Combined Deep Priors With Low-Rank Tensor Factorization for Hyperspectral Image Restoration

Qiang Zhang, *Member, IEEE*, Yushuai Dong, Qiangqiang Yuan, *Member, IEEE*, Meiping Song, *Member, IEEE*, and Haoyang Yu, *Member, IEEE*

***Abstract*— Mixed noise pollution severely disturbs hyperspectral image (HSI) processing and applications. Plenty of algorithms have been developed to address this issue via two strategies: model-driven or data-driven strategy. However, model-driven methods exist in the highly time-consuming weakness of iterative optimization and unstable sensitivity of setting parameters. Data-driven methods usually perform poor due to the overfitting effects. To solve these issues, we combine both the deep denoising priors with low-rank tensor factorization (DP-LRTF) for HSI restoration. The proposed method uses Tucker tensor factorization to depict the global spectral low-rank constraint. Then the spectral orthogonal basis and spatial reduced factor are optimized by two deep denoising priors, respectively. Through this integrated strategy, we can simultaneously exploit the intrinsic low-rank property of HSI, and utilize the powerful feature extraction ability by deep learning for HSI restoration. Compared with model-driven and data-driven methods, DP-LRTF outperforms on HSI mixed noise removal and execution efficiency for various simulated/real experiments.**

***Index Terms*— Deep denoising priors, hyperspectral image (HSI), low-tank, restoration, Tucker-1 tensor factorization.**

I. Introduction

DUE to the atmospheric absorption and sensor defects, hyperspectral image (HSI) inevitably suffers from noise pollution [1]. The noise types may include Gaussian noise, stripe noise, impulse noise, and mixed noise. This issue severely disturbs the application of HSI. Hence, HSI restoration is significant for improving the quality of HSI.

Up to now, plenty of HSI restoration algorithms have been presented [2]. From the perspective of optimization way, these approaches could be divided into two types: model-driven methods and data-driven methods [3]. Detailed descriptions are depicted as follows.

1) *Model-Driven Methods:* By establishing the variation models and defining special constraints, this type of method can take advantage of the prior information in HSI to reduce noise [4]. HSI can be regarded as a three-order tensor, in which there exist obvious sparsity and low-rank property. For instance, Zhang et al. [5] transformed the HSI recovery problem to the low-rank matrix factorization, by unfolding the 3-D spatio-spectral cube into the 2-D matrix. Zheng et al. [6] built a double-factor-regularized tensor factorization model, to flexibly depict the structure of noisy HSI.
2) *Data-Driven Methods:* Different from model-driven methods, data-driven methods adopt the deep learning strategy for HSI restoration. By optimizing trainable parameters from clean labels, deep learning could effectively extract internal features in HSI. For example, Yuan et al. [7] introduced the spatio-spectral convolutional neural network (CNN), to learn the nonlinear map from noisy HSIs to clean HSIs. Furthermore, Zhang et al. [8] utilized spatial and spectral gradient network for mixed noise removal in HSI. Wei et al. [9] developed a quasi-recurrent 3-D CNN for HSI restoration.

In summary, there are pros and cons in the two type of methods for HSI restoration. Model-driven methods can accurately depict the inherent characteristics of HSI. However, these methods are sensitive to the setting parameters, such as rank value and iteration number [10]. Besides, the consuming-time of model-based methods usually take too long, due to the complex iterative optimizations [11]. In terms of data-driven methods, these methods execute more efficiently due to the data-driven strategy. Nevertheless, the simulated noise distribution for clean HSI samples is usually hard to agree with the real noise distribution because of the complicated imaging procedure [12]. Last but not least, different hyperspectral sensors have different spectral scopes and spectral resolution. While data-driven methods usually perform poor, if the spatial scale and spectral range of training HSI samples are dissimilar with testing HSIs. These issues greatly affect the generality of data-driven methods for HSI restoration.

From the above perspectives, we develop a novel method for HSI restoration via combining the model-driven with data-driven strategy. The main contributions are listed as follows.

1) We combine both the deep denoising priors with low-rank tensor factorization (DP-LRTF) for HSI restoration. Tucker tensor factorization is employed to depict the global spectral low-rank constraint.

