

# Hyperspectral image restoration using framelet-regularized low-rank nonnegative matrix factorization



Yong Chen, Ting-Zhu Huang\*, Xi-Le Zhao, Liang-Jian Deng

School of Mathematical Sciences/Resarch Center for Image and Vision Computing, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, PR China

## ARTICLE INFO

### Article history:

Received 11 January 2018

Revised 3 June 2018

Accepted 25 June 2018

Available online 2 July 2018

### Keywords:

Hyperspectral image restoration

Low-rank nonnegative matrix factorization

Framelet regularization

Block successive upper-bound minimization

## ABSTRACT

Hyperspectral image (HSI) restoration is a process to remove a mixture of various kinds of noise, which is a key preprocessing step to improve the performance of subsequent applications. Since the HSI has a large correlation between spectral bands and abundant geometric features in the spatial domain, thus the low-rank prior and spatial structure prior can be introduced to HSI restoration. However, many existing approaches usually directly use the nuclear norm and total variation (TV) regularization to depict the spectral-spatial priors of HSI, which inevitably requires singular value decomposition (SVD) computation and causes staircase artifacts in the image, respectively. To overcome these limitations, in this work, we propose a novel HSI restoration method named framelet-regularized low-rank nonnegative matrix factorization (F-LRNMF), in which the low-rank nonnegative matrix factorization is developed to describe that the HSI lies in a low-rank subspace. Furthermore, to decrease the staircase artifacts caused by TV regularization that directly applies to HSI, we use framelet regularization to constrain the factor whose size is much less than HSI itself. The framelet regularization can effectively preserve the details and geometric features of the restored HSI in the spatial domain. An efficient block successive upper-bound minimization (BSUM) algorithm is designed to solve the proposed optimization model. Meanwhile, we theoretically analyze that the algorithm can converge to the set of coordinate-wise minimizers. Experiments under various cases of simulated and real HSI data demonstrate the effectiveness of the proposed model and the efficiency of the numerical algorithm in terms of both quantitative and qualitative assessments.

© 2018 Elsevier Inc. All rights reserved.

## 1. Introduction

Hyperspectral image (HSI) involves hundreds of spectral bands and has created great research interests in the field of remote sensing in recent years. With the wealthy spatial and spectral knowledge of HSI, it has a wide range of applications in remote sensing, such as environmental studies, agriculture, military, and geography [1,2]. In the real scene, however, due to the equipment limitations in the imaging process, e.g., dark current, sensor sensitivity, transmission errors, and calibration, HSI is unavoidably contaminated by various types of noise during the collection process, mainly including Gaussian noise, impulse noise, stripe noise, and deadlines. The noise in the HSI not only greatly influences the image quality, but also affects

\* Corresponding author.

E-mail addresses: [chenyong1872008@163.com](mailto:chenyong1872008@163.com) (Y. Chen), [tingzhuhuang@126.com](mailto:tingzhuhuang@126.com), [tzhuang@uestc.edu.cn](mailto:tzhuang@uestc.edu.cn) (T.-Z. Huang), [xlzhao122003@163.com](mailto:xlzhao122003@163.com) (X.-L. Zhao), [liangjian1987112@126.com](mailto:liangjian1987112@126.com) (L.-J. Deng).

the subsequent applications, such as classification [3], unmixing [4], object segmentation [5], and target detection [6]. Hence, HSI restoration has become an essential and inevitable issue before the subsequent analysis and applications of HSI.

In the past decades, noise removal for HSI is an active research topic in remote sensing field, and many HSI restoration methods have been proposed. Although HSI data possesses hundreds of spectral bands, each band also can be presented as a 2-D gray image. Moreover, the same pixel of different bands can be considered as a 1-D vector. Therefore, a number of traditional 2-D or 1-D denoising methods were proposed to decrease noise in the HSI band by band [7,8] or pixel by pixel [9]. However, the correlation between different spectral bands of adjacent pixels is not simultaneously considered in these methods, which will perform the spectral distortion. To take this correlation into consideration, a series of methods which join the spatial and spectral prior knowledge of HSI have been proposed for HSI restoration. For example, a hybrid spatial-spectral wavelet shrinkage method was proposed in [10], which benefits from the dissimilarity of the signal regularity in the spatial and spectral dimensions of HSI. Moreover, Chen and Qian [11] further restored the HSI by combining the bivariate wavelet thresholding with principal component analysis to encode both the spatial and spectral knowledge. In [12], the authors solved the HSI denoising problem by sparse nonnegative matrix factorization, which can take the spectral-spatial structures of HSI into account. Zhao and Yang [13] proposed sparse representation and low-rank constraint for removing Gaussian noise. Thereupon, Yang et al. [14] proposed a sparse representation framework that unifies denoising and spectral unmixing in a closed-loop manner.

Regarding the HSI as a 3-D cube data, a number of volumetric denoising methods have been proposed and can be used for HSI noise removal. Chen and Qian [15] extended 2-D image denoising to 3-D cube data denoising by using patulous bivariate wavelet thresholding. Based on the nonlocal similarity and Wiener filtering, an extension of the BM3D method (BM4D) was proposed in [16]. Another type of methods that can efficiently excavate the spatial and spectral information of HSI are tensor decomposition based multidimensional data analysis [17,18]. Most existing tensor-based HSI denoising methods are mainly based on two forms: CP decomposition and Tucker decomposition. For example, a Genetic Kernel Tucker decomposition method was proposed in [17] to develop the spatial and spectral information of HSI.

Actually, under the limitation of prior knowledge, the aforementioned methods are mainly used for additive Gaussian noise. In real-world HSI, however, there is often a mixture of various kinds of noise during the collection process, such as Gaussian noise, impulse noise, deal lines, stripe noise, and many unknown noises [19]. Recently, regarding the mixture noise removal issue of HSI as an ill-posed problem, low-rank optimization models have attracted many attentions and are widely applied to HSI restoration [19–24]. Low-rank based approaches well make use of the spectral correlation of HSI, based on the unmixing hypothesis that the clean HSI usually lives in a low-dimension subspace and can be formulated by a linear combination of few pure endmembers. That is to say, we reshape each band as a vector and then stack all vectors into a matrix, the reshaped HSI matrix can be regarded as a low-rank matrix. Zhang et al. [19] proposed HSI restoration method based on low-rank matrix recovery (LRMR), which can simultaneously remove the mixture noise. However, LRMR method only employs the spectral prior and fails to consider the special structure in the spatial domain. In [20], the authors incorporated TV regularization in the low-rank matrix factorization model (LRTV) to explore the spatial piecewise smooth structure of HSI. To capture the correlation of spectral difference, the spatio-spectral TV (SSTV) model which extends the traditional TV model and accounts for both the spatial and the spectral correlation was proposed in [25]. Moreover, many low-rank tensor recovery based methods are also popular in HSI restoration [26–28]. However, tensor-based approaches may result in some issues, such as computing CP rank is NP-hard problem and the Tucker decomposition is not unique [22].

Despite the satisfactory restoration results of the above methods for removing a mixture of various noise in HSI, these methods still exist some drawbacks. On the one hand, all low-rank based methods usually use the nuclear norm, weighted nuclear norm or nonconvex low-rank approximation function to capture the spectral correlation of HSI, which may introduce some poor problems. They should calculate the singular value decomposition (SVD) in their solving algorithm. As we all know, the size of HSI is very large, thus it requires plenty of computation cost to perform. Moreover, the poor convergence will occur in computing the SVD process. On the other hand, since the low-rank constraint fails to capture the spatial structure of HSI, thus many works have been incorporated other priors into the low-rank model. One of the most popular is TV regularization, which can efficiently preserve the spatial piecewise smooth structure of images [29–32]. However, these TV-regularized methods also emerge some weaknesses. First, the well-known staircase artifacts will be created in the image by using TV regularization [33]. Second, most of the methods directly used TV regularization to HSI itself, which also makes the time-consuming. In summary, most of the above drawbacks are mainly due to the fact that the size of HSI is too large.

To track the above issues, this paper proposes a novel framelet-regularized low-rank nonnegative matrix factorization method (F-LRNMF) for HSI restoration, which can simultaneously capture the high correlation of spectral bands and the geometric features in the spatial domain. Especially, we model the low-rank property of the clean HSI with low-rank nonnegative matrix factorization rather than the nuclear norm or other low-rank approximate functions. The advantages of this strategy can be concluded as: (1) it avoids to conduct SVD in each iteration; (2) it is computationally more efficient to perform in a low-dimension space than in the original large-scale space of HSI; (3) since the unknown variables are significantly decreased, thus the estimation of variables will be more precise than if we directly compute in the original HSI dimensionality [34]. Moreover, to alleviate the staircase artifacts of TV regularization and better preserve the details and geometric features in restored HSI, we incorporate the framelet regularization to reduced-dimensionality factorization coefficient in the low-rank nonnegative matrix factorization model, because the factorization coefficient can be regarded as underlying factor image, and maintains some priors which are related to the original HSI. Other advantages of using framelet regularization to the factorization coefficient are that this is indeed reasonable since the subspace spanned by the factor is

the same as the one where the HSI resides, and it can save plenty of memory by regularizing the factor rather than directly regularizing the HSI itself. It is worth mentioning that some of the previous work is similar to ours [35]. However, these work was mainly used to remove Gaussian noise. In this paper, we mainly focus on the HSI mixture noise removal. The main ideas and contributions of this paper are summarized as follows:

1. A novel low-rank nonnegative matrix factorization model is proposed for HSI restoration. Unlike the previous nuclear norm and low-rank approximation, the low-rank nonnegative matrix factorization avoids SVD computation and is computationally efficient to perform in a low-dimensional space rather than in the original space HSI.
2. The framelet regularization is introduced to constrain the reduced-dimensionality factor rather than directly regularizing the HSI itself in low-rank nonnegative matrix factorization model, which can effectively preserve the details and geometric features of restored HSI in the spatial domain and demand much less calculation and computer memory.
3. The block successive upper-bound minimization (BSUM) is designed to solve the proposed optimization model, and we theoretically analysis that the algorithm can convergence to the set of coordinate-wise minimizers. Both simulated and real data experiments demonstrate that our method outperforms other popular methods for HSI restoration.

The rest of this paper is organized as follows: In [Section 2](#), the HSI degradation model and low-rank matrix factorization model are analyzed in detail. The proposed F-LRNMF model is presented in [Section 3](#). In [Section 4](#), we design an effective BSUM algorithm to solve the proposed model and analyze its convergence. Various cases of simulated and real HSI data experimental results and model discussion are shown in [Section 5](#). Finally, [Section 6](#) concludes this paper.

## 2. Problem formulation

### 2.1. HSI degradation model

Let  $\mathcal{Y} \in R^{m \times n \times b}$  represents an observed HSI cube, where  $m$  and  $n$  are the spatial size of each spectral band, and  $b$  is the spectral dimensions of HSI. For the purpose of discussing, the HSI cube can be reshaped as an  $b \times mn$  matrix  $Y = [Y_1, Y_2, \dots, Y_{mn}]$ , where  $Y_i \in R^b$  is the  $b$ -variate spectral vector denoting the  $i$ th pixel. Assuming that the mixture noise in HSI is additive [19], thus the noise degradation model can be formulated as

$$Y = L + S + N, \quad (1)$$

where  $L \in R^{b \times mn}$  is the clean HSI,  $S \in R^{b \times mn}$  is the sparse noise (impulse noise, deadlines, stripes, and so on), and  $N \in R^{b \times mn}$  is the Gaussian noise. The goal of the HSI restoration is to precise estimate the clean HSI  $L$  from the noisy HSI  $Y$ .

### 2.2. Low-rank matrix factorization for HSI restoration

Since the correlation of the spectra band of neighboring pixels is significant, thus the clean HSI image has the low-rank property. We can depict the low rank property of HSI as

$$\text{rank}(L) \ll \min\{b, mn\}, \quad (2)$$

where  $\text{rank}(\cdot)$  is the rank of a matrix.

The low-rank property can be exploited based on the perspective of a linear mixture model (LMM) perspective [36]. From the perspective of linear mixture model, it is well known that each spectral signature (the column of  $L$ ) can be represented by a linear combination of a small number of pure spectral endmembers. Let the number of pure spectral endmembers is  $r$  in the HSI, then  $L$  can be represented as  $L = MR$ , where  $M$  is an  $b \times r$  matrix of endmember, and  $R$  is an  $r \times mn$  matrix of abundance. Since the number of endmembers in a scene is far smaller than the number of spectral bands and pixels, thus we can obtain  $r \ll \min\{b, mn\}$ . As the matrix  $L$  is factorized into the product of matrix  $M$  and  $R$ , then the rank of  $L$  satisfies  $\text{rank}(L) \leq r \ll \min\{b, mn\}$ , which implies the low-rank property of the matrix  $L$ . Moreover, the sparsity property of the non-Gaussian noise  $S$  is captured [37]. Based on the two priors, the low-rank matrix factorization framework is widely applied to HSI restoration problem.

Having the degraded model (1) to consider, the goal of our work is to recover the clean HSI  $L$ . However, it is an ill-posed problem. To address ill-posed problems, regularization is an effective method to guarantee the stable solution. Thus, the HSI restoration model can be formulated as:

$$\min_{L,S} \frac{1}{2} \|Y - L - S\|_F^2 + \lambda_1 R_1(S) + \lambda_2 R_2(L), \quad (3)$$

where  $\|Y - L - S\|_F^2$  is the data fidelity term, which can be obtained by the maximum *a posteriori* (MAP) estimation theory.  $R_1(S)$  and  $R_2(L)$  are two regularization terms characterizing the prior information of the sparse noise  $S$  and the clean HSI  $L$ , respectively.  $\lambda_1$  and  $\lambda_2$  denote the regularization parameters, which are used to balance the fidelity term and regularization terms.

