Hyperspectral Image Restoration by Tensor Fibered Rank Constrained Optimization and Plug-and-Play Regularization

Yun-Yang Liu, Xi-Le Zhao, Yu-Bang Zheng, *Student Member, IEEE*, Tian-Hui Ma, and Hongyan Zhang, *Senior Member, IEEE*

Abstract—Hyperspectral images (HSIs) are often contaminated by several types of noise, which significantly limits the accuracy of subsequent applications. Recently, low-rank modeling based on tensor singular value decomposition (T-SVD) has achieved great success in HSI restoration. Most of them use the convex and nonconvex surrogates of the tensor rank, which cannot well approximate the tensor singular values and obtain suboptimal restored results. We suggest a novel HSI restoration model by introducing a fibered rank constrained tensor restoration framework with an embedded plug-and-play (PnP)-based regularization (FRCTR-PnP). More precisely, instead of using the convex and nonconvex surrogates to approximate the fibered rank, the proposed model directly constrains the tensor fibered rank of the solution, leading to a better approximation to the original image. Since exploiting the low-fibered-rankness of HSI is mainly to capture the global structure, we further employ an implicit PnP-based regularization to preserve the image details. Particularly, the above two building blocks are complementary to each other, rather than isolated and uncorrelated. Based on the alternating direction multiplier method (ADMM), we propose an efficient algorithm to tackle the proposed model. For robustness, we develop a three-directional randomized T-SVD (3DRT-SVD), which preserves the intrinsic structure of the clean HSI and removes partial noise by projecting the HSI onto a low-dimensional essential subspace. Extensive experimental results including simulated and real data demonstrate that the proposed method achieves superior performance over compared methods in terms of quantitative evaluation and visual inspection.

Index Terms—Hyperspectral image (HSI) restoration, low-fibered-rank, plug and play, three-directional randomized tensor singular value decomposition (3DRT-SVD).

I. INTRODUCTION

WITH the great progress of hyperspectral imaging sensors, hyperspectral images (HSIs) provide abundant spectral information and play an important role in classification [1]–[4], segmentation [5], and unmixing [6]–[11]. Due to hardware limitations and poor conditions, HSIs are inevitably contaminated by various kinds of noise, e.g., Gaussian noise, impulse noise, and deadlines, which degrade the HSI quality and limits the subsequent applications. Therefore, it is important to conduct HSI restoration, which aims to restore an underlying HSI from its corrupted version. For this purpose, a key is to depict the spatial and spectral features of HSI, e.g., the strong correlations among spatial pixels and spectral bands [12]–[32], see Fig. 1.

By exploring the spatial and spectral features of HSI, many restoration methods for specific noise were proposed. Since each band of an HSI can be viewed as a gray image, a direct method is to restore the HSI band-by-band by applying gray image restoration techniques, such as dictionary learning [33], nonlocal means filter [34], and weighted nuclear norm minimization [35]. These methods do not take the correlations among the spectral bands into consideration and thus usually provide unsatisfactory results. Subsequently, many restoration methods consider similar 3-D cubes as a basic unit to exploit nonlocal self-similarity (NSS) in both spatial and spectral domains, such as BM4D [36] and nonlocal tensor dictionary learning [37]. Besides, many works exploit the piecewise smoothness along spatial and spectral dimensions simultaneously, including spectral–spatial adaptive hyperspectral total variation (SSAHTV) [38], anisotropic spectral–spatial total variation (ASSTV) [39], and spatio-spectral total variation (SSTV) [40]. Low-rank modeling has been widely used in the HSI restoration. A representation is Tucker3 decomposition-based methods [41]–[44], which conducts both spatial low-rank approximation and spectral dimensionality reduction to HSIs. Subsequently, Parallel Factor Analysis was applied to HSI restoration in [45]. And then, several different types of regularizations [46]–[48] were considered based on [45] to achieve better restored results. To reduce the computational complexity, low-rank subspace representation is introduced for HSI restoration [49], [50]. The above methods limit the
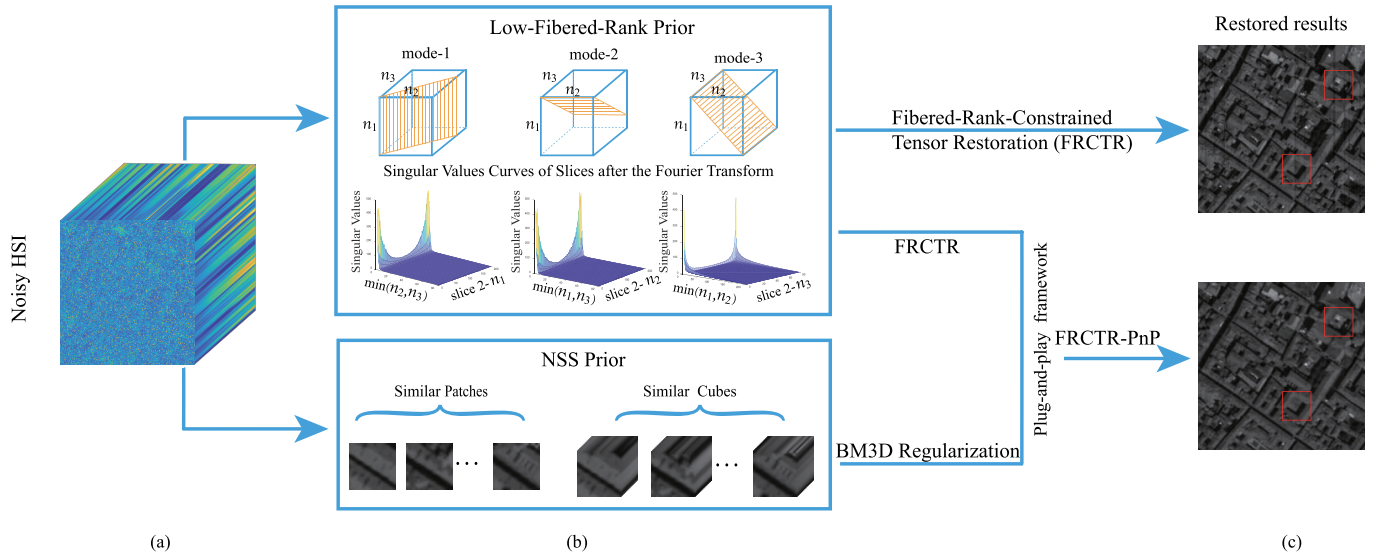
Fig. 1. PnP-based framework for HSI restoration. (a) Noisy HSI. (b) Illustration of low-fibered-rank prior and NSS prior. (c) Restored results of the simulated experiment.

kinds of noise, i.e., Gaussian noise. However, in real-world scenarios, the acquired HSIs are usually corrupted by several kinds of noise.

Recently, extensive works have been proposed on HSI restoration [51]–[57] for the mixed noise case. Exploiting the low-rankness of the clean HSI and the sparsity property of the non-Gaussian noise (including salt and pepper noise, stripe noise, and deadlines), the low-rank matrix recovery framework [51] was proposed to exploit the low-rank property of the unfolding matrix of HSI along the spectral mode. Peng *et al.* [52] suggested a mixed weighted nuclear norm and l_1 norm for enhancing low-rankness and sparsity, respectively. Considering piecewise smoothness, He *et al.* [54] proposed total variation (TV)-regularized low-rank matrix factorization, which used the matrix nuclear norm and TV regularization to characterize the spectral low-rankness and the spatial piecewise smooth structure, respectively. However, unfolding the HSI into a matrix destroys the spatial structure and leads to distortion of some bands in the restored results.

To preserve the global structure, tensor-based techniques have been proposed to characterize the tensorial structure of the HSI, which are motivated by different tensor decomposition schemes. Tucker decomposition considers the low-rankness of unfolding matrices along all three dimensions. Bai *et al.* [58] utilized the NSS to group similar cubes to third-order noisy tensors, and then used the nonnegative Tucker decomposition to solve the task of image restoration. In [59], the low-Tucker-rank model and SSTV regularizer were combined (LRTDTV) to exploit the global spatial-and-spectral correlation and enhance the spatial information, respectively. Chen *et al.* [60] took both Tucker decomposition and a weighted group sparsity term into consideration, which improved the restoration results compared with the previous TV methods. Zhang *et al.* [61] combined nonlocal low-rank Tucker decomposition and TV regularization, which exploited high correlation across the spectral bands and captured the NSS, respectively. But the unbalanced matricization scheme of

the Tucker format causes difficulty in characterizing the global correlation [62]. Based on tensor singular value decomposition (T-SVD) [63], Fan *et al.* [64] incorporated SSTV regularization into a low-rank tensor factorization framework (SSTV-LRTF), which used tensor nuclear norm (TNN) to present the low-rank property of HSIs and SSTV regularization to exploit the piecewise smoothness among spatial and spectral domains. The T-SVD framework lacks flexibility in describing different correlations of all modes of HSIs, resulting in suboptimal restoration performance. Therefore, Zheng *et al.* [65] generalized T-SVD to the mode-*k* T-SVD, proposed the fibered rank, and introduced a convex surrogate (3DTNN) and nonconvex surrogate (3DLogTNN) to approximate the tensor fibered rank. However, the convex and nonconvex surrogates cannot well approximate the singular values and singular vectors of slices after the Fourier transform. Fig. 2 shows the comparison between a clean HSI and the restored results obtained by 3DTNN and 3DLogTNN on singular values and singular vectors, based on the first slice after the Fourier transform along the second mode. We clearly observe that the singular values and singular vectors obtained by 3DTNN and 3DLogTNN do not approximate well to those of the original ones.

To tackle the HSI restoration problem, we introduce a fibered rank constrained tensor restoration framework with an embedded Plug-and-Play (PnP)-based regularization (FRCTR-PnP). The fibered rank mainly helps to characterize the global structure, and the PnP-based regularization helps to preserve fine details and achieves better-restored results. These two building blocks are complementary to each other, rather than isolated and uncorrelated. More specifically, the contributions of this article are threefold.

First, instead of using a convex surrogate or nonconvex surrogate to approximate the fibered rank, we directly constrain the tensor fibered rank of the solution, which achieves a better approximation to the clean HSI in terms of fibered rank, leading tensor singular values, and the corresponding singular
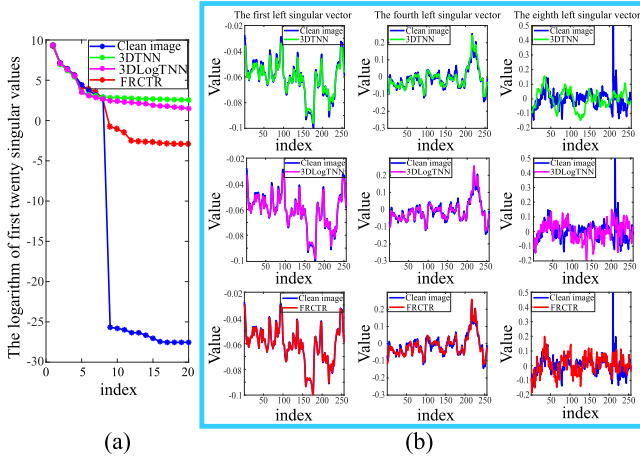
Fig. 2. Comparison among restored results obtained by 3DTNN, 3DLogTNN, and FRCTR on the approximation of singular values and singular vectors. (a) Logarithm of the first twenty singular values of the first slice after the Fourier transform. (b) Three left singular vectors of the first slice after the Fourier transform.

vectors. Fig. 2 illustrates that the fibered rank, singular values, and singular vectors of FRCTR are closer to those of the clean HSI than 3DTNN and 3DLogTNN.

Second, only considering the low-rankness prior usually faces limitations in preserving local details and removing the noise with low-rank property, such as stripe noise and deadlines. Therefore, we introduce an implicit regularizer under the PnP framework [66] to break through the limitations, which is flexible to embed a large number of general denoisers (e.g., classic BM3D denoiser [67] and advanced deep-learning denoisers [68]–[74]). This article employs the classic BM3D denoiser as an example, which exploits the NSS prior. The motivation is that the plugged BM3D denoiser can preserve the image details conveyed by dropping small tensor singular values due to low-fibered-rankness promotion.

Third, to efficiently tackle the proposed FRCTR-PnP model, we develop an algorithm based on alternating direction multiplier method (ADMM) to decompose the original problem into several simple subproblems [66]. Besides, we design a three-directional randomized T-SVD (3DRT-SVD) to tackle the low-fibered-rank subproblem. The 3DRT-SVD can retain the intrinsic information of the clean HSI while removing partial noise by projecting the HSI onto a low-dimensional essential subspace. Extensive experiments on simulated and real data demonstrate that the proposed method yields superior performance over the compared methods in HSI restoration.

The remainder of this article is organized as follows. Section II gives the necessary preliminaries to facilitate our presentation. Section III introduces the proposed FRCTR-PnP model and the ADMM-based solving algorithm. Section IV presents experimental results and discussion. Section V concludes this article.

II. NOTATION AND PRELIMINARIES

In this section, we summarize the minimal and necessary notations in Table I. Next, we introduce some definitions for subsequent discussion.

TABLE I
NOTATIONS

Notation	Explanation
$x, \boldsymbol{x}, X, \mathcal{X}$	scalar, vector, matrix, tensor.
$\|X\|_*$	nuclear norm of X.
$x_{i,j,k}$	the (i, j, k)-th entry of $\mathcal{X}$.
$\mathcal{X}(i, j, :)$	the (i, j)-th tube of $\mathcal{X}$.
$\mathcal{X}(i, :, :)$	the i-th horizontal slice of $\mathcal{X}$.
$\mathcal{X}(:, i, :)$	the i-th lateral slice of $\mathcal{X}$.
$\mathcal{X}(:, :, i)$ or X^i	the i-th frontal slice of $\mathcal{X}$.
$\star$	circular convolution.
$\langle \mathcal{X}, \mathcal{Y} \rangle$	inner product of tenors $\mathcal{X}$ and $\mathcal{Y}$, $\langle \mathcal{X}, \mathcal{Y} \rangle = \sum_{i,j,k} x_{i,j,k} y_{i,j,k}$.
$\|\mathcal{X}\|_F$	Frobenius norm, $\|\mathcal{X}\|_F = \sqrt{\langle \mathcal{X}, \mathcal{X} \rangle}$.
$\|\mathcal{X}\|_1$	L_1 norm, $\|\mathcal{X}\|_1 = \sum_{i,j,k} |\mathcal{X}_{i,j,k}|$.
$*$	t-product, see Definition 1.
$*_k$	mode-k t-product, see Definition 6.
$\mathrm{rank}_t(\mathcal{X})$	tubal rank of $\mathcal{X}$, see Definition 4.
$\mathrm{rank}_f(\mathcal{X})$	fibered rank of $\mathcal{X}$, see Definition 8.

