145	0.0300	0.0694	0.0283	0.0188
Scenario3)	MPSNR (dB)	12.72	35.64	34.99	35.69	***37.89***
	MSSIM	0.16	0.91	0.92	0.98	***0.98***
	MSAM	0.4245	0.0333	0.0356	0.0290	0.0232
Scenario4)	MPSNR (dB)	13.78	34.37	34.78	35.06	***36.16***
	MSSIM	0.2	0.89	0.91	***0.98***	0.97
	MSAM	0.4039	0.0425	0.359	0.0311	0.0266
Scenario5)	MPSNR (dB)	13.54	32.42	34.78	35.05	***35.91***
	MSSIM	0.2	0.87	0.91	***0.98***	0.97
	MSAM	0.4039	0.0435	0.0338	0.0316	0.0276

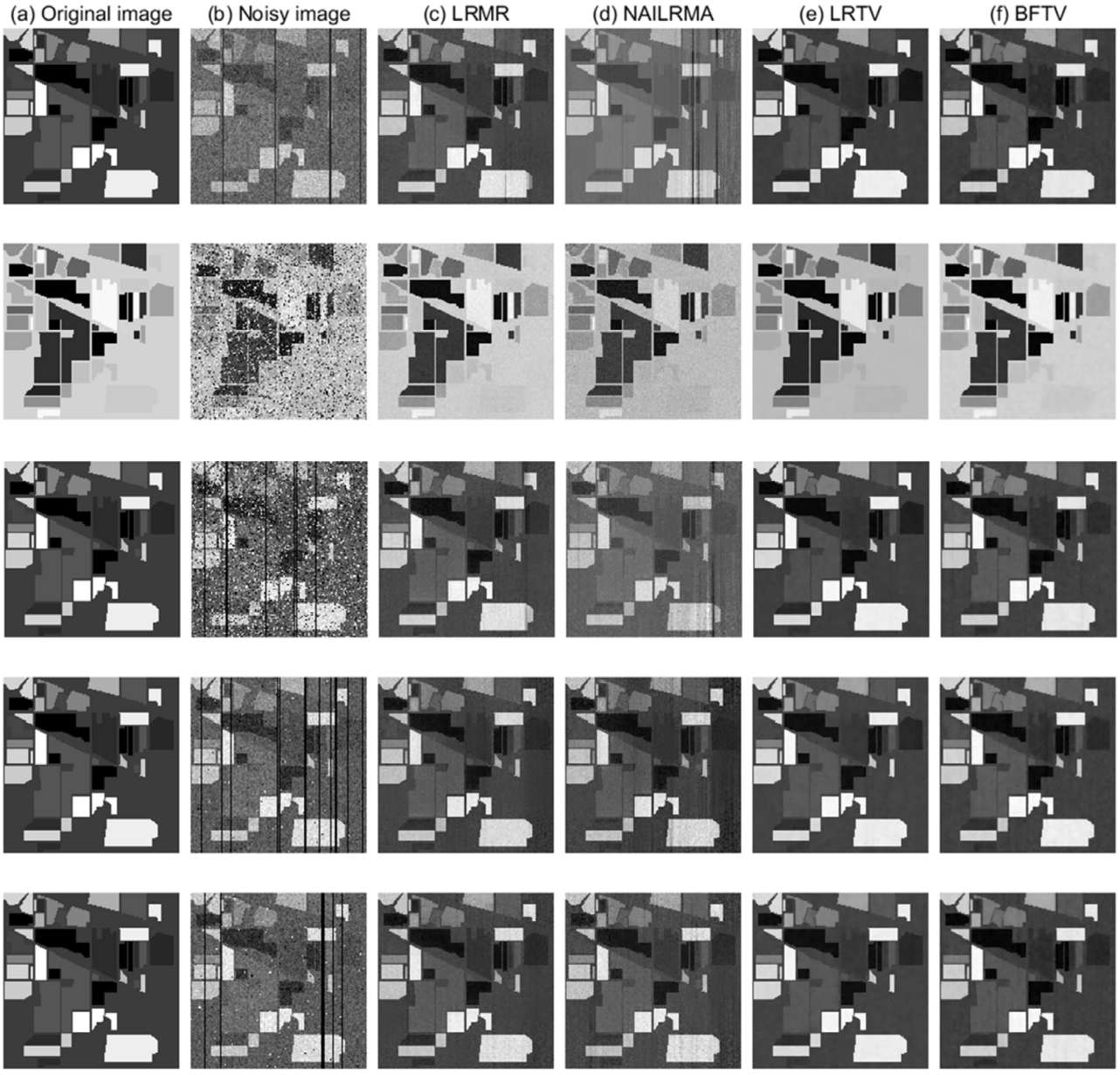

Fig. 3. Denoised results: the row from top to bottom **Scenario 1)** to **Scenario 5)**, (a) Original image (b) Noisy image (c) LRMR (d) NAILRMA (e) LRTV (f) BFTV.