In [19], Zhang et al. used the GoDec model proposed in [38] to capture the “low-rank + sparse” factorization for HSI restoration. Moreover, its extension have been also effectively applied in HSI restoration [20,21,39]. Since the sparse noise (impulse noise, dead lines and stripes) only corrupt some parts or some bands of the HSI, thus  $R_1(S)$  can be modeled as

$\|S\|_1$  [19–21,25,39], where  $\|\cdot\|_1$  represents the sum of absolute value of the matrix elements.  $R_2(L)$  is the regularization term associated to the clean HSI  $L$ , and it can be formulated as  $\text{rank}(L)$ , because the clean HSI possesses strong low-rank property in spectral dimension [13,19]. Due to the NP-hard and intractable problem of the  $\text{rank}$  function, researchers used the nuclear norm [19,20], weighted nuclear norm [21] or nonconvex low-rank approximation function [22] to approximate the  $\text{rank}$  function. Moreover, there are many other regularization that can joint low-rank regularization to depict the prior properties of the clean HSI, such as total variation regularization [20], spatio-spectral total variation regularization [25,39] and cross total variation regularization [40].

### 3. Proposed framelet-regularized low-rank nonnegative matrix factorization for HSI restoration

In this section, we will show our novel F-LRNMF model for HSI restoration. The key issue of the proposed F-LRNMF method is to use low-rank nonnegative matrix factorization, which can depict the low-rank property of the clean HSI better than the nuclear norm or low-rank approximation function. Moreover, to better preserve geometric features and details of HSI, we incorporate framelet regularization into the low-rank nonnegative matrix factorization framework for HSI restoration.

#### 3.1. Low-rank nonnegative matrix factorization

Recently, most of previous HSI restoration methods depict the low-rank prior to the clean HSI by using the nuclear norm [13,20], weighted nuclear norm [39] or nonconvex matrix rank approximation [21,22]. One inevitable problem is that all of these methods need to compute SVD in their optimization process, which becomes increasingly costly as the size of the underlying spatial and spectral dimensions increase. Moreover, the poor convergence will be occurred in the nuclear norm based depictions due to many iterations. Though many methods apply the nuclear norm to a subcube of HSI, they require to compute a lot of patches in their restoration process, thus it also needs plenty of computational costs.

To avoid the SVD computation, the purpose of our model is to propose such a non-SVD method in order to more efficiently restore the large-scale low-rank clean HSI. Inspired by the representation method of any matrix  $D \in R^{h \times w}$  of a rank up to  $r$  can be formulated into a matrix product form  $D = \hat{A}\hat{X}$ , where  $\hat{A} \in R^{h \times r}$  and  $\hat{X} \in R^{r \times w}$ . Thus we can replace the nuclear norm by utilizing the low-rank matrix factorization to approximate the clean HSI. The low-rank matrix factorization framework is extensively used in matrix completion [41], tensor completion [42,43], hyperspectral image denoising [35], and hyperspectral image superresolution [34], which performs a satisfactory result.

We can further explain the low-rank matrix factorization for hyperspectral data  $L$  from other perspective. Following [34,36], the hyperspectral data commonly has a large correlation between bands. All spectral vectors ( $L_i \in R^{b \times 1}$ ,  $i = 1, 2, \dots, mn$ ) are usually spanned by a subspace whose dimension much less than  $b$ . Therefore, we can reformulate  $L$  as:

$$L = AX, \quad (4)$$

where  $A \in R^{b \times r}$  is a matrix whose  $r$  columns span the identical subspace as the columns of  $L$ , and  $X \in R^{r \times mn}$  represents the representation coefficient. Since the small value of  $r$ , i.e.,  $r \ll \min\{b, mn\}$ , the description of data  $L$  translate into a relatively low dimensional space. Therefore, low-rank regularization term  $R_2(L)$  in model (3) can be replaced by the low-rank matrix factorization Eq. (4), and taking  $L = AX$  into the model (3), we can rewrite the model (3) as follows:

$$\min_{A,X,S} \frac{1}{2} \|Y - AX - S\|_F^2 + \lambda_1 \|S\|_1. \quad (5)$$

The model (5) need to simultaneously solve the factors  $A$  and  $X$ , it is similar to HSI blind unmixing problem [44]. In HSI sparse and blind unmixing problem,  $A$  is called endmember matrix and  $X$  denotes the relative endmember abundance. Therefore,  $A$  and  $X$  should satisfy the nonnegative constraint. Motivated by the nonnegative characteristic of  $A$  and  $X$  in HSI unmixing, we also introduce nonnegative condition to these factors  $A$  and  $X$ , which can maintain the physical significance of  $A$  and  $X$ . Thus, the model (5) can be formulated as follows:

$$\begin{aligned} \min_{A,X,S} \quad & \frac{1}{2} \|Y - AX - S\|_F^2 + \lambda_1 \|S\|_1 \\ \text{s.t. } & A \geq 0, \quad X \geq 0, \end{aligned} \quad (6)$$

This low-rank nonnegative matrix factorization has two advantages. One advantage is that it avoids performing SVD in each iteration, and we only use matrix multiplication and small-scale matrix inversion in the computation process. Thus it is computationally more efficient to work in a lower dimensional space than in the original space  $L$ . The other advantage is that since the number of unknown variables is evidently diminished, thus it is more accurate computation than if we directly estimate original dimensionality  $L$ .

#### 3.2. Framelet regularization

Though the low-rank nonnegative matrix factorization is a better choice to approximate the clean HSI, there are also potential drawbacks of the model (6). One drawback is that this factorization in the model only considers the spectral prior

to the HSI, while it does not consider the spatial information of HSI. Moreover, the strongly ill-posed inverse problem in the model may obtain infinite solutions. To address these drawbacks of the model (6), we should introduce additional information, or say prior knowledge, to this model. In addition to the low-rank property for HSI, there is more prior knowledge that may further enhance the potential capacity.

In HSI unmixing,  $A$  can be called as the spectral library or endmember library and  $X$  represents the fractional abundance matrix. In addition to the nonnegative property, if the factor matrixes  $A$  and  $X$  in (6) exist other particular physical properties, we can add its prior knowledge to the model (6) so as to guarantee a better and more stable solution. In reality, the abundance maps (each row of  $X$ ) possess piecewise smooth prior, since the material distribution is locally continuous in nature. Inspired by this physical character of the fractional abundance, researchers successfully introduced the TV regularization into HSI unmixing, including sparse unmixing [4] and blind unmixing [44]. Thus we can use TV regularization to constrain  $X$  so as to preserve the spatial piecewise smoothness of HSI.

However, it is well known that the TV regularization may result in staircase artifacts in the image so that the details and geometric features of the factor images (abundance maps) will not be well preserved. In real world applications, the details and geometric features are most important. To eliminate the staircase artifacts, the framelet regularization is proposed in [45] to promote the factor images sparsity in framelet domain and discriminate the approximation and detail components of hyperspectral data after framelet decomposition. Moreover, the framelet regularization is effectively applied to other image processing fields, such as image deblurring [46], tensor completion [47], and image desriping [48]. Motivated by the effectiveness of the framelet regularization, we introduce the framelet regularization to effectively depict the prior knowledge of the factor  $X$ , thus we can formulate the regularization as follows:

$$R(X) = \|WX\|_1, \quad (7)$$

where  $W$  represents the framelet transform matrix, and  $WX$  denotes the coefficients of factor images in the framelet transform domain. For detailed knowledge of the framelet system, we suggest the reader to refer [45,46,49]. It should be noted that the choice of the framelet decomposition operator is the same as [45] in our restoration work. For the detail, we can refer to Section 3-A in [45].

### 3.3. F-LRNMF model

From the above, the model (6) can effectively decompose the noisy HSI into the low-rank clean HSI and sparse noise image. Moreover, it is more advantages of using low-rank nonnegative matrix factorization to approximate the low-rank clean HSI than the nuclear norm or others. However, it cannot efficiently preserve the details and geometric features of the clean HSI. By incorporating framelet regularization for  $X$  into the model (6), this drawback can be solved. Finally, we can formulate our F-LRNMF restoration model as the following optimization function:

$$\begin{aligned} \min_{A, X, S} f(A, X, S) &= \frac{1}{2} \|Y - AX - S\|_F^2 + \lambda_1 \|S\|_1 + \lambda_2 \|WX\|_1, \\ \text{s.t. } A &\geq 0, \quad X \geq 0, \end{aligned} \quad (8)$$

where  $\lambda_1$  and  $\lambda_2$  are two positive regularization parameters.

The F-LRNMF model (8) can not only capture the spectral information of HSI but also preserve the spatial detail information for HSI. The low-rank nonnegative matrix factorization can translate the original high dimensional HSI into a low dimensional space. Moreover, combined with the sparsity constraint of the sparsity noise, we can capture the sparsity noise information successfully. After the sparsity noise is extracted, the framelet regularization is used to preserve the details and geometric features of each band, so as to remove the Gaussian noise.

In fact, there are other advantages for using the framelet regularization to depict the prior of the factor  $X$  except for the ability of preserving the details and geometric features. In previous work, TV regularization is a powerful method to preserve the piecewise smooth structure of HSI, such as LRTV [20], SSTV [25]. However, it not only well known for producing staircase-like artifacts but also directly applies to  $L$ . We all know that the size of  $L$  is very large in HSI, thus it needs more computation and memory to perform their algorithm. In contrast, we apply the framelet regularization to the reduced-dimensionality factor  $X$  and not to image  $L$  itself. This is more stable and reasonable since the subspace of  $L$  is the same as the one where spanned by  $X$ . Thus by using the framelet regularization to  $X$ , we can avoid the staircase effect and save the memory and computation in each iteration.

## 4. Numerical algorithm and convergence analysis

In this section, we will present the optimization scheme for solving the problem (8), and then give the convergence analysis of the algorithm.

### 4.1. Numerical algorithm

Clearly, the proposed model is not jointly convex for  $(A, X, S)$ , but this model is convex with respect to  $A, X, S$  independently. For solving the non-convex problem efficiency, we introduce the BSUM method [50].

Before using the BSUM method, we first define the proximal operator  $\text{prox}_g$  for a given convex function  $g(x)$  as

$$\text{prox}_g(y) = \arg \min_x g(x) + \frac{\rho}{2} \|x - y\|^2, \quad (9)$$

where  $\rho$  is a positive proximal parameter. The proximal operator problem is strongly convex in  $x$ , provided  $g(x)$  is convex. Since the classical proximal minimization obtains a solution of the problem  $\min_x g(x)$  by solving an equivalent problem  $\min_{x,y} g(x) + \frac{\rho}{2} \|x - y\|^2$  [50], thus we can obtain the minimization of  $g(x)$  by iteratively solving  $\text{prox}_g(x^k)$ , where  $x^k$  represents the latest update of  $x$ .

The nonnegative property of  $A$  and  $X$  can be added to the object function by using indicator function. We rewrite the model (8) as follows:

$$\min_{A,X,S} f(A, X, S) = \frac{1}{2} \|Y - AX - S\|_F^2 + \lambda_1 \|S\|_1 + \lambda_2 \|WX\|_1 + \iota_+(A) + \iota_+(X) \quad (10)$$

where

$$\iota_+(A) = \begin{cases} 0, & \text{all } A_{i,j} \geq 0 \\ +\infty, & \text{otherwise,} \end{cases} \quad (11)$$

and

$$\iota_+(X) = \begin{cases} 0, & \text{all } X_{i,j} \geq 0 \\ +\infty, & \text{otherwise.} \end{cases} \quad (12)$$

Let  $Z = (A, X, S)$ , using the proximal operator, the update of the proposed model (10) can be performed as the following proximal minimization problem

$$\min_Z h(Z, Z^k) = f(Z) + \frac{\rho}{2} \|Z - Z^k\|_F^2, \quad (13)$$

where  $Z^k = (A^k, X^k, S^k)$ . Therefore, to solve the problem of model (10), we convert into a optimization of the proximal minimization problem (13) and then develop the BSUM to iteratively optimize one variable, while fixing the others. By using this method, the problem (13) can be rewritten as three subproblems:

$$\begin{cases} h_1(A, Z_1^k) = f(A, X^k, S^k) + \frac{\rho}{2} \|A - A^k\|_F^2, \\ h_2(X, Z_2^k) = f(A^{k+1}, X, S^k) + \frac{\rho}{2} \|X - X^k\|_F^2, \\ h_3(S, Z_3^k) = f(A^{k+1}, X^{k+1}, S) + \frac{\rho}{2} \|S - S^k\|_F^2, \end{cases} \quad (14)$$

where  $Z_1^k = (A^k, X^k, S^k)$ ,  $Z_2^k = (A^{k+1}, X^k, S^k)$ , and  $Z_3^k = (A^{k+1}, X^{k+1}, S^k)$  in Eq. (14). Then, the subproblems of  $A$ ,  $X$ , and  $S$  can be solved as follows:

$$\begin{cases} A^{k+1} = \arg \min_A h_1(A, Z_1^k), \\ X^{k+1} = \arg \min_X h_2(X, Z_2^k), \\ S^{k+1} = \arg \min_S h_3(S, Z_3^k). \end{cases} \quad (15)$$

Note that  $A$ ,  $X$ , and  $S$  subproblems can be solved alternately and iteratively.

