A Novel Weighted Cross Total Variation Method for Hyperspectral Image Mixed Denoising

LE SUN[1,2], (Member, IEEE), BYEUNGWOO JEON[2], (Senior Member, IEEE),
YUHUI ZHENG[1], (Member, IEEE), AND ZEBIN WU[3,4], (Member, IEEE)
[1]Jiangsu Engineering Center of Network Monitoring, School of Computer and Software, Nanjing University of Information
Science and Technology, Nanjing 210044, China
[2]School of Electronic and Electrical Engineering, Sungkyunkwan University, Suwon 440746, South Korea
[3]School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China
[4]Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks, Nanjing 210003, China

Corresponding author: Byeungwoo Jeon (bjeon@skku.edu)

ABSTRACT Due to atmospheric effects and secondary illumination, hyperspectral images (HSIs) usually
suffer from system noises, stripes, and dead pixels, which greatly degrade the imaging quality and limit the
precision of the subsequent processing. In this paper, a novel HSI mixed denoising method based on 3-D
spectral–spatial cross total variation (TV) is proposed to overcome such problem. First, the HSI is treated
as a 3-D cube, and the TV with 2-D spatial directions on the spectral difference images, which could be
treated as cross TV of HSI cube, is minimized to enhance the spatial smoothness and exploit the spectral
redundancy and correlation. Second, an adaptive mechanism for calculating the spectral–spatial weights is
adopted to balance the fidelity term and the cross TV regularization according to different spatial structures.
Alternating direction method of multipliers is finally extended to solve the proposed model by separating
it into several simpler subproblems. Experimental results on simulated and real-HSI data sets validated the
effectiveness of the proposed method.

INDEX TERMS Hyperspectral denoising, mixed noise removal, cross total variation, ADMM.

I. INTRODUCTION

Hyperspectral images, as a three-dimensional data cube with hundreds or thousands of narrow spectral bands ranging from 0.4 to 2.5 μm, have been widely used in various fields [1], e.g., precise agriculture, mineral detection, environment monitoring, urban planning and so on. However, many factors degrade the quality of hyperspectral image in course of acquiring the data and severely limit the precision of subsequent processing, e.g., unmixing, classification [2] and target detection. Therefore it is critical to develop an effective preprocessing technique for HSI.

To date, a multitude of techniques have been developed for HSI restoration. Maximum noise fraction (MNF) [3] and noise-adjusted principal component (NAPC) [4] transformation are two famous conventional methods for noise reduction of HSI and have been embedded in commercial software, e.g., ENVI. By regarding each band of HSI as a gray-level image, powerful denoising methods, such as block-matching 3-D filtering [5], nonlocal based algorithm [6], can be applied to restore HSI data. However, this kind of method does not take the spectral correlations into consideration, and often results in a less competitive performance.

Due to the good properties of representing signals with few coefficients and the multi-scale analysis [7], wavelet based method is also a powerful instrument for HSI denoising. In [8] and [9], the combination of principle component analysis (PCA) and wavelet shrinkage, which utilizes 2D or 3D wavelet filter to remove the noise in the low-energy PCA channels, was firstly proposed for HSI denosing. Many other methods always combine wavelet transform with sparse regularizations, such as Lasso penalty [10] or l_1 penalty [11]. In [12], a wavelet-based sparse reduced-rank regression (WSRRR) method was put forward for hyperspectral restoration by minimizing the l_1 sparse regularization on

wavelet coefficient of the reduced-rank components. In [13], the 2-D Daubechies wavelet for spatial dimension and 1-D Fourier transform on spectral dimension was utilized for 3-D HSI denoising, which exploited the spatial-spectral information of HSI. The same idea was also applied in [14] using block matching 3-D wavelet filtering [15].

Low rank representation (LRR) is another powerful technique and has achieved good performance in the field of hyperspectral restoration. Lu *et al.* [16] proposed an LRR method for stripe noise removal in HSI by exploiting the high spectral correlation between different bands, and a graph regularization was constructed to preserve the intrinsic local structure. Zhang *et al.* [17] exploited an HSI restoration method based on low-rank matrix recovery (LRMR). He *et al.* [18] then improved the patchwise LRMR by a noise-adjusted iteration strategy, which makes LRMR adaptive for the noise intensity in different bands. Even LRMR was constructed on overlapped patches and achieved good performance for denoising Gaussian and sparse noise simultaneously, but it only considers the local similarity within patches. To overcome such drawback, Zhu *et al.* [19] extended the LRMR to nonlocal sense by a spectral nonlocal methods. Meanwhile, Wang *et al.* [20] proposed a group low rank representation (GLRR) for HSI denoising. GLRR enables the exploitation of both the local similarity within a patch and the nonlocal similarity across the patches in a group simultaneously and achieves better performance. In [21], the total variation regularization was combined with low rank decomposition for the restoration of HSI. In addition, sparse representation is one promising tool for image processing [22] as well as in HSI denoising. In [23], sparse coding based on learned dictionary was exploited to model the global redundancy and correlation (RAC) in spatial domain and local RAC in spectral domain. While in [24], a spectral-spatial adaptive sparse representation method was proposed to further generate better noise free estimation by jointly sparse coding on the shape adaptive local regions [25]. More literatures on low rank and sparse representation for HSI denoising can be found in [26]–[28] and therein references.

Total variation (TV) regularization, which has shown great potentials in natural image denoising [29]–[31], recently draws more attentions in HSI restoration, denoising, and inpainting. In [32], a spatial-spectral adaptive hyperspectral TV (SSAHTV) was proposed for noise removal of HSI, which can adaptively estimate the denoising strength according to different spatial properties and different noise intensity in bands. Due to being insensitive to image details of SSAHTV, Yuan *et al.* [33] employed a spectral-spatial kernel method to maintain the spectral correlations and preserve the spatial structures. Meanwhile, Cheng *et al.* [34] extended the SSAHTV to nonlocal sense, and presented a multichannel nonlocal TV regularization for hyperspectral inpainting. The problem was that it only joined all nonlocal gradients along the spectral dimension and ignored the redundancy information in bands. Then, Li *et al.* [35] calculated the nonlocal gradients in both spectral and spatial dimensions and put forward

a multidimensional nonlocal TV model for hyperspectral image recovery, which could explore more redundancy from the highly correlated bands and lead to a better performance. It is worth mentioning that these three local or nonlocal TV models only consider the gradients in local or nonlocal sense in spatial domain but ignore the spectral gradients, thus result in a non real 3-D TV model. A normal way to extend the conventional 2-D TV to a real 3-D cubic TV (CTV) is to impose the spectral gradients [36], in which CTV was defined along two spatial dimensions and one spectral dimension, thus leading to both spatial and spectral smooth results. Due to the smoothness along both horizontal and vertical directions in spatial domain, CTV does not work well in removing stripe noise or dead lines. By constraining the gradients in both the spatially horizontal and spectral directions, Chang *et al.* [37], [38] employed an anisotropic spectral-spatial total variation for multispectral image destriping and denoising. However, all the above mentioned TV based method can only remove one or two kinds of noise, and the ideal of real cubic TV has not been exhaustively built to explore spatial and spectral information completely.

In this paper, we propose a novel 3-D spatial-spectral TV regularization, called 3-D cross TV (3DCrTV), for simultaneously removing several types of noise, such as Gaussian noise, impulse noise, and stripes in HSI. The flowchart of the proposed method is shown in Figure 1. This 3-D cross TV consists of 2-D spatial TV on the spectral difference images (1-D spectral TV), which can explore the spectral redundancy and spatial smoothness simultaneously and reveal the fact that total variation on the spectral difference images is as smooth as possible, see Fig. 2(d) and (e). The main idea and contributions of this paper are summarized as follows.

1) The HSI is treated as a 3-D cube and the proposed 3DCrTV is defined as a 2-D spatial TV (along horizontal and vertical directions) across the spectral difference image (1-D spectral TV), which enforces the spectral and spatial smoothness simultaneously, thus suppressing the spectral-spatial noises effectively.

2) The spatial-spectral weights of TV regularization are calculated according to the sum of spatial gradients along spectral dimension. This adaptive mechanism for calculating the weights could preserve the edge and texture details in spatial domain.

3) The proposed 3DCrTV is a general form of total variation on the spectral difference image of HSI, it can easily be extended to nonlocal or structural tensor sense, thus exploring much more detail information for restoring HSI.

4) The restoration model is established to remove Gaussian noise and sparse noise (i.e., impulse noise and stripes or dead lines) simultaneously. The alternating direction method of multipliers (ADMM) is utilized to split the proposed model into several simpler subproblems, and experimental evaluations of simulated and real HSI datasets are provided.
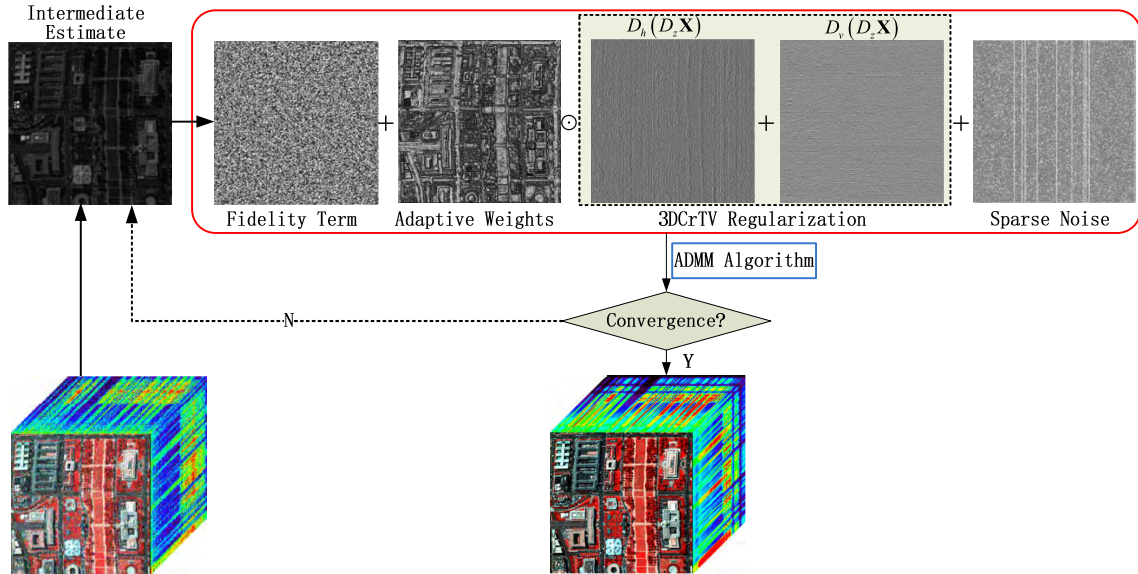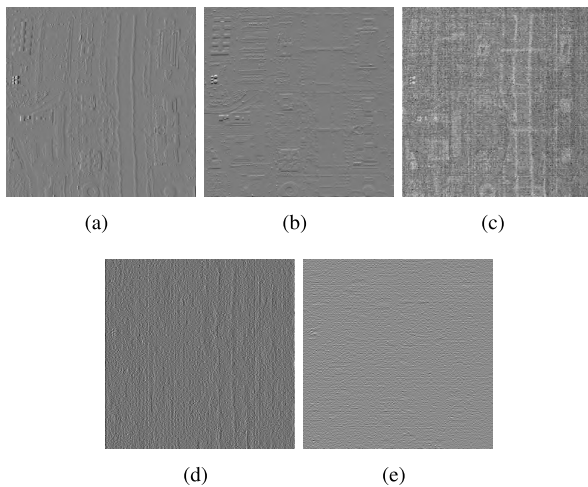
FIGURE 1. Flowchart of the proposed method.

FIGURE 2. Band 10 of different difference images of the clean HSI. (a) $D_h \mathbf{X}$; (b) $D_v \mathbf{X}$; (c) $D_z \mathbf{X}$; (d) $D_h(D_z \mathbf{X})$; (e) $D_v(D_z \mathbf{X})$.

The remainder of the paper is organized as follows. Section II formulates the restoration problem for HSI in variational framework using TV regularization. Section III describes the proposed 3DCrTV model which is defined as 2-D spatial TV on the spectral difference image, as well as the adaptive mechanism of calculating the weights and optimization algorithm. Section IV illustrates the restoration results on the simulated and real hyperspectral datasets in comparison with other state-of-the-art methods. Section V concludes with some remarks.

