

# Hyperspectral Image Restoration: Where Does the Low-Rank Property Exist

Yi Chang, *Member, IEEE*, Luxin Yan, *Member, IEEE*, Bingling Chen, Sheng Zhong,  
and Yonghong Tian, *Member, IEEE*

**Abstract**—Hyperspectral images (HSIs) restoration is to recover the clean image from degraded version, such as the noisy, blurred or damaged. Recent low-rank tensor-based recovery methods have been widely explored in HSIs restoration. Most of previous methods, however, neglect an inconspicuous but important phenomenon that the physical meaning and dimension along the spatial, spectral and non-local mode are markedly different. In this work, we discover the low-rank property **discrepancy** along spatial, spectral and non-local self-similarity mode in the HSIs, and argue that the intrinsic low-rank correlations along each mode contribute different to the final restoration results. Consequently, we figure out that the combination of the spectral and non-local induced low-rank is most beneficial for HSIs modeling, and propose an optimal low-rank tensor model (OLRT) for HSIs restoration. Further, we not only explore the low-rank property in the image component, but also in the sparse error component (stripe noise in HSIs). Thus, we extend OLRT to the OLRT-RPCA with low-rank tensor priors for both the HSIs and sparse error. Besides, previous methods are usually designed for one specific HSIs task, which is less robust to various tasks. We prove that the proposed optimal low-rank prior is very flexible for various HSIs restoration problems including denoising, deblurring, inpainting and destriping. The proposed methods have been extensively evaluated on several benchmarks and tasks, and greatly outperform state-of-the-arts. We show the simple yet effective OLRT strategy is also beneficial to STOA.

**Index Terms**—Hyperspectral images, image restoration, low-rank tensor recovery.

## I. INTRODUCTION

A HSI could provide abundant information with multiple specific frequencies across the electromagnetic spectrum, which facilitates the fine-grained representation of a natural

This work is supported by the projects of the National Natural Science Foundation of China under Grants No. 61571207 and 61971460, in part by the project of China Postdoctoral Science Foundation under Grants No. 2020M672748, in part by the project of the Hubei Provincial Natural Science Foundation of China under Grants No. 2018CFA089, and in part by the project of Key-Area Research and Development Program of Guangdong Province under Grants No. 2019B121204008.

Y. Chang is with the National Key Laboratory of Science and Technology on Multispectral Information Processing, School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan, Hubei, 430074, P.R. China, and also with the AI Center, Pengcheng Laboratory, Shenzhen, P.R. China. (e-mail: owuchangyu@gmail.com);

L. Yan, and S. Zhong are with the National Key Laboratory of Science and Technology on Multispectral Information Processing, School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan, Hubei, 430074, P.R. China. (e-mail:{yanluxin, zhongsheng}@hust.edu.cn);

B. Chen is with the AI Center, Pengcheng Laboratory, Shenzhen, P.R. China. (e-mail: chenbl@pcl.ac.cn).

Y. Tian is with the School of Electrical Engineering and Computer Science, Peking University, Beijing, China, and also AI Center, Pengcheng Laboratory, Shenzhen, P.R. China. (e-mail: yhtian@pku.edu.cn).

scene. Unfortunately, during the imaging procedure, the HSI is usually contaminated by the noises, blurs, and damages, making the HSI unsuitable for subsequent applications. Mathematically, the problem can be generally formulated by a linear degradation model as follow:

$$\mathcal{Y} = \mathcal{T}(\mathcal{X}) + \mathcal{E} + \mathcal{N}, \quad (1)$$

where  $\mathcal{Y} \in \mathbb{R}^{R \times C \times B}$  is an observed HSI,  $\mathcal{X} \in \mathbb{R}^{R \times C \times B}$  represents the desired clean HSI,  $\mathcal{E} \in \mathbb{R}^{R \times C \times B}$  denotes the sparse error (for example stripe noise in HSIs),  $\mathcal{N} \in \mathbb{R}^{R \times C \times B}$  means the additive random noise, and  $\mathcal{T}(\bullet)$  stands for the linear degradation operator. With different settings, Eq. (1) can represent different HSIs restoration problems; When  $\mathcal{T}(\bullet)$  is an identity tensor, the problem (1) becomes HSIs denoising (only consider  $\mathcal{N}$ ) or HSIs destriping (only consider  $\mathcal{E}$ ), or HSIs mixed noise removal (both  $\mathcal{N}$  and  $\mathcal{E}$ ); when  $\mathcal{T}(\bullet)$  is a blur operator, the problem (1) turns into the HSIs deblurring; For HSIs inpainting,  $\mathcal{T}(\bullet)$  is a binary mask tensor, 0 for missing pixels and 1 for fine pixels.