Manuscript received 14 December 2022; accepted 9 January 2023. Date of publication 12 January 2023; date of current version 2 February 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 41922008, Grant 42101350, Grant 61971082, and Grant 61971319. *(Corresponding author: Qiangqiang Yuan.)*

Qiang Zhang, Yushuai Dong, Meiping Song, and Haoyang Yu are with the Center of Hyperspectral Imaging in Remote Sensing (CHIRS), Information Science and Technology College, Dalian Maritime University, Dalian 116026, China (e-mail: whuqzhang@gmail.com; d1120221466@dlmu.edu.cn; smping@163.com; yuhy@dlmu.edu.cn).

Qiangqiang Yuan is with the School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China (e-mail: yqiang86@gmail.com).

Digital Object Identifier 10.1109/LGRS.2023.3236341

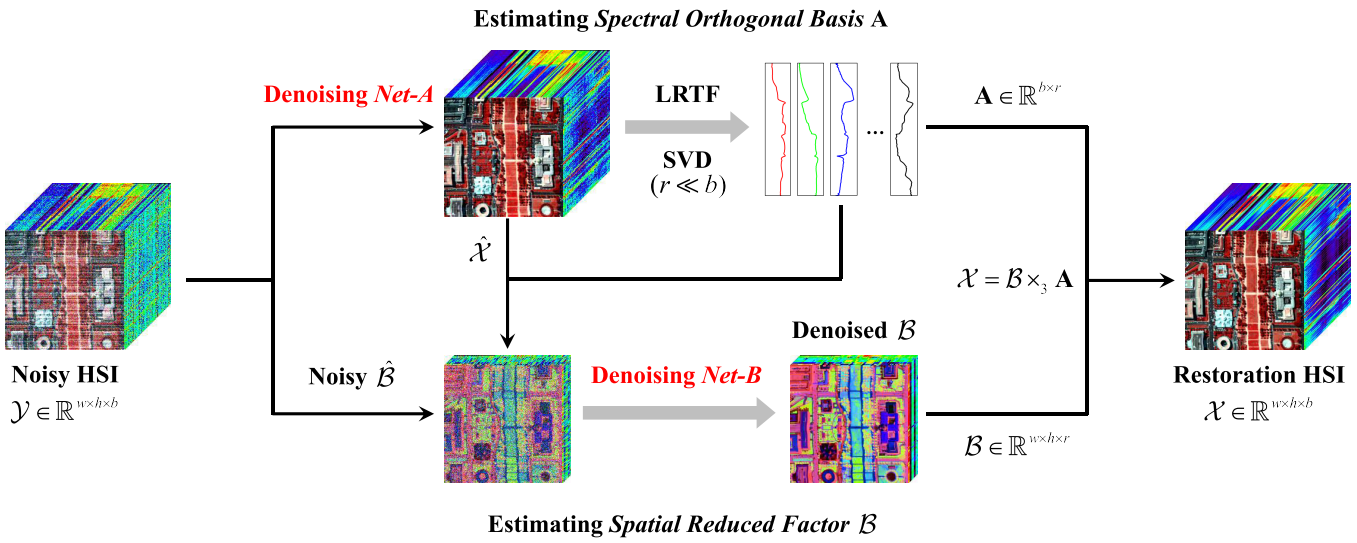

Fig. 1. Flowchart of the proposed DP-LRTF model for HSI restoration.

2) DP-LRTF optimizes the spectral orthogonal basis and spatial reduced factor via two deep denoising networks, without complex and time-consuming iterative optimization.
3) DP-LRTF outperforms on mixed noise removal and execution efficiency for various noisy HSIs (indoor and outdoor), compared with both model-driven and data-driven methods.