Definition 1 (t-Product [63]): The t-product between $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $\mathcal{Y} \in \mathbb{R}^{n_2 \times n_4 \times n_3}$, denoted as $\mathcal{X} * \mathcal{Y}$, is a tensor $\mathcal{Z}$ of size $n_1 \times n_4 \times n_3$ with tubes

$$\mathcal{Z}(m, n, :) = \sum_{t=1}^{n_2} \mathcal{X}(m, t, :) \star \mathcal{Y}(t, n, :) \qquad (1)$$

where "$\star$" denotes circular convolution [75] of two tubes.

Definition 2 (Conjugate Transpose [63]): The conjugate transpose of $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, denoted as $\mathcal{X}^H$, is constructed by conjugate transposing each frontal slice of $\mathcal{X}$ and then reversing the order of transposed frontal slices 2 through n_3.

Definition 3 (T-SVD [76]): Let $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ be a third-order tensor, then $\mathcal{X}$ can be factorized as

$$\mathcal{X} = \mathcal{U} * \mathcal{D} * \mathcal{V}^H \qquad (2)$$

where $\mathcal{U} \in \mathbb{R}^{n_1 \times n_1 \times n_3}$ and $\mathcal{V} \in \mathbb{R}^{n_2 \times n_2 \times n_3}$ satisfy $\mathcal{U} * \mathcal{U}^H = \mathcal{V} * \mathcal{V}^H = \mathcal{I}$, and $\mathcal{U}(:, j, k)$ and $\mathcal{V}(:, j, k)$ are called the left singular vectors and right singular vectors of $\mathcal{X}$, respectively. $\mathcal{D} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is an f-diagonal tensor where each of its frontal slices is diagonal, and the entries in $\mathcal{D}$ are called the tensor singular values of $\mathcal{X}$.

Definition 4 (Tensor Tubal Rank [76]): The tubal rank of $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, denoted as $\mathrm{rank}_t(\mathcal{X})$, is the number of nonzero tubes of $\mathcal{D}$ from the decomposition $\mathcal{X} = \mathcal{U} * \mathcal{D} * \mathcal{V}^H$, that is

$$\mathrm{rank}_t(\mathcal{X}) = \#\{i : \mathcal{D}(i, i, :) \neq 0\} = \max_k \mathrm{rank}(\bar{X}^k) \qquad (3)$$

where $\bar{\mathcal{X}}$ is computed by taking the Fast Fourier Transform (FFT) along the third mode of $\mathcal{X}$, and $\bar{X}^k$ is the kth frontal slide of $\bar{\mathcal{X}}$.

Definition 5 (Tensor Mode-k Permutation [65]): The mode-k permutation of $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, denoted as $\mathcal{X}^{\bar{k}}$ or $\mathrm{permute}(\mathcal{X}, \mathrm{k})$, satisfies that $\mathcal{X}_{i,j,k} = \mathcal{X}_{j,k,i}^{\bar{1}} = \mathcal{X}_{k,i,j}^{\bar{2}} = \mathcal{X}_{i,j,k}^{\bar{3}}$. We define the inverse operator as $\mathrm{ipermute}(\mathcal{X}, \mathrm{k})$, such that $\mathcal{X} = \mathrm{ipermute}(\mathrm{permute}(\mathcal{X}, k), k)$.

The tensor tubal rank lacks flexibility when dealing with the different relationships of all modes of HSIs, which leads to inadequate characterization for the spectral mode. Therefore, Zheng *et al.* [65] extended mode-3 operations to mode-k operations, such as mode-k conjugate transpose, mode-k t-product, and mode-k T-SVD.

Definition 6: The mode-1, mode-2, and mode-3 t-product [65] are defined as follows:

The mode-1 t-product between $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $\mathcal{Y} \in \mathbb{R}^{n_1 \times n_3 \times n_4}$, denoted as $\mathcal{X} *_1 \mathcal{Y}$, is a tensor $\mathcal{Z}$ of size $n_1 \times n_2 \times n_4$ with tubes

$$\mathcal{Z}(:, j, k) = \sum_{t=1}^{n_3} \mathcal{X}(:, j, t) \star \mathcal{Y}(:, t, k). \tag{4}$$

The mode-2 t-product between $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $\mathcal{Y} \in \mathbb{R}^{n_4 \times n_2 \times n_1}$, denoted as $\mathcal{X} *_2 \mathcal{Y}$, is a tensor $\mathcal{Z}$ of size $n_4 \times n_2 \times n_3$ with tubes

$$\mathcal{Z}(i, :, k) = \sum_{t=1}^{n_1} \mathcal{X}(t, :, k) \star \mathcal{Y}(i, :, t). \tag{5}$$

The mode-3 t-product between $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $\mathcal{Y} \in \mathbb{R}^{n_2 \times n_4 \times n_3}$, denoted as $\mathcal{X} *_3 \mathcal{Y}$, is a tensor $\mathcal{Z}$ of size $n_1 \times n_4 \times n_3$ with tubes

$$\mathcal{Z}(i, j, :) = \sum_{t=1}^{n_2} \mathcal{X}(i, t, :) \star \mathcal{Y}(t, j, :). \tag{6}$$

Based on mode-k t-product, we have the mode-k T-SVD.

Definition 7 (Mode-k T-SVD [65]): Let $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ be a third-order tensor, then it can be factorized as

$$\mathcal{X} = \mathcal{U}_k *_k \mathcal{D}_k *_k \mathcal{V}_k^{H_k} \tag{7}$$

where $\mathcal{U}_k^{\bar{k}}$ and $\mathcal{V}_k^{\bar{k}}$ satisfy $\mathcal{U}_k^{\bar{k}} * (\mathcal{U}_k^{\bar{k}})^H = \mathcal{V}_k^{\bar{k}} * (\mathcal{V}_k^{\bar{k}})^H = \mathcal{I}$, $\mathcal{D}_k^{\bar{k}}$ is a f-diagonal tensor.

Definition 8 (Mode-k Tensor Fibered Rank and Tensor Fibered Rank [65]): The mode-k fibered rank of $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, denoted as $\text{rank}_{f_k}(\mathcal{X})$, is defined as the number of nonzero mode-k fibers of $\mathcal{D}_k$. Then, the fibered rank of $\mathcal{X}$, denoted as $\text{rank}_f(\mathcal{X})$, is defined as a vector, and its kth element is the mode-k tensor fibered rank.

Theorem 1: For $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, its tensor fibered rank and tensor tubal rank have the following relationship:

$$\text{rank}_{f_k}(\mathcal{X}) = \text{rank}_t(\mathcal{X}^{\bar{k}}). \tag{8}$$

III. PNP-REGULARIZED FIBERED RANK-CONSTRAINED TENSOR RESTORATION

Assuming that the noise is independent additive noise, we consider the following degradation model in our work:

$$\mathcal{Y} = \mathcal{X} + \mathcal{N} + \mathcal{S} \tag{9}$$

where the tensors $\mathcal{Y}$, $\mathcal{X}$, $\mathcal{N}$, and $\mathcal{S}$ denote the corrupted HSI, the clean HSI, the Gaussian noise, and the sparse noise (including salt and peppers, stripes, and deadlines).

Based on this degradation model, we propose an FRCTR-PnP to restore the global structural information of the clean HSI by the tensor fibered rank constraint and preserve the fine details by the PnP-based regularization. Following the above discussions, we formulate the FRCTR-PnP model as:

$$\min_{\mathcal{X}, \mathcal{S}} \quad \|\mathcal{S}\|_1 + \lambda \Phi(\mathcal{X})$$
$$\text{s.t. } \|\mathcal{Y} - \mathcal{X} - \mathcal{S}\|_F^2 \le \epsilon, \quad \text{rank}_f(\mathcal{X}) \le [r_1, r_2, r_3] \tag{10}$$

where λ is a tuning parameter and $[r_1, r_2, r_3]$ is the upper bound of the fibered rank of $\mathcal{X}$. $\Phi(\mathcal{X})$ is an implicit regularizer exploiting certain priors of the HSI, which can be chosen from a large number of denoisers (e.g., BM3D denoiser [67] and deep-learning denoisers [68]–[74]). By introducing auxiliary variables $\mathcal{F}_k$ ($k = 1, 2, 3$) and $\mathcal{L}$, (10) can be rewritten as

$$\min_{\mathcal{X}, \mathcal{S}, \mathcal{F}_k} \quad \|\mathcal{S}\|_1 + \lambda \Phi(\mathcal{L})$$
$$\text{s.t. } \|\mathcal{Y} - \mathcal{X} - \mathcal{S}\|_F^2 \le \epsilon, \quad \mathcal{X} = \mathcal{L}$$
$$\mathcal{X} = \mathcal{F}_k, \quad \text{rank}_{f_k}(\mathcal{F}_k) \le r_k \quad (k = 1, 2, 3) \tag{11}$$

where $\text{rank}_{f_k}(\mathcal{X})$ denotes the mode-k tensor fibered rank of $\mathcal{X}$. We design an ADMM-based algorithm under PnP framework [77] by introducing the Lagrangian multipliers. We have the augmented Lagrangian function of (11) as

$$\mathcal{L}_{\beta_k}(\mathcal{X}, \mathcal{F}_k, \mathcal{L}, \mathcal{S}, \mathcal{P}_k)$$
$$= \sum_{k=1}^{3} \left\{ \langle \mathcal{X} - \mathcal{F}_k, \mathcal{P}_k \rangle + \frac{\beta_k}{2} \|\mathcal{X} - \mathcal{F}_k\|_F^2 \right\} + \|\mathcal{S}\|_1$$
$$+ \langle \mathcal{Y} - \mathcal{X} - \mathcal{S}, \mathcal{P}_4 \rangle + \frac{\beta_4}{2} \|\mathcal{Y} - \mathcal{X} - \mathcal{S}\|_F^2$$
$$+ \lambda \Phi(\mathcal{L}) + \langle \mathcal{X} - \mathcal{L}, \mathcal{P}_5 \rangle + \frac{\beta_5}{2} \|\mathcal{X} - \mathcal{L}\|_F^2$$
$$\text{s.t. } \text{rank}_{f_k}(\mathcal{F}_k) \le r_k \quad (k = 1, 2, 3) \tag{12}$$

where $\mathcal{P}_k$ ($k = 1, \ldots, 5$) are Lagrangian multipliers, and β_k ($k = 1, \ldots, 5$) are penalty parameters. Within the framework of ADMM, $\mathcal{X}$, $\mathcal{F}_k$, $\mathcal{L}$, and $\mathcal{S}$ can be divided into two groups.

The first group is

$$\left(\mathcal{F}_k^{t+1}, \mathcal{L}^{t+1}, \mathcal{S}^{t+1} \right) = \underset{\mathcal{F}_k, \mathcal{L}, \mathcal{S}}{\arg\min} \ L_{\beta_k}\left(\mathcal{X}^t, \mathcal{F}_k, \mathcal{L}, \mathcal{S}, \mathcal{P}_k^t \right). \tag{13}$$

These variables are decoupled from each other, so they can be solved separately

$$\begin{cases} \mathcal{F}_k^{t+1} = \underset{\text{rank}_{f_k}(\mathcal{F}_k) \le r_k}{\arg\min} \ L_{\beta_k}\left(\mathcal{X}^t, \mathcal{F}_k, \mathcal{L}^t, \mathcal{S}^t, \mathcal{P}_k^t \right) \\ \mathcal{L}^{t+1} = \underset{\mathcal{L}}{\arg\min} \ L_{\beta_k}\left(\mathcal{X}^t, \mathcal{F}_k^{t+1}, \mathcal{L}, \mathcal{S}^t, \mathcal{P}_k^t \right) \\ \mathcal{S}^{t+1} = \underset{\mathcal{S}}{\arg\min} \ L_{\beta_k}\left(\mathcal{X}^t, \mathcal{F}_k^{t+1}, \mathcal{L}^{t+1}, \mathcal{S}, \mathcal{P}_k^t \right). \end{cases} \tag{14}$$

The second group is

$$\mathcal{X}^{t+1} = \underset{\mathcal{X}}{\arg\min} \ L_{\beta_k}\left(\mathcal{X}, \mathcal{F}_k^{t+1}, \mathcal{L}^{t+1}, \mathcal{S}^{t+1}, \mathcal{P}_k^t \right). \tag{15}$$

Specifically, the variables are alternately updated as follows:
1) *Update $\mathcal{F}_k$ ($k = 1, 2, 3$)*

$$\mathcal{F}_k^{t+1} = \underset{\text{rank}_{f_k}(\mathcal{F}_k) \le r_k}{\arg\min} \ \frac{\beta_k}{2} \left\| \mathcal{X}^t - \mathcal{F}_k + \frac{\mathcal{P}_k^t}{\beta_k} \right\|_F^2. \tag{16}$$

Invoking Theorem 1 in [65]

$$\mathcal{Z} = \mathcal{X} *_k \mathcal{Y} \Leftrightarrow \mathcal{Z}^{\bar{k}} = \mathcal{X}^{\bar{k}} * \mathcal{Y}^{\bar{k}}. \tag{17}$$

We use the mode-k permutation operator to conveniently calculate the mode-k t-product and mode-k T-SVD. Let $\hat{\mathcal{F}} = \mathcal{X}^t + \mathcal{P}_k^t / \beta_k$, the objective is to find low fibered rank tensors $\mathcal{F}_k$ ($\text{rank}_{f_k}(\mathcal{F}_k) \le r_k$, $k = 1, 2, 3$) while the residual tensors $\|\mathcal{F}_k - \hat{\mathcal{F}}\|_F^2$ are minimized. The best "mode-k tensor fibered rank"

Algorithm 1 3DTT-SVD

Input: $\hat{\mathcal{F}} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, target fibered rank $[r_1, r_2, r_3]$.
1: $\hat{\mathcal{F}}_1 = \texttt{permute}(\hat{\mathcal{F}}, 1) \in \mathbb{R}^{n_2 \times n_3 \times n_1}$, $\hat{\mathcal{F}}_2 = \texttt{permute}(\hat{\mathcal{F}}, 2) \in \mathbb{R}^{n_3 \times n_1 \times n_2}$, and $\hat{\mathcal{F}}_3 = \texttt{permute}(\hat{\mathcal{F}}, 3) \in \mathbb{R}^{n_1 \times n_2 \times n_3}$;
2: Compute T-SVD of $\hat{\mathcal{F}}_1, \hat{\mathcal{F}}_2, \hat{\mathcal{F}}_3$, truncate them with target rank r_1, r_2, r_3, and obtain $\mathcal{U}_1^{r_1}, \mathcal{S}_1^{r_1}, \mathcal{V}_1^{r_1}, \mathcal{U}_2^{r_2}, \mathcal{S}_2^{r_2}, \mathcal{V}_2^{r_2}, \mathcal{U}_3^{r_3}, \mathcal{S}_3^{r_3}, \mathcal{V}_3^{r_3}$;
3: $\mathcal{F}_1 = \mathcal{U}_1^{r_1} * \mathcal{S}_1^{r_1} * \mathcal{V}_1^{r_1\,H}$, $\mathcal{F}_2 = \mathcal{U}_2^{r_2} * \mathcal{S}_2^{r_2} * \mathcal{V}_2^{r_2\,H}$, and $\mathcal{F}_3 = \mathcal{U}_3^{r_3} * \mathcal{S}_3^{r_3} * \mathcal{V}_3^{r_3\,H}$;
4: $\mathcal{F}_1 = \texttt{ipermute}(\mathcal{F}_1, 1)$, $\mathcal{F}_2 = \texttt{ipermute}(\mathcal{F}_2, 2)$, and $\mathcal{F}_3 = \texttt{ipermute}(\mathcal{F}_3, 3)$.
Output: $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$.