II. PROBLEM FORMULATION

A. OBSERVATION MODEL

First, we make some notations for HSI restoration problem. Let $\mathbf{X} \in \mathbb{R}^{m \times n \times l}$ represent a clean HSI cube with $m \times n$ spatial dimension and l spectral dimension. Then, the observed HSI

cube $\mathbf{Y} \in \mathbb{R}^{m \times n \times l}$ corrupted by Gaussian noise $\mathbf{N} \in \mathbb{R}^{m \times n \times l}$ and sparse noise $\mathbf{S} \in \mathbb{R}^{m \times n \times l}$ can be formulated as

$$\mathbf{Y} = \mathbf{X} + \mathbf{S} + \mathbf{N} \tag{1}$$

This mixed noise formulation was first utilized for hyperspectral restoration in [17] using LRMR. It is used in removing the gaussian noise brought by the imaging sensors and model system, as well as the sparse noise like dead lines or stripes brought by pushbroom sensors.

B. RESTORATION MODEL

Having the model (1) in mind, the restoration model of reducing mixed Gaussian and sparse noises for HSI can be formulated as

$$\min_{\mathbf{X}, \mathbf{S}} ||\mathbf{Y} - \mathbf{X} - \mathbf{S}||_F^2 + \lambda_1 J_1(\mathbf{S}) + \lambda_2 J_2(\mathbf{X}), \tag{2}$$

where, $J_1(\mathbf{S})$ is usually modeled as sparse regularization, i.e. $||\mathbf{S}||_1 = \Sigma_i |s_i|$ where s_i is the element of $\mathbf{S}$ and $J_2(\mathbf{X})$ is the regularization associated to the clean HSI $\mathbf{X}$, which always models the prior properties of the clean HSI. A good choice of $J_2(\mathbf{X})$ is the total variation regularization. λ_1 and λ_2 are two parameters, which make tradeoff between the fidelity term and regularizations.

C. TV RELATED REGULARIZATION

TV regularization is usually used to model the image as piecewise smooth, which is one of the intrinsic priors for natural images. The band-by-band hyperspectral total variation (HTV) can be expressed as [39], [40]

$$\text{HTV}(\mathbf{X}) = \sum_k \left\{ \sum_{i,j} \sqrt{(D_h \mathbf{X}(i,j,k))^2 + (D_v \mathbf{X}(i,j,k))^2} \right\} \tag{3}$$

where D_h and D_v denote the horizontal and vertical difference operators along the spatial dimension, respectively. $\mathbf{X}(i, j, k)$ denotes a pixel at the spatal location (i, j) in the kth band. This HTV model is an extension of the conventional TV by summing the gradient magnitudes of each band of HSI, and it only exploits the spatial smoothness of HSI. Moreover, HTV enforces the same strength in each band, thus producing a worse denoising performance.

To define a hyperspectral TV model that can adaptively adjust the denoising strength of different bands, Yuan *et al.* [32] extended the vectorial TV (VTV) model for hyperspectral image.

$$\text{VTV}(\mathbf{X}) = \sum_{i,j} \left\{ \sqrt{\sum_k (D_h \mathbf{X}(i, j, k))^2 + (D_v \mathbf{X}(i, j, k))^2} \right\} \tag{4}$$

The VTV model couples the gradients of all bands at every spatial location of HSI. By doing so, the bands with higher noise intensities will be smoothed stronger, and vice versa. However, both HTV and VTV models only exploit spatial information and ignore the spectral smoothness of HSI, they could not remove the heavy Gaussian noise and sparse noise very well.

A straightforward way to extend the TV model to a 3-D version is to impose the spatial-spectral total variation into a unit system and it can be formulated as follows.

$$\text{3DTV}(\mathbf{X}) = \sum_{i,j,k}$$
$$\left\{ \sqrt{(D_h \mathbf{X}(i, j, k))^2 + (D_v \mathbf{X}(i, j, k))^2 + \rho(D_z \mathbf{X}(i, j, k))^2} \right\} \tag{5}$$

where D_z is a linear operator corresponding to the spectral difference, ρ is the parameter which controls the proportion of spectral difference. Now, the 3DTV in (5) exploits both spatial and spectral smoothness for HSI, and the definitions of its three operators are expressed as

$$\begin{cases} D_h \mathbf{X}(i, j, k) = \mathbf{X}(i + 1, j, k) - \mathbf{X}(i, j, k) \\ D_v \mathbf{X}(i, j, k) = \mathbf{X}(i, j + 1, k) - \mathbf{X}(i, j, k) \\ D_z \mathbf{X}(i, j, k) = \mathbf{X}(i, j, k + 1) - \mathbf{X}(i, j, k) \end{cases} \tag{6}$$

By enforcing the constraints on two spatial dimensions and one spectral dimension, the piecewise smoothness in both domains can be pursued.

III. PROPOSED 3D CROSS TV MODEL FOR HSI DENOISING

A. PROPOSED 3DCRTV MODEL
We note that $D_z(\mathbf{X})$ still contains lots of detail information of the clean HSI (see Fig. 2 (c)), and minimizing 3DTV regularization defined in (5) will result in an estimate which does not converge to the clean HSI, meanwhile, weaken the spectral features (i.e., absorption peak) of the original spectrum.

To overcome aforementioned drawbacks, in this paper, we propose a 3D cross TV model which exploits the spectral-spatial information by enforcing spatial smoothness on the spectral difference space, for HSI mixed denoising. The model can be formulated as

$$\min_{\mathbf{X}, \mathbf{S}} ||\mathbf{Y} - \mathbf{X} - \mathbf{S}||_F^2 + \lambda_1 ||\mathbf{S}||_1 + \lambda_2 \text{CrTV}(\mathbf{X}), \tag{7}$$

where

$$\text{CrTV}(\mathbf{X}) = \text{HTV}(D_z \mathbf{X}) = \sum_k \sum_{i,j}$$
$$\left\{ \sqrt{(D_h(D_z \mathbf{X}(i, j, k)))^2 + (D_v(D_z \mathbf{X}(i, j, k)))^2} \right\} \tag{8}$$

is a 2D spatial TV defined on the spectral difference images, which can exploit the spatial information as well as the spectral information at the same time. The advantages of the proposed 3DCrTV model are twofold.

1) It exploits the intrinsic properties of HSI cube that the total variation in the spectral difference space is as smooth as possible, see Fig. 2 (d) and (e).
2) It depicts a new neighborhood system for 3-D images, which can exploit more information for restoration. Clearly speaking, for a given pixel (i, j, k), the CrTV regularization can be expressed as:

$$D_h(D_z \mathbf{X}(i, j, k)) = D_h(\mathbf{X}(i, j, k + 1) - \mathbf{X}(i, j, k))$$
$$= \mathbf{X}(i + 1, j, k + 1) - \mathbf{X}(i, j, k)$$
$$- (\underbrace{D_h \mathbf{X}(i, j, k) + D_z \mathbf{X}(i, j, k)}) \tag{9}$$

From (9), the $D_h(D_z \mathbf{X}(i, j, k))$ term not only contains $D_h \mathbf{X}(i, j, k)$ and $D_z \mathbf{X}(i, j, k)$ terms, but also involves extra information exploited by the diagonal pixel in the neighbourhood system. Analogous to $D_h(D_z \mathbf{X}(i, j, k))$, $D_v(D_z \mathbf{X}(i, j, k))$ has the similar properties.

To make the CrTV regularization have the adaptive property with the spatial structure distribution, we adopt the mechanism in [32] to involve a spatial weight $\mathbf{W}$ for each pixel. The weight $\mathbf{W}$ is calculated by summing the gradient of the clean HSI along the spectral direction, and can be expressed as

$$\mathbf{W}(i, j) = \frac{1}{1 + \mu_s \sum_{k=1}^l \sqrt{(D_h \mathbf{X}(i, j, k))^2 + (D_v \mathbf{X}(i, j, k))^2}} \tag{10}$$

where, μ_s is a constant parameter, and $\mathbf{W} \in \mathbb{R}^{m \times n}$. In order to fit different HSI dataset, the weight is normalized as $\mathbf{W} = \mathbf{W}/\overline{w}$, where $\overline{w}$ is the mean value of $\mathbf{W}$. By treating the equal strength in all bands, we finally replicate $\mathbf{W}$ to be a $m \times n \times l$ matrix.

Having (7) and (10) in mind, the proposed spatially weighted 3DCrTV (termed as 3DCrWTV) model can be

formulated in the constraint form as

$$\min_{\mathbf{X},\mathbf{S}} ||\mathbf{Y} - \mathbf{X} - \mathbf{S}||_F^2 + \lambda_1 ||\mathbf{S}||_1 + \lambda_2 \mathbf{W} \odot (\text{HTV}(\mathbf{V}_1)),$$
$$s.t. \ \mathbf{V}_1 = D_z \mathbf{X} \tag{11}$$

where $\odot$ denotes the multiplication in component-wise.

B. OPTIMIZATION ALGORITHM

Due to non-differential property of the model (11), the alternating direction of multipliers (ADMM) method is extended to separate it into several simpler subproblems. By involving the variants $\mathbf{V}_1 = D_z \mathbf{X}$, $\mathbf{V}_2 = D_h \mathbf{V}_1$ and $\mathbf{V}_3 = D_v \mathbf{V}_1$, the model (11) can be rewritten as

$$\min_{\mathbf{X},\mathbf{S}} ||\mathbf{Y} - \mathbf{X} - \mathbf{S}||_F^2 + \lambda_1 ||\mathbf{S}||_1 + \lambda_2 \mathbf{W} \odot \varphi(\mathbf{V}_2, \mathbf{V}_3)$$
$$subject \ to \ \mathbf{V}_1 = D_z \mathbf{X}$$
$$\mathbf{V}_2 = D_h \mathbf{V}_1$$
$$\mathbf{V}_3 = D_v \mathbf{V}_1. \tag{12}$$

where

$$\varphi(\mathbf{V}_2, \mathbf{V}_3) = \sum_k \sum_{i,j} \left\{ \sqrt{(\mathbf{V}_2(i,j,k))^2 + (\mathbf{V}_3(i,j,k))^2} \right\} \tag{13}$$

The augmented Lagrangian function of (12) can be formulated as

$$\mathcal{L}(\mathbf{X}, \mathbf{S}, \mathbf{V}_1, \ldots, \mathbf{V}_3, \mathbf{D}_1, \ldots, \mathbf{D}_3)$$
$$= ||\mathbf{Y} - \mathbf{X} - \mathbf{S}||_F^2 + \lambda_1 ||\mathbf{S}||_1$$
$$+ \lambda_2 \mathbf{W} \odot \varphi(\mathbf{V}_2, \mathbf{V}_3) + \mu ||\mathbf{V}_1 - D_z \mathbf{X} - \mathbf{D}_1||_F^2$$
$$+ \mu ||\mathbf{V}_2 - D_h \mathbf{V}_1 - \mathbf{D}_2||_F^2 + \mu ||\mathbf{V}_3 - D_v \mathbf{V}_1 - \mathbf{D}_3||_F^2, \tag{14}$$

where $\mathbf{D}_1, .., \mathbf{D}_3$ are augmented multipliers related to variables $\mathbf{V}_1, .., \mathbf{V}_3$, and μ is the regularization parameters. Here, we employ the same parameters μ for the three regularization terms.

Since it is still difficult to solve problem (14) simultaneously, the alternating minimization scheme is utilized to solve it by optimizing one variable while fixing the others. Thus, the problem (14) can be converted into several simpler subproblems.

- The subproblem related to $\mathbf{X}$ is

$$\min_{\mathbf{X}} ||\mathbf{Y} - \mathbf{X} - \mathbf{S}||_F^2 + \mu ||\mathbf{V}_1 - D_z \mathbf{X} - \mathbf{D}_1||_F^2 \tag{15}$$

It is a convex problem and equals to the following linear problem

$$(\mathbf{I} + \mu D_z^T D_z)\mathbf{X} = (\mathbf{Y} - \mathbf{S}) + \mu D_z^T (\mathbf{V}_1 - \mathbf{D}_1) \tag{16}$$

which has a closed form solution by the fast Fourier transform (FFT) in (17), where $\mathcal{F}(\cdot)$ is the fast Fourier transform and $\mathcal{F}^{-1}(\cdot)$ is the inverse transform.