LRTV and BFTV both integrate the total variation regularizer with the low-rank matrix approximation model to describe another essential HSI characteristic.

2) Visual Criteria: We also give the visual comparison among all methods by selecting the representative band of each scenario. The first row of Fig. 3 shows the HSI denoising results of band 120 in **Scenario1)**. It is obvious that all HSI denoising methods can remove the Gaussian noise effectively. For the LRMR and NAILRMA methods, there still remain some deadlines in the right area of images. However, the proposed BFTV provides the more satisfying results over its competitors. Fig. 3 present the visual comparison results of the remaining four scenarios. From these figures, we can deduce that all of these HSI denoising models can remove the mixed noise to a certain extent and they can suppress the Gaussian noise component effectively. However, in the **Scenario 2, Scenario 4** and **Scenario 5**, we can notice that the NAILRMA method still remains the mixed noise to a certain extent. The proposed BFTV model outperforms the other three popular HSI denoising methods.

4.2. Real data experiments

One real-world HSI is used to investigate the performance of the proposed BFTV when handling the real noisy scenario. The HYDICR urban images are selected as the testing data. Different from [16] which discarded some bands with heavy noises, we adopt all 210 bands to evaluate the performance of LRMR, NAILRMA, LRTV and the proposed LRTV. The size of the HYDICR urban images is $307 \times 307 \times 210$.

The denoising results of all methods on the HYDICE urban data are shown in Fig. 4, in which the first row corresponds to the 109th band while the second row is the 206th band. The HYDICE urban data are corrupted by Gaussian noise, horizontal stripes, and deadlines as shown in the first column of Fig. 4. Generally, the proposed BFTV has achieved the best visual performance. Compared to all the competitors, BFTV removes most of the horizontal stripes as shown in the first row of Fig. 4. For the 206th band, LRMR, NAILRMA and LRTV still cannot remove the horizontal stripes as shown in the highlighted red rectangle in Fig. 4. These observa-

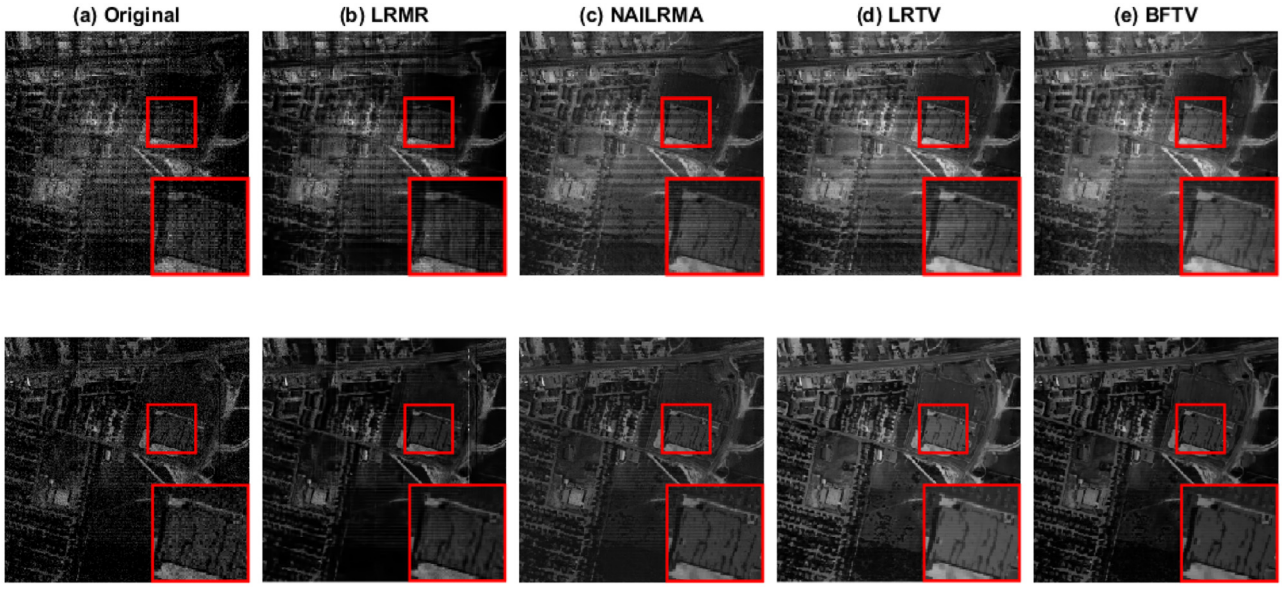

Fig. 4. Denoised results of **HYDICE urban data set**): (a) Original bands 109 and 206 (b) LRMR (c) NAILRMA (d) LRTV (e) BFTV.