Algorithm 2 3DRT-SVD

Input: $\hat{\mathcal{F}} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, target fibered rank $[r_1, r_2, r_3]$, and oversampling parameter $[p_1, p_2, p_3]$.
1: Compute $\hat{\mathcal{F}}_1 = \texttt{permute}(\hat{\mathcal{F}}, 1)$, $\hat{\mathcal{F}}_2 = \texttt{permute}(\hat{\mathcal{F}}, 2)$, and $\hat{\mathcal{F}}_3 = \texttt{permute}(\hat{\mathcal{F}}, 3)$;
2: Generate Gaussian random tensors $\mathcal{G}_1 \in \mathbb{R}^{n_3 \times (r_1+p_1) \times n_1}$, $\mathcal{G}_2 \in \mathbb{R}^{n_1 \times (r_2+p_2) \times n_2}$, and $\mathcal{G}_3 \in \mathbb{R}^{n_2 \times (r_3+p_3) \times n_3}$;
3: Compute $\mathcal{Z}_1 = \hat{\mathcal{F}}_1 * \mathcal{G}_1$, $\mathcal{Z}_2 = \hat{\mathcal{F}}_2 * \mathcal{G}_2$, and $\mathcal{Z}_3 = \hat{\mathcal{F}}_3 * \mathcal{G}_3$;
4: Compute $\mathcal{Q}_1 \in \mathbb{R}^{n_2 \times (r_1+p_1) \times n_1}$, $\mathcal{Q}_2 \in \mathbb{R}^{n_3 \times (r_2+p_2) \times n_2}$, and $\mathcal{Q}_3 \in \mathbb{R}^{n_1 \times (r_3+p_3) \times n_3}$ by using t-QR factorization of $\mathcal{Z}_1$, $\mathcal{Z}_2$, $\mathcal{Z}_3$;
5: Compute $\mathcal{C}_1 = \mathcal{Q}_1^T * \hat{\mathcal{F}}_1$, $\mathcal{C}_2 = \mathcal{Q}_2^T * \hat{\mathcal{F}}_2$, and $\mathcal{C}_3 = \mathcal{Q}_3^T * \hat{\mathcal{F}}_3$;
6: Compute $\mathcal{B}_k$ as the r_k truncation T-SVD of $\mathcal{C}_k$, where $k = 1, 2, 3$;
7: $\hat{\mathcal{F}}_1 = \mathcal{Q}_1 * \mathcal{B}_1$, $\hat{\mathcal{F}}_2 = \mathcal{Q}_2 * \mathcal{B}_2$, $\hat{\mathcal{F}}_3 = \mathcal{Q}_3 * \mathcal{B}_3$;
8: $\mathcal{F}_1 = \texttt{ipermute}(\hat{\mathcal{F}}_1, 1)$, $\mathcal{F}_2 = \texttt{ipermute}(\hat{\mathcal{F}}_2, 2)$, and $\mathcal{F}_3 = \texttt{ipermute}(\hat{\mathcal{F}}_3, 3)$.
Output: $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$.

---

approximation has an explicit solution obtained by truncated T-SVD. The specific algorithm is described in Algorithm 1.

Computing the T-SVD in step 2 of Algorithm 1 needs to compute the SVDs in the Fourier domain, which is extremely time-consuming. To solve this issue, we introduce the following lemma.

Lemma 1: For $\hat{\mathcal{F}} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the outputs $\mathcal{F}_1$, $\mathcal{F}_2$, and $\mathcal{F}_3$ of Algorithm 1 have the following relationships:

$$\begin{cases} \mathcal{F}_1 = \mathcal{U}_1^{r_1} * \mathcal{U}_1^{r_1\,H} * \hat{\mathcal{F}}_1 \\ \mathcal{F}_2 = \mathcal{U}_2^{r_2} * \mathcal{U}_2^{r_2\,H} * \hat{\mathcal{F}}_2 \\ \mathcal{F}_3 = \mathcal{U}_3^{r_3} * \mathcal{U}_3^{r_3\,H} * \hat{\mathcal{F}}_3. \end{cases} \tag{18}$$

Here, the lemma shows that $\mathcal{F}_1$, $\mathcal{F}_2$, and $\mathcal{F}_3$ can be regarded as the projection of $\hat{\mathcal{F}}_1$, $\hat{\mathcal{F}}_2$, and $\hat{\mathcal{F}}_3$ on the image subspace along the first, second, and third mode, respectively, which motivates us to design a random projection algorithm (3DRT-SVD) to replace three-directional truncated T-SVD (3DTT-SVD). By projecting the HSI to a low-dimensional essential subspace, the proposed 3DRT-SVD preserves the intrinsic information of the clean HSI and removes partial noise. Meanwhile, 3DRT-SVD reduces the time cost since the size

of the subspace is much smaller than that of the original data. The details of 3DRT-SVD are shown in Algorithm 2.

We compare the complexity of Algorithms 1 and 2. The main computation burden of the two algorithms comes from T-SVD. By introducing the random projection method, the computation complexity of calculating SVDs is reduced from $\mathcal{O}(n_1 n_2 n_3 \sum_{i=1}^{3} \min(n_i, n_{i+1}))$, $n_4 = n_1$ for 3DTT-SVD to $\mathcal{O}(\sum_{i=1}^{3} n_i n_{i-1} (r_i + p_i)^2)$, $n_0 = n_3$ for 3DRT-SVD. Since the rank of spectral bands is small, e.g., $r_1 + p_1 \ll n_1$, $r_2 + p_2 \ll n_2$, the compression of first and second modes is significant.

According to the discussion above, we use the 3DRT-SVD algorithm to compute the following closed-form solution of $\mathcal{F}_k$-subproblems:

$$\mathcal{F}_k^{t+1} = \text{3DRT-SVD}\left(\mathcal{X}^t + \frac{\mathcal{P}_k^t}{\beta_k}, [r_1, r_2, r_3], [p_1, p_2, p_3]\right). \tag{19}$$

The computational complexity of updating $\mathcal{F}_k^{t+1}$ ($k = 1, 2, 3$) is $\mathcal{O}(n_1 n_2 n_3 \log(n_1 n_2 n_3) + n_1 n_2 n_3 \sum_{i=1}^{3} (r_i + p_i))$.

2) Update $\mathcal{L}$

$$\mathcal{L}^{t+1} = \operatorname*{argmin}_{\mathcal{L}} \ \lambda \Phi(\mathcal{L}) + \frac{\beta_5}{2} \left\| \mathcal{X}^t - \mathcal{L} + \frac{\mathcal{P}_5^t}{\beta_5} \right\|_F^2. \tag{20}$$

Letting $\sigma = (\lambda/\beta_5)^{1/2}$, (20) can be rewritten as

$$\text{prox}_\Phi(\mathcal{L}^{t+1}) = \operatorname*{argmin}_{\mathcal{L}} \Phi(\mathcal{L}) + \frac{1}{2\sigma^2} \left\| \mathcal{X}^t - \mathcal{L} + \frac{\mathcal{P}_5^t}{\beta_5} \right\|_F^2. \tag{21}$$

Based on the PnP framework, the proximal operator of regularization $\text{prox}_\Phi : \mathbb{R}^{n_1 \times n_2 \times n_3} \to \mathbb{R}^{n_1 \times n_2 \times n_3}$ is replaced by the state-of-the-art denoiser, which reconstructs the clean image from the noisy image. Here, the denoiser acts as an implicit regularizer to express HSIs priors, e.g., piecewise smoothness and NSS.

BM3D is a denoiser achieving an enhanced sparse representation of natural images by performing a 3-D transform-domain collaborative filtering on groups of similar patches. More specifically, in BM3D, similar patches are stacked into 3-D groups by block matching, and the 3-D groups are transformed into the transform domain. Then, hard thresholding or Wiener filtering with coefficients is employed in the transform domain. Finally, after an inverse transform of coefficients, all estimated patches are aggregated to reconstruct the whole image. BM3D can achieve promising denoising performance by exploiting simultaneously the sparsity and the NSS of natural images. It also has shown good generalization ability and efficient implementation, thus we choose BM3D as the example denoiser in our FRCTR-PnP framework. The solution is given by

$$\mathcal{L}^{t+1} = \text{BM3D}\left(\mathcal{X}^t + \frac{\mathcal{P}_5^t}{\beta_5}, \sigma\right) \tag{22}$$

where σ is a parameter related to the noise level. The computational complexity of updating $\mathcal{L}^{t+1}$ is $\mathcal{O}(n_1 n_2 n_3)$ [67] when the preset parameters in BM3D are fixed.

3) Update $\mathcal{S}$

$$\mathcal{S}^{t+1} = \operatorname*{argmin}_{\mathcal{S}} \|\mathcal{S}\|_1 + \frac{\beta_4}{2} \left\| \mathcal{Y} - \mathcal{X}^t - \mathcal{S} + \frac{\mathcal{P}_4^t}{\beta_4} \right\|_F^2 \tag{23}$$

Algorithm 3 ADMM-Based Algorithm for Solving FRCTR-PnP

Input: The noisy HSI $\mathcal{Y}$, parameter λ, fibered rank $[r_1, r_2, r_3]$, oversampling parameter $[p_1, p_2, p_3]$, stopping criteria ϵ, and acceleration parameter δ.

Initialization: $t = 0$, let $\mathcal{X}^0, \mathcal{S}^0, \mathcal{L}^0 \mathcal{F}_k^0, (k = 1, 2, 3)$ and Lagrangian multiplies $\mathcal{P}_k^0 (k = 1, 2, \ldots, 5)$ be zeros tensors, parameter $\beta_k (k = 1, 2, \ldots, 5)$, and $\beta_{max} = 10^{10}$;

1: Update $\mathcal{F}_k^{t+1}$ via (19);

2: Update $\mathcal{L}^{t+1}$ via (22);

3: Update $\mathcal{S}^{t+1}$ via (24);

4: Update $\mathcal{X}^{t+1}$ via (27);

5: Update the multiplies via (28);

6: $\beta_k = \min(\delta\beta_k, \beta_{max})$;

7: Check convergence criteria: $\dfrac{\|\mathcal{X}^{t+1} - \mathcal{X}^t\|_F}{\|\mathcal{X}^t\|_F} \leq \epsilon$;

8: If the convergence criteria is not meet, set $t := t + 1$ and go to Step 1.

Output: The restored HSI $\mathcal{X}$.

---

which has the following closed solution:

$$\mathcal{S}^{t+1} = \text{soft}\left(\mathcal{Y} - \mathcal{X}^t + \frac{\mathcal{P}_4^t}{\beta_4}, \frac{1}{\beta_4}\right) \qquad (24)$$

where

$$\text{soft}(\mathcal{X}, \gamma) = \begin{cases} x_{i,j,k} - \gamma, & \text{if } x_{i,j,k} > \gamma \\ x_{i,j,k} + \gamma, & \text{if } x_{i,j,k} < -\gamma \\ 0, & \text{otherwise.} \end{cases} \qquad (25)$$

The computational complexity of updating $\mathcal{S}^{t+1}$ is $\mathcal{O}(n_1 n_2 n_3)$.

4) Update $\mathcal{X}$

$$\mathcal{X}^{t+1} = \sum_{k=1}^{3} \frac{\beta_k}{2} \left\| \mathcal{X} - \mathcal{F}_k^{t+1} + \frac{\mathcal{P}_k^t}{\beta_k} \right\|_F^2$$
$$+ \frac{\beta_4}{2} \left\| \mathcal{Y} - \mathcal{X} - \mathcal{S}^{t+1} + \frac{\mathcal{P}_4^t}{\beta_4} \right\|_F^2$$
$$+ \frac{\beta_5}{2} \left\| \mathcal{X} - \mathcal{L}^{t+1} + \frac{\mathcal{P}_5^t}{\beta_5} \right\|_F^2. \qquad (26)$$

This is a least squares problem whose solution can be exactly calculated as

$$\mathcal{X}^{t+1} = \sum_{k=1}^{3} \beta_k \left(\mathcal{F}_k^{t+1} - \frac{\mathcal{P}_k^t}{\beta_k}\right) \Big/ \sum_{k=1}^{5} \beta_k + \beta_4\left(\mathcal{Y} - \mathcal{S}^t + \frac{\mathcal{P}^4}{\beta_4}\right)$$
$$\Big/ \sum_{k=1}^{5} \beta_k + \beta_5\left(\mathcal{L}^{t+1} - \frac{\mathcal{P}_5^t}{\beta_5}\right) \Big/ \sum_{k=1}^{5} \beta_k. \qquad (27)$$

The computational complexity of updating $\mathcal{X}^{t+1}$ is $\mathcal{O}(n_1 n_2 n_3)$.

5) Update Multipliers: The Lagrangian multipliers are updated as follows:

$$\begin{cases} \mathcal{P}_k^{t+1} = \mathcal{P}_k^t + \beta_k(\mathcal{X}^{t+1} - \mathcal{F}_k^{t+1}), & k = 1, 2, 3 \\ \mathcal{P}_4^{t+1} = \mathcal{P}_4^t + \beta_4(\mathcal{Y} - \mathcal{X}^{t+1} - \mathcal{S}^{t+1}) \\ \mathcal{P}_5^{t+1} = \mathcal{P}_5^t + \beta_5(\mathcal{X}^{t+1} - \mathcal{L}^{t+1}). \end{cases} \qquad (28)$$

The computational complexity of updating multipliers is $\mathcal{O}(n_1 n_2 n_3)$.