The superscripts T and H denote the operator of transpose and complex conjugate, respectively.

$$\mathbf{X} = \mathcal{F}^{-1} \left(\frac{\mathcal{F}(\mathbf{Y} - \mathbf{S} + \mu D_z^T(\mathbf{V}_1 - \mathbf{D}_1))}{1 + \mu \mathcal{F}(D_z)^H \mathcal{F}(D_z)} \right) \tag{17}$$

- The subproblem related to $\mathbf{S}$ is

$$\min_{\mathbf{S}} \lambda_1 ||\mathbf{S}||_1 + ||\mathbf{Y} - \mathbf{X} - \mathbf{S}||_F^2 \tag{18}$$

which can be easily solved by the *soft threshold* operator

$$\mathbf{S} \equiv \text{soft}(\boldsymbol{\eta}_1, \lambda) = \text{sign}(\boldsymbol{\eta}_1) \times \max \left\{ 0, |\boldsymbol{\eta}_1| - \frac{\lambda}{2} \right\} \tag{19}$$

where $\boldsymbol{\eta}_1 = \mathbf{Y} - \mathbf{X}$.

- The subproblems related to $\mathbf{V}_2$, $\mathbf{V}_3$ are summarized as

$$\begin{cases} \min_{\mathbf{V}_2} & \lambda_2 \mathbf{W} \odot \varphi(\mathbf{V}_2, \mathbf{V}_3) + \mu ||\mathbf{V}_2 - D_h \mathbf{V}_1 - \mathbf{D}_2||_F^2 \\ \min_{\mathbf{V}_3} & \lambda_2 \mathbf{W} \odot \varphi(\mathbf{V}_2, \mathbf{V}_3) + \mu ||\mathbf{V}_3 - D_v \mathbf{V}_1 - \mathbf{D}_3||_F^2 \end{cases} \tag{20}$$

These two variables should be solved jointly and the solution is given by a band-by-band *vector-soft threshold* function.

$$\{\mathbf{V}_2(:,:,k), \mathbf{V}_3(:,:,k)\} = \frac{\max(||\mathbf{C}||_F - \frac{\lambda}{\mu}, 0)}{\max(||\mathbf{C}||_F - \frac{\lambda}{\mu}, 0) + \frac{\lambda}{\mu}} \mathbf{C} \tag{21}$$

where

$$\mathbf{C} = \{\mathbf{W}(:,:,k) \odot \boldsymbol{\eta}_2(:,:,k), \mathbf{W}(:,:,k) \odot \boldsymbol{\eta}_3(:,:,k)\}, \tag{22}$$

$\boldsymbol{\eta}_2 = D_h \mathbf{V}_1 + \mathbf{D}_2$, $\boldsymbol{\eta}_3 = D_v \mathbf{V}_1 + \mathbf{D}_3$, and $(:,:,k)$ denotes the kth band of HSI.

- The subproblem related to $\mathbf{V}_1$ is

$$\min_{\mathbf{V}_1} \mu ||\mathbf{V}_1 - D_z \mathbf{X} - \mathbf{D}_1||_F^2 + \mu ||\mathbf{V}_2 - D_h \mathbf{V}_1 - \mathbf{D}_2||_F^2$$
$$+ \mu ||\mathbf{V}_3 - D_v \mathbf{V}_1 - \mathbf{D}_3||_F^2 \tag{23}$$

which has the similar form as subproblem (15), and equals to the following linear system

$$(\mathbf{I} + D_h^T D_h + D_v^T D_v)\mathbf{V}_1 = D_z \mathbf{X} + \mathbf{D}_1 + D_h^T(\mathbf{V}_2 - \mathbf{D}_2)$$
$$+ D_v^T(\mathbf{V}_3 - \mathbf{D}_3) \tag{24}$$

which also has the closed form solution by FFT and the solution is given as

$$\mathbf{V}_1 = \mathcal{F}^{-1} \left(\frac{\mathcal{F}(\boldsymbol{\xi}_1 + D_h^T \boldsymbol{\xi}_2 + D_v^T \boldsymbol{\xi}_3)}{1 + \mathcal{F}(D_h)^H \mathcal{F}(D_h) + \mathcal{F}(D_v)^H \mathcal{F}(D_v)} \right) \tag{25}$$

where $\boldsymbol{\xi}_1 = D_z \mathbf{X} + \mathbf{D}_1$, $\boldsymbol{\xi}_2 = \mathbf{V}_2 - \mathbf{D}_2$, $\boldsymbol{\xi}_3 = \mathbf{V}_3 - \mathbf{D}_3$.

The aforementioned process can be summarized as pseudocode in Algorithm 1.

The convergence of the Algorithm 1 using ADMM could be guaranteed theoretically in [41]. Moreover, the convergence rate of ADMM algorithm depends on parameter μ.

Algorithm 1 Extended ADMM Method for TV Model on Spectral Difference Image

---

1: **Initialization** set $k = 0$, $\lambda_1 > 0$, $\lambda_1 > 0$, $\mu > 0$, $\mathbf{V}_1^{(0)}$, $\mathbf{V}_2^{(0)}$, $\mathbf{V}_3^{(0)}$, $\mathbf{D}_1^{(0)}$, $\mathbf{D}_2^{(0)}$, $\mathbf{D}_3^{(0)}$, $\mathbf{W} = ones(m, n, l)$.
2: **Repeat**
3:     calculate $\mathbf{X}^{(k+1)}$ according to Eq. (17)
4:     calculate $\mathbf{V}_1^{(k+1)}$ according to Eq. (25)
5:     $\mathbf{S}^{(k+1)} = \text{soft}(\mathbf{Y} - \mathbf{X}^{(k+1)}, \frac{\lambda_1}{\mu})$
6:     $\mathbf{V}_2$ and $\mathbf{V}_3$ are calculated by Eq. (21)
7:     **Update multipliers**
8:     $\mathbf{D}_1^{(k+1)} \leftarrow \mathbf{D}_1^{(k)} + D_z \mathbf{X}^{(k+1)} - \mathbf{V}_1^{(k+1)}$
9:     $\mathbf{D}_2^{(k+1)} \leftarrow \mathbf{D}_2^{(k)} + D_h \mathbf{V}_1^{(k+1)} - \mathbf{V}_2^{(k+1)}$
10:    $\mathbf{D}_3^{(k+1)} \leftarrow \mathbf{D}_3^{(k)} - D_v \mathbf{V}_1^{(k+1)} - \mathbf{V}_3^{(k+1)}$
11:   **Update weight**
12:     calculate $\mathbf{W}$ according to Eq. (10)
13:   **Update iteration** $k = k + 1$
14: **Until some stopping criterion is satisfied**

We experimentally verify that the proposed algorithm has an absolutely fast convergence rate with $\mu = 0.8$. It is worth mentioning that all matrices involved in Algorithm 1 are calculated in 3-D dimension.

Finally, the computational complexity of the proposed algorithm is provided. According to previous section, we recall that l is the number of the bands of the input HSI cube, and $q = m \times n$ is the total number of pixels in each band. Then, in Algorithm 1, the computational complexity for accelerating $\mathbf{X}$ and $\mathbf{V}_1$ by FFT is $O(lq \log q)$, while the $\mathbf{S}$, $\mathbf{V}_2$, $\mathbf{V}_3$ subproblems are solved by efficient *soft threshold* and *vector-soft threshold* operators respectively, with a linear computational complexity of $O(lq)$. Taking all parts into consideration, the overall order of complexity per iteration for Algorithm 1 is $O(lq \log q)$.

IV. EXPERIMENT

Experiments with simulated and real data are conducted to demonstrate the effectiveness of our proposed method for HSI restoration. For comparison, five different state-of-the-art HSI restoration methods are employed as the benchmark in the experiments, i.e., BM4D [14], PCABM4D [9], LRMR [17], 3DTV in (5) and the LRTV [21]. BM4D is the block matching 4-D filtering method, which is the extension of BM3D to volumetric data and has been demonstrated as one of the best denoising methods for natural images while PCABM4D is the method combining principal component analysis and BM4D, which utilizes BM4D filter on the low energy principal components for noise removal. LRMR is one of the representative methods with low rank matrix recovery via GoDec algorithm, while LRTV is a total variation regularized low rank method, and both of them have achieved promising performance in removing the mixed noise of HSI.

As in [17] and [32], the mean peak signal-to-noise ratio (MPSNR) index and the mean structural similarity (MSSIM) index [42], mean feature similarity (MFSIM)

index [43], erreur relative globale adimensionnelle de synthe 'se (ERGAS) [44] and mean spectral angle (MSA) [45] are employed to give a quantitative assessment of HSI restoration results. Three of the indices are defined as:

$$\text{MPSNR} = \frac{1}{l} \sum_{i=1}^{l} \text{PSNR}_i$$

$$= \frac{1}{l} \sum_{i=1}^{l} 10 \log_{10} \left(\frac{\max(\mathbf{X}_i)}{\text{MSE}(\mathbf{X}_i, \hat{\mathbf{X}}_i)} \right) \qquad (26)$$

$$\text{MSSIM} = \frac{1}{l} \sum_{i=1}^{l} \text{MSSIM}_i$$

$$= \frac{1}{l} \sum_{i=1}^{l} 10 \log_{10}$$

$$\times \left(\frac{(2\mu_{\mathbf{X}_i}\mu_{\hat{\mathbf{X}}_i} + C_1)(2\sigma_{\mathbf{X}_i\hat{\mathbf{X}}_i} + C_2)}{(\mu_{\hat{\mathbf{X}}_i}^2 + \mu_{\hat{\mathbf{X}}_i}^2 + C_1)(\sigma_{\mathbf{X}_i}^2 + \sigma_{\hat{\mathbf{X}}_i}^2 + C_2)} \right) \qquad (27)$$

$$\text{MSA} = \frac{1}{m \times n} \sum_{i=1}^{m \times n} cos^{-1}(x_i, \hat{x}_i) \qquad (28)$$

where, $\max(\mathbf{X}_i)$ stands for the maximum value of the ith band image of input HSI $\mathbf{X}$, and MSE stands for the mean square error. $\mu_{\mathbf{X}_i}$ and $\mu_{\hat{\mathbf{X}}_i}$ stand for the average values of ith band image of clean HSI $\mathbf{X}$ and reconstructed HSI $\hat{\mathbf{X}}$, respectively, while $\sigma_{\mathbf{X}_i}$ and $\sigma_{\hat{\mathbf{X}}_i}$ stand for the variances and $\sigma_{\mathbf{X}_i\hat{\mathbf{X}}_i}$ is the covariance between $\mathbf{X}_i$ and $\hat{\mathbf{X}}_i$. Generally, the larger MPSNR, MSSIM and MFSIM values denote the better results whereas the smaller ERGAS and MSA values denote the better results.
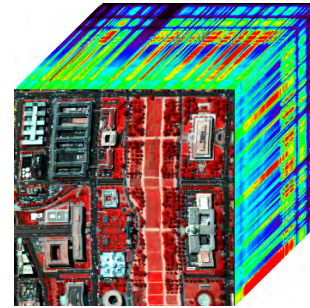
FIGURE 3. HYDICE Washington DC dataset used in the simulated experiment (false color image composed of bands 60, 27, 17 for the red, green and blue wavelength, respectively).

A. SIMULATED EXPERIMENT ON HYDICE

In the first simulated experiment, the Hyperspectral Digital Imagery Collection Experiment (HYDICE) image of the Washington DC Mall is adopted. A subimage of size $256 \times 256 \times 191$ is used in the experiment, see Fig 3. Before the simulated process, the gray values of each band of the HSI are normalized to the range of [0, 1], and after the restoration, they will be stretched back to the original range. In simulated experiments, three kinds of noise are added to the Washington DC image according to [17].
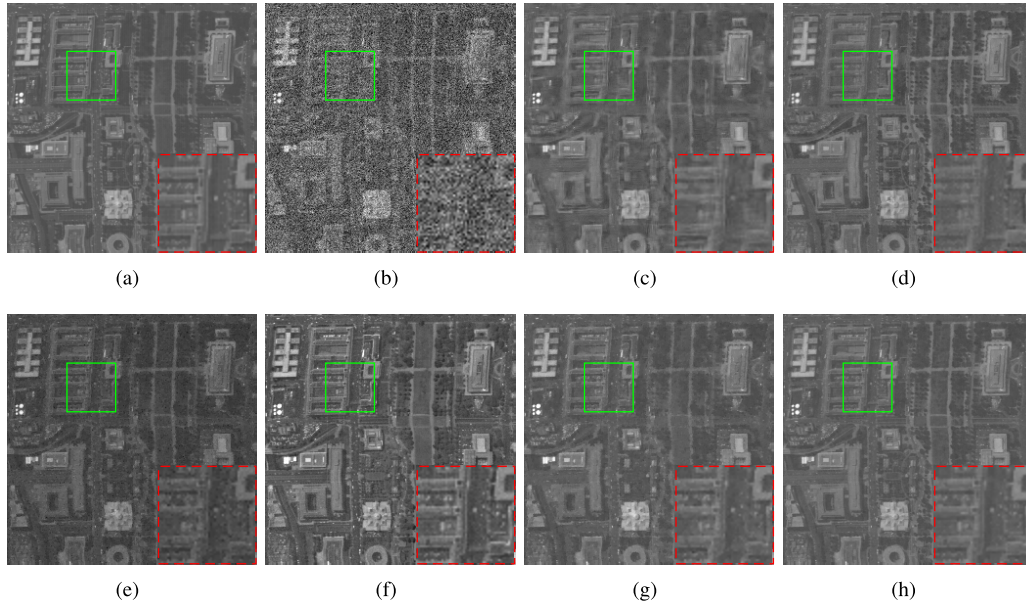
FIGURE 4. Band 2 image of the restoration results in Washington DC dataset. (a) Original band 2; (b) noisy band (PSNR = 19.50 dB); (c) BM4D (PSNR = 31.16 dB); (d) PCABM4D (PSNR = 32.43 dB); (e) 3DTV (PSNR = 21.33 dB); (f) LRMR (PSNR = 27.09 dB); (g) LRTV (PSNR = 32.47 dB); (h) 3DCrTV (PSNR = 35.29 dB).