II. Proposed HSI Restoration Model

A. Problem Formulation

Considering mixed noise as the additive noise elements in HSI, the general HSI degraded model can be simplified as

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

where $\mathcal{Y} \in \mathbb{R}^{w\times h\times b}$ stands for the acquired noisy HSI. $\mathcal{X}$ denotes the clean HSI. $\mathcal{N}$ represents the random-distribution noise and $\mathcal{S}$ refers to the sparse-distribution noise. Apparently, the solving procedure from $\mathcal{Y}$ to $\mathcal{X}$ is an ill-posed problem. Taking the low-rank characteristic in HSI, an extensive HSI restoration framework could be formulated as

$$\min_{\mathcal{X},\mathcal{S}} \frac{1}{2}\|\mathcal{Y} - \mathcal{X} - \mathcal{S}\|_F^2 + \alpha \cdot \mathcal{T}(\mathcal{X}) + \beta \cdot \|\mathcal{S}\|_1 \tag{2}$$

where $\mathcal{T}(\cdot)$ can be represented as the low-rank tensor factorization (LRTF) model to impose the low-rank prior for HSI. $\|\mathcal{S}\|_1$ constraints the sparsity prior. α and β are the balancing factors of the two regularization prior terms. In addition, the global spectral low-rank property in subsapce field could be denoted as the Tucker decomposition

$$\mathcal{X} = \mathcal{B} \times_3 \mathbf{A} \tag{3}$$

where $\times_k$ stands for mode-k tensor-matrix product [13]. $\mathbf{A} \in \mathbb{R}^{b\times r}$ $(r \ll b)$ denotes the spectral orthogonal basis matrix. $\mathcal{B} \in \mathbb{R}^{w\times h\times r}$ represents the spatial reduced factor. Then the HSI restoration framework in (2) could be rewritten as

$$\min_{\mathbf{A},\mathcal{B},\mathcal{S}} \frac{1}{2}\|\mathcal{Y} - \mathcal{B} \times_3 \mathbf{A} - \mathcal{S}\|_F^2 + \alpha \cdot \mathcal{T}(\mathcal{B}) + \beta \cdot \|\mathcal{S}\|_1. \tag{4}$$

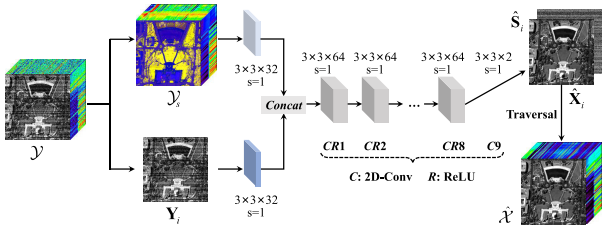

Fig. 2. Deep denoising prior Net-A in the proposed DP-LRTF framework.

B. DP-LRTF HSI Restoration Model

Based on (4), the flowchart of the proposed DP-LRTF model is shown in Fig. 1. The general structure of DP-LRTF can be divided as two parts: estimating spectral orthogonal basis $\mathbf{A}$ and estimating spatial reduced factor $\mathcal{B}$ through two deep denoising priors, respectively. Details of these two parts are described below.

1) Estimating Spectral Orthogonal Basis: Before estimating the spectral orthogonal basis $\mathbf{A}$, we firstly obtain the preliminary HSI denoising result $\hat{\mathcal{X}}$ via deep denoising prior Net-A

$$\hat{\mathcal{X}} = \overset{b}{\underset{i=1}{\Gamma}}(\hat{\mathbf{X}}_i) = \overset{b}{\underset{i=1}{\Gamma}}(\mathbf{Y}_i - \text{Net-A}(\mathbf{Y}_i, \mathcal{Y}_s)) \tag{5}$$

where $\Gamma(\cdot)$ represents the band traversal procedure in Net-A. $\mathbf{Y}_i$ stands for the ith band in noisy HSI $\mathcal{Y}$. $\mathcal{Y}_s$ refers to the corresponding adjacent spectral gradients of $\mathbf{Y}_i$.

As shown in Fig. 2, Net-A fully extracts spatio-spectral features through eight-layer 2-D convolution layers and ReLU layers. The final denoising result $\hat{\mathbf{X}}_i$ and estimating sparse noise $\hat{\mathbf{S}}_i$ are simultaneously exported. In terms of the Net-A's optimizing, specific descriptions are accounted in Section II-C.