Algorithm 3 summarizes the ADMM algorithm to solve the proposed FRCTR-PnP model (10). The computational complexity at each iteration of the proposed algorithm is $\mathcal{O}(n_1 n_2 n_3 \log(n_1 n_2 n_3) + n_1 n_2 n_3 \sum_{i=1}^{3}(r_i + p_i))$. In the FRCTR-PnP solver, the inputs include noisy image $\mathcal{Y}$, the fibered rank $[r_1, r_2, r_3]$, oversampling parameter $[p_1, p_2, p_3]$, and the stopping tolerance ϵ. The penalty parameters $\beta_k (k = 1, \ldots, 5)$ are updated as $\min(\delta\beta_k, \beta_{max}) (k = 1, \ldots, 5)$ in each iteration. This strategy has been widely used in ADMM-based algorithm to accelerate convergence.

IV. EXPERIMENTAL RESULT AND DISCUSSION

We conduct experiments on both simulated and real HSI data sets to demonstrate the effectiveness of the proposed FRCTR-PnP. To adequately evaluate the restoration performance of FRCTR-PnP, we consider the following six HSI restoration methods for comparison: wavelet-based method FORPDN [78], subspace-based method SNLRSF [79], T-SVD-based method SSTV-LRTF [64], Tucker decomposition-based method LRTDTV [59], and mode-k T-SVD-based method 3DTNN and 3DLogTNN [65].

In all experiments, each band of the HSI is normalized into [0, 1], and parameters in compared methods are manually adjusted to the optimal performance.

A. Simulated Experiments

We use a $256 \times 256 \times 191$ subimage of Washington dc Mall (WDC Mall) data set[1] and a $200 \times 200 \times 80$ subimage of Pavia City Center data set[2] in simulated experiments. The mean of peak signal-to-noise ratio (MPSNR) over all bands, the mean of structural similarity (MSSIM) over all bands, and the spectral angle mapping (SAM) [65] are adopted to give a quantitative assessment for restored results of the simulated experiments. We generate noisy data by adding the following five types of noise.

Case 1 (Gaussian Noise): This case includes four subcases. Gaussian noise with zero-mean is added to all bands, and the noise standard deviation in each band is G = 0.1, 0.01, 0.005, and 0.001, respectively.

Case 2 (Gaussian Noise + Salt and Pepper Noise): This case includes four subcases. Gaussian noise with zero mean is added to all bands, and the noise standard deviation is G = 0.1. Furthermore, salt and pepper noise is added to all bands, and the noise proportion in each band is S = 0.25, 0.2, 0.15, and 0.1, respectively.

Case 3 (Gaussian Noise + Salt and Pepper Noise + Stripe Noise): Each band is corrupted by zero-mean Gaussian noise, and the noise standard deviation is randomly generated from [0.1, 0.2]. Then, each band is corrupted by salt and pepper noise with the density randomly generated from [0.1, 0.3]. Furthermore, ten selected bands of WDC Mall and five selected bands of Pavia City Center are further corrupted by 20 stripes and 15 stripes, respectively. Especially, the elements of the

---

TABLE II
QUANTITATIVE COMPARISON OF DIFFERENT METHODS ON THE DATA SET WDC MALL

Case	Gaussian noise(SNR)	Salt and pepper noise	Indicators	Noise	FORPDN	SNLRSF	SSTV-LRTF	LRTDTV	3DTNN	3DLog-TNN	FRCTR-PnP
Case 1	G=0.1 (6.9450 dB)		MPSNR	20.001	34.057	**40.029**	35.505	36.720	35.832	37.532	<u>38.168</u>
			MSSIM	0.4018	0.9302	**0.9812**	0.9499	0.9550	0.9646	0.9688	<u>0.9738</u>
			SAM	30.695	6.8732	**2.5679**	4.8184	4.5197	3.8600	3.5857	<u>3.1812</u>
	G=0.01 (26.945 dB)		MPSNR	40.001	45.685	**56.399**	48.967	52.744	49.369	52.721	<u>53.876</u>
			MSSIM	0.9751	0.9939	**0.9994**	0.9979	0.9989	0.9984	0.9990	<u>0.9992</u>
			SAM	3.6956	1.9990	**0.4862**	0.9974	0.7774	0.8510	0.6549	<u>0.5715</u>
	G=0.005 (32.966 dB)		MPSNR	46.018	49.885	**61.508**	54.250	56.266	52.745	55.908	<u>57.399</u>
			MSSIM	0.9935	0.9976	**0.9998**	0.9992	0.9995	0.9986	0.9994	<u>0.9996</u>
			SAM	1.8529	1.2351	**0.2806**	0.6091	0.4560	0.8380	0.5086	<u>0.3882</u>
	G=0.001 (46.945 dB)		MPSNR	60.001	60.708	**73.876**	63.448	68.677	64.712	67.208	<u>69.934</u>
			MSSIM	0.9997	0.9998	**0.9999**	0.9999	0.9999	0.9999	0.9999	<u>0.9999</u>[1]
			SAM	0.3707	0.3439	**0.0720**	0.2205	0.1371	0.1993	0.1576	<u>0.1104</u>
Case 2	G=0.1	S=0.25	MPSNR	9.9576	20.601	32.661	33.226	34.239	32.264	**35.121**	<u>35.082</u>
			MSSIM	0.0805	0.6838	0.8855	0.9208	0.9243	0.9222	<u>0.9515</u>	**0.9574**
			SAM	50.221	13.501	7.9959	6.1487	7.0389	6.0787	<u>4.4192</u>	**4.1588**
	G=0.1	S=0.2	MPSNR	10.823	22.231	33.351	33.717	34.934	33.195	<u>35.613</u>	**36.021**
			MSSIM	0.1024	0.7309	0.9021	0.9280	0.9353	0.9359	<u>0.9557</u>	**0.9592**
			SAM	49.176	12.342	8.1513	5.8205	6.4623	5.4231	<u>4.1690</u>	**3.8865**
	G=0.1	S=0.15	MPSNR	11.899	24.194	34.351	34.244	35.480	34.045	<u>36.104</u>	**36.788**
			MSSIM	0.1340	0.7794	0.9151	0.9352	0.9418	0.9480	<u>0.9601</u>	**0.9659**
			SAM	47.579	11.081	7.3345	5.5097	5.8730	4.8059	<u>3.9446</u>	**3.6896**
	G=0.1	S=0.1	MPSNR	13.326	26.611	35.215	34.698	36.048	34.692	<u>36.543</u>	**37.287**
			MSSIM	0.1815	0.8269	0.9266	0.9406	0.9472	0.9545	<u>0.9631</u>	**0.9693**
			SAM	45.052	9.8579	6.5287	5.2385	5.4705	4.4116	<u>3.7741</u>	**3.5196**
Case 3			MPSNR	10.161	21.841	30.181	29.527	<u>30.620</u>	26.052	28.086	**32.207**
			MSSIM	0.0873	0.7110	0.8358	0.8404	<u>0.8639</u>	0.7798	0.8204	**0.9121**
			SAM	51.121	14.008	11.624	7.8388	<u>7.4715</u>	12.797	10.823	**7.3728**
Case 4			MPSNR	10.211	22.079	30.379	30.712	<u>31.551</u>	27.430	29.597	**33.374**
			MSSIM	0.0876	0.7141	0.8370	0.8759	0.8823	0.8310	<u>0.8884</u>	**0.9336**
			SAM	51.364	13.446	10.413	7.5820	7.7430	9.5372	<u>7.2102</u>	**5.8634**
Case 5			MPSNR	10.181	21.929	<u>30.371</u>	28.569	29.428	24.099	26.531	**32.258**
			MSSIM	0.0870	0.7115	<u>0.8342</u>	0.8135	0.8214	0.6720	0.7573	**0.9177**
			SAM	51.252	13.705	11.503	11.183	<u>8.3036</u>	13.535	11.699	**7.2065**

[1] The four decimal places of the SSIM values of these methods are the same, so the best and the second-best results are determined based on the value of the fifth decimal place.

whole column will set to a certain value randomly generated from the range of [0.6, 0.8].

Case 4 (Gaussian Noise + Salt and Pepper Noise + Deadlines): We add Gaussian noise and salt and pepper noise as Case 3. Ten selected bands of WDC Mall and five selected bands of Pavia City Center are further corrupted by deadlines and the number of deadlines in each selected band is randomly sampled from the set $\{6, 7, \ldots, 10\}$. Especially, the width of the deadline is randomly sampled from the set $\{1, 2, 3\}$, and the elements of the whole column are set to zero.

Case 5 (Gaussian Noise + Salt and Pepper Noise + Stripe Noise + Deadlines): We consider a challenging case. We add Gaussian noise and salt and pepper noise as Case 3. Besides,

stripe noise and deadlines are added as described in Case 3 and Case 4, respectively.

Tables II and III give the MPSNR, MSSIM, and SAM values obtained by all compared restoration methods on the data sets WDC Mall and Pavia City Center, where the best and the second-best results are highlighted by bold and underline, respectively. We observe that the proposed FRCTR-PnP obtains an overall better performance than the compared methods for mixed noise removal, and the MSSIM values have been greatly improved. We can see that SNLRSF obtains the best-restored results for case 1. The reason is twofold. First, SNLRSF can learn a good subspace under pure Gaussian noise scenarios. Second, SNLRSF uses advanced nonlocal

TABLE III

QUANTITATIVE COMPARISON OF DIFFERENT METHODS ON THE DATA SET PAVIA CITY CENTER

Case	Gaussian noise(SNR)	Salt and pepper noise	Indicators	Noise	FORPDN	SNLRSF	SSTV-LRTF	LRTDTV	3DTNN	3DLog-TNN	FRCTR-PnP
Case 1	G=0.1 (8.3100 dB)		MPSNR	20.006	34.995	**37.624**	34.478	34.269	33.691	35.133	<u>36.416</u>
			MSSIM	0.4211	0.9501	**0.9726**	0.9424	0.9352	0.9506	0.9588	<u>0.9630</u>
			SAM	25.436	3.8987	**2.4868**	3.5297	4.7384	2.9441	2.8292	<u>2.7362</u>
	G=0.01 (28.310 dB)		MPSNR	40.002	49.476	**52.619**	48.787	49.947	45.743	50.113	<u>51.131</u>
			MSSIM	0.9803	0.9980	**0.9989**	0.9977	0.9978	0.9970	0.9986	<u>0.9987</u>
			SAM	2.9248	0.9112	**0.5804**	0.8912	0.8629	0.8890	0.7155	<u>0.7136</u>
	G=0.005 (34.331 dB)		MPSNR	46.022	53.671	**57.312**	53.365	55.453	52.263	54.859	<u>56.147</u>
			MSSIM	0.9949	0.9992	**0.9996**	0.9991	0.9994	0.9988	0.9994	<u>0.9995</u>
			SAM	1.4647	0.5787	**0.3649**	0.5564	0.4298	0.6849	0.4629	<u>0.4211</u>
	G=0.001 (48.310 dB)		MPSNR	60.001	63.241	**70.038**	65.206	67.748	63.011	66.554	<u>68.595</u>
			MSSIM	0.9998	0.9999	**0.9999**	0.9999	0.9999	0.9999	0.9999	<u>0.9999</u>[1]
			SAM	0.2931	0.2012	**0.0881**	0.1569	0.1187	0.2056	0.1312	<u>0.1043</u>
Case 2	G=0.1	S=0.25	MPSNR	10.338	21.846	29.661	32.312	32.318	31.390	<u>33.185</u>	**34.127**
			MSSIM	0.0816	0.7043	0.8450	0.9106	0.9032	0.9172	<u>0.9371</u>	**0.9434**
			SAM	44.604	6.2974	13.421	4.3375	5.7230	3.7873	<u>3.3146</u>	**3.3068**
	G=0.1	S=0.2	MPSNR	11.182	23.403	30.493	32.848	33.001	31.989	<u>33.655</u>	**34.771**
			MSSIM	0.1042	0.7539	0.8476	0.9199	0.9121	0.9275	<u>0.9430</u>	**0.9493**
			SAM	43.251	5.8296	10.494	3.8684	5.8269	3.5190	<u>3.1991</u>	**3.1938**
	G=0.1	S=0.15	MPSNR	12.258	25.353	31.582	33.319	33.450	32.397	<u>34.009</u>	**35.182**
			MSSIM	0.1374	0.8047	0.8922	0.9274	0.9213	0.9347	<u>0.9478</u>	**0.9537**
			SAM	41.345	5.3633	7.4462	3.9167	5.4246	3.3526	<u>3.0917</u>	**3.0847**
	G=0.1	S=0.1	MPSNR	13.673	27.713	32.799	33.774	34.072	32.863	<u>34.399</u>	**35.663**
			MSSIM	0.1870	0.8513	0.9093	0.9335	0.9281	0.9407	<u>0.9517</u>	**0.9584**
			SAM	38.443	4.8925	6.6433	3.7744	5.0549	3.1965	<u>3.0135</u>	**2.9589**
Case 3			MPSNR	10.518	23.000	27.539	29.698	<u>30.091</u>	26.793	28.545	**31.314**
			MSSIM	0.0890	0.7316	0.7986	0.8512	0.8503	0.7977	<u>0.8533</u>	**0.9048**
			SAM	45.743	8.0550	10.859	**4.3844**	6.0703	5.7330	6.0844	<u>5.6842</u>
Case 4			MPSNR	10.561	23.289	28.034	29.545	<u>30.259</u>	25.960	27.977	**31.514**
			MSSIM	0.0901	0.7391	0.8006	0.8474	<u>0.8538</u>	0.7552	0.8369	**0.9147**
			SAM	45.974	7.5838	14.441	<u>4.8701</u>	7.2458	5.8701	5.6180	**4.7961**
Case 5			MPSNR	10.523	22.922	28.004	29.239	<u>29.807</u>	25.214	26.919	**31.233**
			MSSIM	0.0896	0.7310	0.8021	<u>0.8679</u>	0.8485	0.6955	0.7974	**0.9099**
			SAM	45.773	7.9605	16.274	<u>5.6743</u>	10.417	6.8965	6.0100	**5.1164**

tensor denoiser while our method uses the basic BM3D denoiser. Matrix-based method FORPDN and subspace-based method SNLRSF perform worse than tensor-based methods SSTV-LRTF and LRTDTV in most cases. Besides, 3DTNN and 3DLogTNN only consider the low-fibered-rankness, and the stripes and deadlines are more likely to be regarded as the low fibered rank part, so 3DTNN and 3DLogTNN perform not well in cases 3–5. To compare the effects of all the restoration methods, we show the peak signal-to-noise ratio (PSNR) values of each band under different cases in Fig. 3. It can be seen that the proposed method achieves the highest PSNR values in most cases and bands.