For  $A$  subproblem, we can deduce that

$$A^{k+1} = \arg \min_A h_1(A, Z_1^k) = \arg \min_A \frac{1}{2} \|Y - AX^k - S^k\|_F^2 + \frac{\rho}{2} \|A - A^k\|_F^2 + \iota_+(A). \quad (16)$$

To constrain the nonnegativity of  $A$ , we introduce auxiliary variable  $C$  and rewrite the problem (16) as follows:

$$\begin{aligned} \arg \min_{A,C} \frac{1}{2} \|Y - AX^k - S^k\|_F^2 + \frac{\rho}{2} \|A - A^k\|_F^2 + \iota_+(C), \\ \text{s.t. } C = A. \end{aligned} \quad (17)$$

This problem can be solved by alternating direction method of multipliers (ADMM) [51–53], and the augmented Lagrangian function of (17) is

$$L(A, C, P) = \frac{1}{2} \|Y - AX^k - S^k\|_F^2 + \frac{\rho}{2} \|A - A^k\|_F^2 + \iota_+(C) + \langle P, A - C \rangle + \frac{\beta}{2} \|A - C\|_F^2, \quad (18)$$

where  $P$  is Lagrange multiplier, and  $\beta$  is positive penalty parameter. Then, the minimization problem of (18) can be updated via alternating direction as follows:

$$\begin{cases} A^{k+1} = ((Y - S^k)(X^k)^T + \rho A^k + \beta(C - \frac{P^k}{\beta}))(X^k(X^k)^T + (\rho + \beta)I)^{-1}, \\ C^{k+1} = \max(A^{k+1} + \frac{P^k}{\beta}, 0), \\ P^{k+1} = P^k + \beta(A^{k+1} - C^{k+1}). \end{cases} \quad (19)$$

As for the subproblem of  $X$ , we have the following:

$$\begin{aligned} X^{k+1} &= \arg \min_X h_2(X, Z_2^k) = \arg \min_X \frac{1}{2} \|Y - A^{k+1}X\|_F^2 \\ &\quad - S^k \|_F^2 + \lambda_2 \|WX\|_1 + \frac{\rho}{2} \|X - X^k\|_F^2 + \iota_+(X). \end{aligned} \quad (20)$$

Since the  $\ell_1$ -norm term is nondifferentiable and inseparable, we similar use ADMM to solve this subproblem. In order to separate the  $\ell_1$ -norm and  $\ell_2$  norm, two auxiliary variables  $U$  and  $V$  are introduced to transform the unconstrained minimization problem (20) into the following constrained problem

$$\begin{aligned} \arg \min_{X, U, V} \frac{1}{2} \|Y - A^{k+1}X - S^k\|_F^2 + \lambda_2 \|U\|_1 + \frac{\rho}{2} \|X - X^k\|_F^2 + \iota_+(V), \\ \text{s.t. } U = WX, V = X. \end{aligned} \quad (21)$$

Then, the augmented Lagrangian for problem (21) can be written as follows:

$$\begin{aligned} L(X, U, V, \Lambda_1, \Lambda_2) &= \frac{1}{2} \|Y - A^{k+1}X - S^k\|_F^2 + \lambda_2 \|U\|_1 + \frac{\rho}{2} \|X - X^k\|_F^2 + \langle \Lambda_1, WX - U \rangle \\ &\quad + \langle \Lambda_2, X - V \rangle + \frac{\beta}{2} (\|WX - U\|_F^2 + \|X - V\|_F^2) + \iota_+(V), \end{aligned} \quad (22)$$

where  $\Lambda_1$  and  $\Lambda_2$  are two Lagrangian multiplies, and  $\beta$  is the positive penalty parameter. Therefore, the problem of (20) can be solved by three easy subproblems and the update of multiplies as follows:

$$\begin{cases} X^{k+1} = \arg \min_X \frac{1}{2} \|Y - A^{k+1}X - S^k\|_F^2 + \frac{\rho}{2} \|X - X^k\|_F^2 \\ \quad + \frac{\beta}{2} (\|WX - U^k + \frac{\Lambda_1^k}{\beta}\|_F^2 + \|X - V^k + \frac{\Lambda_2^k}{\beta}\|_F^2), \\ U^{k+1} = \arg \min_U \lambda_2 \|U\|_1 + \frac{\beta}{2} (\|WX^{k+1} - U + \frac{\Lambda_1^k}{\beta}\|_F^2), \\ V^{k+1} = \arg \min_V \frac{\beta}{2} (\|X^{k+1} - V + \frac{\Lambda_2^k}{\beta}\|_F^2) + \iota_+(V), \\ \Lambda_1^{k+1} = \Lambda_1^k + \beta (WX^{k+1} - U^{k+1}), \\ \Lambda_2^{k+1} = \Lambda_2^k + \beta (X^{k+1} - V^{k+1}). \end{cases} \quad (23)$$

The  $X$  subproblem is a linear system, and the  $U$  subproblem can be solved by using the soft-thresholding shrinkage operator [54]. Moreover, the projection operator is performed to solve  $V$  subproblem. Therefore, the problem in (23) can be updated through alternating direction as:

$$\begin{cases} X^{k+1} = ((A^{k+1})^T A^{k+1} + (\rho + \beta W^T W + \beta I)^{-1} ((A^{k+1})^T \\ \quad (Y - S) + \rho X^k + \beta (W^T (U^k - \frac{\Lambda_1^k}{\beta}) + (V^k - \frac{\Lambda_2^k}{\beta}))), \\ U^{k+1} = \max\{|WX^{k+1} + \frac{\Lambda_1^k}{\beta}| - \frac{\lambda_2}{\beta}, 0\} \frac{WX^{k+1} + \frac{\Lambda_1^k}{\beta}}{|WX^{k+1} + \frac{\Lambda_1^k}{\beta}|}, \\ V^{k+1} = \max\{X^{k+1} + \frac{\Lambda_2^k}{\beta}, 0\}, \\ \Lambda_1^{k+1} = \Lambda_1^k + \beta (WX^{k+1} - U^{k+1}), \\ \Lambda_2^{k+1} = \Lambda_2^k + \beta (X^{k+1} - V^{k+1}). \end{cases} \quad (24)$$

Finally, the  $S$  subproblem is formulated as follows:

$$S^{k+1} = \arg \min_S h_3(S, Z_3^k) = \arg \min_S \frac{1}{2} \|Y - A^{k+1}X^{k+1} - S\|_F^2 + \lambda_1 \|S\|_1 + \frac{\rho}{2} \|S - S^k\|_F^2, \quad (25)$$

which can be obtained the close form solution by using soft-thresholding shrinkage operator as follows:

$$S^{k+1} = \max\{|\hat{S}| - \frac{\lambda_1}{1 + \rho}, 0\} \frac{\hat{S}}{|\hat{S}|}, \quad (26)$$

where  $\hat{S} = \frac{(Y - A^{k+1}X^{k+1}) + \rho S^k}{1 + \rho}$ .

We summarize the detailed numerical procedure for the restoration of HSI data is presented in Algorithm 1. It is worth noting that we use a warm-start strategy in ADMM for  $A$  and  $X$  subproblems to speed up convergence.

#### 4.2. Convergence analyses

In this section, we will give the convergence analysis of the proposed algorithm. In recent, the BSUM for non-smooth optimization problem was proposed by Razaviyayn et al. [50]. It is an alternative inexact block coordinate descent approach which updates the variable blocks by successively minimizing a sequence of approximations of  $f$ . In the following, the convergence result in [50] is restated for convenience. In particular, we also can refer to [43] for the convergence analysis.

**Algorithm 1** The optimization procedure for the proposed model.

---

**Input:** The noisy HSI matrix  $Y \in R^{b \times mn}$ , the estimated rank  $r$ , the parameters  $\lambda_1, \lambda_2, \rho, \beta, \varepsilon, k_{\max}, i_{\max}$ , and  $j_{\max}$

- 1: Initialize: Set  $A^0 \in R^{b \times r}$  and  $X^0 \in R^{r \times mn}$  by randomly selecting entries in the interval  $[0, 1]$ ,  $S^0 = P^0 = C^0 = U^0 = V^0 = \Lambda_1^0 = \Lambda_2^0 = 0$ , and  $k = 0$ .
- 2: **while** stopping criterion is not satisfied **do**
- 3:   Initialize: Set  $A^{k+1,0} = A^k, C^{k+1,0} = C^k, P^{k+1,0} = P^k$ , and  $i = 0$ .
- 4:   **while** stopping criterion is not satisfied **do**
- 5:     Compute  $A^{k+1,i+1}$  by (19).
- 6:     Compute  $C^{k+1,i+1}$  by (19).
- 7:     Update  $P^{k+1,i+1}$  by (19).
- 8:     Set  $i = i + 1$ .
- 9:   **end while** and output  $A^{k+1} = A^{k+1,i}, C^{k+1} = C^{k+1,i}$  and  $P^{k+1} = P^{k+1,i}$
- 10:   Initialize: Set  $X^{k+1,0} = X^k, U^{k+1,0} = U^k, V^{k+1,0} = V^k, \Lambda_1^{k+1,0} = \Lambda_1^k, \Lambda_2^{k+1,0} = \Lambda_2^k$  and  $j = 0$ .
- 11:   **while** stopping criterion is not satisfied **do**
- 12:     Compute  $X^{k+1,j+1}$  by (24).
- 13:     Compute  $U^{k+1,j+1}$  by (24).
- 14:     Compute  $V^{k+1,j+1}$  by (24).
- 15:     Update  $\Lambda_1^{k+1,j+1}$  by (24).
- 16:     Update  $\Lambda_2^{k+1,j+1}$  by (24).
- 17:     Set  $j = j + 1$ .
- 18:   **end while** and output  $X^{k+1} = X^{k+1,j}, U^{k+1} = U^{k+1,j}, V^{k+1} = V^{k+1,j}, \Lambda_1^{k+1} = \Lambda_1^{k+1,j}$ , and  $\Lambda_2^{k+1} = \Lambda_2^{k+1,j}$ .
- 19:   Compute  $S^{k+1}$  by (26).
- 20:   Check the convergence conditions:  

$$\|A^{k+1}X^{k+1} - A^kX^k\|_F^2 / \|A^{k+1}X^{k+1}\|_F^2 < \varepsilon$$
- 21:   Set  $k = k + 1$
- 22: **end while**

**Output:** Restored HSI image  $L = AX$ .

---

**Lemma 1.** Give the problem  $\min_x f(x)$  s.t.  $x \in \mathcal{X}$ , where  $\mathcal{X}$  is the feasible set. Assume  $u(x, x^{k+1})$  is an approximation of  $f(x)$  at the  $(k+1)$ -th iteration, which satisfied the following conditions:

- (1)  $u_i(y_i, y) = f(y), \forall y \in \mathcal{X}, \forall i$ ,
  - (2)  $u_i(x_i, y) \geq f(y_1, \dots, y_{i-1}, x_i, y_{i+1}, \dots, y_n), \forall x_i \in \mathcal{X}_i, \forall y \in \mathcal{X}, \forall i$ ,
  - (3)  $u'_i(x_i, y; d_i)|_{x_i=y_i} = f'(y; d), \forall d = (0, \dots, d_i, \dots, 0), \text{ s.t. } y_i + d_i \in \mathcal{X}_i, \forall i$ ,
  - (4)  $u_i(x_i, y)$  is continuous in  $(x_i, y), \forall i$ .
- (27)

where  $u_i(x_i, y)$  is a subproblem with respect to the  $i$ th block and  $f'(y; d)$  is the direction derivative of  $f$  at the point  $y$  in direction  $d$ . Suppose  $u(x_i, y)$  is quasi-convex in  $x_i$  for  $i = 1, \dots, n$ . Moreover, assume that each subproblem  $\arg \min u_i(x_i, x^{k+1})$ , s.t.  $x \in \mathcal{X}_i$  has a unique solution for any point  $x^{k+1} \in \mathcal{X}$ . Then, every limit point  $z$  of the iterates generated by the BSUM algorithm converge to the set of coordinatewise minimum of  $f$ . In addition, if  $f(\cdot)$  is regular at  $z$ , then  $z$  is a stationary point of  $f$ .

Next, we will show that the proposed algorithm fits the framework of the BSUM method, thus the convergence of the proposed algorithm is guaranteed for the optimization problem (8).

**Theorem 1.** The iterated sequences generated by (13) converge to the set of coordinate-wise minimizers.

**Proof.** It is evident that  $h(Z, Z^k)$  is an approximation and a global upper bound of  $f(Z)$  at the  $k$ -th iteration, which satisfies the following conditions:

- (1)  $h_i(Z_i, Z) = f(Z), \forall Z, i = 1, 2, 3$ ,
  - (2)  $h_i(\bar{Z}_i, Z) \geq f(Z_1, \dots, \bar{Z}_i, \dots, Z_3), \forall \bar{Z}_i, \forall Z, i = 1, 2, 3$ ,
  - (3)  $h'_1(\bar{Z}_1, Z; D_1)|_{\bar{Z}_1=Z_1} = f'(Z; D^1), \forall D^1 = (D_1, 0, 0)$ ,
  - (4)  $h'_2(\bar{Z}_2, Z; D_2)|_{\bar{Z}_2=Z_2} = f'(Z; D^2), \forall D^2 = (0, D_2, 0)$ ,
  - (5)  $h'_3(\bar{Z}_3, Z; D_3)|_{\bar{Z}_3=Z_3} = f'(Z; D^3), \forall D^3 = (0, 0, D_3)$ ,
  - (6)  $h_i(\bar{Z}_i, Z)$  is continuous in  $(\bar{Z}_i, Z), i = 1, 2, 3$ .
- (28)

where  $Z_i$  is  $A, X, S$  for  $i = 1, 2, 3$ , respectively, and  $Z = (A, X, S)$ . Moreover,  $h_i$  ( $i = 1, 2, 3$ ) is strictly convex with respect to  $A, X$ , and  $S$ , respectively. In other words, each  $h_i$  subproblem has a unique solution. Thus, the assumptions in **Lemma 1** are all satisfied.