- Zero-mean Gaussian noise is added to all the bands of input HSI. For different bands, the noise intensity is different. The SNR value of each band varies from 10 to 20 dB randomly.
- Impulse noise is added to 20 bands randomly. The percentage of impulse noise is 20%.
- Stripes are simulated for 10 bands, in which 5 bands are the same as the impulse noise bands, and 5 bands are chosen randomly from the rest, and the mean SNR value of the simulated data is 12.92 dB.

Due to the complexity and variety of noise, the parameters of our method are chosen empirically. For BM4D, PCABM4D, 3DTV, LRMR and LRTV algorithms, we set the parameters according to the references and then tune them slightly to obtain the optimal results. For the first simulated dataset, the parameters are set as follows. For BM4D, the only parameter is the standard deviation, which is set to 0.15. For PCABM4D, the number of high energy channel is set to 2. For 3DTV, the regularization parameter for spatial TV is set to 0.005, while for spectral TV it is set to 0.15 due to the heavy stripe noise, and the parameter for sparse regularization is set to 0.1. For LRMR, the patch size is set as $q = 26$, the step size is set to 8, and the value of low rank parameter is set as $r = 4$, and sparse parameter is set to $k = 6000$. For LRTV method, the low rank parameter is set as $r = 8$, and the parameter for total variation regularization is set as $\tau = 0.005$, while the other associated paramors are set as described in the reference paper. For the proposed 3DCrTV and 3DCrWTV algorithms, there are only two parameters, which are set as $\lambda_1 = 0.05$, $\lambda_2 = 0.1$, respectively.

In order to thoroughly illustrate the restoration results, images contaminated by different combination of these three

kinds of noises are presented to show the visual effect. Fig. 4 shows the image of band 2 before and after restoration, which is only corrupted by Gaussian noise with PSNR value of 19.50 dB. From Fig. 4, it is obvious that the proposed 3DCrTV algorithm performs the best among those six methods since it simultaneously removes the heavy Gaussian noise and preserves the fine details of the original image. BM4D can also suppress the Gaussian noise in an extent, but many fine details of the image are oversmoothed obviously. PCABM4D can produce better results than BM4D, but some fine details are also lost [see the zoom-in portion for details]. 3DTV performs the worst in those six methods, it loses a lot of fine details. LRMR can effectively remove the Gaussian noise, but in order to pursue the low rank properties of the whole HSI cube, LRMR performs almost as worst as 3DTV in terms of PSNR. The LRTV could remove the Gaussian noise effectively and produce the acceptable results, but it is still slightly lower than 3DCrTV.

Fig. 5 illustrates the restoration results of band 32, which is contaminated by Gaussian noise, impulse noise, and stripe noise simultaneously. From Fig. 5, it is obviously observed that the proposed 3DCrTV method can remove the mixed Gaussian noise, impulse noise and stripes effectively, and lead to the best performance among those six methods. BM4D can suppress the Gaussian noise very well, but it fails to remove the impulse noise and the fine details of the pixels surrounding the impulse noise are severely distorted. PCABM4D can successfully remove the Gaussian noise and impulse noise in an extent. However, both BM4D and PCABM4D fail to remove the stripes in the image. 3DTV can remove the mixed Gaussian noise, impulse noise and stripes simultaneously by strengthening TV constraint along spectral direction, and it
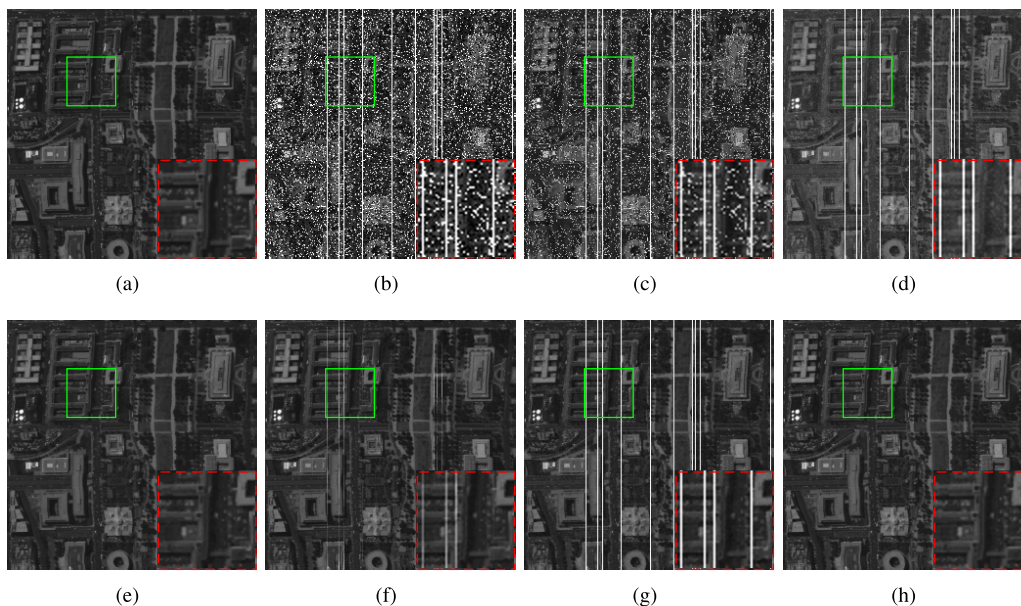
FIGURE 5. Band 32 image of the restoration results in Washington DC dataset. (a) Original band 32; (b) noisy band (PSNR = 10.27 dB); (c) BM4D (PSNR = 12.77 dB); (d) PCABM4D (PSNR = 14.81 dB); (e) 3DTV (PSNR = 38.61 dB); (f) LRMR (PSNR = 28.38 dB); (g) LRTV (PSNR = 16.06 dB); (h) 3DCrTV (PSNR = 39.80 dB).
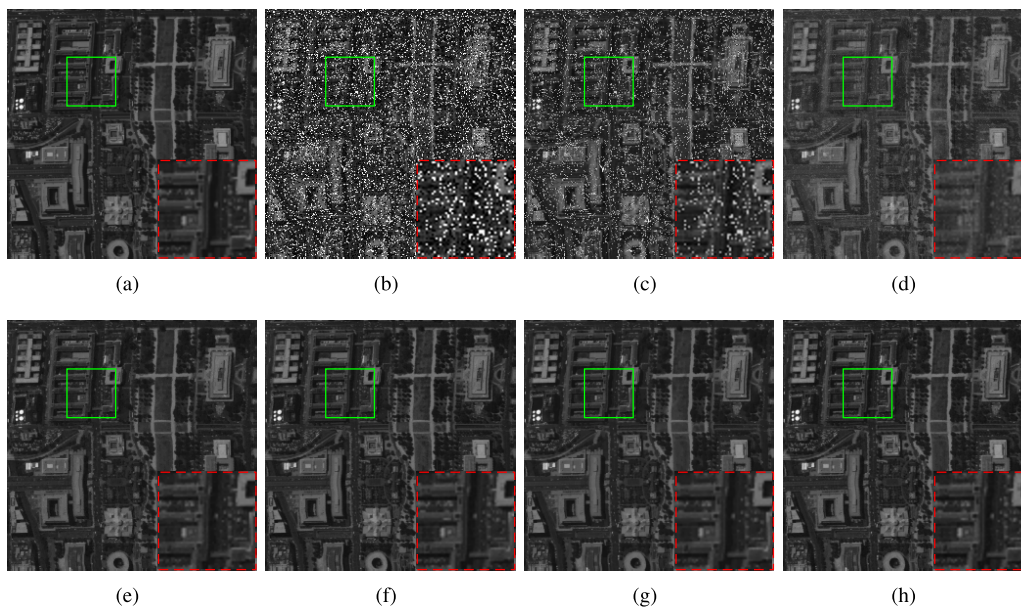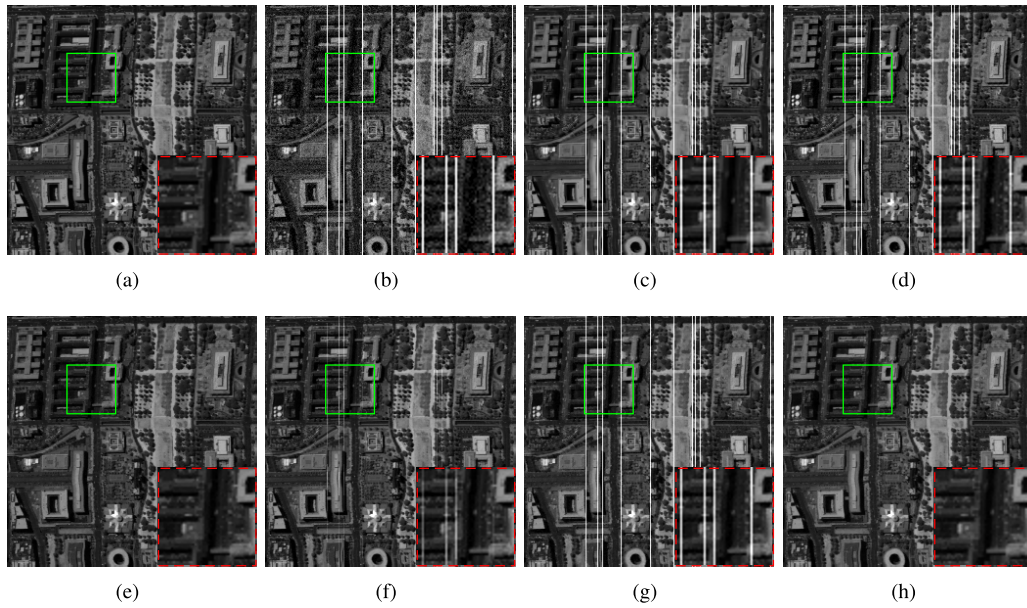

FIGURE 6. Band 37 image of the restoration results in Washington DC dataset. (a) Original band 37; (b) noisy band (PSNR = 11.52 dB); (c) BM4D (PSNR = 16.17 dB); (d) PCABM4D (PSNR = 24.01 dB); (e) 3DTV (PSNR = 37.33 dB); (f) LRMR (PSNR = 36.65 dB); (g) LRTV (PSNR = 38.90 dB); (h) 3DCrTV (PSNR = 40.37 dB).

almost leads to a result as better as 3DCrTV. LRMR can also effectively remove the mixed three noises in an extent, and lead to a relatively good result. But for the stripes, due to the different noise intensity in each band and the stripe lines in the same 20 bands, LRMR regards the stripes as a low rank component and fails to remove them completely. For LRTV, it performs even worse than LRMR in removing the stripe noise. The reason is that when the low rank regularization could not successfully remove the stripe noise and treat it as

one low rank component, TV regularization will enhance the stripe noise instead of getting rid of it.