By reason of the complexity for blended noise distribution, residual noise especially on spectral dimension in $\hat{\mathcal{X}}$ is obviously not removed in Fig. 1. This issue is also the limitation and weakness of deep learning-based methods for HSI restoration. Therefore, to overcome this problem, we introduce the LRTF strategy to better exploit the low-rank prior of three-order tensor. As displayed in Fig. 1, the spectral orthogonal

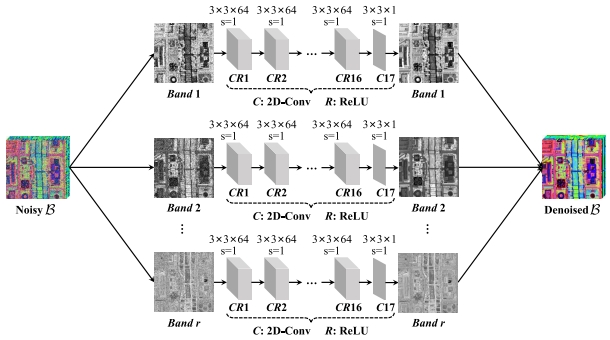

Fig. 3. Denoising Net-B in the proposed DP-LRTF.

basis $\mathbf{A}$ is generated from preliminary HSI denoising result $\hat{\mathcal{X}}$ in (5) by singular value decomposition (SVD)

$$\mathbf{A} = \mathrm{SVD}_r(\hat{\mathcal{X}}_{(3)}), \quad \text{s.t. } \mathbf{A}^\top \mathbf{A} = \mathbf{I} \tag{6}$$

where $\mathcal{X}_{(n)}$ stands for n-order folding format of tensor $\mathcal{X}$. $\mathrm{SVD}_r(\cdot)$ denotes the interception result with rank parameter r under matrix SVD function. We transfer three-order tensor as a mode-3 matrix and use matrix SVD to calculate $\mathbf{A}$.

2) Estimating Spatial Reduced Factor: After getting the spectral orthogonal basis $\mathbf{A}$ in (6), we utilize another deep denoising prior to estimating spatial reduced factor $\mathcal{B}$. Firstly, the initialized factor $\hat{\mathcal{B}}$ is generated as follows:

$$\hat{\mathcal{B}} = \mathrm{reshape}_{h,w,r}\left(\mathbf{A}^\top \hat{\mathcal{X}}_{(3)}\right)^\top \tag{7}$$

where $\mathrm{reshape}_{n_1,n_2,n_3}(\cdot)$ function represents the dimension reshaping operation with the new dimension n_1–n_3. Depending on the powerful feature extraction and expression ability of CNN, Net-B is utilized to eliminate the residual noise in $\hat{\mathcal{B}}$. As shown in Fig. 3, a 17-layer CNN model is employed for spatial reduced image $\hat{\mathcal{B}}$ denoising

$$\mathcal{B} = \prod_{i=1}^{r}(\mathbf{B}_i) = \mathrm{IN}\left\{\prod_{i=1}^{r}(\mathrm{N}(\hat{\mathbf{B}}_i) - \text{Net-B}(\mathrm{N}(\hat{\mathbf{B}}_i)))\right\} \tag{8}$$

where $\mathrm{N}(\cdot)$ function represents the normalized operation for each band of $\hat{\mathcal{B}}$. $\mathrm{IN}(\cdot)$ function refers to the inverse normalized operation of $\mathrm{N}(\cdot)$. The denoising procedure for factor $\hat{\mathcal{B}}$ obeys band by band mode through Net-B, as shown in Fig. 3. In terms of the Net-B's training and optimizing, specific descriptions are given in Section II-C.

C. Network Optimization

Taking both the spatial and spectral information into consideration, the loss function ξ_{A} of Net-A is defined below

$$\xi_{\mathrm{A}} = \xi_{\mathrm{spatial}} + \lambda \cdot \xi_{\mathrm{sparse}} \tag{9}$$

$$\xi_{\mathrm{spatial}} = \frac{1}{2\rho}\sum_{i=1}^{\rho}\left\|\text{Net-A}\left(\mathbf{Y}_k^i, \mathcal{Y}_s^i\right) - \left(\mathbf{Y}_k^i - \mathbf{X}_k^i\right)\right\|_F \tag{10}$$

$$\xi_{\mathrm{sparse}} = \frac{1}{2\rho}\sum_{i=1}^{\rho}\left\|\nabla\text{Net-A}\left(\mathbf{Y}_k^i, \mathcal{Y}_s^i\right) - \mathbf{S}_k^i\right\|_1 \tag{11}$$

where ξ_{spatial} denotes the spatial loss term to constrain the global spatial texture information. ξ_{sparse} represents the sparse loss term to suppress the sparse noise. λ stands for the penalty coefficient to equilibrate spatial and sparse terms. ρ refers to the number of HSI training patches. ∇Net-A points to the sparse gradient output of Net-A. $\mathbf{S}_k$ is the simulated additive sparse noise of kth band in training HSI samples.