Figs. 4 and 5 display visual performance of restored results by different methods on the data sets WDC Mall and Pavia City Center, respectively. We mark the same subregion of each subfigure and enlarge it in a red box. All the compared methods remove such mixed noise to some extent. For FORPDN and SNLRSF, they unfold HSIs into the matrices, which destroy the intrinsic structure of HSIs and leads to some bands of the restored HSIs distortion. SSTV-LRTF and LRTDTV take use of the whole structure of tensors and SSTV regularization. However, they fail to well recover the local details. 3DTNN and 3DLogTNN take advantage of the structure information in three directions but fail to remove the stripes and deadlines. Comparatively, FRCTR-PnP outperforms the compared methods, efficiently removing the mixed noise and preserving the essential structures and local details of the clean HSIs.

To further compare the performance of spectral curve recovery, a subcase (the standard deviation of zero-mean Gaussian noise is G = 0.1, the noise proportion of salt and pepper noise is S = 0.2) of case 2 and case 5 are selected as two repre-
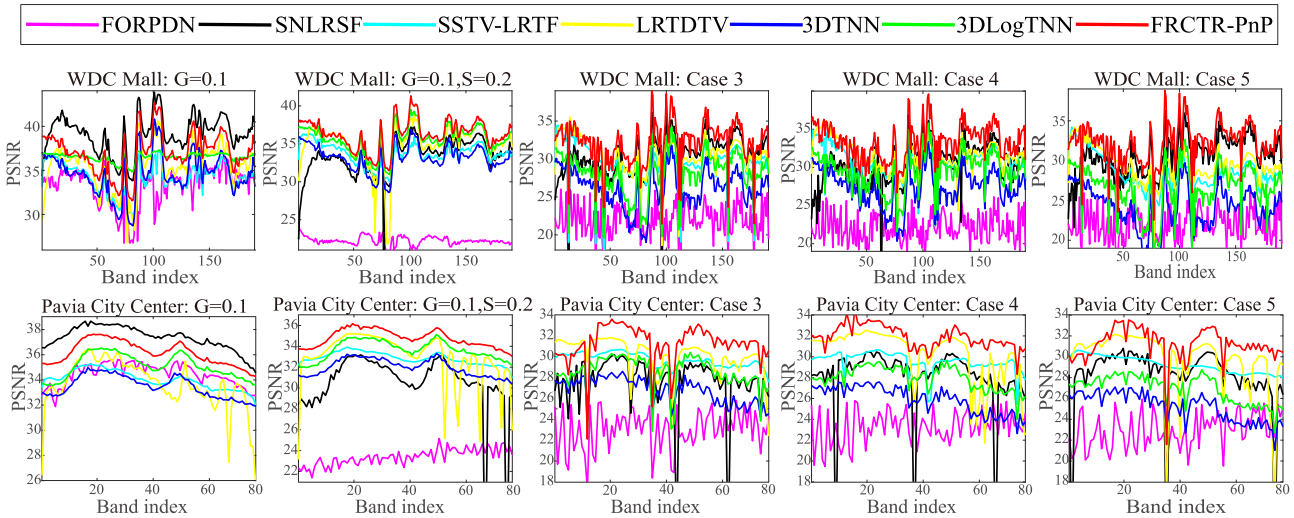
Fig. 3. PSNR values of restored results by different methods on WDC Mall and Pavia City Center.
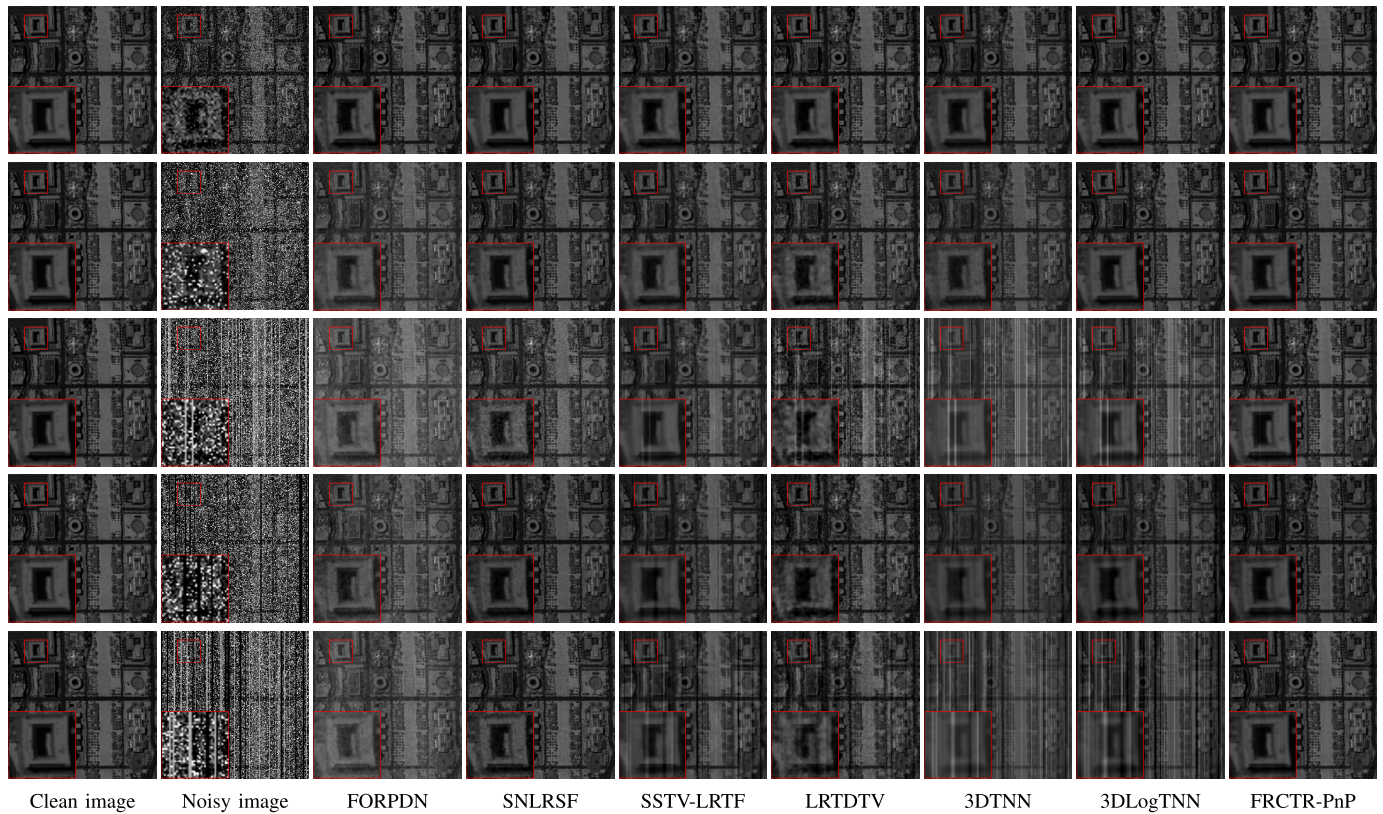
Fig. 4. Restored results of band 96 on WDC Mall by different methods. From top to bottom: the results under a subcase (the standard deviation of zero-mean Gaussian noise is $G = 0.1$) of case 1, a subcase (the standard deviation of zero-mean Gaussian noise is $G = 0.1$, the noise proportion of salt and pepper noise is $S = 0.2$) of case 2, and cases 3–5, respectively.

sentative cases. Fig. 6 shows the spectral curves at one spatial location of the restored results by different compared methods. We clearly observe that the spectral curves obtained by the proposed FRCTR-PnP better approximate the original ones than those obtained by the compared methods. In summary, the above observations illustrate that the proposed FRCTR-PnP achieves the best performance on mixed noise removal, fine details preservation, and spectral signatures restoration, among all compared methods.

The reason why the proposed FRCTR-PnP performs well is that it uses the fibered rank to capture the global structure information and PnP regularization to preserve the image details. On one hand, in the noisy HSIs, the leading tensor singular values mainly correspond to the overall structure and image subspace of the HSIs, and the small tensor singular values mainly correspond to fine details and noise subspace. So we directly eliminate the small singular values and noise subspace through the rank constraint, which helps to restore
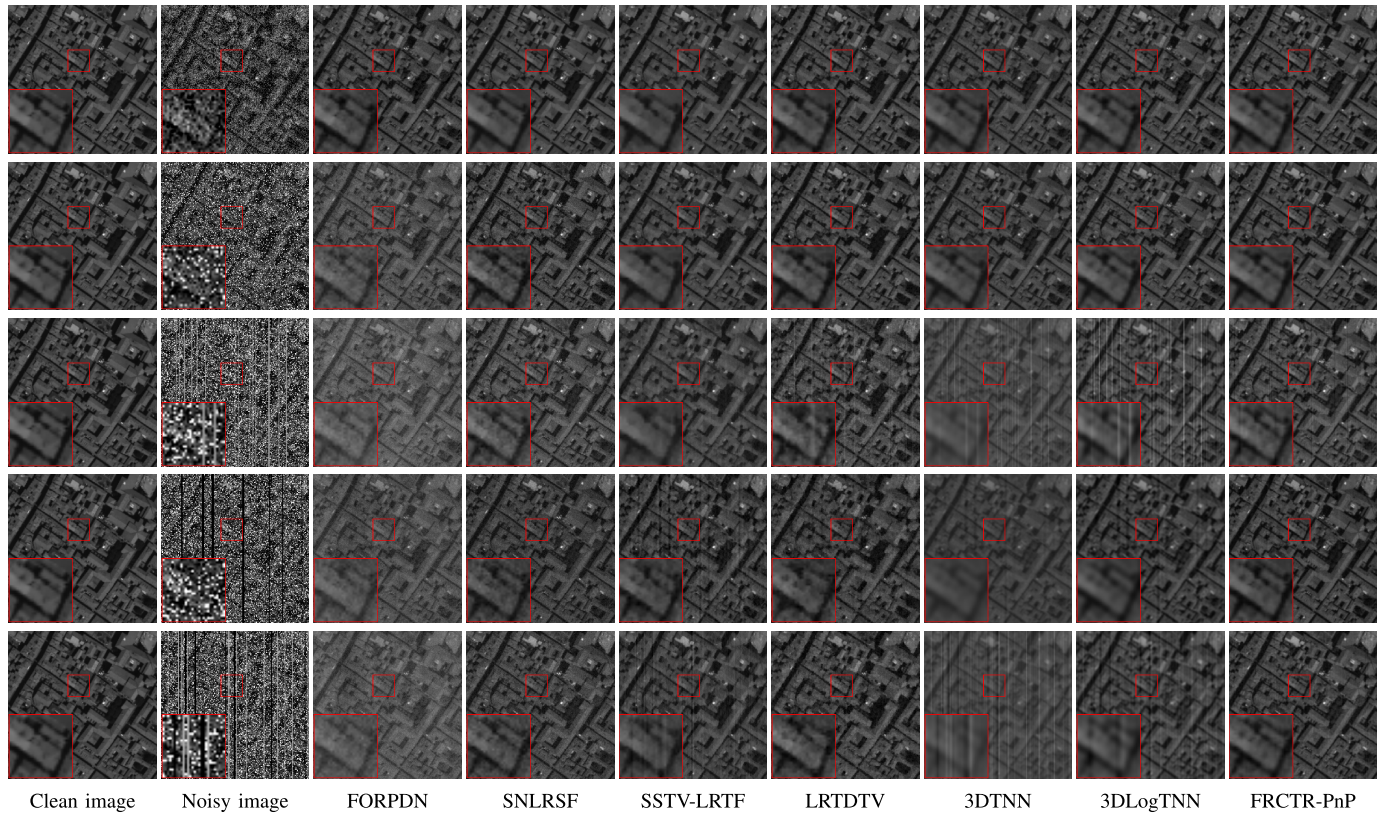
Fig. 5.   Restored results of band 45 on Pavia City Center by different methods. From top to bottom: the results under a subcase (the standard deviation of zero-mean Gaussian noise is G = 0.1) of case 1, a subcase (the standard deviation of zero-mean Gaussian noise is G = 0.1, the noise proportion of salt and pepper noise is S = 0.2) of case 2, and cases 3–5, respectively.
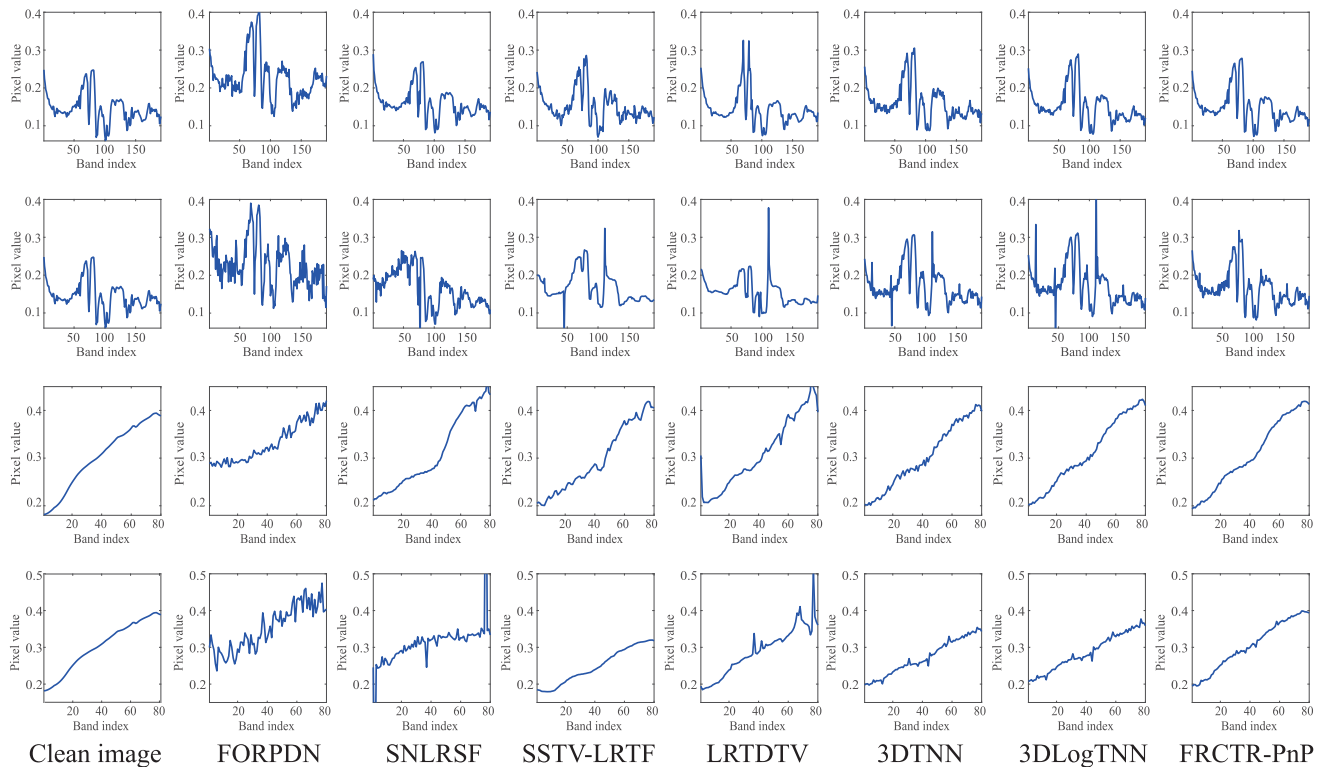