Band 37 is corrupted by Gaussian noise and impulse noise, while band 114 is corrupted by Gaussian noise and stripe noise. The restoration results of these two bands are illustrated in Fig. 6 and Fig. 7, respectively. From Fig. 6 and Fig. 7, it can be concluded that BM4D can effectively remove Gaussian noise, but fail to remove the impulse noise and the stripe noise. PCABM4D can almost totally remove

FIGURE 7. Band 114 image of the restoration results in Washington DC dataset. (a) Original band 114; (b) noisy band (PSNR = 16.11 dB); (c) BM4D (PSNR = 16.67 dB); (d) PCABM4D (PSNR = 17.51 dB); (e) 3DTV (PSNR = 33.59 dB); (f) LRMR (PSNR = 29.30 dB); (g) LRTV (PSNR = 16.92 dB); (h) 3DCrTV (PSNR = 35.92 dB).
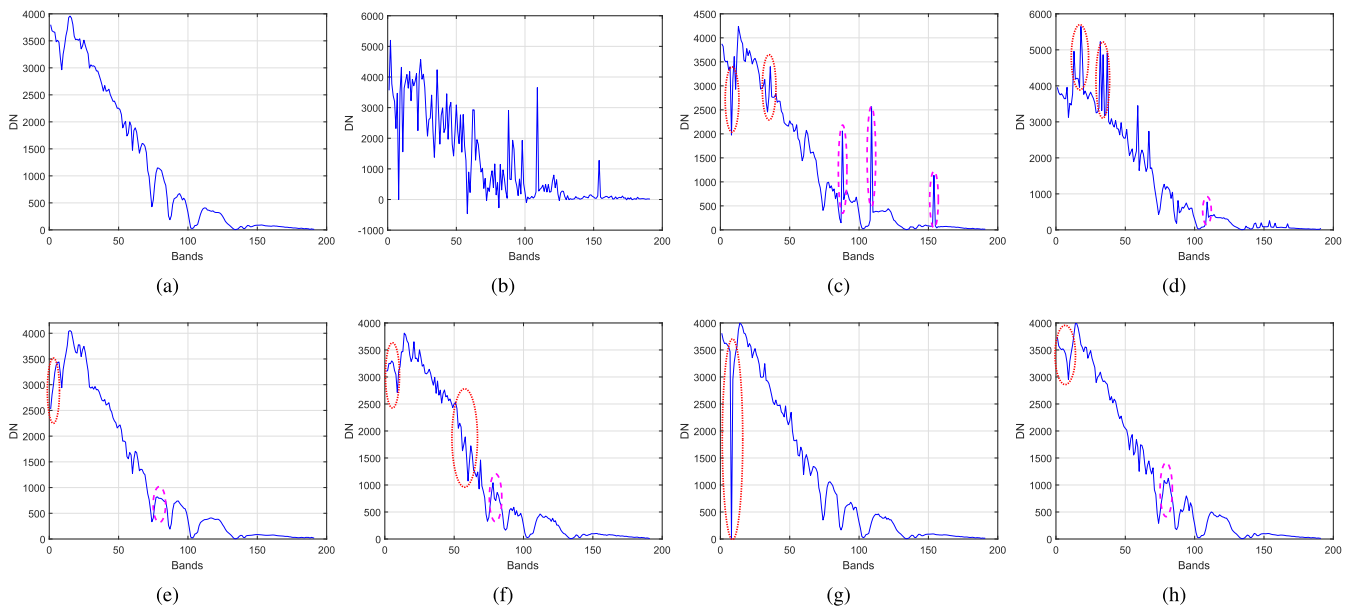
FIGURE 8. Spectrum of pixel (110, 206) which belongs to the class of building roof in the restoration results. (a) Original; (b) Noise-corrupted; (c) BM4D; (d) PCABM4D; (e) 3DTV; (f) LRMR; (g) LRTV; (h) 3DCrTV.

the Gaussian noise and the impulse noise, but also fail to remove the stripe noise. 3DTV can remove the Gaussian, impulse noises and stripes simultaneously due to the TV constraint along spectral direction, but its performance in removing the heavy Gaussian noise is not good. LRMR can effectively remove the mixed Gaussian noise, impulse noise, and strip noise. However, when the HSI cube are heavily corrupted by these three kinds of noise in different intensity, it fails to balance the reduction of all noises simultaneously.

For example, LRMR can completely remove the stripe noise in band 37 but fail to completely remove the stripe noise in band 32 and band 114. In addition, for the bands with heavy stripe noise, LRMR regards the stripes as a low rank component and fails to remove them. LRTV usually could perform slightly better than LRMR, except that when low rank regularization could not totally remove the stripe lines and treat them as one of the low rank components. In this case, TV regularization would enhance the stripes because it
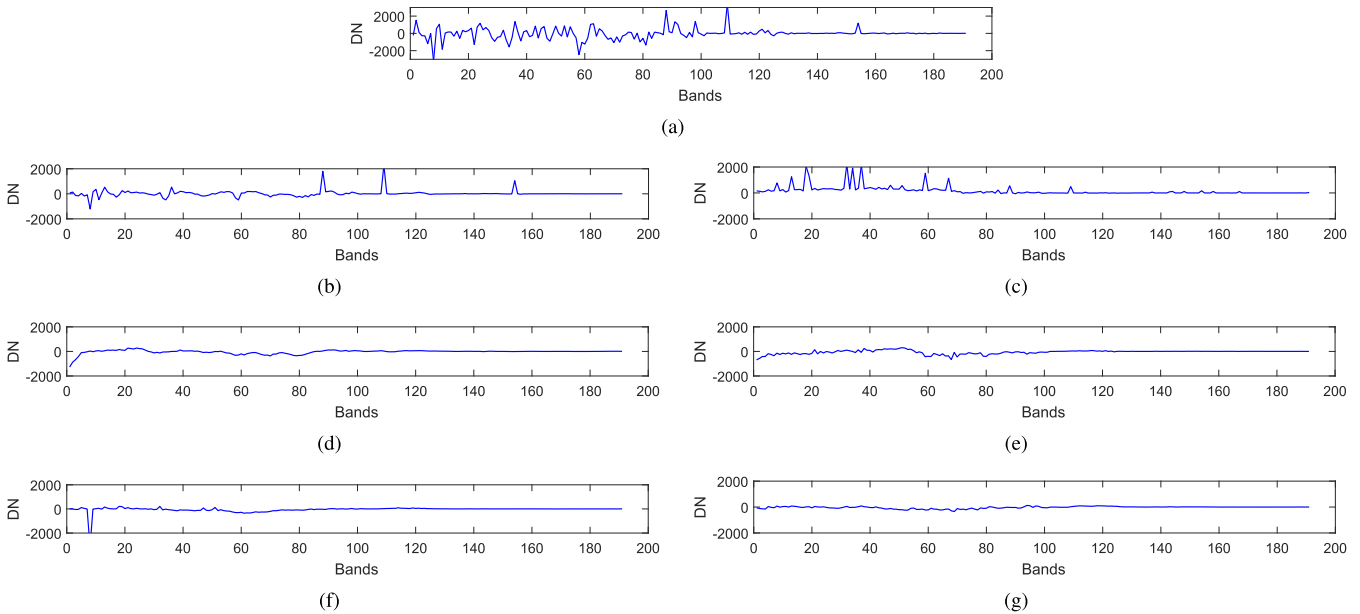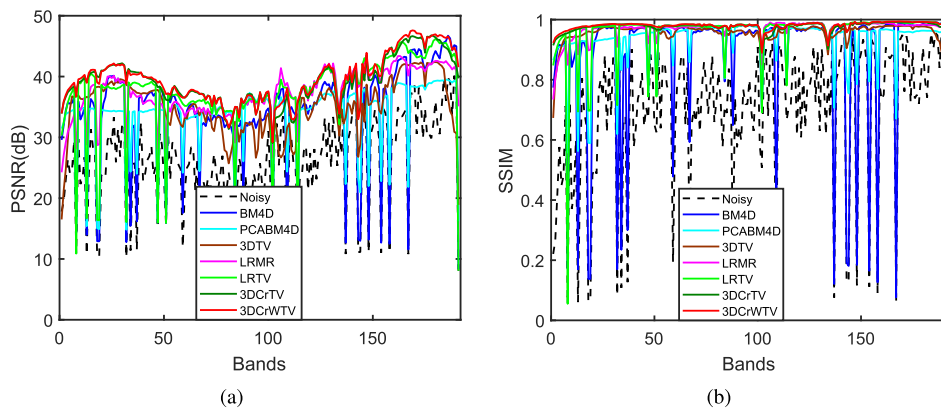
FIGURE 9. Difference spectrum of original spectrum and the reconstructed spectrum in the pixel (110, 206) (a) Noise-corrupted; (b) BM4D; (c) PCABM4D; (d) 3DTV; (e) LRMR; (f) LRTV; (g) 3DCrTV.

FIGURE 10. PSNR and SSIM values in each band of the reconstructed Washington DC dataset with different algorithms. (a) PSNR; (b) SSIM.

enforces the difference along one of directions of the stripe noise. For example, LRTV performs better than LRMR in band 2 and band 37 but performs worse than LRMR in band 32 and band 114. The proposed 3DCrTV algorithm can deal with the most complicated situation and lead to the best results among those six methods.

To further compare the restoration performance of those six algorithms, the spectral signature of pixel (110, 206), which belongs to the class of building roof, is plotted and presented in Fig. 8. The differences of the spectral signature between the original spectrum and the reconstructed spectrum of pixel (110, 206) is presented in Fig. 9. From Fig. 8, it is obvious that the proposed 3DCrTV reconstructed the best spectral signatures among those restoration algorithms. For 3DTV, even it produces a relatively smooth spectrum, some details (*i.e.* absorption peak) circled in the pink ellipse in Fig. 8 (d)

are lost. For BM4D, PCABM4D, LRMR, LRTV methods, it is obvious to see that there are more or less fluctuations (circled by the ellipses) in the restored spectra. It means that those method could not remove the stripe noise completely at the pixel (110,206). The similar phenomenon could also be found in Fig. 9, it also leads us to conclude that the proposed 3DCrTV produces the best results among all those six methods.

In addition, the PSNR and SSIM values in each band of the reconstructed results with different restoration algorithms are illustrated in Fig. 10. From it, we can conclude that the PSNR and SSIM values in almost all of the bands obtained by our algorithm are higher than those of the other five algorithms. Furthermore, even the noise intensity has a slight impact of the results constructed by our algorithm, it leads to an extremely good results in the bands corrupted by mixed
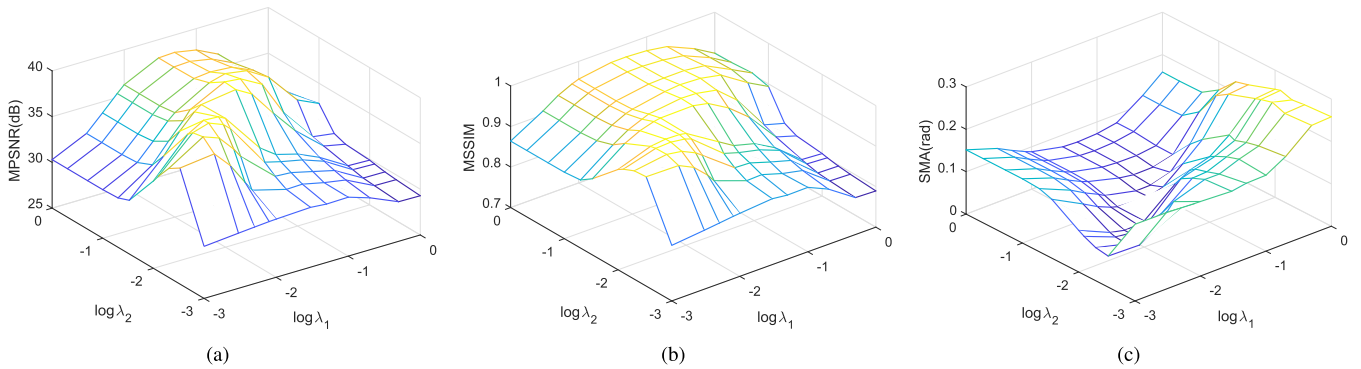
FIGURE 11. MPSNR and MSSIM values as a function of parameter λ_1 and λ_2. (a) MPSNR versus λ_1 and λ_2; (b) MSSIM versus λ_1 and λ_2; (c) MSA versus λ_1 and λ_2.

TABLE 1. MPSNR and MSSIM values of the restoration results in Washington DC dataset.

Method	Noisy	BM4D[5]	3DTV	PCABM4D[9]	LRMR[17]	LRTV [21]	3DCrTV	3DCrWTV
MPSNR (dB)	26.13	34.14	35.06	33.71	37.32	36.00	38.91	39.16
MSSIM	0.6957	0.8818	0.9541	0.9283	0.9696	0.9716	0.9755	0.9772
FSSIM	0.8526	0.9388	0.9771	0.9601	0.9799	0.9793	0.9861	0.9865
ERGAS	404.00	307.34	93.78	164.86	57.04	65.61	46.75	45.01
SMA (rad)	0.4434	0.2998	0.1205	0.1551	0.0704	0.0710	0.0625	0.0598
Time (s)	-	686	50	731	350	130	366	440

Gaussian noise, impulse noise and strip noise. Our method has a better ability in balancing the restoration results in bands with heavy and different noise intensity than LRMR. Benefiting from the weights at every spatial location, the proposed 3DCrWTV performances slightly better than 3DCrTV in almost all of the bands in terms of PSNR and SSIM.