For Net-B, we estimate the denoising spatial reduced factor $\mathcal{B}$ from the noisy factor $\hat{\mathcal{B}}$. A deep residual learning strategy is employed for optimizing this training procedure and the loss function ξ_{B} is determined as follows:

$$\xi_{\mathrm{B}} = \frac{1}{2\mu}\sum_{i=1}^{\mu}\left\|\text{Net-B}(\hat{\mathbf{B}}^i) - (\hat{\mathbf{B}}^i - \mathbf{B}^i)\right\|_F \tag{12}$$

where μ represents the number of natural image training patches in Net-B. The whole training procedure of the detailed parameters, training samples, and related operations in both Net-A and Net-B can be found in Section III-A.

TABLE I
EVALUATION INDEXES ON THE SIMULATED W. DC MALL HSI DATASET

Case	Index	Noisy	LRMR	NGMeet	SSGN	Proposed
1	MPSNR	20.287	29.938	33.657	33.248	**35.196**
	MSSIM	0.7464	0.9617	0.9794	0.9715	**0.9823**
	MSAM	19.883	7.2454	4.3681	5.0693	**3.5762**
2	MPSNR	19.742	25.863	30.646	28.842	**32.469**
	MSSIM	0.7065	0.9342	0.9658	0.9527	**0.9826**
	MSAM	20.138	8.6775	5.4913	6.3862	**3.9753**
	MTime	—	448.32	216.78	**7.8629**	13.496

D. Analysis of Complexity

The computational complexity consists of three steps: low-rank tensor Tucker factorization, Net-A, and Net-B. The total complexity is $O(whb + wh\min wh, r + whr)$.

III. Experimental Results

A. Network Training

Net-A and Net-B are trained separately in this work. we employ the Xiongan Matiwan Village HSI [12] as the training data set. This HSI is captured by aircraft-based hyperspectral instrument from the wavelength 400–1000 nm. Both the simulated non-i.i.d. Gaussian noise and stripe noise are imposed on the clean HSI patches as noisy HSI samples Net-A. The penalty coefficient λ is fixed as 0.01 in (9). The number μ of image training patches in loss term (12) is equal to 160 000. The simulated non-i.i.d. Gaussian noise is imposed on the clean image patches as noisy image samples for Net-B.

B. Simulated Experiments

To validate the effectiveness of the proposed DP-LRTF method for HSI restoration, we simulate two cases for both W. DC Mall and Pavia University HSI data sets.

Case 1 (Gaussian Noise): Every band in the two HSIs is contaminated by additive Gaussian noise. The variance of simulated Gaussian noise is equal for each band.

Case 2 (Mixed Noise): Every band in the two HSIs is contaminated by additive stripe noise, which is simultaneously contaminated by Gaussian noise in case 1.