Fig. 6.   Spectral curves of the restored results by different compared methods. The first two rows are the results at spatial location (200, 200) of the data set WDC Mall under a subcase (the standard deviation of zero-mean Gaussian noise is G = 0.1, the noise proportion of salt and pepper noise is S = 0.2) of case 2 and case 5, respectively. The last two rows are the results at spatial location (42, 20) of the Pavia City Center under a subcase (the standard deviation of zero-mean Gaussian noise is G = 0.1, the noise proportion of salt and pepper noise is S = 0.2) of case 2 and case 5, respectively.
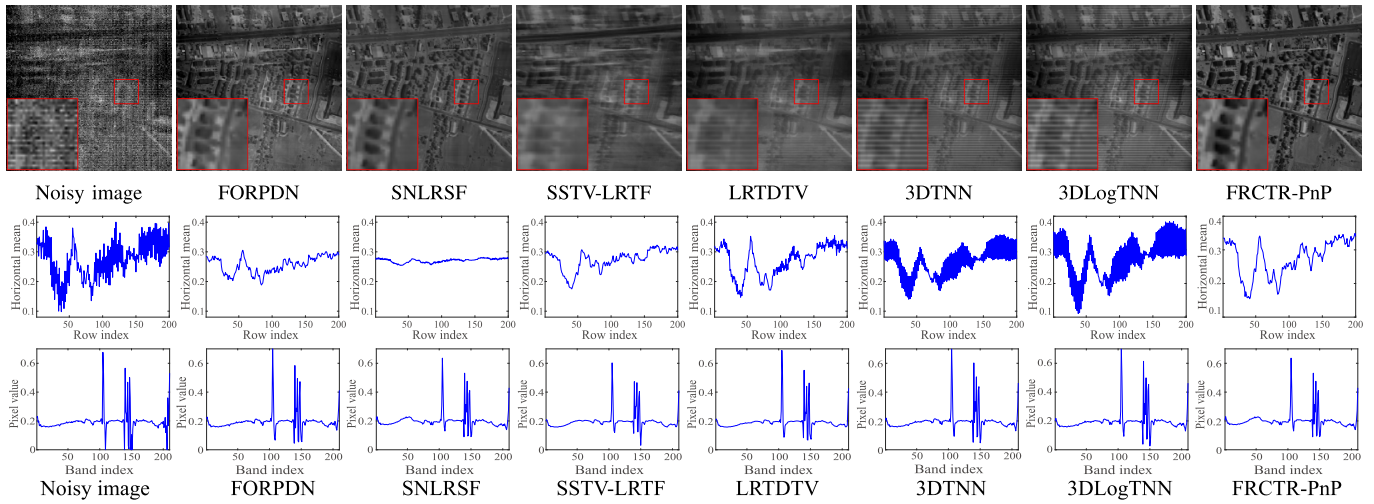
Fig. 7. Visual results (first row) and the horizontal mean profiles (second row) of band 108, and the spectral curves at spatial location (44,44) (third row) by different methods on real HSI Urban.

Fig. 8. Visual results (first row) and the vertical mean profiles (second row) of band 96, and the spectral curves at spatial location (66,66) (third row) by different methods on real HSI EO-1.

the global structure and remove the noise. On the other hand, the elimination of small singular values and singular vectors can also lose a lot of details, which can be restored by the plugged BM3D regularization.

B. Real Experiments

We consider two real-world HSI data sets in our experiments: the hyperspectral digital imagery collection experiment (HYDICE) Urban data set[3] and Earth Observing-1 (EO-1) Hyperion data set.[4]

1) HYDICE Urban Data Set: The original image is of size $307 \times 307 \times 210$, and we use a $200 \times 200 \times 210$ subimage of the data set for our experiment. The full urban image is corrupted by stripes, deadlines, water absorption, and other unknown noise.

[3]http://www.erdc.usace.army.mil/Media/Fact-Sheets/Fact-Sheet-Article-View/Article/610433/hypercube/

[4]http://www.lmars.whu.edu.cn/prof_web/zhanghongyan/resource/noise_EOI.zip

Fig. 7 shows the visual results (first row) and the horizontal mean profiles (second row) of band 108, and the spectral curve (third row) by different methods. We observe from Fig. 7 that 3DTNN and 3DLogTNN cannot effectively remove the stripes. This is mainly because bands 104–109 and bands 199–210 have a large number of stripes and deadlines in the same position, which have low-rank structures and thus are assumed to be the clean part. As a result, the horizontal mean profiles of band 108 rapidly fluctuates in the noisy image and the restoration images obtained by 3DTNN and 3DLogTNN. SSTV-LRTF and LRTDTV can remove lots of noise, but the restored images are smooth and lose a lot of details. FORPDN and SNLRSF remove almost all the noise, but cannot reconstruct the edges of the image well. Comparatively, the proposed FRCTR-PnP obtains the best visual result, especially in noise removal and restoration of global structure and fine details.

2) EO-1 Hyperion Data set [51]: The size of the original image is $400 \times 1000 \times 242$, and we use a $200 \times 200 \times 166$ subimage of the data set for our experiment. The EO-1 image

PSNR:15.436

(a)   (b)

PSNR:17.921   PSNR:18.474

(c)   (d)

(e)

Fig. 9. Comparison of two restored results. (a) Clean image. (b) Noisy image. (c) Restored by truncated SVD. (d) Restored by randomized SVD. (e) Residual vectors of restored images and clean image.

is mainly corrupted by stripes, water atmosphere, and other complex noise.

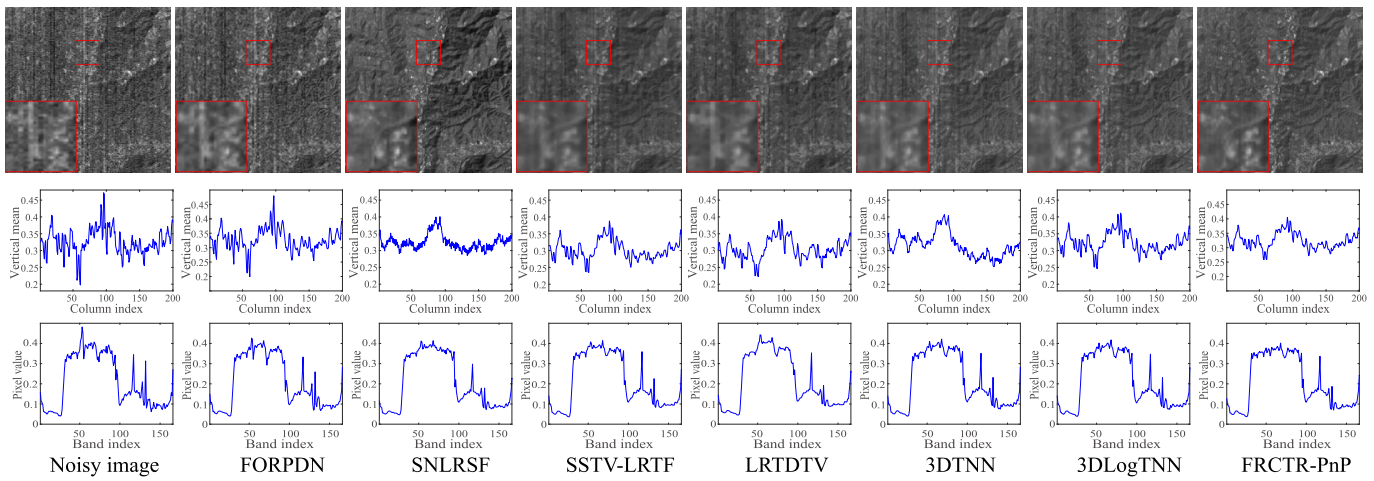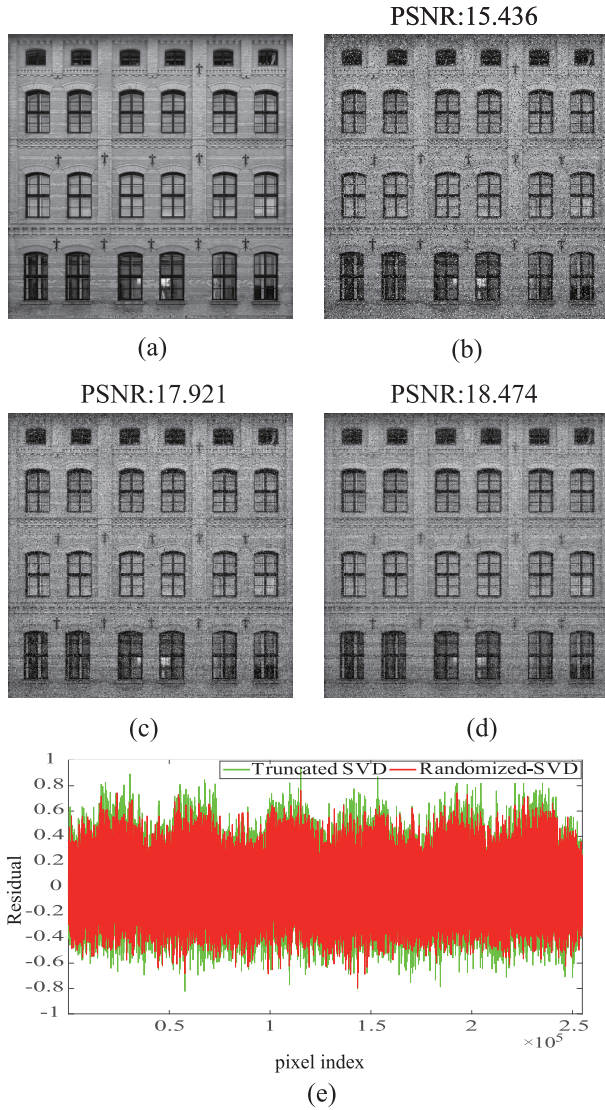Fig. 8 shows the visual results (first row) and the vertical mean profiles (second row) of band 96, and the spectral curves at spatial location (66,66) (third row) by different methods. There are many stripes and deadlines in bands 1–6, bands 94–97, bands 129–139, and bands 160–166. It is clear that the values of the stripes in band 96 are not very different from the surrounding values, and the vertical mean profile fluctuates, but not very much. Here, we can clearly observe that there are still some stripes and distortion in the magnified areas restored by FORPDN, SSTV-LRTF, LRTDTV, and 3DTNN. Compared to SNLRSF, the proposed FRCTR-PnP can restore more details. The spectral curve obtained by the proposed FRCTR-PnP is smoother than the compared methods.

C. Discussion

1) Contribution of Two Building Blocks: The building blocks of the proposed method are 3DRT-SVD, and BM3D.
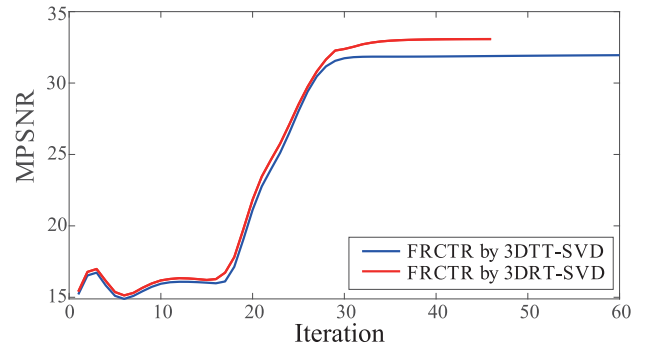
Fig. 10. MPSNR values with respect to the iteration number.
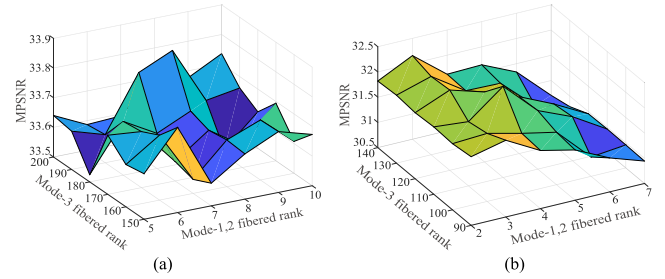
Fig. 11. Influence of the fibered rank. (a) Values of MPSNR for WDC Mall. (b) Values of MPSNR for Pavia City Center.

The building blocks are complementary to each other, rather than isolated and uncorrelated. We discuss the contributions of the two building blocks, based on the simulated data set WDC Mall. Each band of WDC Mall data set is corrupted by zero-mean Gaussian noise with the noise standard deviation is randomly generated from $[0.1, 0.2]$ and salt and pepper noise with the density randomly generated from $[0.1, 0.3]$. First, we introduce the following FRCTR model:

$$\min_{\mathcal{X}, \mathcal{S}} \quad \|\mathcal{S}\|_1$$
$$\text{s.t.} \quad \|\mathcal{Y} - \mathcal{X} - \mathcal{S}\|_F^2 \leq \epsilon, \quad \text{rank}_f(\mathcal{X}) \leq [r_1, r_2, r_3]. \quad (29)$$

In Table IV, we show the indicators of restored results obtained by five models BM3D, 3DTNN, 3DLogTNN, FRCTR, and FRCTR-PnP. Based on this empirical study, we give some discussion.

a) Contribution of 3DRT-SVD: In Table IV, the restored results of the 3DRT-SVD method are better than the 3DTT-SVD method. We explain this phenomenon by the following two points.