## 5. Experimental results and discussions

In this section, to evaluate the efficiency and effectiveness of the proposed method, we conduct extensive experiments on both simulated and real HSI data. Moreover, we show the results in comparison with five different popular HSI restoration methods, i.e., block matching 4-D filtering method (BM4D) [16], low-rank matrix recovery method (LRMR) [19], multi-task sparse nonnegative matrix factorization method (MTSNMF) [12], TV-regularized low-rank matrix factorization method (LRTV) [20], and spatio-spectral TV method (SSTV) [25]. All implementation codes can be obtained from the authors' homepages. In the following experiments, the parameters in those five compared methods are manually tuned according to the authors' rules in their paper to guarantee their possibly good performance. For the parameters in our method, we would like to show the detailed discussion in Section 5.3. To better the numerical calculation and visualization, the gray values of each HSI band are normalized into  $[0, 1]$  before restoration. All of the experiments are implemented in MATLAB R2017a on a desktop of 16GB RAM, Intel (R) Core (TM) i5-4590 CPU, @3.30GHz.

### 5.1. Simulated data experiments

In this experiments, two simulated HSI data sets are used to evaluate the performance of the proposed method, i.e., the Hyperspectral Digital Imagery Collection Experiment (HYDICE) image of the Washington DC Mall and the reflective optics system imaging spectrometer (ROSIS) images of the Pavia University, Italy. The original Washington DC Mall contains 191 spectral bands with  $1208 \times 307$  pixels for each band. We extract a subimage of size  $256 \times 256 \times 191$  from the Washington DC Mall for our experiment. Moreover, Pavia University data consists of  $610 \times 340$  pixels with 103 bands, a patch of size  $340 \times 340 \times 103$  is considered for our experiment. In order to demonstrate the robustness of the proposed method, we empirically fix the parameter values  $\lambda_1 = 0.05$ ,  $\lambda_2 = 0.3$ ,  $\rho = 0.1$ , and  $\beta = 200\lambda_2$  in all of the simulated experiments. For the estimated rank, we fix it as the number of singular values larger than 1% of the largest one [42].

HSI is usually contaminated by a mixture of various noise in the acquisition process, including Gaussian noise, impulse noise, deadlines, stripes, and so on [19]. To simulate these complicated scenarios, we add five types of noise to the original clean HSI data. The detailed noise types we list in the following:

Case (1) Gaussian noise: In practice, noise intensity in each band is also different in real HSI. To simulate this case, we add zero-mean Gaussian noise with different intensity to the clean HSI. The signal noise ratio (SNR) value of each band is generated from the uniform distribution with the value varies from 5 to 15 dB. The mean SNR value of all of the bands is 9.81 dB for Washington DC Mall, and 9.89 dB for Pavia University.

Case (2) Gaussian + impulse noise: The Gaussian noise is added to each band the same as Case (1). In addition, all bands are added impulse noise with different percentages, which are randomly chosen from 0 to 0.2.

Case (3) Gaussian + deadline noise: All bands are degraded by Gaussian noise as Case (1). Besides, 40 bands in the clean HSI data are randomly selected to add deadline noise, and the number of deadlines in each band randomly changes from 4 to 10.

Case (4) Gaussian + stripe noise: In this case, the Gaussian noise is added to each band just as Case (1). The stripe noise is simulated for 40 bands, the locations of the band are randomly selected from the original HSI, and the number of stripes in each band randomly changes from 4 to 10.

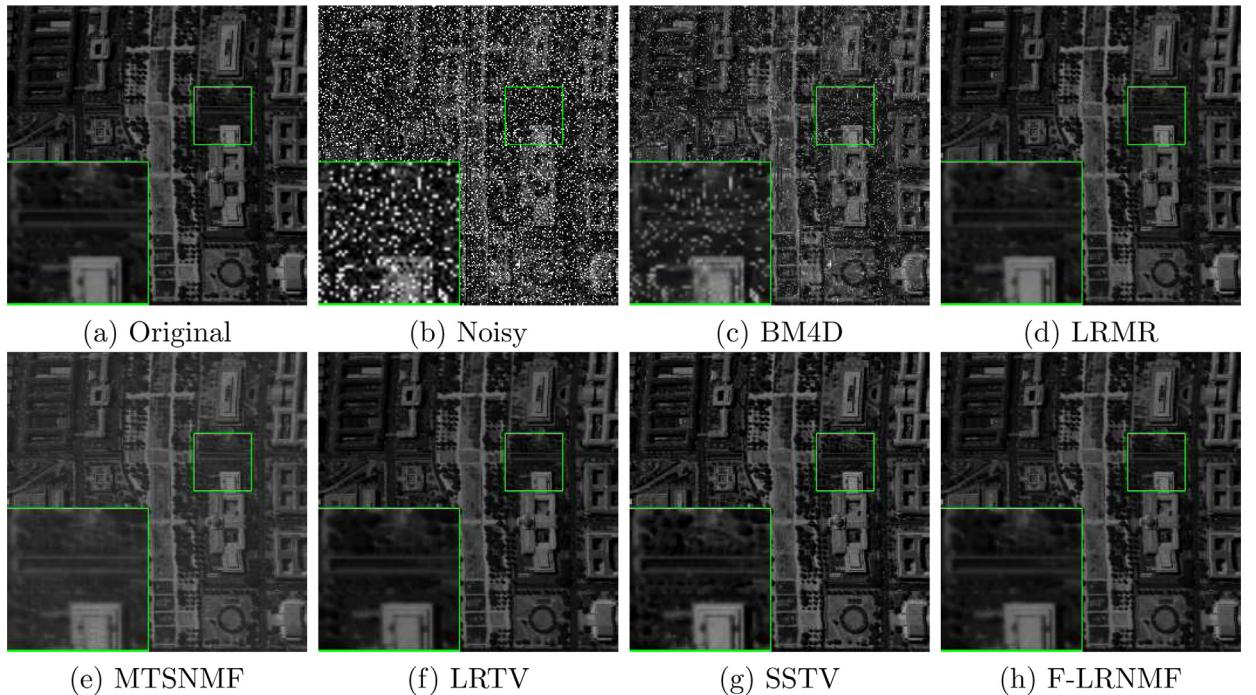
Case (5) Mixture noise: In this case, the Gaussian noise, impulse noise, deadline, and stripe noise are added to the original image. The Gaussian noise is added to each band just as Case (1). Moreover, we randomly choose 40 bands to add impulse noise, deadline and stripe noise, respectively.

To give an overall evaluation, we adopt the visual quality comparison and quantitative comparison for the six compared methods.

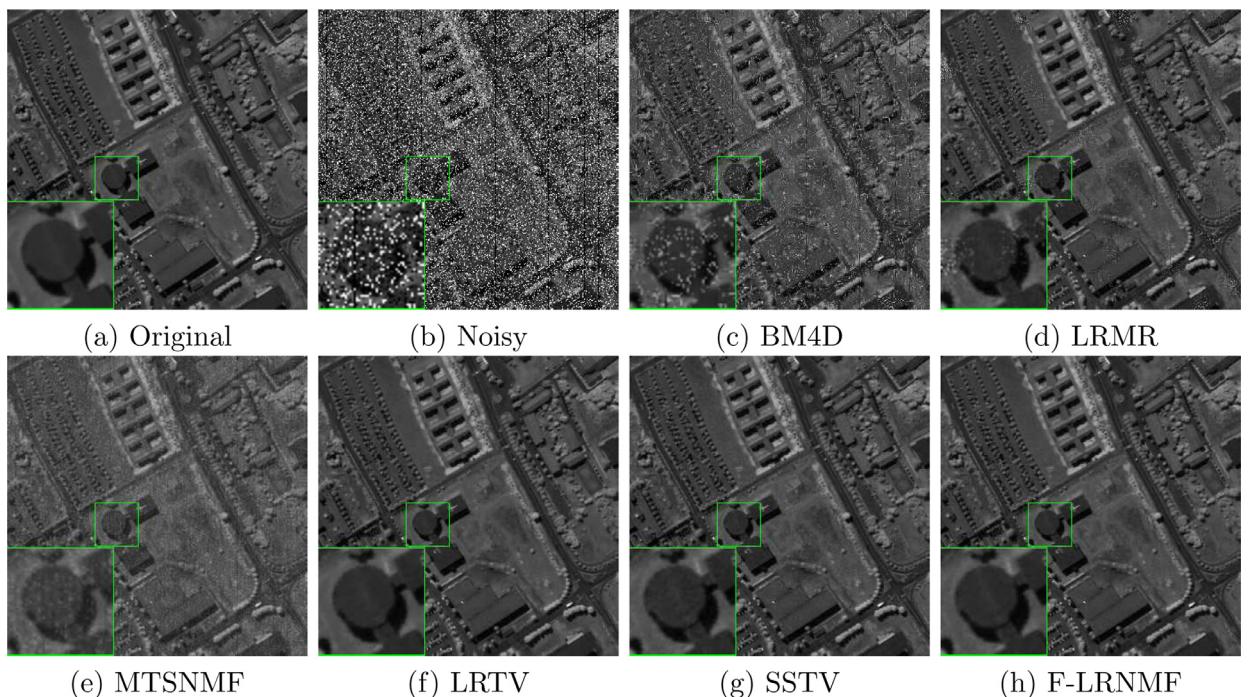
(1) Visual quality comparison: Since there are many bands in the HSI data and the page limitation, we select one band from the Washington DC Mall and Pavia University data for comparison, respectively. Figs. 1 and 2 show the restoration results of Washington DC Mall band 107 and Pavia University band 92, respectively. To better visual comparison, the same subregion is marked by a green box and then enlarged it.

From Figs. 1–2(b), it is evident that the original clean HSIs are greatly degraded by mixture noise. After denoising by the six methods, most of the noises are removed. However, some methods cannot remove all of the noises from the noisy image. For example, the denoised results by BM4D method as shown in Figs. 1–2(c), we can see that BM4D method can remove most of the Gaussian noise, but fails to remove the impulse noise and deadlines. MTSNMF method can remove all of the Gaussian noise, but some sparse noise also left in the image, and the input images are distorted as shown in Figs. 1–2(e). Although LRMR obtains much better results than BM4D and MTSNMF, some deadline and impulse noise still exist in the image (see the enlarged green box). The results of SSTV are shown in Figs. 1–2(g), it is clear that the mixture noise is eliminated on the whole, but some residual deadline noise can be observed in the restored image as shown in Fig. 1(g). The results by LRTV and F-LRNMF are visually similar, all of the noise can be removed from the noisy image, which agrees with the quantitative comparison in the next. In summary, the proposed F-LRNMF method performs competitive results among all of the compared methods, effectively removing all of the mixture noise while preserving the more edges and details in the restored HSI.

(2) Quantitative comparison: To further compare the effectiveness of our F-LRNMF method, the peak signal-to-noise ratio (PSNR), structural similarity index (SSIM) [55], and feature similarity (FSIM) [56] are calculated on each spectral band; then



**Fig. 1.** Denoising results of band 107 in Washington DC Mall band 107 under mixture noise case 5: (a) original band 107, (b) noisy band, the results of (c) BM4D, (d) LRMR, (e) MTSNMF, (f) LRTV, (g) SSTV, (h) F-LRNMF.



**Fig. 2.** Denoising results of band 92 in Pavia University under mixture noise case 5: (a) original band 92, (b) noisy band, the results of (c) BM4D, (d) LRMR, (e) MTSNMF, (f) LRTV, (g) SSTV, (h) F-LRNMF.

**Table 1**

The quantitative comparison of MPSNR, MSSIM and MFSIM with different methods on the simulated data of different cases.

Data	Noise case	Index	Noisy	BM4D	LRMR	MTSNMF	LRTV	SSTV	F-LRNMF
Washington DC Mall	Case 1	MPSNR	23.19	34.16	36.59	34.89	35.67	36.31	<b>38.54</b>
		MSSIM	0.615	0.944	0.969	0.963	0.967	0.966	<b>0.980</b>
		MFSIM	0.824	0.967	0.981	0.978	0.978	0.979	<b>0.987</b>
	Case 2	MPSNR	14.57	24.19	35.67	26.44	34.93	35.51	<b>37.46</b>
		MSSIM	0.291	0.672	0.963	0.853	0.968	0.960	<b>0.975</b>
		MFSIM	0.674	0.869	0.978	0.944	0.979	0.977	<b>0.984</b>
	Case 3	MPSNR	22.95	33.24	36.45	34.24	35.55	36.21	<b>38.39</b>
		MSSIM	0.608	0.937	0.969	0.958	0.964	0.965	<b>0.980</b>
		MFSIM	0.821	0.963	0.981	0.975	0.976	0.979	<b>0.987</b>
	Case 4	MPSNR	23.08	33.83	36.51	34.90	35.64	36.23	<b>38.52</b>
		MSSIM	0.612	0.941	0.969	0.963	0.967	0.966	<b>0.980</b>
		MFSIM	0.822	0.965	0.981	0.978	0.978	0.979	<b>0.987</b>
	Case 5	MPSNR	21.67	31.36	36.19	32.75	35.08	36.05	<b>38.26</b>
		MSSIM	0.559	0.890	0.967	0.942	0.969	0.965	<b>0.979</b>
		MFSIM	0.799	0.945	0.980	0.969	0.979	0.979	<b>0.987</b>
Pavia University	Case 1	MPSNR	24.03	36.58	36.56	37.45	37.39	37.59	<b>38.23</b>
		MSSIM	0.494	0.935	0.934	0.951	0.950	0.940	<b>0.954</b>
		MFSIM	0.771	0.965	0.972	0.978	0.976	0.975	<b>0.979</b>
	Case 2	MPSNR	14.88	25.36	35.82	27.41	36.77	36.79	<b>37.40</b>
		MSSIM	0.206	0.598	0.923	0.812	0.938	0.930	<b>0.946</b>
		MFSIM	0.597	0.839	0.969	0.929	0.969	0.973	<b>0.976</b>
	Case 3	MPSNR	23.70	35.05	36.38	36.80	37.30	37.38	<b>37.95</b>
		MSSIM	0.486	0.922	0.932	0.947	0.950	0.939	<b>0.953</b>
		MFSIM	0.767	0.959	0.972	0.976	0.975	0.975	<b>0.979</b>
	Case 4	MPSNR	23.87	35.90	36.47	37.36	37.35	37.48	<b>38.19</b>
		MSSIM	0.490	0.928	0.933	0.951	0.950	0.940	<b>0.954</b>
		MFSIM	0.769	0.962	0.972	0.977	0.976	0.975	<b>0.979</b>
	Case 5	MPSNR	20.62	30.01	35.54	31.78	36.78	37.05	<b>37.62</b>
		MSSIM	0.385	0.760	0.920	0.883	0.939	0.936	<b>0.951</b>
		MFSIM	0.707	0.896	0.966	0.947	0.969	0.975	<b>0.977</b>

the MPSNR, MSSIM, and MFSIM which computed by the mean of all bands are utilized to measure the performance, with higher MPSNR, MSSIM and MFSIM reflecting better denoising and restoration performance.