The relationship between MPSNR, MSSIM and MSA values produced by the proposed method with logarithm value of the sparse parameter λ_1 and CrTV regularization parameter λ_2 are presented in Fig. 11, where, λ_1 and λ_2 are chosen from [0.001, 0.003, 0.005, 0.008, 0.01, 0.03, 0.05, 0.1, 0.2, 0.5, 1]. From Fig. 11, it is observed that sparse parameter λ_1 really has a relatively strong impact on the performance. However, when the value of λ_1 is in the range of [0.01 − 0.2], our method could produce relatively better results. In addition, the proposed method is really robust with regard to the CrTV regularization parameter λ_2, even the value of λ_2 is in the whole range of [0.01 − 1], the proposed method can lead to a good results with PSNR up to almost 38 dB.

Table 1 exhibits the quantitative values of MSPNR and MSSIM for the above six restoration methods as well as the proposed 3DCrWTV method. From Table 1, it can be observed that the values of all accuracy metrics (*i.e.*, MPSNR, MSSIM, FSSIM, ERGAS, MSA) show great consistency with the visual evaluations. Moreover, the proposed 3DCrWTV method could produce better performance than 3DCrTV due to the adaptive spatial weights for every pixels.

B. SIMULATED EXPERIMENT ON ROSIS
Reflective optics system imaging spectrometer (ROSIS) images of the Pavia University, Italy, is used for the second simulated experiment. The size of the subimage is 610×340,
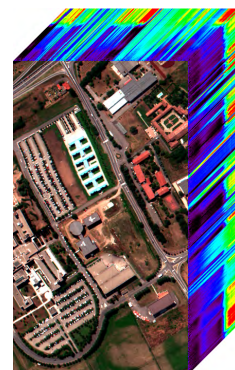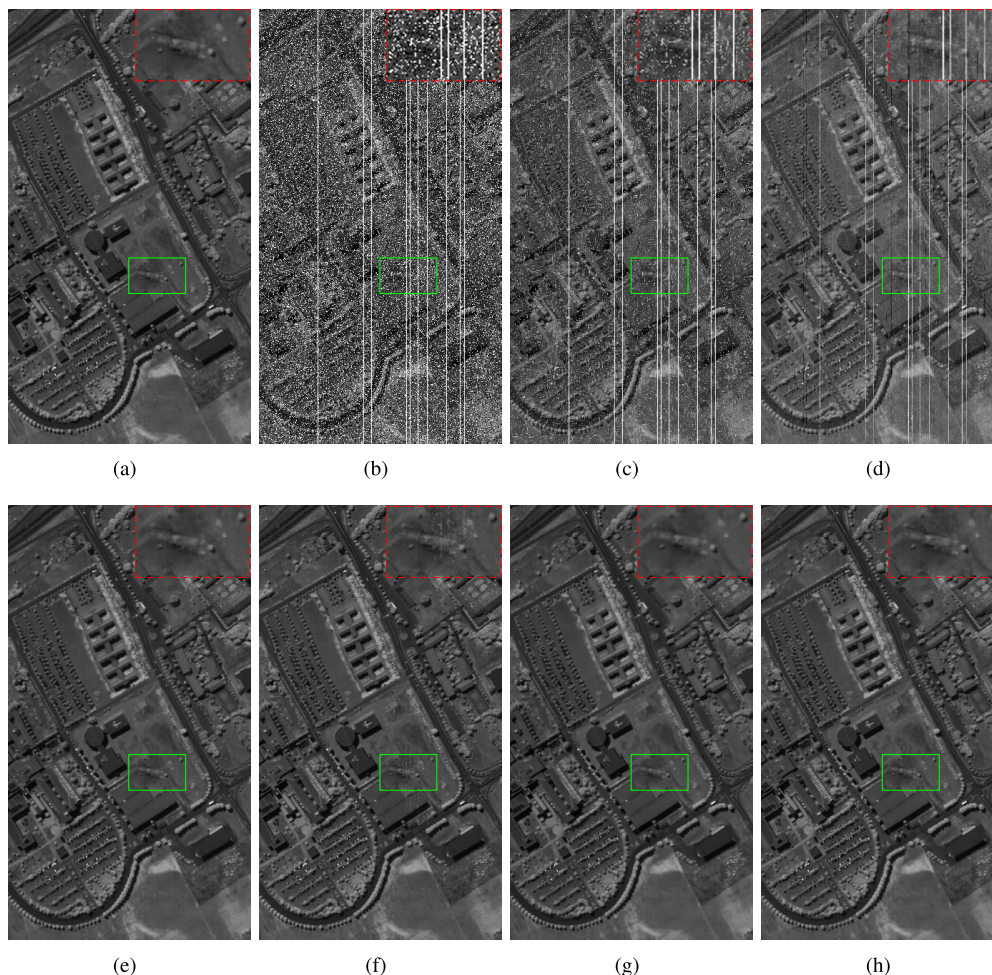

FIGURE 12. ROSIS University of Pavia data set used in the simulated experiment (false color image composed of bands 60, 27, 17 for the red, green and blue wavelength, respectively).

with 103 bands in the spectral range of 0.43-0.86 μm, and the false color cube is illustrated in Fig. 12.

In the second simulated experiment, the original data are added with Gaussian noise, salt-and-pepper noise, and stripe noise simultaneously. Gaussian noise is added to all the bands with $\sigma = 5\%$, salt-and-pepper noise and stripe noise are randomly added to 10 bands, the percentage of salt-and-pepper noise is 20%, and the number of stripe lines is 10 in each band. Furthermore, there are two bands that are added with all the three kinds of noises, one of which is shown in Fig. 13. The final mean SNR of the simulated data is 9.15 dB.

The restoration results of band 85, which are corrupted with all three kinds of noises, are illustrated in Fig. 13. From it, we can draw the same conclusion that our proposed 3DCrTV algorithm produces the best performance in terms of PSNR and visual effect among those six methods. BM4D

FIGURE 13. Band 85 image of the restoration results in University of Pavia dataset. (a) Original band; (b) noisy band (PSNR = 11.09 dB); (c) BM4D (PSNR = 18.42 dB); (d) PCABM4D (PSNR = 18.42 dB); (e) 3DTV (PSNR = 34.99 dB); (f) LRMR (PSNR = 37.16 dB); (g) LRTV (PSNR = 37.56 dB); (h) 3DCrTV (PSNR = 37.99 dB).
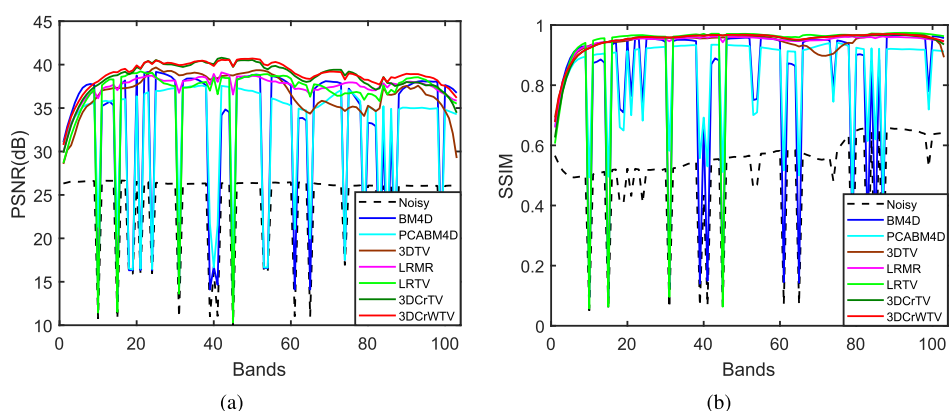
FIGURE 14. PSNR and SSIM values in each band of the reconstructed University of Paiva dataset with different algorithms. (a) PSNR; (b) SSIM.

can remove Gaussian noise effectively, but fails to remove the impulse noise and stripe noise. Besides, BM4D distorts the details of the pixels surrounding the impulse noise seriously. PCABM4D can do better than BM4D in removing the

Gaussian noise and impulse noise, but it also fails to get rid of the strip lines. 3DTV can remove these three kinds of noises simultaneously, but due to the strong strength of the TV constraint along the spectral direction, it makes the spectra

TABLE 2. MPSNR and MSSIM values of the restoration results in University of Pavia dataset.

Method	Noisy	BM4D [5]	PCABM4D [9]	3DTV	LRMR [17]	LRTV [21]	3DCrTV	3DCrWTV
MPSNR (dB)	23.65	32.66	32.37	36.73	37.46	36.39	38.58	**38.93**
MSSIM	0.5023	0.8269	0.8480	0.9409	0.9443	0.9201	0.9476	**0.9498**
FSSIM	0.7740	0.9107	0.9359	0.9734	0.9769	0.9550	0.9784	**0.9808**
ERGAS	471.75	336.49	235.74	64.60	57.65	247.24	51.88	**49.14**
MSA (rad)	0.5095	0.3524	0.2456	0.0837	0.0710	0.2234	0.0632	**0.0608**
Time (s)		1071	1001	167	740	376	922	1038

**TABLE 3.** Classification accuracy (%) for the University of Pavia image using training and testing samples.

Class	Train/Test	Noisy	Clean	BM4D[5]	PCABM4D[9]	3DTV	LRMR[17]	LRTV [21]	3DCrTV	3DCrWTV
Asphalt	50/6581	23.44	75.49	61.44	55.63	68.95	78.81	58.54	84.18	83.16
Meadows	50/18599	21.22	85.56	65.76	63.35	66.06	81.94	52.37	90.49	93.34
Gravel	50/2049	13.71	80.00	77.85	65.26	61.18	80.36	47.13	87.71	91.41
Trees	50/3014	54.11	93.02	79.49	79.78	84.84	95.56	78.24	93.77	92.78
Metal	50/1295	75.47	98.78	96.12	95.37	96.74	98.66	96.03	98.90	98.76
Bare soil	50/4979	18.35	84.76	72.52	62.42	66.38	91.68	56.35	93.39	95.43
Bitumen	50/1280	28.91	93.55	87.72	78.02	83.80	93.63	70.52	94.55	94.36
Bricks	50/3632	27.27	81.36	73.18	67.75	67.82	78.50	58.88	88.04	87.62
Shadows	50/897	99.93	99.79	99.63	98.26	99.75	99.88	99.11	99.82	99.63
OA (%)		27.28	84.75	70.39	65.84	69.98	84.44	58.82	90.32	91.71
std (%)		1.90	1.36	1.82	1.58	1.63	1.46	1.47	1.40	0.72
AA (%)		40.27	88.04	79.30	73.98	77.28	88.78	68.57	92.32	92.94
std (%)		0.93	0.56	0.84	1.03	1.10	0.52	0.72	0.67	0.60
Kappa		0.2017	0.8025	0.6318	0.5773	0.6211	0.8006	0.4934	0.8737	0.8903
std		0.0138	0.0163	0.0118	0.0171	0.0179	0.0175	0.0142	0.0176	0.0092

oversmoothed, this fact will be validated in the following classification experiment. LRMR can effectively remove all three kinds noises simultaneously, and almost produce the results as good as that of the proposed 3DCrTV algorithm. However, it fails to preserve some fine details of spatial structures [see the zoom-in portion in Fig. 13(f)]. LRTV could remove the noise of band 85 better than LRMR due to the smoothness enforced by TV regularization in spatial domain.

Fig. 14 presents the PSNR and SSIM values in each band of the reconstructed University of Pavia data set with different restoration methods. It is obvious that the proposed 3DCrTV and 3DCrWTV algorithm has an absolutely superior restoration performance in all of the bands. Furthermore, for the bands contaminated by these three kinds of noises simultaneously, 3DCrTV and 3DCrWTV can still produce results that are as good as that of the other bands corrupted by Gaussian noise only. These two phenomena both indicate that our methods have an effective and stable performance in simultaneously removing mixed Gaussian, impulse, and stripe noises.

Table 2 exhibits the MPSNR and MSSIM values obtained by the above six denoising algorithms as well as the proposed 3DCrWTV algorithm. In all indices, the 3DCrWTV method outperforms the other six restoration methods. As mentioned above, even LRTV could remove the mixed noise of band 85 very well, but the metrics are not better than LRMR, especially for the MSA.