First, to intuitively compare the two methods, we perform matrix versions truncated SVD and randomized SVD [80] of them on the noisy image: the Gaussian noise is of variance 0.02 and the percentage of impulse noise is 0.1. An example is shown in Fig. 9. By comparing (b) and (c), we can see that the truncated SVD can remove partial noise by dropping small singular values and the corresponding singular vectors. From the results of (d) and (e), we can see that the restoration effect of randomized SVD is better than that of truncated SVD, and the residual vector of randomized SVD is smaller than that of truncated SVD. The above phenomena indicate that randomized SVD preserves the intrinsic information of

Indicators	Noise	BM3D	3DTNN	3DLogTNN	FRCTR		FRCTR-PnP	
					3DTT-SVD	3DRT-SVD	3DTT-SVD	3DRT-SVD
MPSNR	10.194	20.855	29.672	32.885	31.997	**33.034**	32.449	**33.597**
MSSIM	0.0876	0.4217	0.8816	0.9282	0.8793	**0.9229**	0.8879	**0.9338**
time (s)	–	20.928	274.14	398.48	449.73	254.38	915.65	975.56

Fig. 12. Influence of the tuning parameter on the data sets WDC Mall and Pavia City Center under cases 1 and 2.
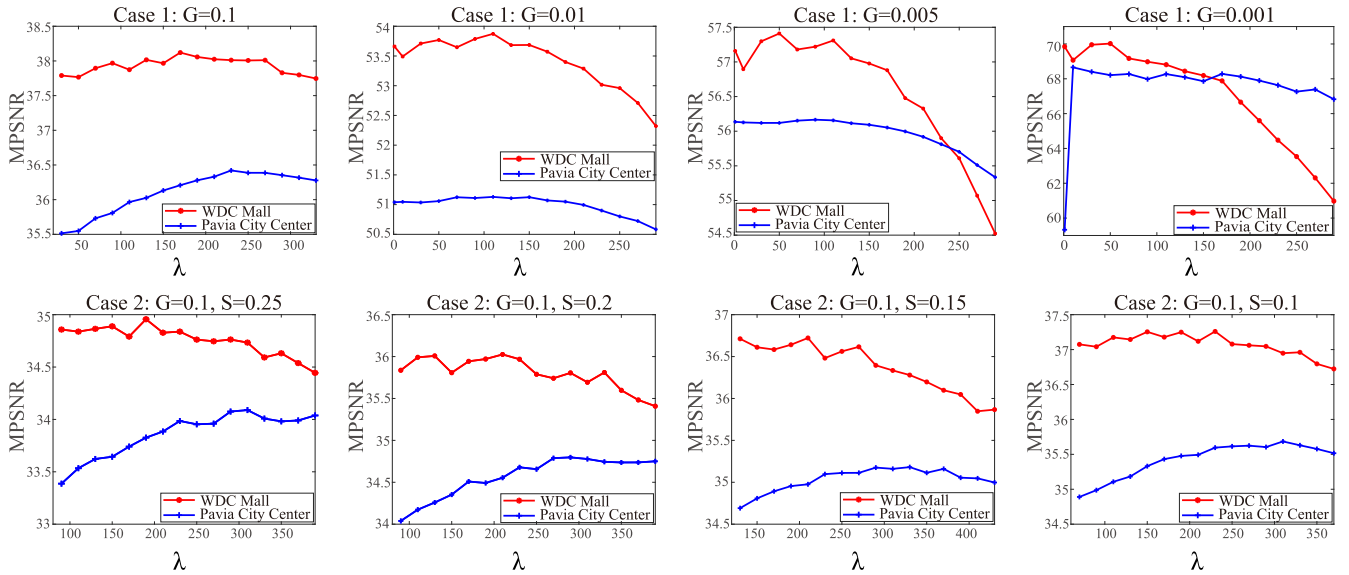
the image and removes part of noise compared with truncated SVD.

Second, we show the change of MPSNR values in the FRCTR model with respect to the iteration number in Fig. 10. We can see from the figure that the MPSNR values of 3DRT-SVD are higher than those of 3DTT-SVD in all iteration numbers; with each step of accumulation, the final MPSNR value of 3DRT-SVD is higher than that of the 3DTT-SVD; in the FRCTR model, using 3DRT-SVD needs fewer iterations than 3DTT-SVD to reach the convergence condition.

b) Contribution of BM3D: From Table IV, we can observe that compared with BM3D denoising alone, our model has obvious advantages in mixed noise removal. Compared with the restored results of FRCTR and FRCTR-PnP, we can see that FRCTR-PnP shows significant improvement in MPSNR and MSSIM. This is because when we integrate BM3D into the PnP framework, BM3D and low-rankness constraint are complementary to each other, rather than isolated and uncorrelated, which lead to better-restored results.

2) Parameter Analysis: We discuss the influence of parameters, based on the simulated data sets under cases 1–5. Specifically, in the FRCTR-PnP method, these parameters can be roughly divided into four categories.

a) Influence of the fibered rank: The low-rankness prior with fibered rank $[r_1, r_2, r_3]$ captures the global structure of HSIs. Before running the algorithm, we should give an estimate of the fibered rank along the three modes.

The mode-1 and mode-2 fibered rank related to the spectral directions should be small, since the spectral bands are highly correlated. In case 1 and case 2, we assume that mode-1 fibered rank and mode-2 fibered rank are equal. Fig. 11 shows the sensitivity analysis of the fibered rank. It can be easily observed that the MPSNR values first increase and then decrease with the growth of the estimated fibered rank. This is because when the estimated fibered rank is small, global information cannot be explored well. When the estimated fibered rank is large, some singular values and singular vectors corresponding to noise subspace are also included. Besides, considering that the larger fibered rank leads to high computational complexity, we set the decided fibered rank [8, 8, 180] and [4, 4, 110] for WDC Mall and Pavia City Center, respectively. As for case 3, case 4, and case 5, it is changed to restore the noisy HSIs since there are many stripes and deadlines, so we can reduce the fibered rank properly, then BM3D further preserves local details.

b) Influence of the oversampling parameters: In [81], Theorem 4 shows that the larger the oversampling parameters $[p_1, p_2, p_3]$, the better the approximation effect. We finally choose [30, 30, 50] for all the experiments for balancing the running time and accuracy,

c) Influence of the regularization parameter λ: The tuning parameter actually regularizes the parameter σ in the denoiser. Fig. 12 shows the influence of the tuning parameter on the data sets WDC Mall and Pavia City Center under

TABLE V
ANALYSIS OF FIVE TERMS IN (27)

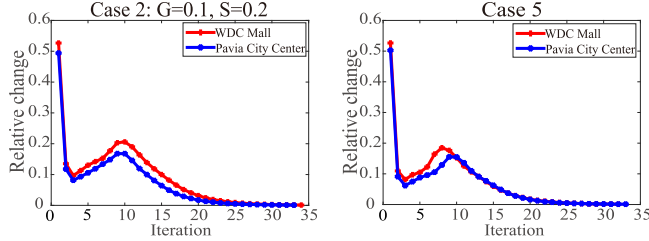| Parameters	Variables	Representation.
β_1	$\mathcal{F}_1 - \mathcal{P}_1/\beta_1$	Mode-1 fibered rank approximation
β_2	$\mathcal{F}_2 - \mathcal{P}_2/\beta_2$	Mode-2 fibered rank approximation
β_3	$\mathcal{F}_3 - \mathcal{P}_3/\beta_3$	Mode-3 fibered rank approximation
β_4	$\mathcal{Y} - \mathcal{S} + \mathcal{P}_4/\beta_4$	Removing sparse noise
β_5	$\mathcal{L} - \mathcal{P}_5/\beta_5$	Clean image obtained by BM3D denoiser

Fig. 13. History of relative change curves for a subcase (the standard deviation of zero mean Gaussian noise is G = 0.1, the noise proportion of salt and pepper noise is S = 0.2) of case 2 and case 5 on the data sets WDC Mall and Pavia City Center.

cases 1 and 2. We can clearly observe that the optimal value of λ is increasing as the Gaussian noise standard deviation increases. When the parameter λ is too small or large, we cannot obtain satisfying restored results. It achieves the optimal performance with a proper λ, which well balances noise removal and details preservation.

d) Influence of the penalty parameters: β_k $(k = 1, \ldots, 5)$ are introduced in the augmented Lagrangian function. In (27), they are used to balance the proportion between the five terms, as shown in Table V. In [65], $(1, 1, 0.001)/2.001$ is used as the weighted coefficient of 3DTNN. Therefore, considering the meaning of the five terms, β_k $(k = 1, 2, 4, 5)$ are fixed at 0.1, and β_3 is selected from the set $\{10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}\}$.

3) Convergence Analysis: Since the proposed model is a nonconvex optimization problem with an implicit regularization, its theoretical convergence is still an open problem. Thus, we discuss the numerical convergence behavior. Fig. 13 displays the relative change curves of the proposed method for a subcase (the standard deviation of zero mean Gaussian noise is G = 0.1, the noise proportion of salt and pepper noise is S = 0.2) of case 2 and case 5 on the data sets WDC Mall and Pavia City Center. We clearly observe that although there have fluctuations in the middle of the convergence curves, the overall trend is decreasing steadily, which illustrates the numerical convergence of the proposed algorithm. Therefore, it can be applied to practical situations.

V. CONCLUSION

In this article, we propose a tensor-based HSI restoration method FRCTR-PnP. Specifically, the low fibered rank is utilized to characterize the global spatial–spectral correlation among all HSI bands, and a PnP-based regularization is introduced to further exploit the NSS of HSI. Then, we develop the ADMM algorithm to tackle the proposed model. For robustness, we suggest the 3DRT-SVD to solve the subproblem

of low-fibered rank approximation, which preserves the intrinsic structure of the clean HSI and removes partial noise by projecting the HSI onto a low-dimensional essential subspace. The simulated and real experiments and discussions demonstrate that the proposed method achieves superior performance over compared methods quantitatively and qualitatively. This is because the FRCTR can help to restore the global information of the target HSI, while the embedded BM3D is beneficial to preserve the image details and remove the low-rank structure noise.

APPENDIX

In this appendix, we give detailed proof for Lemma 1.
Proof:

1) *For the Case of Matrix:* The singular value decomposition (SVD) of matrix F can be described as

$$F = USV^H$$

where U and V are orthogonal matrices, and $S = \text{diag}(\sigma_1, \ldots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \sigma_n \geq 0$. Writing U and V in terms of their columns

$$U = (u_1, \ldots, u_n) \quad \text{and} \quad V = (v_1, \ldots, v_n)$$

then u_j and v_j are the left and right singular vectors corresponding to σ_j. Then the rank-k truncated SVD of F is

$$\begin{aligned}
\hat{F} &= (u_1, \ldots, u_k)\text{diag}(\sigma_1, \ldots, \sigma_k)(v_1, \ldots, v_k)^H \\
&= U_k S_k V_k^H \\
&= U_k U_k^H [U_k \ U_s] \begin{bmatrix} S_k \\ S_s \end{bmatrix} \begin{bmatrix} V_k^H \\ V_s^H \end{bmatrix} = U_k U_k^H F \quad (30)
\end{aligned}$$

since $U_k^H U_k$ is the identity matrix, and $U_k^H U_s$ is the zero matrix.

2) *For the Case of Tensor:* The tensor SVD (T-SVD) of tensor $\hat{\mathcal{F}}$ is

$$\hat{\mathcal{F}} = \mathcal{U} * \mathcal{S} * \mathcal{V}^H.$$

Then the truncated T-SVD of $\hat{\mathcal{F}}$ is

$$\mathcal{F} = \mathcal{U}^k * \mathcal{S}^k * (\mathcal{V}^k)^H = \mathcal{U}^k * (\mathcal{U}^k)^H * \hat{\mathcal{F}}.$$

According to the truncated T-SVD algorithm, we first transform tensor $\hat{\mathcal{F}}$ into the Fourier domain. Next, we compute rank-k truncated SVD of each band. Here, we can use (30) of matrix truncated SVD. Last, we take the inverse Fourier transform and transform the obtained tensors back to the original domain. □

REFERENCES

[1] R. Archibald and G. Fann, "Feature selection and classification of hyperspectral images with support vector machines," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 4, pp. 674–677, Oct. 2007.

[2] H.-C. Li, W.-Y. Wang, L. Pan, W. Li, Q. Du, and R. Tao, "Robust capsule network based on maximum correntropy criterion for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 738–751, 2020.

[3] D. Hong *et al.*, "More diverse means better: Multimodal deep learning meets remote-sensing imagery classification," *IEEE Trans. Geosci. Remote Sens.*, early access, Aug. 24, 2020, doi: 10.1109/TGRS.2020.3016820.

[4] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, and J. Chanussot, "Graph convolutional networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, early access, Aug. 18, 2020, doi: 10.1109/TGRS.2020.3015157.

[5] J. Ma, C. Li, Y. Ma, and Z. Wang, "Hyperspectral image denoising based on low-rank representation and superpixel segmentation," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 3086–3090.

[6] J. M. Bioucas-Dias *et al.*, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.

[7] X. Fu, W.-K. Ma, T.-H. Chan, and J. M. Bioucas-Dias, "Self-dictionary sparse regression for hyperspectral unmixing: Greedy pursuit and pure pixel search are related," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 6, pp. 1128–1141, Sep. 2015.

[8] X. Fu, W.-K. Ma, J. M. Bioucas-Dias, and T.-H. Chan, "Semiblind hyperspectral unmixing in the presence of spectral library mismatches," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 9, pp. 5171–5184, Sep. 2016.

[9] W. He, H. Zhang, and L. Zhang, "Total variation regularized reweighted sparse nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 3909–3921, Jul. 2017.

[10] D. Hong, N. Yokoya, J. Chanussot, and X. X. Zhu, "An augmented linear mixing model to address spectral variability for hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1923–1938, Apr. 2019.

[11] J. Yao, D. Meng, Q. Zhao, W. Cao, and Z. Xu, "Nonconvex-sparsity and nonlocal-smoothness-based blind hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 6, pp. 2991–3006, Jun. 2019.

[12] B. Rasti, P. Scheunders, P. Ghamisi, G. Licciardi, and J. Chanussot, "Noise reduction in hyperspectral imagery: Overview and application," *Remote Sens.*, vol. 10, pp. 482–509, Mar. 2018.

[13] J.-H. Yang, X.-L. Zhao, J.-J. Mei, S. Wang, T.-H. Ma, and T.-Z. Huang, "Total variation and high-order total variation adaptive model for restoring blurred images with cauchy noise," *Comput. Math. Appl.*, vol. 77, no. 5, pp. 1255–1272, Mar. 2019.

[14] Y. Qian and M. Ye, "Hyperspectral imagery restoration using nonlocal spectral-spatial structured sparse representation with noise estimation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 499–515, Apr. 2013.

[15] T.-X. Jiang, M. K. Ng, X.-L. Zhao, and T.-Z. Huang, "Framelet representation of tensor nuclear norm for third-order tensor completion," *IEEE Trans. Image Process.*, vol. 29, pp. 7233–7244, 2020.

[16] Q. Xie, Q. Zhao, D. Meng, and Z. Xu, "Kronecker-Basis-Representation based tensor sparsity and its applications to tensor recovery," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 8, pp. 1888–1902, Aug. 2018.

[17] W. He, N. Yokoya, L. Yuan, and Q. Zhao, "Remote sensing image reconstruction using tensor ring completion and total variation," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 8998–9009, Nov. 2019.