The quantitative results of all compared methods in Washington DC Mall and Pavia University data on different cases are presented in Table 1. The highest results for each quality index are labeled in bold. From Table 1, it is clear that the proposed F-LRNMF obtains the best results over the other compared methods under all cases, which confirms the advantage of our proposed method over others. It is worth noting that the proposed method achieves nearly 2 dB improvement than the second-best results in Washington DC Mall data. We further calculate the PSNR, SSIM, and FSIM values of different bands in all simulated data cases. Figs. 3–4 show the quantitative values of each band for Washington DC Mall and Pavia University in five simulated noise cases, respectively. As displayed in Fig. 3, it is observed that the proposed method performs higher PSNR, SSIM, and FSIM values than other methods for almost all bands in Washington DC Mall data. For Pavia University data, the proposed F-LRNMF also can achieve the better value than other ones in most of the bands as shown in Fig. 4.

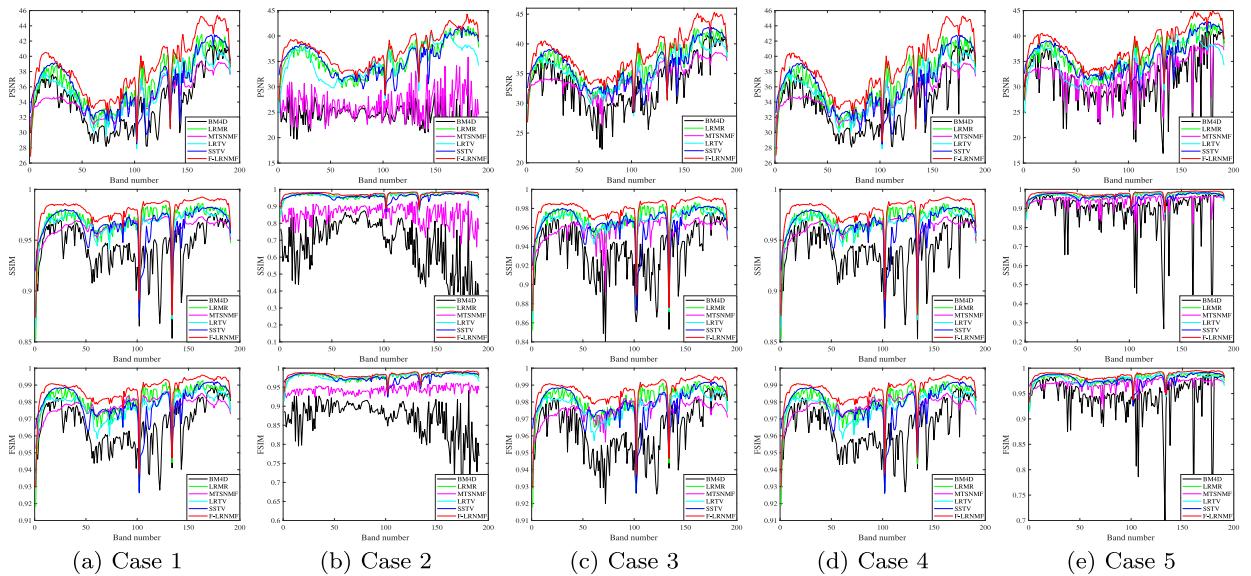
Moreover, to further compare the performances of the six compared methods, we show the spectral signature between the clear image with the denoised image. Figs. 5–6 show the comparisons of spectral signature of different compared methods on Washington DC Mall pixel (200,50) and Pavia University pixel (200,250) in Case 5, respectively. These figures give the spectral reflectance difference curves between the reconstructed pixel and the original noise-free pixel. In Figs. 5–6(a), the noisy spectral signatures are greatly degraded by the mixture noise. After denoising, the fluctuations are mainly suppressed. Comparing with the original curves (red curve in the figure), we can find that the compared methods fail to reconstruct the original spectral signature, which indicates that the denoised images obtained by these methods left in some noise or cause the undesirable distortion. In contrast, as displayed in Figs. 5–6(g), it can be observed that the spectral signatures produced by the proposed F-LRNMF method perform best and just hold the almost same curves with original image curves, which illustrates that the spectral distortion caused by our method is less than others. In summary, the proposed method performs the better restoration results among compared methods in terms of visual quality and quantitative indices.

## 5.2. Real data experiments

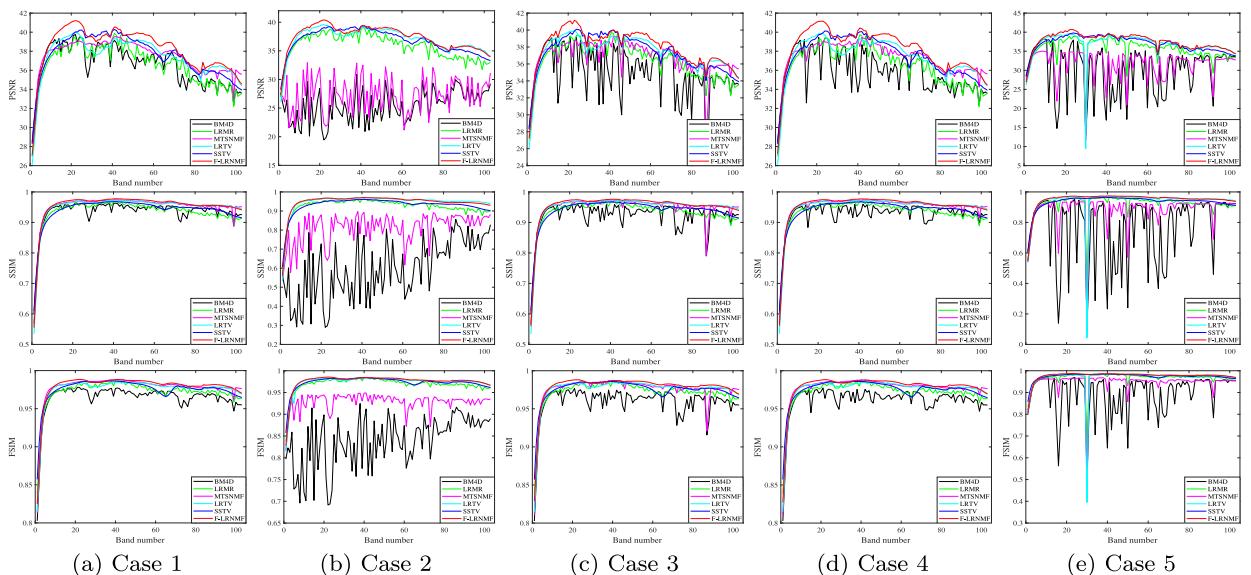
To show the proposed method can be applied to real HSI data restoration, three real HSI data sets are employed to evaluate the performance, including the HYDICE Urban data<sup>1</sup>, the AVIRIS Indian Pines data<sup>2</sup>, and the EO-1 Hyperion Australia

<sup>1</sup> <http://www.tec.army.mil/hypercube>

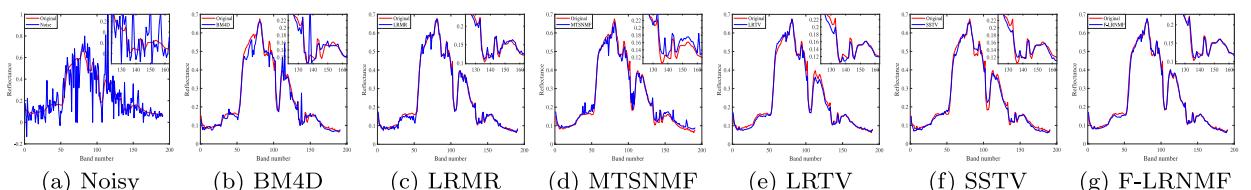
<sup>2</sup> <https://engineering.purdue.edu/~biehl/MultiSpec/hyperspectral.html>



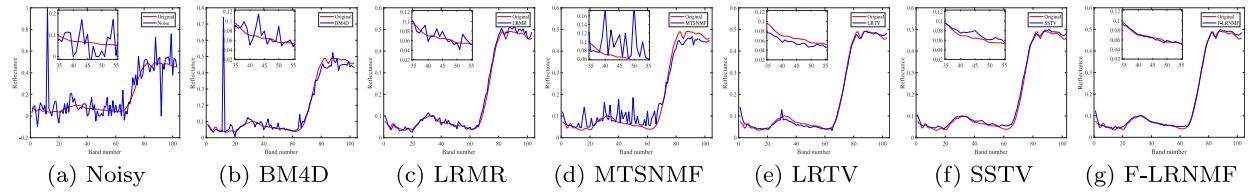
**Fig. 3.** Each column shows the PSNR, SSIM, and FSIM values of each band of all of the methods under different noise cases in Washington DC Mall: (a) case 1, (b) case 2, (c) case 3, (d) case 4, (e) case 5. Please zoom in for better viewing.



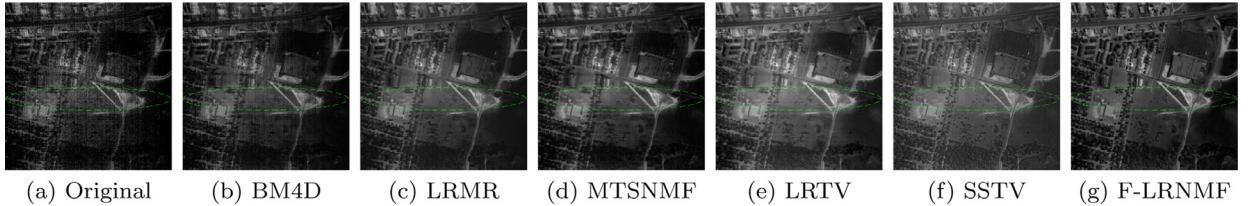
**Fig. 4.** Each column shows the PSNR, SSIM, and FSIM values of each band of all of the methods under different noise cases in Pavia University: (a) case 1, (b) case 2, (c) case 3, (d) case 4, (e) case 5. The figure is better viewed in a zoomed-in PDF.



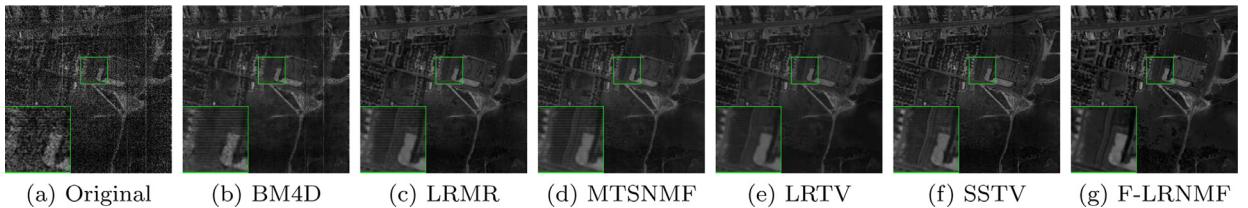
**Fig. 5.** The reflectance of special pixel in Washington DC Mall under mixture noise case 5. Pixel (200,50): (a) Noisy, (b) BM4D, (c) LMRM, (d) MTSNMF, (e) LRTV, (f) SSTV, (g) F-LRNMF. The subfigure represents the local amplification in the figure. Please zoom in for better viewing.



**Fig. 6.** The reflectance of special pixel in Pavia University under mixture noise case 5. Pixel (200,250): (a) Noisy, (b) BM4D, (c) LRMR, (d) MTSNMF, (e) LRTV, (f) SSTV, (g) F-LRNMF. The subfigure represents the local amplification in the figure. Please zoom in for better viewing.



**Fig. 7.** Denoised results of band 109 in Urban HSI data: (a) Original band 109, the results of (b) BM4D, (c) LRMR, (d) MTSNMF, (e) LRTV, (f) SSTV, (g) F-LRNMF. Please zoom in for better viewing.



**Fig. 8.** Denoised results of band 207 in Urban HSI data: (a) Original band 207, the results of (b) BM4D, (c) LRMR, (d) MTSNMF, (e) LRTV, (f) SSTV, (g) F-LRNMF. Please zoom in for better viewing.