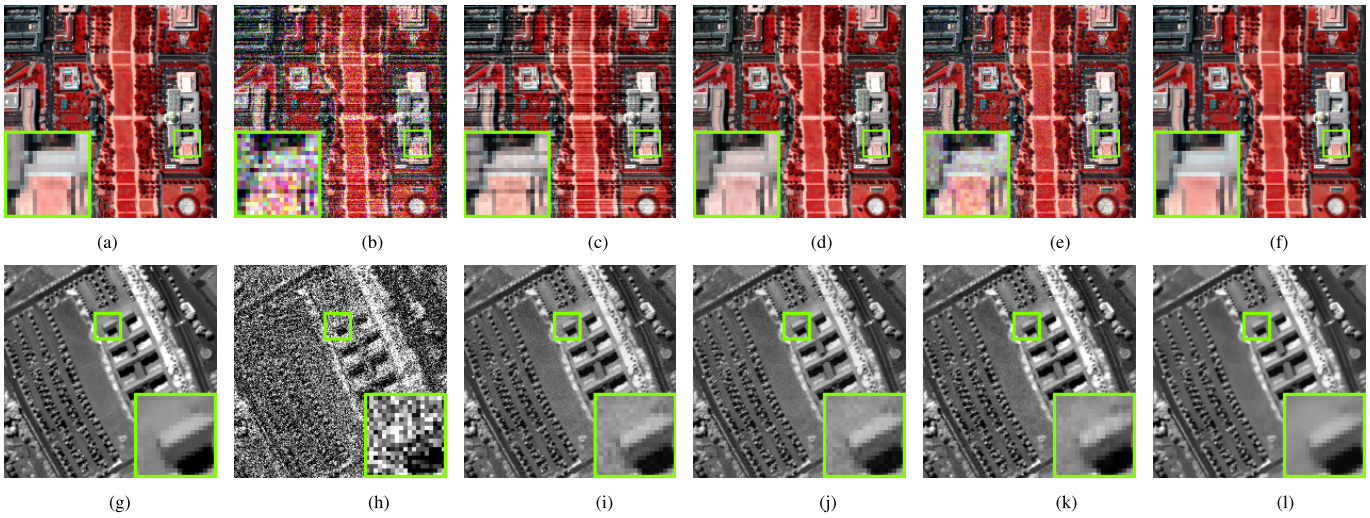

Fig. 4. (Top line) Simulated denoising results (a)–(f) for bands (57, 27, 17) of W. DC Mall HSI dataset in case 2. (Bottom line) Simulated denoising results (g)–(l) for band 3 of Pavia University HSI dataset in case 1.

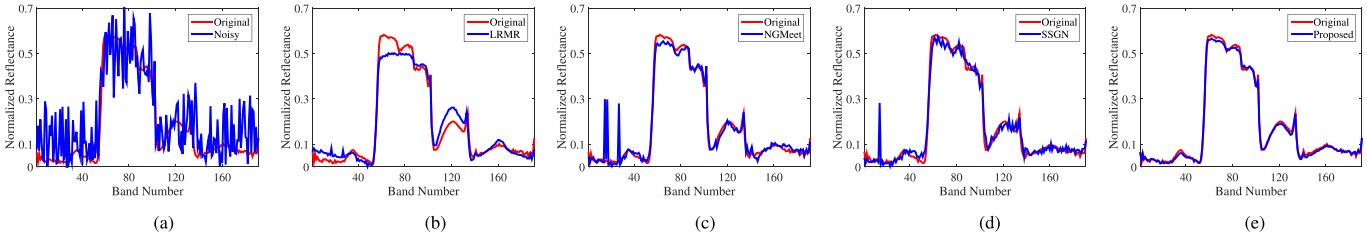

Fig. 5. Spectral curves in position (179, 64) of simulated W. DC Mall HSI dataset in case 2. (a) Noisy. (b) LRMR. (c) NGMeet. (d) SSGN. (e) Proposed.

TABLE II
EVALUATION INDEXES ON THE SIMULATED PAVIA UNIVERSITY HSI DATASET

Case	Index	Noisy	LRMR	NGMeet	SSGN	Proposed
1	MPSNR	18.487	27.291	31.845	30.726	**32.943**
	MSSIM	0.7258	0.9496	0.9613	0.9579	**0.9704**
	MSAM	20.469	7.818	4.9542	5.6768	**4.0359**
2	MPSNR	17.924	24.068	29.139	27.654	**30.487**
	MSSIM	0.6871	0.9273	0.9602	0.9438	**0.9712**
	MSAM	21.245	8.9624	5.687	6.1724	**4.5648**
	MTime	—	514.82	339.64	**7.5727**	18.641

Besides, four state-of-the-art HSI denoising algorithms are regarded as the comparisons. These methods include both model-driven methods: LRMR [5], NGMeet [13] and FGSLR [14]; data-driven method: SSGN [8].

In terms of the quantitative evaluation indexes for HSI restoration, mean peak-signal-to-noise-ratio (MPSNR), mean structural-similarity-index-measure (MSSIM), mean spectral angle (MSA) and mean running-time (MTime) for each algorithm are given in the two simulated HSI datasets, as listed in Tables I and II. The optimal index for each case is marked in bold format.