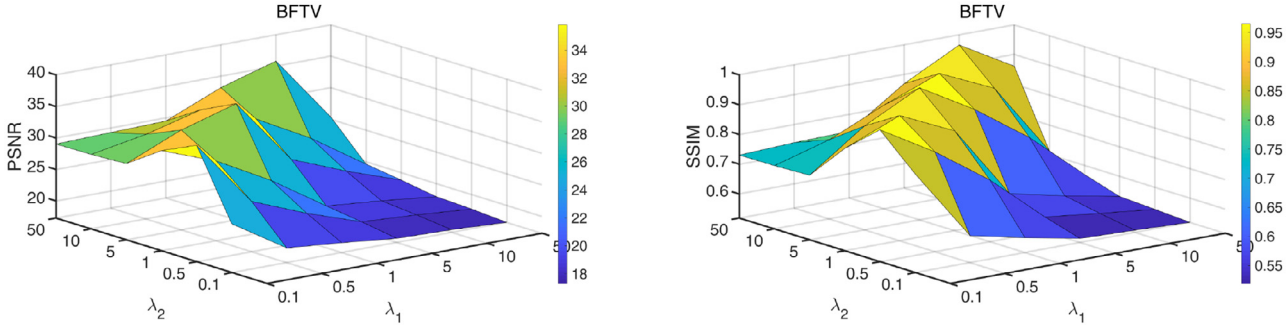

Fig. 5. PSNR and SSIM values of BFTV with different combinations of λ_1 and λ_2 on **Scenario 6**.

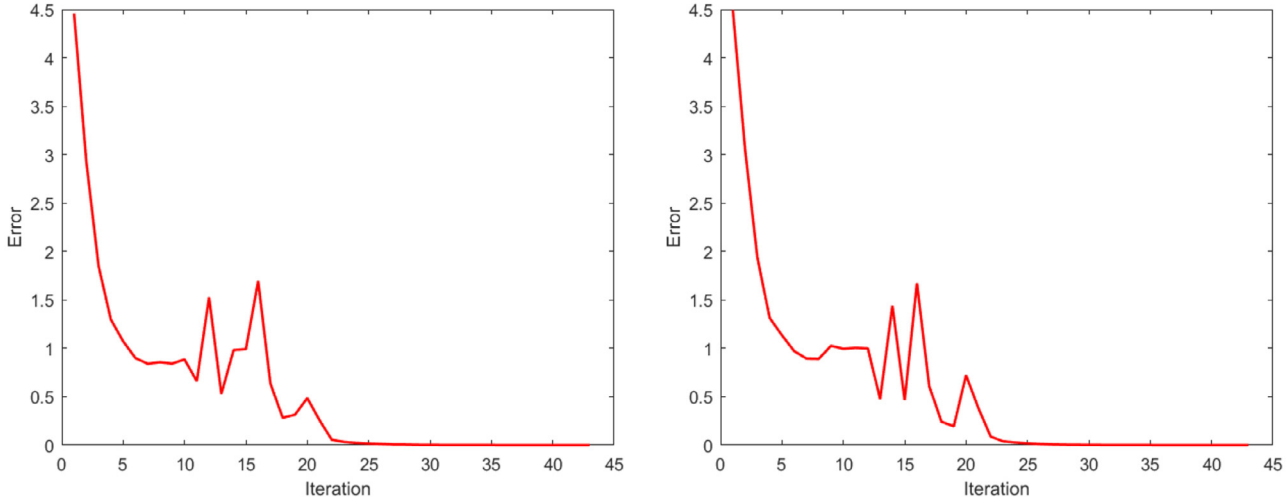

Fig. 6. Error versus iteration of BFTV on **Scenario 5** (Left) and **Scenario 6** (Right).

tions further demonstrate the high efficiency of the proposed BFTV in the real noisy scenario.

4.3. Discussion

In this section, we discuss the parameter selection and empirical convergence of the proposed BFTV.