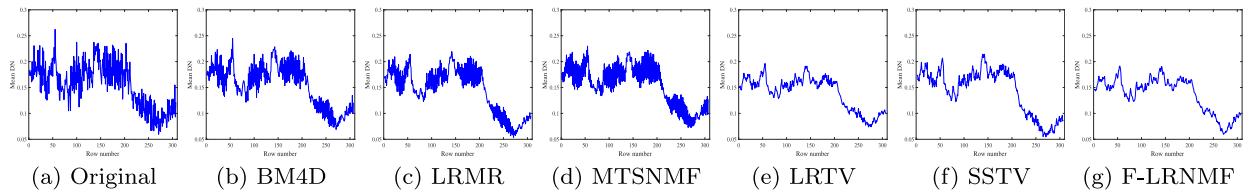
datasets<sup>3</sup>. In real data experiments, the same as simulated experiments, the parameters choice of compared methods has been finely tuned by the related references' suggestion to perform a possibly well visual performance. For the parameters of the proposed method, we will give it in Section 5.3. Moreover, before testing the restoration experiment, the gray values of each image band are scaled into [0,1].

(1) HYDICE Urban data set: The first Urban data set is acquired by the HYDICE sensor with the size of  $307 \times 307 \times 210$ , and it is seriously degraded by mixture noise (stripes, deadlines, atmosphere, water absorption, and other unknown). Although some bands are largely polluted by the atmosphere and water absorption and cannot provide any useful information, we still use all bands in the experiment to show the robustness of the proposed method.

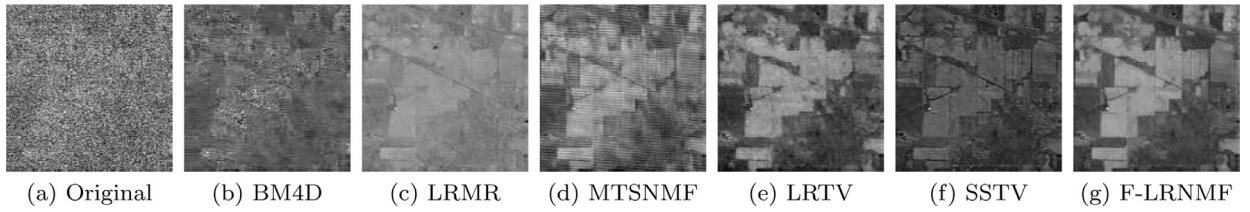
Figs. 7–8 show the restoration results of band 109 and band 207 for six methods, respectively. From Figs. 7–8(a), we can find that original images are degraded by mixture noise, such as the stripe noise, deadline, and random noise. After denoising, BM4D method can remove random noise, but it fails to remove sparsity noise, including stripe noise and deadline noise as shown in Figs. 7–8(b). The reason is that BM4D mainly assumes that the noise is Gaussian noise. Thus once the image suffers from heavy stripe noise and deadline shown in Figs. 7–8(a), BM4D method fails to restore well. MTSNMF also only considers the Gaussian noise in hyperspectral image denoising task, thus the stripe noise cannot be removed shown in Figs. 7–8(d). Moreover, it is clear to see that nuclear norm based low-rank matrix methods LRMR and LRTV can effectively remove random noise and deadline, but they cannot remove dense stripes shown in Figs. 7–8(c) and (e), respectively. As seen in Fig. 7(f), SSTV can remove mixture noise in the original image, but image contrast has changed, which results in image distortion. In Fig. 8(f), the stripe noise still exist in the restoration image. The proposed F-LRNMF method, in contrast, can effectively remove all of the mixture noise and preserve more edges and details as shown in Figs. 7–8(g). This is mainly because the nonnegative matrix factorization depicts the low-rank better than the nuclear norm to the clear HSI, and the framelet regularization can efficiently preserve the detail and texture information.

Since without ground-truth image as the reference, we show the quantitative comparison. Fig. 9 displays the horizontal mean profiles of band 207 in this data as an example. The horizontal axis denotes the row number, and the vertical axis is the mean digital number value of each row in the figure. There are rapid fluctuations in the curve shown in Fig. 9(a), since the mixture noise, especially the stripe noise, exist in the image shown in Fig. 8(a). From Fig. 9(b)–(d), it is observed that BM4D, LRMR, and MTSNMF fail to remove stripes in the horizontal direction. After restoration, LRTV and SSTV can suppress

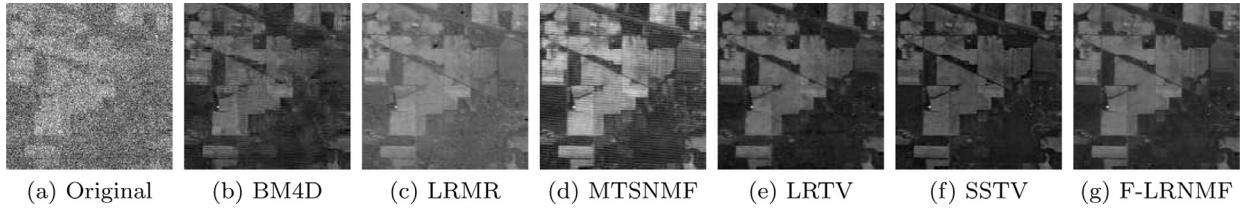
<sup>3</sup> <http://remote-sensing.nci.org.au/>



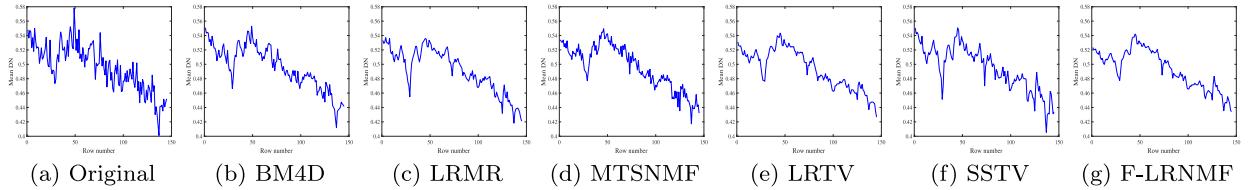
**Fig. 9.** Horizontal mean profiles of band 207 in Urban HSI data: (a) Original band 207, the results of (b) BM4D, (c) LRMR, (d) MTSNMF, (e) LRTV, (f) SSTV, (g) F-LRNMF. Please zoom in for better viewing.



**Fig. 10.** Denoised results of band 150 in Indian Pines HSI data: (a) Original band 150, the results of (b) BM4D, (c) LRMR, (d) MTSNMF, (e) LRTV, (f) SSTV, (g) F-LRNMF. Please zoom in for better viewing.



**Fig. 11.** Denoised results of band 220 in Indian Pines HSI data: (a) Original band 220, the results of (b) BM4D, (c) LRMR, (d) MTSNMF, (e) LRTV, (f) SSTV, (g) F-LRNMF. Please zoom in for better viewing.



**Fig. 12.** Horizontal mean profiles of band 220 in Indian Pines HSI data: (a) Original band 220, the results of (b) BM4D, (c) LRMR, (d) MTSNMF, (e) LRTV, (f) SSTV, (g) F-LRNMF. Please zoom in for better viewing.

the fluctuations in Fig. 9(e)–(f), respectively. In Fig. 9(g), it is clear that the proposed F-LRNMF performs smoother curve, indicating that F-LRNMF can effectively remove all of the noise in the image. This is in accordance with the visual result presented in Fig. 8(g).

(2) *AVIRIS Indian Pines dataset*: The second data was acquired by the NASA AVIRIS instrument over the Indian Pines test site in Northwestern Indiana in 1992, and the size is  $145 \times 145$  with 220 bands. This data is polluted by the mixture of Gaussian noise, salt and pepper noise, and unknown noise. Although some bands cannot provide any useful information, we also remain these bands as the test data, which makes the restoration experiment more difficult.

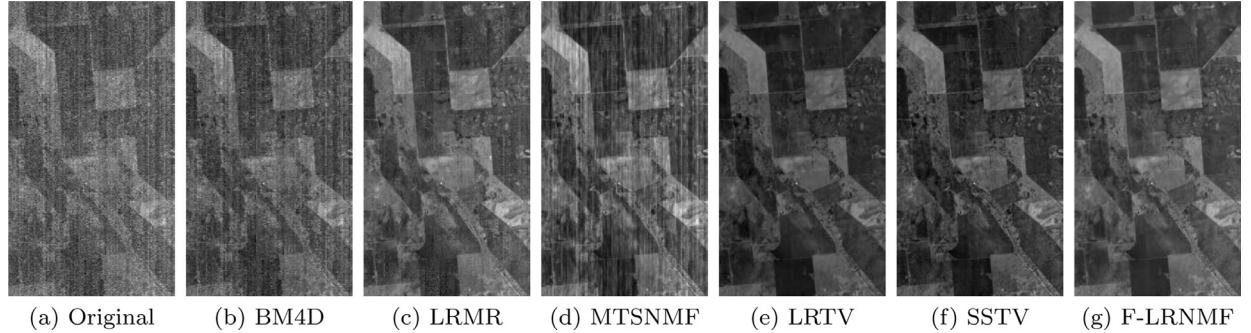
Figs. 10–11 show the restoration images of band 150 and 220 by different methods. Fig. 10(a) shows that the image is completely covered by noise, and the texture and detail are completely invisible. BM4D completely fails to restore the image shown in Fig. 10(b), since it is not suitable for the removal of heavy impulse noise. LRMR can remove the noise shown in Fig. 10(c), but some texture details are lost and distortion in the restored result. MTSNMF, LRTV, and SSTV can restore the image texture, but they cannot remove other noise and cause artifacts shown in Figs. 10(c)–(f). As to Fig. 11, although the image is seriously degraded by noise, the image texture can be seen. The denoising results obtained by the compared methods are not complete, and some texture details are lost in the restored results shown in Figs. 11(b)–(f). As display in Figs. 10–11(g), the performance of F-LRNMF can be easily observed in both detail and texture preservation and all of the noise is removed, which again indicates the superiority of the proposed method than other methods.

We also show the horizontal mean profile of band 220 before and after restoration as an example. From the result shown in Fig. 12, it can be clearly found that the proposed F-LRNMF method gives the better curve among the restored

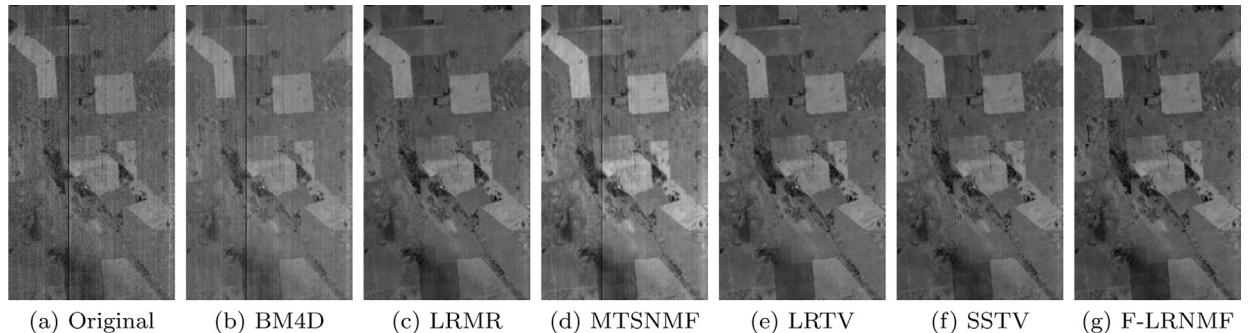
**Table 2**

No-reference HSI quality assessment on the AVIRIS Indian Pines data.

Method	BM4D	LRMR	MTSNMF	LRTV	SSTV	F-LRNMF
Score	14.24	16.74	16.68	16.38	16.51	16.11



**Fig. 13.** Denoised results of band 51 in Hyperion Australia HSI data: (a) Original band 51, the results of (b) BM4D, (c) LRMR, (d) MTSNMF, (e) LRTV, (f) SSTV, (g) F-LRNMF. Please zoom in for better viewing.



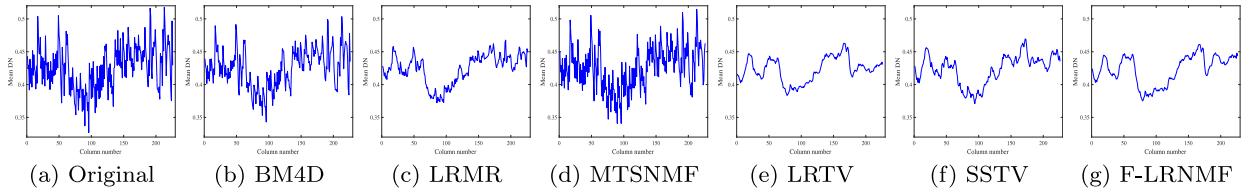
**Fig. 14.** Denoised results of band 123 in Hyperion Australia HSI data: (a) Original band 123, the results of (b) BM4D, (c) LRMR, (d) MTSNMF, (e) LRTV, (f) SSTV, (g) F-LRNMF. Please zoom in for better viewing.

horizontal mean profiles, because the proposed method performs noticeable smoother curve, while the compared methods existing some fluctuations in their results. Moreover, we use a no-reference HSI quality assessment, as introduced in [57], to evaluate the restored images. In general, a lower assessment score indicates the better restored result. As shown in Table 2, BM4D method obtains the lowest score, but the image structure information is destroyed in some bands (see Figs. 10(b)). Comparing with other optimization based methods, the proposed method achieves the lower score, which further demonstrates the efficiency of our method.