Another main purpose of the second experiment is to evaluate the performance of the six restoration approaches according to the classification results. Table 3 exhibits the classification results of the University of Pavia data after restoration by BM4D, PCABM4D, 3DTV, LRMR, LRTV and
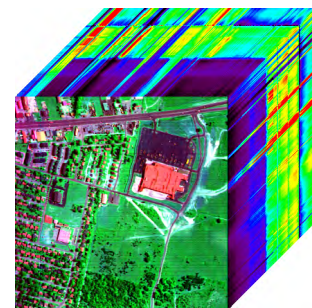
FIGURE 15. HYDICE Urban dataset used in the real experiment (false color image composed of bands 2, 103, 187 for the red, green and blue wavelength, respectively).

the proposed 3DCrTV and 3DCrTV methods, respectively. In the experiment of classification, the sparse logistic regression via variable splitting and augmented Lagrangian (LORSAL) algorithm proposed in [46] is adopted as the classifier, and 50 samples of each class are randomly selected as the training samples. The classification results are measured by the overall accuracy (OA), average accuracy (AA), and kappa statistic (Kappa). All of the measurements listed in Table 3 are achieved by averaging the results of ten Monte Carlo runs. It can be observed that classification results of the proposed 3DCrWTV method achieves the greater performance than the other six approaches in terms of OA, AA and Kappa statistic. Even 3DTV can produce restoration results as good as LRMR in terms of PSNR, but it leads to relatively poor classification results due to the oversmoothness of the spectra. LRTV usually could produce slightly better results than LRMR in removing the mixed noise, however, it is difficult to keep the balance between the low rank regularization and
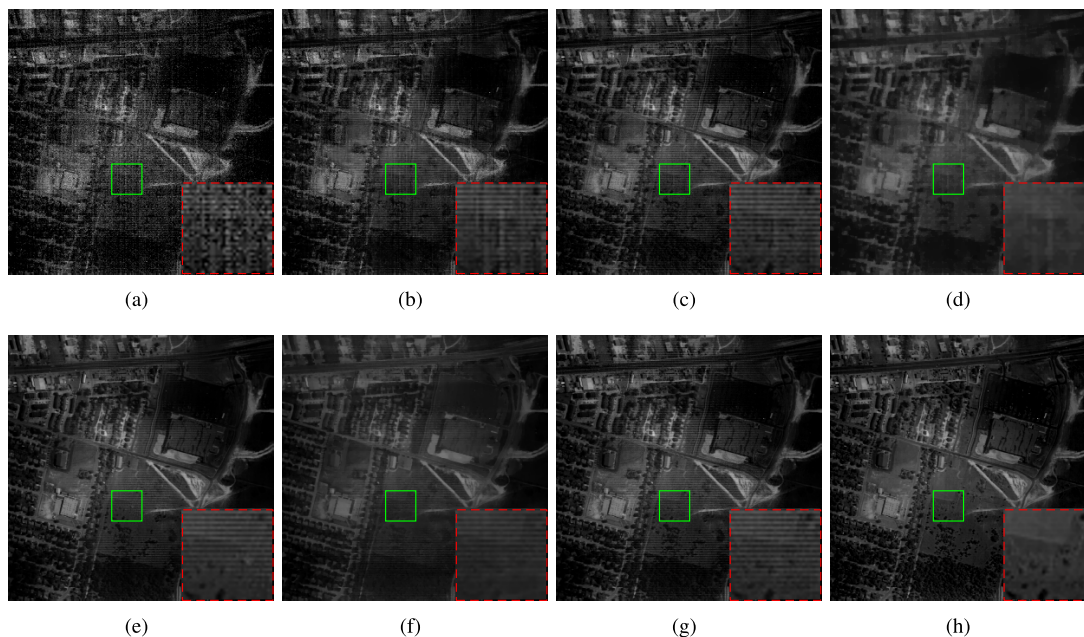
FIGURE 16. Image of band 109 of the restoration results in Urban dataset. (a) Original image; (b) BM4D; (c) PCABM4D; (d) 3DTV; (e) LRMR; (f) LRTV; (g) NAIRLMA; (h) 3DCrTV.

band-by-band TV regularization, thus distorting the spectra a lot. It is the reason LRTV produces the worst classification results among all competing methods.

In summary, the experiments of restoration and classification on the simulated datasets all indicate that the proposed 3DCrTV and 3DCrWTV methods both perform much better than the state-of-the-art methods, *i.e.*, BM4D, PCABM4D, 3DTV, LRMR and LRTV, in terms of quantitative assessment and visual effect.

C. REAL DATA EXPERIMENT

The hyperspectral digital collection experiment (HYDICE) dataset, Urban of Copperas Cove, Texas [47], is adopted for the real data experiment. The size of the scene is $307 \times 307 \times 220$. Among all bands, bands 104-108, 139-151, and 207-210 are severely polluted by the atmosphere and water absorption. The false color image composed of band 2, 103, and 187 is shown in Fig. 15. From Fig. 15, it is observed that the Urban data is corrupted by heavy Gaussian noise, strips and dead lines. In this experiment, we add a new state-of-the-art denoising method, *i.e.*, noise-adjusted iterative low-rank matrix approximation (NAILRMA) [18], for comparison, due to its promising performance for hyperspectral restoration.

Fig. 16 illustrates the image of band 109 and the reconstructed results with BM4D, PCABM4D, 3DTV, LRMR, LRTV, NAIRLMA and 3DCrTV methods, while Fig. 17 presents the reconstructed results of band 207. It is obviously seen that the BM4D and PCABM4D methods both perform poorly in removing the strip noise. For the parts contaminated by heavy Gaussian noise, lots of fine details are

oversmoothed, and for the edges between trees and grasses, the structures are distorted, see Fig. 16 (b) and (c). The 3DTV method can reduce the Gaussian noise and stripe noise to a certain extent, but it loses many fine details in the spatial domain. For low rank based methods, *i.e.*, LRMR, LRTV and NAIRLMA, they could remove the mixed noise to an extent, however, all of them fail to remove the heavy stripes in band 109 and 207 (see the zoom-in portions of the corresponding images in Fig. 16 and Fig. 17). The main reason is the fact that the stripes and dead lines of this scene locate at the same place in several adjacent bands. For example, the stripes exist at the same place from band 100 to 109 and also band 201 to 220. In this case, all low rank based methods would treat the stripes or dead lines as one of the low rank components and fail to get rid of them. For the proposed method, it could effectively remove all mixed noise (including Gaussian noise, impulse noise and structured stripes or dead lines) and preserve the fine details very well, due to the HTV regularization on the spectral difference space.

Fig. 18 plots the horizontal mean profiles of band 109, which is a good way to evaluate the performance of different restoration approaches in the ability of preserving spectrum. Due to the existence of stripes, there are rapid fluctuations in the results of BM4D, PCABM4D, LRMR, LRTV and NAIRLMA methods. It validates that all those methods fail to remove stripe noise completely. In addition, the 3DTV and the proposed 3DCrTV can effectively remove the stripes and produce much smoother mean profiles than other methods. However, more smoother mean profiles do not stand for a better restoration result, it should be jointly analyzed with the spatial details shown in Fig. 16 and 17. As in Fig. 18 (c), the
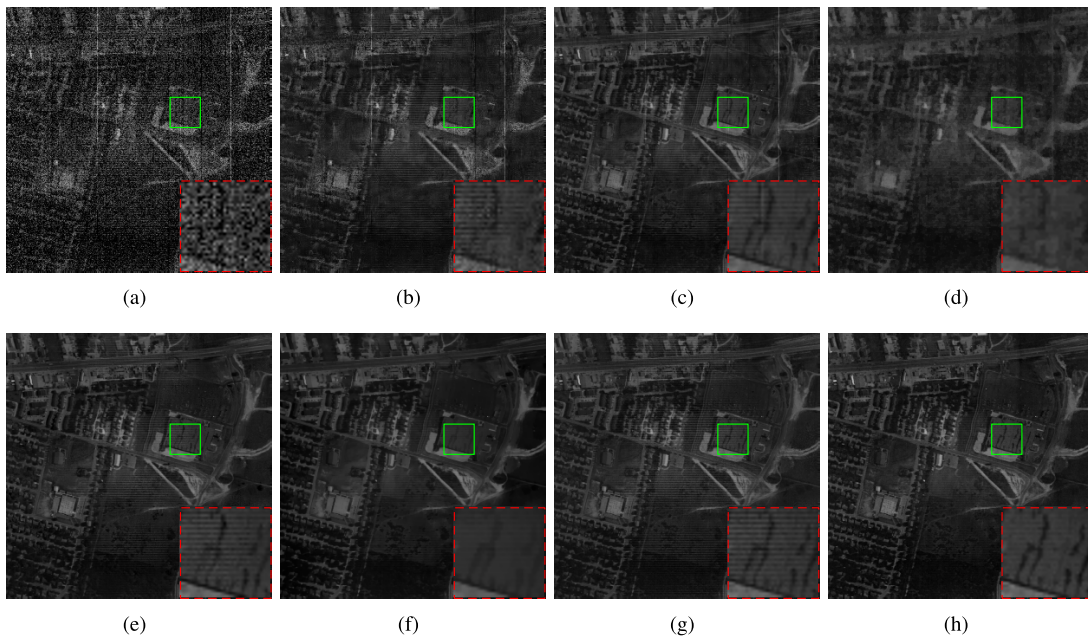
FIGURE 17. Image of band 207 of the restoration results in Urban dataset. (a) Original image; (b) BM4D; (c) PCABM4D; (d) 3DTV; (e) LRMR; (f) LRTV; (g) NAIRLMA; (h) 3DCrTV.
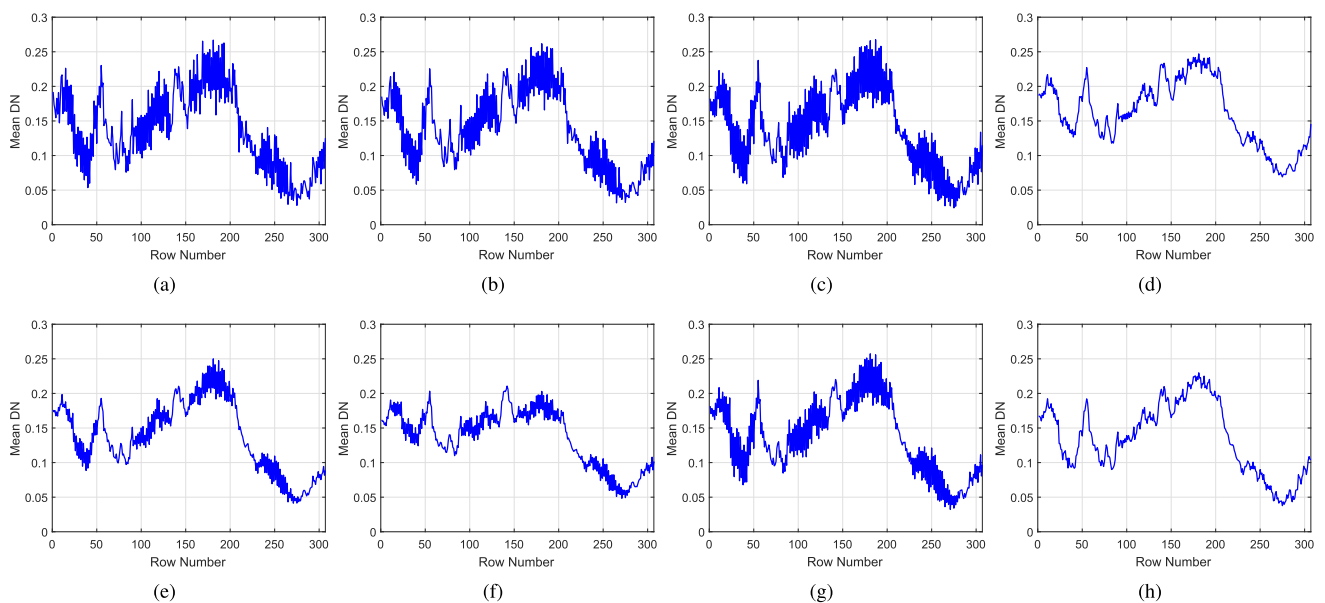
FIGURE 18. Horizontal mean profiles of Band 109 of the restoration results in Urban dataset. (a) Original image; (b) BM4D; (c) PCABM4D; (d) 3DTV; (e) LRMR; (f) LRTV; (g) NAIRLMA; (h) 3DCrTV.

3DTV algorithm can produce the mean profiles as smooth as that of 3DCrTV, due to the spectral TV constraint, but in the first two peaks of the spectrum, the DN values go up to almost 0.25, it is far from the linear regression of the original spectrum. Moreover, 3DTV loses a lot of details in spatial domain.

In a word, the proposed 3DCrTV can not only effectively remove Gaussian noise, impulse noise and structured stripes in the real HSI, but also has potentials in preserving the fine details and spectrum.