1) Parameter Selection: Two parameters λ_1 and λ_2 are involved in the proposed BFTV. In all experiments, we select λ_1 and λ_2 from the same candidates $[0.1, 0.5, 1, 5, 10, 50]$. Fig. 5 reports the PSNR and SSIM values of BFTV with different combinations of λ_1 and λ_2 on **Scenario 6**. One can see that when $\lambda_1 = 50$ and $\lambda_2 = 10$, BFTV achieves the best denoising performance.

Table 2
Running time (in seconds) on all databases.

Method	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	HYDICE urban
LRMR	52.93	45.89	43.28	44.58	46.52	353.28
NAILRMA	32.98	35.93	33.64	36.75	34.56	188.32
LRTV	41.60	45.77	44.64	46.90	45.62	202.79
BFTV	54.00	55.57	60.79	54.12	55.77	262.33

2) Empirical Convergence:
To investigate the empirical convergence of BFTV, we show the error of BFTV with respect to iteration in Fig. 6, where the y-axis is the error defined in Eq. (21). It is easy to see that BFTV will converge after 25 iterations.

3) Running Time: In this section, we aim to compare the running time of all HSI denoising methods. All methods are conducted in Matlab R2018 at the server with Intel(R) CPU E3-1241 3.50Hz with 16GB memory. Table 2 reports the running time of all methods on the both simulated and real HSIs. From Table 2, we can observe that NAILRMA is the fastest method. The proposed BFTV has competitive running time with LRMR on the five scenarios, while costs less time than LRMR on the real HYDICE urban data. Because BFTV exploits the globally low-rank property and locally piecewise smoothness, the running time of BFTV could be longer than that of the other methods.

5. Conclusion

In this paper, we proposed a novel hyperspectral image denoising method, called total variation-regularized bilinear factorization method (BFTV). It took two intrinsic characteristics of hyperspectral images, i.e., the globally low-rank property and locally piecewise smoothness into consideration. To solve the proposed model, we designed an efficient algorithm based on the augmented Lagrangian multiplier. Expensive results as well as the empirical convergence analysis have demonstrated the superior performance of BFTV over three state-of-the-art hyperspectral image denoising methods. There still has room for the performance improvement. One possible solution is that we may consider the tensor factorization strategy insteowad of the bilinear factorization to explore the low-rankness.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

CRediT authorship contribution statement

Yongyong Chen: Conceptualization, Methodology, Software, Writing - original draft, Investigation. **Jiaxue Li:** Conceptualization, Data curation, Supervision, Writing - review & editing. **Yicong Zhou:** Funding acquisition, Writing - review & editing.

Acknowledgments

This work was funded by The Science and Technology Development Fund, Macau SAR (File no. 189/2017/A3), and by University of Macau (File no. MYRG2018-00136-FST).