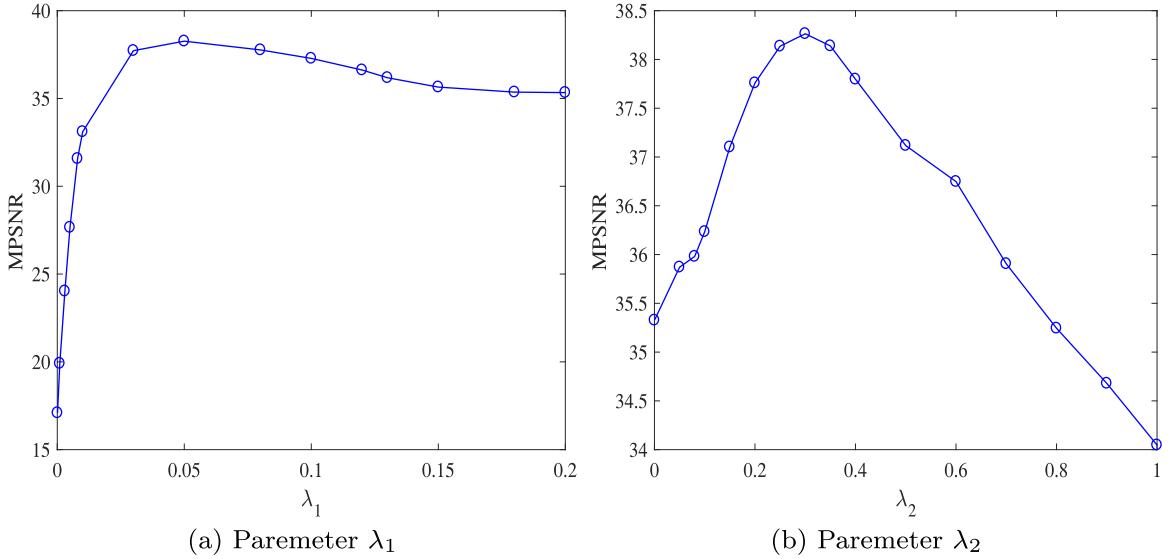
(3) *EO-1 Hyperion Australia dataset*: The EO-1 Hyperion Australia data set is employed to our third real data experiment. The original image was captured on December 4, 2010 with the location at Australia, and the size of this HSI is  $3858 \times 256 \times 242$ . Following [21], the overlapping bands between visual near-infrared and shortwave infrared ranges are removed, and we select a subregion of size  $400 \times 226 \times 150$  in our restoration.

Figs. 13–14 show the restoration results of band 51 and band 123 by different methods. Similarly the above experiments, BM4D and MTSNMF cannot work for sparsity noise, especially stripe noise and deadline. LRMR can partly deal with most of the noise, but some noises also exist in the results shown in Figs. 13–14(c). LRTV and SSTV can remove all of the noise, since these two methods use TV regularization prior to the image component, and this data possesses the piecewise smooth structure in the spatial domain. Comparing with TV regularization, the framelet regularization can better preserve the detail and texture information, and we can see the results obtained by the proposed F-LRNMF method shown in Figs. 13–14(g). Moreover, the vertical mean profiles of band 51 before and after restoration are shown in Fig. 15, and it can be seen that LRTV and our method obtain evidently smoother curve than others.

In summary, extensive experiments on both simulated and real HSI data show that the proposed F-LRNMF method outperforms other popular methods because the mixture noise is completely removed and more useful structures and details are preserved.



**Fig. 15.** Horizontal mean profiles of band 51 in Hyperion Australia HSI data: (a) Original band 51, the results of (b) BM4D, (c) LRMN, (d) MTSNMF, (e) LRTV, (f) SSTV, (g) F-LRNMF. Please zoom in for better viewing.



**Fig. 16.** The MPSNR curves as function of the regularization parameters: (a) the relationship between MPSNR and the parameter  $\lambda_1$ , (b) the relationship between PSNR and the parameter  $\lambda_2$ .

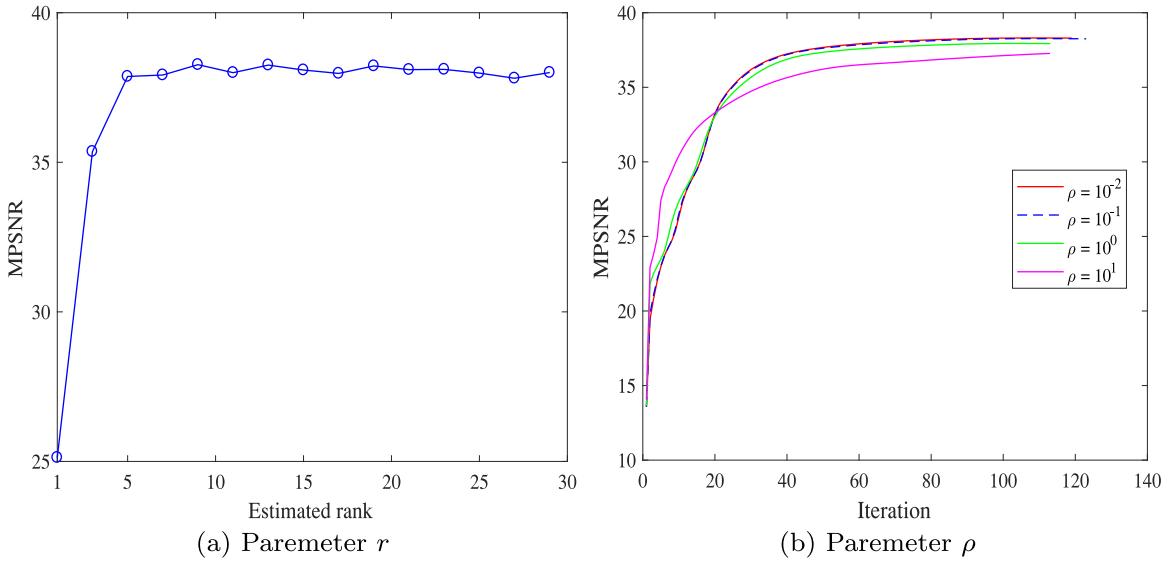
### 5.3. Discussions

In this section, we will analyze the sensitivity of parameters of the proposed F-LRNMF method. There are five parameters involved in F-LRNMF, including two regularization parameters  $\lambda_1$  and  $\lambda_2$ , the proximal parameter  $\rho$ , the penalty parameter  $\beta$ , and the estimated rank parameter  $r$ . It is worth noting that to simplify the numbers of the parameter adjustment, we empirically set  $\beta = 200\lambda_2$  in all experiments. In the following, we need to consider how to choose other four parameters, and we shall show the process for selecting such parameters in our experiments. To evaluate and analyze the impact and optimal values of these parameters, we employ simulated Washington DC Mall case 5 as an example and use the MPSNR values as the evaluation measure in all of the analysis. Moreover, we display the convergence curves of the proposed solved algorithm from numerical experiment perspective, though we have theoretically analyzed the convergence in Section 4.2. Finally, we will report the running time for the three real data to demonstrate the effectiveness of the proposed method.

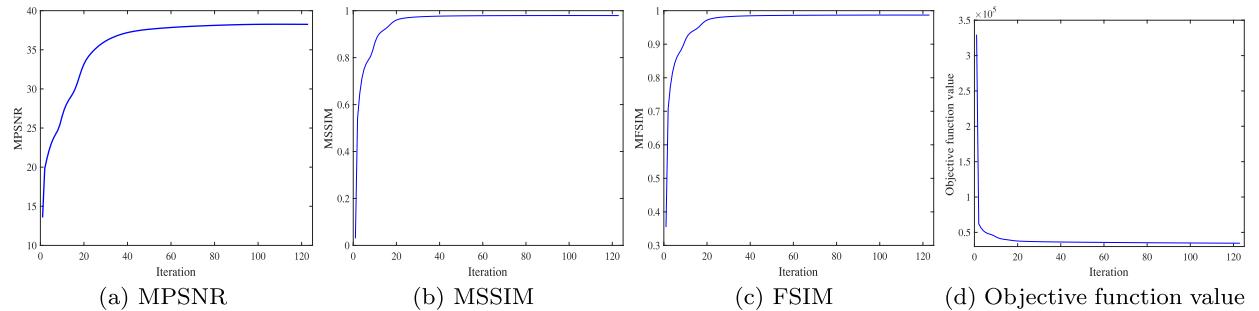
(1) *Sensitivity analysis of parameter  $\lambda_1$ :* The parameter  $\lambda_1$  is involved in the model (8), which is used to control the sparsity of the sparse noise. The experimental results of MPSNR values as the function of the regularization parameters  $\lambda_1$  is displayed in Fig. 16(a). From the Figure to see, we can find that MPSNR values achieve significant improvement when  $\lambda_1$  is changed from 0 to 0.05. Following  $\lambda_1$  further increasing, we also see that MPSNR emerges a slight reduction, and the highest MPSNR value is achieved with  $\lambda_1$  equal 0.05. In general, the result of MPSNR remains relatively stable with the value of  $\lambda_1$  is near 0.05. Therefore, we empirically fix the parameter  $\lambda_1 = 0.05$  in all simulated experiments. In real data experiments, we can select  $\lambda_1$  with the near of 0.05 to obtain the best restoration results.

(2) *Sensitivity analysis of parameter  $\lambda_2$ :* The parameter  $\lambda_2$  is used to balance the factor regularization in the framelet domain. When  $\lambda_2$  is larger, the sparsity degree of the image in the framelet domain is stronger, and vice versa. The F-LRNMF solver is low-rank nonnegative matrix factorization if  $\lambda_2$  is set to zero. Fig. 16(b) shows the PSNR value performance versus parameter  $\lambda_2$ , from which we can observe that the best restoration result is  $\lambda_2 = 0.3$ . Thus, we suggest that we can choose it with the scope of 0.1 to 0.5 in all of the experiments.

(3) *Sensitivity analysis of estimated rank  $r$ :* In our F-LRNMF optimization, we use the low-rank nonnegative matrix to depict low-rank prior of the clean HSI. Thus the estimated rank should be given for initializing the  $A$  and  $X$  before performing Algorithm 1. The MPSNR curve of F-LRNMF solver with the varied rank-constrained values is shown in Fig. 17(a). It can be



**Fig. 17.** The MPSNR curves as function of the regularization parameters: (a) the relationship between MPSNR and the parameter  $r$ , (b) the relationship between PSNR and the parameter  $\rho$ .



**Fig. 18.** The MPSNR, MSSIM, MFSIM and objective function values curves versus the iteration number of F-LRNMF: (a) change in the MPSNR value, (b) change in the MSSIM value, (c) change in the FSIM value, (d) objective function value.

clearly observed that the MPSNR is not sensitive to estimated rank when  $r$  is not less than 7. In practice, the performance complies with the intuition that the estimated rank is related to the number of endmembers. From Figure to see, the values of MPSNR are basically equal when  $r = 5$  and  $r = 29$ , since  $r = 29$  indicates that the information of matrix  $A$  is redundancy. However, it requires an additional amount of computation to perform [Algorithm 1](#) when the rank  $r$  (i.e., the column number of matrix  $A$ ) is relatively large.

Though the MPSNR results are not very sensitive to  $r$ , we also need to give a strategy to determine the estimated rank. In our simulated experiments, since existing the ground-truth data, thus we select the rank to equal the number of singular values larger than 1% of the larger one. This method of approximating low-rank is also used in the matrix factorization for low-rank tensor completion [42]. In real data experiments, we cannot estimate the rank following simulated experiments since without the ground-truth data. In [20,28], the authors use the HSI subspace estimation method (e.g., HySime [58]) to estimate the desired rank value. Therefore, we also apply this method in our real data experiments to determine the rank  $r$ .

(4) *Sensitivity analysis of proximal parameter  $\rho$ :* In our optimization process, we introduce the proximal term to guarantee the convergence of the proposed algorithm. The proximal parameter can influence the performance and the rate of convergence of the proposed method. [Fig. 17\(b\)](#) displays the MPSNR value with respect to iteration number for different proximal parameters. The plot suggests that the proposed method can perform the similar MPSNR and iteration number for  $\rho = \{10^{-2}, 10^{-1}\}$ . With increasing the proximal parameter, the iteration number is reduced, but the performance also can get relatively poor. Therefore, we fix  $\rho = 0.1$  in our experiments.

(5) *Numerical convergence of the F-LRNMF solver:* In [Section 4.2](#), we have illustrated the convergence of the proposed method in theoretical analysis. In the next, we will show the numerical convergence curve in our experiments process. [Fig. 18](#) displays the MPSNR, MSSIM, MFSIM and objective function value versus the iteration number of the F-LRNMF solver. As display in [Fig. 18](#), we can observe that the MPSNR, MSSIM, MFSIM and objective function value all converge to the optimal solutions with the increasing of iterations, which indicates the convergence property of the F-LRNMF method.

**Table 3**

Comparison of running time (in second) on three real data by using six compared methods.

Data	BM4D	LRMR	MTSNMF	LRTV	SSTV	F-LRNMF
Urban data	2529.75	<b>112.50</b>	372.41	410.11	802.79	<u>219.98</u>
Indian Pines data	247.65	<b>27.60</b>	133.01	84.39	187.37	<u>53.17</u>
Hyperion Australia data	1169.04	<b>75.90</b>	523.59	275.08	553.36	<u>107.57</u>

**Table 4**

The parameters selection in simulated and real experiments.