V. CONCLUSION

In this paper, a novel weighted cross total variation method has been proposed for HSI restoration in the framework of regularization variation. This CrTV consists of a 2-D spatial TV on the spectral difference image, and exploits the intrinsic property of smoothness and minimum energy of the spectral difference images by treating the HSI as 3-D cube. In order to solve the proposed model, which takes all mixed Gaussian, impulse and strip noises into consideration, 3DCrTV algorithm is proposed and accelerated by

nFFT. Experiments on two simulated HSI datasets and one real HSI dataset were conducted and validated that the proposed 3DCrWTV algorithm outperforms other state-of-the-art methods, *i.e.* BM4D, PCABM4D, 3DTV, LRMR, LRTV and NAIRLMA in removing the mixed Gaussian, impulse and strip noises simultaneously and effectively. In addition, the proposed 3DCrTV and 3DCrWTV algorithms are quite stable to the parameters λ_1 and λ_2.

As future work, we will consider a new spatial-spectral adaptive mechanism for the proposed model, which can automatically calculate the spatial-spectral weights according to the spatial structures and spectral noise intensity. Furthermore, in the framework of cross TV, many other conventional TV could be extended to define different cross TV and the CrTV model can be also easily extended to the nonlocal sense.

REFERENCES

[1] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. M. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.

[2] Z. Wu *et al.*, "GPU parallel implementation of spatially adaptive hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, to be published, doi: 10.1109/JSTARS.2017.2755639.

[3] A. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Trans. Geosci. Remote Sens.*, vol. 26, no. 1, pp. 65–74, Jan. 1988.

[4] J. B. Lee, A. S. Woodyatt, and M. Berman, "Enhancement of high spectral resolution remote-sensing data by a noise-adjusted principal components transform," *IEEE Trans. Geosci. Remote Sens.*, vol. 28, no. 3, pp. 295–304, May 1990.

[5] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[6] X. Zhang, M. Burger, X. Bresson, and S. Osher, "Bregmanized nonlocal regularization for deconvolution and sparse reconstruction," *SIAM J. Imag. Sci.*, vol. 3, no. 3, pp. 253–276, 2010.

[7] J. Wang, S. Lian, and Y. Q. Shi, "Hybrid multiplicative multi-watermarking in DWT domain," *Multidimensional Syst. Signal Process.*, vol. 28, no. 2, pp. 617–636, Apr. 2017.

[8] G. Chen and S.-E. Qian, "Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 3, pp. 973–980, Mar. 2014.

[9] G. Chen, T. D. Bui, K. G. Quach, and S.-E. Qian, "Denoising hyperspectral imagery using principal component analysis and block-matching 4D filtering," *Can. J. Remote Sens.*, vol. 40, no. 1, pp. 60–66, Mar. 2014.

[10] B. Rasti, J. R. Sveinsson, M. O. Ulfarsson, and J. A. Benediktsson, "Hyperspectral image denoising using 3D wavelets," in *Proc. IEEE IGARSS*, Jul. 2012, pp. 1349–1352.

[11] A. Zelinski and V. Goyal, "Denoising hyperspectral imagery and recovering junk bands using wavelets and sparse approximation," in *Proc. IEEE IGARSS*, Jul./Aug. 2006, pp. 387–390.

[12] B. Rasti, J. R. Sveinsson, and M. O. Ulfarsson, "Wavelet-based sparse reduced-rank regression for hyperspectral image restoration," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6688–6698, Oct. 2014.

[13] H. K. Aggarwal and A. Majumdar, "Exploiting spatiospectral correlation for impulse denoising in hyperspectral images," *J. Electron. Imag.*, vol. 24, no. 1, p. 013027, Feb. 2015.

[14] M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi, "Nonlocal transform-domain filter for volumetric data denoising and reconstruction," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 119–133, Jan. 2013.

[15] J. Wang, T. Li, Y. Q. Shi, S. Lian, and J. Ye, "Forensics feature analysis in quaternion wavelet domain for distinguishing photographic images and computer graphics," *Multimeida Tools Appl.*, vol. 76, no. 22, pp. 23721–23737, Nov. 2017.

[16] X. Lu, Y. Wang, and Y. Yuan, "Graph-regularized low-rank representation for destriping of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 4009–4018, Jul. 2013.

[17] H. Zhang, W. He, L. Zhang, H. Shen, and Q. Yuan, "Hyperspectral image restoration using low-rank matrix recovery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4729–4743, Aug. 2014.

[18] W. He, H. Zhang, L. Zhang, and H. Shen, "Hyperspectral image denoising via noise-adjusted iterative low-rank matrix approximation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sen.*, vol. 8, no. 6, pp. 3050–3061, Jun. 2015.

[19] R. Zhu, M. Dong, and J.-H. Xue, "Spectral nonlocal restoration of hyperspectral images with low-rank property," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 3062–3067, Jun. 2015.

[20] M. Wang, J. Yu, J.-H. Xue, and W. Sun, "Denoising of hyperspectral images using group low-rank representation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 9, pp. 4420–4427, Sep. 2016.

[21] W. He, H. Zhang, L. Zhang, and H. Shen, "Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 178–188, Jan. 2016.

[22] Q. Tian and S. Can, "Cross-heterogeneous-database age estimation through correlation representation learning," *Neurocomputing*, vol. 238, pp. 286–295, May 2017.

[23] Y.-Q. Zhao and J. Yang, "Hyperspectral image denoising via sparse representation and low-rank constraint," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 296–308, Jan. 2015.

[24] T. Lu, S. Li, L. Fang, Y. Ma, and J. A. Benediktsson, "Spectral–spatial adaptive sparse representation for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 373–385, Jan. 2016.

[25] Y. Zheng, J. Byeungwoo, D. Xu, Q. M. J. Wu, and H. Zhang, "Image segmentation by generalized hierarchical fuzzy C-means algorithm," *J. Intell. Fuzzy Syst.*, vol. 28, no. 2, pp. 961–973, Mar. 2015.

[26] C. Li, Y. Ma, J. Huang, X. Mei, and J. Ma, "Hyperspectral image denoising using the robust low-rank tensor recovery," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 32, no. 9, pp. 1604–1612, Sep. 2015.

[27] J. Yang, Y.-Q. Zhao, J. C.-W. Chan, and S. G. Kong, "Coupled sparse denoising and unmixing with low-rank constraint for hyperspectral image," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1818–1833, Mar. 2016.

[28] L. Sun, B. Jeon, Y. Zheng, and Z. Wu, "Hyperspectral image restoration using low-rank representation on spectral difference image," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 7, pp. 1151–1155, Jul. 2017.

[29] Y. W. Wen, M. K. Ng, and Y. M. Huang, "Efficient total variation minimization methods for color image restoration," *IEEE Trans. Image Process.*, vol. 17, no. 11, pp. 2081–2088, Nov. 2008.

[30] M. Jung, X. Bresson, T. F. Chan, and L. A. Vese, "Nonlocal Mumford-Shah regularizers for color image restoration," *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1583–1598, Jun. 2011.

[31] A. Chopra and H. Lian, "Total variation, adaptive total variation and non-convex smoothly clipped absolute deviation penalty for denoising blocky images," *Pattern Recognit.*, vol. 43, no. 8, pp. 2609–2619, Aug. 2010.

[32] Q. Yuan, L. Zhang, and H. Shen, "Hyperspectral image denoising employing a spectral–spatial adaptive total variation model," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 3660–3677, Oct. 2012.

[33] Y. Yuan, X. Zheng, and X. Lu, "Spectral–spatial kernel regularized for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3815–3832, Jul. 2015.

[34] Q. Cheng, H. Shen, L. Zhang, and P. Li, "Inpainting for remotely sensed images with a multichannel nonlocal total variation model," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 175–187, Jan. 2014.

[35] J. Li, Q. Yuan, H. Shen, and L. Zhang, "Hyperspectral image recovery employing a multidimensional nonlocal total variation model," *Signal Process.*, vol. 111, pp. 230–248, Jun. 2015.

[36] H. Zhang, "Hyperspectral image denoising with cubic total variation model," in *Proc. ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. I-7. 2012, pp. 95–98.

[37] Y. Chang, L. Yan, H. Fang, and H. Liu, "Simultaneous destriping and denoising for remote sensing images with unidirectional total variation and sparse representation," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 6, pp. 1051–1055, Jun. 2014.

[38] Y. Chang, L. Yan, H. Fang, and C. Luo, "Anisotropic spectral-spatial total variation model for multispectral remote sensing image destriping," *IEEE Trans. Image Process.*, vol. 24, no. 6, pp. 1852–1866, Jun. 2015.

[39] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4484–4502, Nov. 2012.

[40] L. Sun, Z. Wu, J. Liu, L. Xiao, and Z. Wei, "Supervised spectral–spatial hyperspectral image classification with weighted Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1490–1503, Mar. 2015.

[41] J. Eckstein and W. Yao, "Understanding the convergence of the alternating direction method of multipliers: Theoretical and computational perspectives," *Pacific J. Optim.*, vol. 11, no. 4, pp. 619–644, 2015.

[42] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[43] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.

[44] Q. Xie *et al.*, "Multispectral images denoising by intrinsic tensor sparsity regularization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, NV, USA, Jun. 2016, pp. 1692–1700.

[45] L. Sun, Y. Zheng, and B. Jeon, "Hyperspectral restoration employing low rank and 3D total variation regularization," in *Proc. IEEE Int. Conf. Prog. Inform. Comput.*, Shanghai, China, Dec. 2016, pp. 1–4.

[46] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Semisupervised hyperspectral image segmentation using multinomial logistic regression with active learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 4085–4098, Nov. 2010.

[47] U.S. Army Topographic Engineering Center. *HyperCube, Urban*. Accessed: 2015. [Online]. Available: http://www.tec.army.mil/Hypercube/

BYEUNGWOO JEON received the B.S. degree *(magna cum laude)* and the M.S. degree in electronics engineering from Seoul National University, Seoul, South Korea, in 1985 and 1987, respectively, and the Ph.D. degree in electrical engineering from Purdue University, West Lafayette, in 1992. From 1993 to 1997, he was with the Signal Processing Laboratory, Samsung Electronics, South Korea, where he conducted research and development into video compression algorithms, the design of digital broadcasting satellite receivers, and other MPEG-related research for multimedia applications. Since 1997, he has been with the Faculty of the School of Electronic and Electrical Engineering, Sungkyunkwan University, South Korea, where he is currently a Full Professor. He served as the Project Manager of Digital TV and Broadcasting with the Korean Ministry of Information and Communications from 2004 to 2006, where he supervised all digital TV-related R&D in South Korea. He has authored many papers in the areas of video compression, pre/post processing, and pattern recognition. His research interests include multimedia signal processing, video compression, statistical pattern recognition, and remote sensing. He is a member of Tau Beta Pi and Eta Kappa Nu. He is also a member of SPIE, IEEK, KICS, and KSOBE. He was a recipient of the 2005 IEEK Haedong Paper Award from the Signal Processing Society, South Korea.
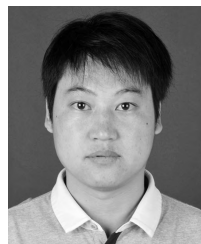
YUHUI ZHENG was born in China, in 1982. He received the B.S. degree in chemistry and the Ph.D. degree in computer science from the Nanjing University of Science and Technology, Nanjing, China, in 2004 and 2009, respectively. From 2014 to 2015, he was a Visiting Scholar at the Digital Media Laboratory, School of Electronic and Electrical Engineering, Sungkyunkwan University, South Korea. He is currently an Associate Professor at the School of Computer and Software, Nanjing University of Information Science and Technology. His research interests cover image processing, pattern recognition, and remote sensing information system.

LE SUN was born in China, in 1987. He received the B.S. degree from the School of Science, Nanjing University of Science and Technology (NJUST), Nanjing, China, in 2009, and the Ph.D. degree from the School of Computer Science and Engineering, NJUST, in 2014. He currently holds a postdoctoral position at the Digital Media Laboratory, School of Electronic and Electrical Engineering, Sungkyunkwan University, South Korea, where he is involved in research in the field of multi-images fusion based on sparse dictionary learning and compressive sensing. His research interests include hyperspectral image processing, sparse representation, and compressive sensing.

ZEBIN WU was born in China, in 1981. He received the B.S. and Ph.D. degrees from the School of Computer Science and Engineering, Nanjing University of Science and Technology, in 2003 and 2008, respectively. He is currently with the School of Computer Science and Engineering, Nanjing University of Science and Technology as a Full Professor. His research interests include virtual reality and system simulation, remote sensing information processing, and distributed computing.

• • •