References

[1] J. Li, Q. Yuan, H. Shen, L. Zhang, Hyperspectral image recovery employing a multidimensional nonlocal total variation model, Signal Process. 111 (2015) 230–248.
[2] H. Zhang, W. He, L. Zhang, H. Shen, Q. Yuan, Hyperspectral image restoration using low-rank matrix recovery, IEEE Trans. Geosci. Remote Sens. 52 (8) (2013) 4729–4743.
[3] Y. Chen, Y. Guo, Y. Wang, D. Wang, C. Peng, G. He, Denoising of hyperspectral images using nonconvex low rank matrix approximation, IEEE Trans. Geosci. Remote Sens. 55 (9) (2017) 5366–5380.
[4] X. Liu, S. Bourennane, C. Fossati, Nonwhite noise reduction in hyperspectral images, IEEE Geosci. Remote Sens. Lett. 9 (3) (2012) 368–372.
[5] A. Buades, B. Coll, J.-M. Morel, A non-local algorithm for image denoising, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2, IEEE, 2005, pp. 60–65.
[6] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Image denoising with block–matching and 3d filtering, in: Image Processing: Algorithms and Systems, Neural Networks, and Machine Learning, 6064, International Society for Optics and Photonics, 2006, p. 606414.
[7] M. Aharon, M. Elad, A. Bruckstein, et al., K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation, IEEE Trans. Signal Process. 54 (11) (2006) 4311.
[8] S. Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, L. Zhang, Weighted nuclear norm minimization and its applications to low level vision, Int. J. Comput. Vis. 121 (2) (2017) 183–208.
[9] Z. Wu, Q. Wang, J. Jin, Y. Shen, Structure tensor total variation-regularized weighted nuclear norm minimization for hyperspectral image mixed denoising, Signal Process. 131 (2017) 202–219.
[10] Q. Wang, Z. Wu, J. Jin, T. Wang, Y. Shen, Low rank constraint and spatial spectral total variation for hyperspectral image mixed denoising, Signal Process. 142 (2018) 11–26.
[11] Y. Xie, Y. Qu, D. Tao, W. Wu, Q. Yuan, W. Zhang, Hyperspectral image restoration via iteratively regularized weighted Schatten p-norm minimization, IEEE Trans. Geosci. Remote Sens. 54 (8) (2016) 4642–4659.
[12] L.I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D 60 (1–4) (1992) 259–268.
[13] Y. Wang, J. Peng, Q. Zhao, Y. Leung, X.-L. Zhao, D. Meng, Hyperspectral image restoration via total variation regularized low-rank tensor decomposition, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 11 (4) (2017) 1227–1243.
[14] E.J. Candès, B. Recht, Exact matrix completion via convex optimization, Found. Comput. Math. 9 (6) (2009) 717.
[15] M. Fazel, Matrix Rank Minimization With Applications, PhD thesis, Stanford University, 2002 Ph.D. thesis.
[16] W. He, H. Zhang, L. Zhang, H. Shen, Hyperspectral image denoising via noise-adjusted iterative low-rank matrix approximation, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8 (6) (2015) 3050–3061.
[17] W. He, H. Zhang, L. Zhang, H. Shen, Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration, IEEE Trans. Geosci. Remote Sens. 54 (1) (2015) 178–188.
[18] M.J. Black, A. Rangarajan, On the unification of line processes, outlier rejection, and robust statistics with applications in early vision, Int. J. Comput. Vis. 19 (1) (1996) 57–91.
[19] F. Xu, Y. Chen, C. Peng, Y. Wang, X. Liu, G. He, Denoising of hyperspectral image using low-rank matrix factorization, IEEE Geosci. Remote Sens. Lett. 14 (7) (2017) 1141–1145.
[20] H. Fan, J. Li, Q. Yuan, X. Liu, M. Ng, Hyperspectral image denoising with bilinear low rank matrix factorization, Signal Process. 163 (2019) 132–152.
[21] Y. Chang, L. Yan, H. Fang, S. Zhong, W. Liao, HSI-DENET: hyperspectral image restoration via convolutional neural network, IEEE Trans. Geosci. Remote Sens. 57 (2) (2018) 667–682.
[22] Q. Yuan, Q. Zhang, J. Li, H. Shen, L. Zhang, Hyperspectral image denoising employing a spatial–spectral deep residual convolutional neural network, IEEE Trans. Geosci. Remote Sens. 57 (2) (2018) 1205–1218.
[23] Q. Zhang, Q. Yuan, J. Li, X. Liu, H. Shen, L. Zhang, Hybrid noise removal in hyperspectral imagery with a spatial–spectral gradient network, IEEE Trans. Geosci. Remote Sens. 57 (10) (2019) 7317–7329.
[24] M. Golbabaee, P. Vandergheynst, Joint trace/tv norm minimization: a new efficient approach for spectral compressive imaging, in: 2012 19th IEEE International Conference on Image Processing, IEEE, 2012, pp. 933–936.
[25] Q. Yuan, L. Zhang, H. Shen, Hyperspectral image denoising employing a spectral–spatial adaptive total variation model, IEEE Trans. Geosci. Remote Sens. 50 (10) (2012) 3660–3677.
[26] F. Shang, Y. Liu, H. Tong, J. Cheng, H. Cheng, Robust bilinear factorization with missing and grossly corrupted observations, Inf. Sci. 307 (2015) 53–72.
[27] F. Shang, J. Cheng, Y. Liu, Z.-Q. Luo, Z. Lin, Bilinear factor matrix norm minimization for robust pca: algorithms and applications, IEEE Trans. Pattern. Anal. Mach. Intell. 40 (9) (2017) 2066–2080.

[28] H.K. Aggarwal, A. Majumdar, Hyperspectral image denoising using spatio-spectral total variation, IEEE Geosci. Remote Sens. Lett. 13 (3) (2016) 442–446.
[29] Z. Lin, M. Chen, Y. Ma, The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices, Technical Report UILU-ENG-09-2215, UIUC Technical Report, 2009.
[30] R. Tibshirani, Regression shrinkage and selection via the lasso, J. R. Stat. Soc. 58 (1) (1996) 267–288.
[31] R.H. Yuhas, J.W. Boardman, A.F. Goetz, Determination of semi-arid landscape endmembers and seasonal trends using convex geometry spectral unmixing techniques (1993).