Data	$\lambda_1$	$\lambda_2$	$\rho$	$\beta$	$r$
Simulated experiments data	0.05	0.3	0.1	$200\lambda_2$	9
Urban data	0.01	0.5	0.1	$200\lambda_2$	8
Indian Pines data	0.01	0.1	0.1	$200\lambda_2$	14
Hyperion Australia data	0.05	0.3	0.1	$200\lambda_2$	27

(6) *Running time*: To show the efficiency of the proposed method, we report the running time of six compared method on three real data experiments in **Table 3**. From **Table 3**, we can observe that LRMR is the fastest method than other approaches in three cases, the reason is that we perform LRMR with SSGoDec algorithm rather than GoDec algorithm, which also can avoid SVD computation. Moreover, LRMR is utilized to restore each patch sequentially. However, its restoration results are limited. Compared with other four methods, the proposed method obtains the shortest time. This mainly due to the fact that we use low-rank nonnegative matrix factorization for the clean HSI and apply the regularizer to the reduced-dimensionality factor rather than to image itself.

In summary, the proposed method is relatively insensitive to most of the parameters involving in [Algorithm 1](#). Moreover, the convergence of the proposed method can be theoretical and numerical guaranteed, and the implementation efficiency also can be guaranteed. The parameter selection of our proposed F-LRNMF method tested for simulated and real data experiments is presented in **Table 4**.

## 6. Conclusions

In this paper, we have proposed a novel framelet-regularized low-rank nonnegative matrix factorization approach to address the HSI restoration problem. In particular, the low-rank nonnegative matrix factorization is utilized to depict the spectral correlation among the HSI bands, which can avoid expensive SVD computation and make the dimension of HSI decrease. The framelet regularization can depict the extra spatial structure prior to preserve the details and geometric features of HSI. Moreover, we use commonly  $\ell_1$  norm to constrain the sparse noise including impulse noise, deal lines, stripe noise, and many unknown noises, which can efficiently separate the sparse noise from the clean HSI. An efficient optimization solver, BSUM, is developed to solve the proposed model. Both the qualitative and quantitative assessments (MPSNR, MSSIM, MFSIM, visual inspection, and mean profile) of simulated and real data experiments have indicated the proposed F-LRNMF method can remove the various mixture noise effectively, as well as preserve the spatial structure, comparing with other state-of-the-art methods.

In our future work, we will try to incorporate more useful factor prior terms into our low-rank nonnegative matrix factorization model to further enhance its capability for HSI restoration. Moreover, regarding the HSI as a third-order tensor, the low-rank tensor factorization will be considered in the future work. Besides, adaptively determining the regularization parameters should be improved.

## Acknowledgments

The authors would like to express their great thankfulness to the editor and reviewers for their much helpful suggestions for revising this paper. The authors would like to thank the authors [12,16,25] for providing the free download of the source code and Prof. F. Fang providing the code for framelet regularization. We would also like to thank Dr. W. He for providing the source codes for LRMR [19] and LRTV [20] and discussions. This work was supported by the NSFC (61772003, 61702083) and the Fundamental Research Funds for the Central Universities (ZYGX2016J132, ZYGX2016KYQD142).

## References

- [1] S.A. Lewis, A.T. Hudak, R.D. Ottmar, P.R. Robichaud, L.B. Lentile, S.M. Hood, J.B. Cronan, P. Morgan, Using hyperspectral imagery to estimate forest floor consumption from wildfire in boreal forests of Alaska, USA, *Int. J. Wildland Fire* 20 (2) (2011) 255–271.
- [2] K.C. Tiwari, M.K. Arora, D. Singh, An assessment of independent component analysis for detection of military targets from hyperspectral images, *Int. J. Appl. Earth Observ. Geoinf.* 13 (5) (2011) 730–740.
- [3] J. Xia, P. Ghamisi, N. Yokoya, A. Iwasaki, Random forest ensembles and extended multextinction profiles for hyperspectral image classification, *IEEE Trans. Geosci. Remote Sens.* 56 (1) (2018) 202–216.
- [4] M.D. Iordache, J. Bioucas-Dias, A. Plaza, Total variation spatial regularization for sparse hyperspectral unmixing, *IEEE Trans. Geosci. Remote Sens.* 50 (11) (2012) 4484–4502.

- [5] Y. Tarabalka, J. Chanussot, J.A. Benediktsson, Segmentation and classification of hyperspectral images using watershed transformation, *Pattern Recognit.* 43 (7) (2010) 2367–2379.
- [6] N. Yokoya, A. Iwasaki, Object detection based on sparse representation and hough voting for optical remote sensing imagery, *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* 8 (5) (2015) 2053–2062.
- [7] M. Elad, M. Aharon, Image denoising via sparse and redundant representations over learned dictionaries, *IEEE Trans. Image Process.* 15 (12) (2006) 3736–3745.
- [8] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Image denoising by sparse 3-D transform-domain collaborative filtering, *IEEE Trans. Image Process.* 16 (8) (2007) 2080–2095.
- [9] A.A. Green, M. Berman, P. Switzer, M.D. Craig, A transformation for ordering multispectral data in terms of image quality with implications for noise removal, *IEEE Trans. Geosci. Remote Sens.* 26 (1) (1988) 65–74.
- [10] H. Othman, S.E. Qian, Noise reduction of hyperspectral imagery using hybrid spatial-spectral derivative-domain wavelet shrinkage, *IEEE Trans. Geosci. Remote Sens.* 44 (2) (2006) 397–408.
- [11] G. Chen, S.E. Qian, Simultaneous dimensionality reduction and denoising of hyperspectral imagery using bivariate wavelet shrinking and principal component analysis, *Can. J. Remote Sens.* 34 (5) (2008) 447–454.
- [12] M. Ye, Y. Qian, J. Zhou, Multitask sparse nonnegative matrix factorization for joint spectral-spatial hyperspectral imagery denoising, *IEEE Trans. Geosci. Remote Sens.* 53 (5) (2015) 2621–2639.
- [13] Y.Q. Zhao, J. Yang, Hyperspectral image denoising via sparse representation and low-rank constraint, *IEEE Trans. Geosci. Remote Sens.* 53 (1) (2015) 296–308.
- [14] J. Yang, Y.Q. Zhao, C.W. Chan, S.G. Kong, Coupled sparse denoising and unmixing with low-rank constraint for hyperspectral image, *IEEE Trans. Geosci. Remote Sens.* 54 (3) (2016) 1818–1833.
- [15] G. Chen, S.E. Qian, Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage, *IEEE Trans. Geosci. Remote Sens.* 49 (3) (2011) 973–980.
- [16] M. Maggioni, V. Katkovnik, K. Egiazarian, A. Foi, Nonlocal transform-domain filter for volumetric data denoising and reconstruction, *IEEE Trans. Image Process.* 22 (1) (2013) 119–133.
- [17] A. Karami, M. Yazdi, A.Z. Asli, Noise reduction of hyperspectral images using kernel non-negative tucker decomposition, *IEEE J. Sel. Top. Signal Process.* 5 (3) (2011) 487–493.
- [18] X. Guo, X. Huang, L. Zhang, L. Zhang, Hyperspectral image noise reduction based on rank-1 tensor decomposition, *J. Photogramm. Remote Sens.* 83 (9) (2013) 50–63.
- [19] 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) (2014) 4729–4743.
- [20] 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) (2016) 178–188.
- [21] 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.
- [22] 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.
- [23] 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 (1) (2017) 202–219.
- [24] Y. Chen, X. Cao, Q. Zhao, D. Meng, Z. Xu, Denoising hyperspectral image with non-i.i.d. noise structure, *IEEE Trans. Cybern.* 48 (3) (2018) 1054–1066.
- [25] H.K. Aggarwal, A. Majumdar, Hyperspectral image denoising using spatio-spectral total variation, *IEEE Geosci. Remote Sens. Lett.* 13 (3) (2016) 442–446.
- [26] C. Li, Y. Ma, J. Huang, X. Mei, J. Ma, Hyperspectral image denoising using the robust low-rank tensor recovery, *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 32 (9) (2015) 1604–1612.
- [27] H. Fan, Y. Chen, Y. Guo, H. Zhang, G. Kuang, Hyperspectral image restoration using low-rank tensor recovery, *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* 10 (10) (2017) 4589–4604.
- [28] 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. Topics Appl. Earth Observ. Remote Sens.* 11 (4) (2018) 1227–1243.
- [29] X.L. Zhao, W. Wang, T.Y. Zeng, T.Z. Huang, M.K. Ng, Total variation structured total least squares method for image restoration, *SIAM J. Sci. Comput.* 35 (6) (2013) B1304–B1320.
- [30] Y. Chen, T.Z. Huang, X.L. Zhao, L.J. Deng, J. Huang, Stripe noise removal of remote sensing images by total variation regularization and group sparsity constraint, *Remote Sens.* 9 (6) (2017) 559.
- [31] X.L. Zhao, F. Wang, M.K. Ng, A new convex optimization model for multiplicative noise and blur removal, *Siam J. Imaging Sci.* 7 (1) (2014) 456–475.
- [32] J.J. Mei, Y. Dong, T.Z. Huang, W. Yin, Cauchy noise removal by nonconvex admm with convergence guarantees, *J. Sci. Comput.* 74 (2) (2018) 1–24.
- [33] J. Liu, T.Z. Huang, X.G. Lv, S. Wang, High-order total variation-based poissonian image deconvolution with spatially adapted regularization parameter, *Appl. Math. Model.* 45 (2017) 516–529.
- [34] M. Simões, J. Bioucas-Dias, L.B. Almeida, J. Chanussot, A convex formulation for hyperspectral image superresolution via subspace-based regularization, *IEEE Trans. Geosci. Remote Sens.* 53 (6) (2015) 3373–3388.
- [35] L. Zhuang, J.M. Bioucas-Dias, Fast hyperspectral image denoising based on low rank and sparse representations, in: Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2016, pp. 1847–1850.
- [36] J.M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, J. Chanussot, Hyperspectral unmixing overview: geometrical, statistical, and sparse regression-based approaches, *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* 5 (2) (2012) 354–379.
- [37] H. Ji, C. Liu, Z. Shen, Y. Xu, Robust video denoising using low rank matrix completion, in: Proceedings of the IEEE Conference CVPR, 2010, pp. 1791–1798.
- [38] T. Zhou, D. Tao, Godec: randomized low-rank and sparse matrix decomposition in noisy case, in: Proceedings of the International Conference Machine Learning, 2011, pp. 1–8.
- [39] 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.
- [40] L. Sun, B. Jeon, Y. Zheng, Z. Wu, A novel weighted cross total variation method for hyperspectral image mixed denoising, *IEEE Access* 5 (2017) 27172–27188.
- [41] Z. Wen, W. Yin, Y. Zhang, Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm, *Math. Prog. Comp.* 4 (4) (2012) 333–361.
- [42] Y. Xu, R. Hao, W. Yin, Z. Su, Parallel matrix factorization for low-rank tensor completion, *Inverse Pro. Imaging* 9 (2) (2015) 601–624.
- [43] T.Y. Ji, T.Z. Huang, X.L. Zhao, T.H. Ma, G. Liu, Tensor completion using total variation and low-rank matrix factorization, *Inf. Sci.* 326 (2016) 243–257.
- [44] W. He, H. Zhang, L. Zhang, Total variation regularized reweighted sparse nonnegative matrix factorization for hyperspectral unmixing, *IEEE Trans. Geosci. Remote Sens.* 55 (99) (2017) 1–13.
- [45] G. Zhang, Y. Xu, F. Fang, Framelet-based sparse unmixing of hyperspectral images, *IEEE Trans. Image Process.* 25 (4) (2016) 1516–1529.
- [46] J.F. Cai, H. Ji, C. Liu, Z. Shen, Framelet-based blind motion deblurring from a single image, *IEEE Trans. Image Process.* 21 (2) (2012) 562–572.
- [47] T.X. Jiang, T.Z. Huang, X.L. Zhao, T.Y. Ji, L.J. Deng, Matrix factorization for low-rank tensor completion using framelet prior, *Inf. Sci.* 436 (2018) 403–417.
- [48] Y. Chang, H. Fang, L. Yan, H. Liu, Robust destriping method with unidirectional total variation and framelet regularization, *Opt. Expr.* 21 (20) (2013) 23307–23323.

- [49] R.H. Chan, S.D. Riemenschneider, L. Shen, Z. Shen, Tight frame: an efficient way for high-resolution image reconstruction, *Appl. Comput. Harmon. Anal.* 17 (1) (2004) 91–115.
- [50] M. Razaviyayn, M. Hong, Z.Q. Luo, A unified convergence analysis of block successive minimization methods for nonsmooth optimization, *SIAM J. Optim.* 23 (2) (2013) 1126–1153.
- [51] Y. Chen, T.Z. Huang, L.J. Deng, X.L. Zhao, M. Wang, Group sparsity based regularization model for remote sensing image stripe noise removal, *Neurocomputing* 267 (6) (2017) 95–106.
- [52] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, *Found. Trends Mach. Learn.* 3 (1) (2010) 1–122.
- [53] L.J. Deng, T.Z. Huang, X.L. Zhao, A directional global sparse model for single image rain removal, *Appl. Math. Model.* 59 (2018) 662–679.
- [54] D.L. Donoho, De-noising by soft-thresholding, *IEEE Trans. Inf. Theory* 41 (3) (1995) 613–627.
- [55] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, *IEEE Trans. Image Process.* 13 (4) (2004) 600–612.
- [56] L. Zhang, L. Zhang, X. Mou, D. Zhang, FSIM: a feature similarity index for image quality assessment, *IEEE Trans. Image Process.* 20 (8) (2011) 2378–2386.
- [57] J. Yang, Y. Zhao, C. Yi, C.W. Chan, No-reference hyperspectral image quality assessment via quality-sensitive features learning, *Remote Sens.* 9 (4) (2017) 305.
- [58] J.M. Bioucas-Dias, J.M. Nascimento, Hyperspectral subspace identification, *IEEE Trans. Geosci. Remote Sens.* 46 (8) (2008) 2435–2445.