Tables I and II display the four objective evaluation indexes of three contrast algorithms, under the two simulated cases. The top line in Fig. 4 shows the simulated denoising results (a)–(f) for bands (57, 27, 17) of the W. DC Mall HSI dataset in case 2. The bottom line in Fig. 4 shows simulated denoising results (g)–(l) for band 3 of the Pavia University HSI dataset in case 1. For better comparisons, the enlarged regions for local details are also given in Fig. 4.

As shown in Tables I and II, the proposed DP-LRTF outperforms on MPSNR, MSSIM, and MSAM indexes, compared with both model-driven methods (LRMR and NGMeet) and data-driven methods (SSGN). Besides, the consuming time of the proposed DP-LRTF also performs efficiently compared with model-driven methods, benefiting from the deep denoising priors. In case 2, the proposed method can simultaneously remove the mixed noise without obvious residual stripe in Fig. 4(f). While the other HSI restoration methods exist spectral distortion (NGMeet) or residual noise in different degrees (LRMR and SSGN). This also verifies the availability of the proposed DP-LRTF model for HSI restoration.

In addition, spectral preservation is extremely significant for HSI restoration. The spectral curves in position (179, 64) of the simulated W. DC Mall HSI dataset in case 2 are also depicted in Fig. 5. The proposed method can preserve spectral information, compared with the other three HSI restoration methods. This also validates the effectiveness of the global spectral low-rankness prior through the LRTF framework in the proposed method.

C. Real Experiments

To further testify the reliability and compatibility of DP-LRTF, HYDICE Urban and Zhuhai-1 datasets are employed.

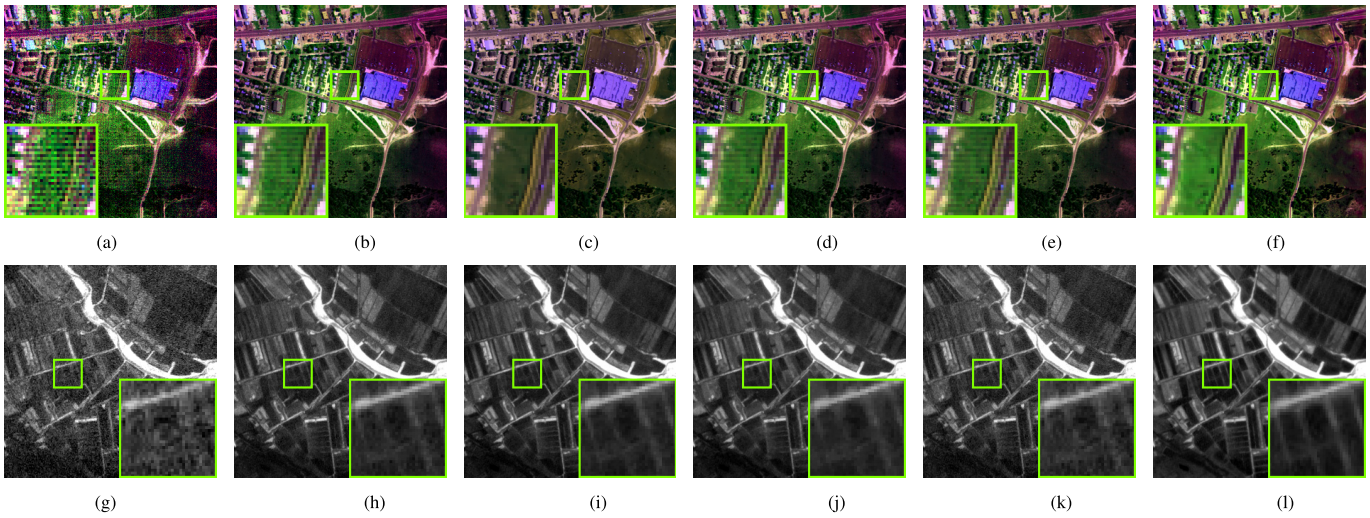

Fig. 6. (Top line) Real denoising results for bands (187, 104, 24) of Urban HSI dataset. (Bottom line) Real denoising results for band 1 of Zhuhai-1 HSI dataset. (a) Noisy (Urban). (b) LRMR. (c) NGMeet. (d) SSGN. (e) FGSLR. (f) Proposed. (g) Noisy (Zhuhai-1). (h) LRMR. (i) NGMeet. (j) SSGN. (k) FGSLR. (l) Proposed.