[18] Y.-T. Wang, X.-L. Zhao, T.-X. Jiang, L.-J. Deng, Y. Chang, and T.-Z. Huang, "Rain streaks removal for single image via kernel-guided convolutional neural network," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Aug. 21, 2020, doi: 10.1109/TNNLS.2020.3015897.

[19] Y. Chen, W. He, N. Yokoya, and T.-Z. Huang, "Blind cloud and cloud shadow removal of multitemporal images based on total variation regularized low-rank sparsity decomposition," *ISPRS J. Photogramm. Remote Sens.*, vol. 157, pp. 93–107, Nov. 2019.

[20] J.-H. Yang, X.-L. Zhao, T.-H. Ma, Y. Chen, T.-Z. Huang, and M. Ding, "Remote sensing images destriping using unidirectional hybrid total variation and nonconvex low-rank regularization," *J. Comput. Appl. Math.*, vol. 363, pp. 124–144, Jan. 2020.

[21] L.-B. Cui, X.-Q. Zhang, and S.-L. Wu, "A new preconditioner of the tensor splitting iterative method for solving multi-linear systems with $\mathcal{M}$-tensors," *Comput. Appl. Math.*, vol. 39, no. 3, p. 173, Sep. 2020, doi: 10.1007/s40314-020-01194-8.

[22] Y. Chang, L. Yan, T. Wu, and S. Zhong, "Remote sensing image stripe noise removal: From image decomposition perspective," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7018–7031, Dec. 2016.

[23] J.-L. Wang, T.-Z. Huang, X.-L. Zhao, J. Huang, T.-H. Ma, and Y.-B. Zheng, "Reweighted block sparsity regularization for remote sensing images destriping," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 12, pp. 4951–4963, Dec. 2019.

[24] J. Li, Q. Yuan, H. Shen, and L. Zhang, "Noise removal from hyperspectral image with joint Spectral–Spatial distributed sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 9, pp. 5425–5439, Sep. 2016.

[25] L.-B. Cui, M.-H. Li, and Y. Song, "Preconditioned tensor splitting iterations method for solving multi-linear systems," *Appl. Math. Lett.*, vol. 96, pp. 89–94, Oct. 2019.

[26] T.-H. Ma, Y. Lou, and T.-Z. Huang, "Truncated l_{1-2} models for sparse recovery and rank minimization," *SIAM J. Imag. Sci.*, vol. 10, no. 3, pp. 1346–1380, Jan. 2017.

[27] R. Dian and S. Li, "Hyperspectral image super-resolution via subspace-based low tensor multi-rank regularization," *IEEE Trans. Image Process.*, vol. 28, no. 10, pp. 5135–5146, Oct. 2019.

[28] Y.-B. Zheng, T.-Z. Huang, X.-L. Zhao, Y. Chen, and W. He, "Double-factor-regularized low-rank tensor factorization for mixed noise removal in hyperspectral image," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 12, pp. 8450–8464, Dec. 2020.

[29] X.-L. Zhao, W. Wang, T. Zeng, T.-Z. Huang, and M. K. Ng, "Total variation structured total least squares method for image restoration," *SIAM J. Sci. Comput.*, vol. 35, no. 6, pp. 1304–1320, 2013.

[30] J. Peng, Q. Xie, Q. Zhao, Y. Wang, L. Yee, and D. Meng, "Enhanced 3DTV regularization and its applications on HSI denoising and compressed sensing," *IEEE Trans. Image Process.*, vol. 29, pp. 7889–7903, 2020.

[31] M. Ding, T.-Z. Huang, T.-Y. Ji, X.-L. Zhao, and J.-H. Yang, "Low-rank tensor completion using matrix factorization based on tensor train rank and total variation," *J. Sci. Comput.*, vol. 81, no. 2, pp. 941–964, Nov. 2019.

[32] J.-L. Wang, T.-Z. Huang, T.-H. Ma, X.-L. Zhao, and Y. Chen, "A sheared low-rank model for oblique stripe removal," *Appl. Math. Comput.*, vol. 360, pp. 167–180, Nov. 2019.

[33] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.

[34] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2005, pp. 60–65.

[35] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2014, pp. 2862–2869.

[36] M. Maggioni, G. Boracchi, A. Foi, and K. Egiazarian, "Video denoising, deblocking, and enhancement through separable 4-D nonlocal spatiotemporal transforms," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 3952–3966, Sep. 2012.

[37] Y. Peng, D. Meng, Z. Xu, C. Gao, Y. Yang, and B. Zhang, "Decomposable nonlocal tensor dictionary learning for multispectral image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2014, pp. 2949–2956.

[38] 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.

[39] 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.

[40] H. Kumar Aggarwal and A. Majumdar, "Hyperspectral image denoising using spatio-spectral total variation," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 3, pp. 442–446, Mar. 2016.

[41] L. D. Lathauwer, B. D. Moor, and J. Vandewalle, "On the best rank-1 and rank-$(R_1, R_2, \cdots, R_N)$ approximation of higher-order tensors," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1324–1342, 2000.

[42] N. Renard, S. Bourennane, and J. Blanc-Talon, "Denoising and dimensionality reduction using multilinear tools for hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 138–142, Apr. 2008.

[43] A. Karami, M. Yazdi, and A. Zolghadre Asli, "Noise reduction of hyperspectral images using Kernel non-negative tucker decomposition," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 487–493, Jun. 2011.

[44] D. Letexier and S. Bourennane, "Noise removal from hyperspectral images by multidimensional filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 7, pp. 2061–2069, Jul. 2008.

[45] X. Liu, S. Bourennane, and C. Fossati, "Denoising of hyperspectral images using the PARAFAC model and statistical performance analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 3717–3724, Oct. 2012.

[46] B. Rasti, J. R. Sveinsson, M. O. Ulfarsson, and J. A. Benediktsson, "Hyperspectral image denoising using a new linear model and sparse regularization," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2013, pp. 457–460.

[47] B. Rasti, J. R. Sveinsson, M. O. Ulfarsson, and J. A. Benediktsson, "Hyperspectral image restoration using wavelets," *Proc. SPIE*, vol. 8892, pp. 37–45, Oct. 2013.

[48] B. Rasti, J. R. Sveinsson, and M. O. Ulfarsson, "Total variation based hyperspectral feature extraction," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2014, pp. 4644–4647.

[49] L. Zhuang and J. M. Bioucas-Dias, "Fast hyperspectral image denoising and inpainting based on low-rank and sparse representations," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 730–742, Mar. 2018.

[50] W. He, Q. Yao, C. Li, N. Yokoya, and Q. Zhao, "Non-local meets global: An integrated paradigm for hyperspectral denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 6861–6870.

[51] 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.

[52] Y. Peng, J. Suo, Q. Dai, and W. Xu, "Reweighted low-rank matrix recovery and its application in image restoration," *IEEE Trans. Cybern.*, vol. 44, no. 12, pp. 2418–2430, Dec. 2014.

[53] H. Ji, C. Liu, Z. Shen, and Y. Xu, "Robust video denoising using low rank matrix completion," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2010, pp. 1791–1798.

[54] 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.

[55] H. Ji, S. Huang, Z. Shen, and Y. Xu, "Robust video restoration by joint sparse and low rank matrix approximation," *SIAM J. Imag. Sci.*, vol. 4, no. 4, pp. 1122–1142, Jan. 2011.

[56] X. Cao, Q. Zhao, D. Meng, Y. Chen, and Z. Xu, "Robust low-rank matrix factorization under general mixture noise distributions," *IEEE Trans. Image Process.*, vol. 25, no. 10, pp. 4677–4690, Oct. 2016.

[57] Y. Chen, Y. Guo, Y. Wang, D. Wang, C. Peng, and G. He, "Denoising of hyperspectral images using nonconvex low rank matrix approximation," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 5366–5380, Sep. 2017.

[58] X. Bai, F. Xu, L. Zhou, Y. Xing, L. Bai, and J. Zhou, "Nonlocal similarity based nonnegative tucker decomposition for hyperspectral image denoising," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 701–712, Mar. 2018.

[59] Y. Wang, J. Peng, Q. Zhao, Y. Leung, X.-L. Zhao, and D. Meng, "Hyperspectral image restoration via total variation regularized low-rank tensor decomposition," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1227–1243, Apr. 2018.

[60] Y. Chen, W. He, N. Yokoya, and T.-Z. Huang, "Hyperspectral image restoration using weighted group sparsity-regularized low-rank tensor decomposition," *IEEE Trans. Cybern.*, vol. 50, no. 8, pp. 3556–3570, Aug. 2020.

[61] H. Zhang, L. Liu, W. He, and L. Zhang, "Hyperspectral image denoising with total variation regularization and nonlocal low-rank tensor decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 5, pp. 3071–3084, May 2020.

[62] J. A. Bengua, H. N. Phien, H. D. Tuan, and M. N. Do, "Efficient tensor completion for color image and video recovery: low-rank tensor train," *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2466–2479, May 2017.

[63] M. E. Kilmer and C. D. Martin, "Factorization strategies for third-order tensors," *Linear Algebra Appl.*, vol. 435, no. 3, pp. 641–658, Aug. 2011.

[64] H. Fan, C. Li, Y. Guo, G. Kuang, and J. Ma, "Spatial–spectral total variation regularized low-rank tensor decomposition for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 6196–6213, Oct. 2018.

[65] Y.-B. Zheng, T.-Z. Huang, X.-L. Zhao, T.-X. Jiang, T.-H. Ma, and T.-Y. Ji, "Mixed noise removal in hyperspectral image via low-fibered-rank regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 734–749, Jan. 2020.

[66] S. H. Chan, X. Wang, and O. A. Elgendy, "Plug-and-play ADMM for image restoration: Fixed-point convergence and applications," *IEEE Trans. Comput. Imag.*, vol. 3, no. 1, pp. 84–98, Mar. 2017.

[67] 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.

[68] X.-L. Zhao, W.-H. Xu, T.-X. Jiang, Y. Wang, and M. K. Ng, "Deep plug-and-play prior for low-rank tensor completion," *Neurocomputing*, vol. 400, pp. 137–149, Aug. 2020.

[69] Y. Chang, L. Yan, H. Fang, S. Zhong, and W. Liao, "HSI-DeNet: Hyperspectral image restoration via convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 667–682, Feb. 2019.

[70] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep CNN denoiser prior for image restoration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2808–2817.

[71] Q. Zhang, Q. Yuan, C. Zeng, X. Li, and Y. Wei, "Missing data reconstruction in remote sensing image with a unified spatial–temporal–spectral deep convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4274–4288, Aug. 2018.

[72] Q. Zhang, Q. Yuan, J. Li, X. Liu, H. Shen, and L. Zhang, "Hybrid noise removal in hyperspectral imagery with a spatial–spectral gradient network," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 7317–7329, Oct. 2019.

[73] Q. Yuan, Q. Zhang, J. Li, H. Shen, and L. Zhang, "Hyperspectral image denoising employing a spatial–spectral deep residual convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 1205–1218, Feb. 2019.

[74] Y. Chang, M. Chen, L. Yan, X.-L. Zhao, Y. Li, and S. Zhong, "Toward universal stripe removal via wavelet-based deep convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2880–2897, Apr. 2020.

[75] J. W. Pierre, "A novel method for calculating the convolution sum of two finite length sequences," *IEEE Trans. Educ.*, vol. 39, no. 1, pp. 77–80, Feb. 1996.

[76] M. E. Kilmer, K. Braman, N. Hao, and R. C. Hoover, "Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 1, pp. 148–172, Jan. 2013.

[77] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model based reconstruction," in *Proc. IEEE Global Conf. Signal Inf. Process.*, Dec. 2013, pp. 945–948.

[78] B. Rasti, J. R. Sveinsson, M. O. Ulfarsson, and J. A. Benediktsson, "Hyperspectral image denoising using first order spectral roughness penalty in wavelet domain," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2458–2467, Jun. 2014.

[79] C. Cao, J. Yu, C. Zhou, K. Hu, F. Xiao, and X. Gao, "Hyperspectral image denoising via subspace-based nonlocal low-rank and sparse factorization," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 3, pp. 973–988, Mar. 2019.

[80] M. Gu, "Subspace iteration randomization and singular value problems," *SIAM J. Sci. Comput.*, vol. 37, no. 3, pp. 1139–1173, 2015.

[81] J. Zhang, A. K. Saibaba, M. E. Kilmer, and S. Aeron, "A randomized tensor singular value decomposition based on the t-product: Randomized t-SVD based on t-product," *Numer. Linear Algebra Appl.*, vol. 25, no. 5, Oct. 2018, Art. no. e2179.

Yun-Yang Liu received the B.S. degree in information and computing science from Henan Normal University, Xinxiang, China, in 2019. She is pursuing the M.S. degree with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, China.

Her research interests include model-driven methods for high-dimensional image processing problems.

Xi-Le Zhao received the M.S. and Ph.D. degrees from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2009 and 2012, respectively.

He is a Professor with the School of Mathematical Sciences, UESTC. His research interests mainly focus on model-driven and data-driven methods for image processing problems. More information can be found on his homepage https://zhaoxile.github.io/.

Yu-Bang Zheng (Student Member, IEEE) received the B.S. degree in information and computing science from the Anhui University of Finance and Economics, Bengbu, China, in 2017. He is pursuing the Ph.D. degree with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, China.

His research interests include sparse and low-rank modeling for high-dimensional image processing problems. More information can be found on his homepage https://yubangzheng.github.io.

Tian-Hui Ma received the B.S. and Ph.D. degrees in applied mathematics from the University of Electronic Science and Technology of China, Chengdu, China, in 2011 and 2017, respectively.

During 2016, he was a Visiting Student with the University of Texas at Dallas, Dallas, TX, USA. He is a Post-Doctoral Researcher with the School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, China. His research interests include inverse problems in image processing and sparse optimization.

Hongyan Zhang (Senior Member, IEEE) received the B.S. degree in geographic information system and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2005 and 2010, respectively.

He has been a Full Professor with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, since 2016. He is a Young Chang-Jiang Scholar appointed by the Ministry of Education of China. He scored first in the Pairwise Semantic Stereo Challenge of the 2019 Data Fusion Contest organized by the IEEE Image Analysis and Data Fusion Technical Committee. He has authored or coauthored more than 90 research articles and 8 talents. His research interests include image reconstruction for quality improvement, hyperspectral information processing, and agricultural remote sensing.

Dr. Zhang serves as an Associate Editor of *Photogrammetric Engineering and Remote Sensing* and *Computers and Geosciences* and the Session Chair of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) Conference and the 2015 IEEE WHISPERS Conference. He is a Reviewer of more than 30 international academic journals, including IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, and IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.