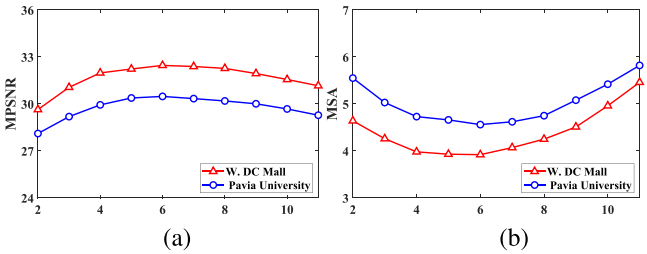

Fig. 7. Parameter sensitivity analysis of the rank value r for the proposed method (in case 2). (a) MPSNR. (b) MSA.

As shown in Fig. 6(a)–(f), the denoising results for bands (187, 104, 24) of the Urban HSI dataset are listed. Especially for the magnified regions, the proposed method outperforms on mixed noise removal and spectral information preservation. While other methods exist obvious residual stripe (LRMR, SSGN and FGSLR) or spectral distortion (NGMeet), to a different degree. For the Zhuhai-1 HSI dataset, the proposed method outperforms random noise removal and spatial details recovery in Fig. 6(h)–(k). Other methods still contain residual noise.

D. Parameter Sensitivity

We discussed the variation of key parameters (rank value) in different data sets. As shown in Fig. 7, although the rank value varies from 2 to 11, the proposed method also performs stably for HSI denoising, without large change. Based on combining model-driven with data-driven strategy, the proposed method manifests the effectiveness to restrain the parameter sensitivity.

IV. Conclusion

In this letter, we combine both the DP-LRTF for HSI restoration. Tucker tensor factorization is employed to depict the global spectral low-rank constraint. DP-LRTF respectively optimizes the spectral orthogonal basis and spatial reduced factor via two deep denoising networks. Experimental results show that DP-LRTF outperforms mixed noise removal for HSI restoration.

References

[1] B. Rasti, Y. Chang, E. Dalsasso, L. Denis, and P. Ghamisi, "Image restoration for remote sensing: Overview and toolbox," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 2, pp. 201–230, Jun. 2022.

[2] B. Rasti, P. Scheunders, P. Ghamisi, G. Licciardi, and J. Chanussot, "Noise reduction in hyperspectral imagery: Overview and application," *Remote Sens.*, vol. 10, no. 3, p. 482, 2018.

[3] B. Rasti, P. Ghamisi, and J. A. Benediktsson, "Hyperspectral mixed Gaussian and sparse noise reduction," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 3, pp. 474–478, Mar. 2020.

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

[5] 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.

[6] Y. Zheng, T. Huang, X. 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.

[7] 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.

[8] 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.

[9] K. Wei, Y. Fu, and H. Huang, "3-D quasi-recurrent neural network for hyperspectral image denoising," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 363–375, Jan. 2021.

[10] Q. Zhang, Q. Yuan, M. Song, H. Yu, and L. Zhang, "Cooperated spectral low-rankness prior and deep spatial prior for HSI unsupervised denoising," *IEEE Trans. Image Process.*, vol. 31, pp. 6356–6368, 2022.

[11] A. Mahmood and M. Sears, "Per-pixel noise estimation in hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[12] Q. Zhang, Q. Yuan, J. Li, F. Sun, and L. Zhang, "Deep spatio–spectral Bayesian posterior for hyperspectral image non-i.I.d. noise removal," *ISPRS J. Photogramm. Remote Sens.*, vol. 164, pp. 125–137, Jun. 2020.

[13] W. He et al., "Non-local meets global: An integrated paradigm for hyperspectral image restoration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 4, pp. 2089–2107, Apr. 2022.

[14] Y. Chen, T.-Z. Huang, W. He, X.-L. Zhao, H. Zhang, and J. Zeng, "Hyperspectral image denoising using factor group sparsity-regularized nonconvex low-rank approximation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5515916.