The task is to estimate clean HSIs  $\mathcal{X}$  and sparse error  $\mathcal{E}$  from the given degradation  $\mathcal{Y}$ . This ill-posed nature implies that additional constraints on  $\mathcal{X}$  and  $\mathcal{E}$  have to be enforced. Thus the key is to find appropriate prior of HSIs. Various regularizations have been proposed for HSIs restoration, such as image processing-based filtering methods [1], model-based optimization methods [2]–[45], and deep learning-based convolutional neural network [46]–[51]. The model-based optimization methods can further be classified into 1-D vector-based sparse representation methods [2]–[11], 2-D matrix-based low-rank matrix recovery methods [12]–[22], and 3-D low-rank tensor approximation methods [23]–[45].

In recent years, the low-rank tensor HSIs restoration methods have received significant attention. The tensor-based methods could naturally utilize both the spatial-spectral information, and at the same time well preserve the high-dimension spatial-spectral structural correlation in 3D HSIs. Interested readers could refer to [52] for more details about the tensor decomposition. The core idea of the tensor-based restoration method is to transform the HSIs via low-rank decomposition, and perform the shrinkage in the transformed domain so as to satisfactorily decouple the noises from the image structure.

For example, in [23], Xie *et al.* proposed a multispectral image denoising method by taking both fine-grained tensor sparsity insights of Tucker (sum of the ranks along each mode of a 3-order tensor) and CP low-rank decompositions into consideration. Ji *et al.* [32] proposed a nonlocal-based 4-order tensor low-rank decomposition method for remotely sensed images inpainting. On the contrary, Peng *et al.* [26]

proposed a decomposable nonlocal tensor dictionary learning for multispectral image denoising, which did not enforce the sparsity constraint on the decomposition coefficient.

Most of the previous low-rank tensor modeling methods either add up the ranks along all tensor modes, or do not enforce any constraint on the ranks along each tensor modes. Nevertheless, they neglect that the spatial, spectral, non-local mode are with different physical meaning and different dimension. The rank along each mode is a reflection of its intrinsic structure correlation. That is to say, the different modes along the tensor possess different intrinsic manifold, namely different low-rank property. It has been proved that *a major difference with the matrix case, however, is the fact that the different n-ranks of a higher-order tensor are not necessarily the same* [52]. Thus, we argue that it is unreasonable to ignore the low-rank property discrepancy along each mode.

In this work, we first give a detailed analysis about the rank properties along each mode of the constructed 3-order tensor via the high-order singular value decomposition (HOSVD) [53]. We discover that low-rank property of the non-local self-similarity and spectral correlation is usually superior to that of the spatial (Section III-A). This suggests that the intrinsic rank among each mode is different and motivates us to take the discrepancy of the structural correlation along each mode into consideration. We demonstrate that, counterintuitively, the most common practice of previous methods by equally adding up all the ranks among each mode would negatively affect the final performance. Consequently, we figure out a simple yet effective solution for HSIs low-rank modeling (Section III-B).

In addition, for the sparse error modeling, most of the previous HSIs mixed noise removal methods mainly utilize the  $L_1$  sparsity. The sparse noise in HSIs caused by the different response of neighborhood detectors, is usually non-periodical horizontal or vertical lines due to the push-broom imaging mechanism [54]. Such a simple  $L_1$  constraint has unexpectedly neglected the intrinsic structural correlation characteristic of the line pattern sparse noise. In this work, we show that the low-rank property exists not only in the HSIs, but also in the sparse error component namely the stripe noise. Thus, we extend our HSIs denoising model to the robust tensor principal component analysis RPCA with low-rank constraints for both the HSIs and stripe noise (Section III-C).

Finally, the previous HSIs methods are suitable for specific tasks, such as the denoising [4], inpainting [7], deblurring [6], destriping [2]. In this work, we show that the OLRT is applicable for general HSIs restoration and present an efficient optimization method to handle various low-level HSIs restoration tasks in a unified model (Section IV). Compared with the state-of-the-art HSIs restoration methods, the contributions of the proposed work are three-folds:

- We explore two very important fundamental problems in HSI restoration community: HSIs modeling and sparse noise modeling from the low-rank tensor perspective, and provide a guidance about how much each low-rank property contributes to the restoration. Consequently, we propose an optimal low-rank tensor (OLRT) model by joint spectral and non-local low-rank tensor prior for HSIs restoration, and further extend the OLRT to the RPCA situation.

- Compared with previous methods designed for the specific task, the proposed low-rank tensor prior is very flexible for general HSIs restoration tasks, such as denoising, destriping, deblurring, and inpainting. We believe that the OLRT could be further employed for other HSIs restoration tasks.
- The proposed method has been extensively tested on various HSIs datasets and tasks and significantly outperforms the state-of-the-arts both quantitative and qualitative. Moreover, we show our OLRT strategy is also beneficial to other state-of-the-arts HSIs restoration methods.

## II. RELATED WORK

In Table I, we list the representative HSIs restoration methods. We mainly consider the methodology, year, tasks, information utilization, and brief description. The interested reader could refer to survey work HSIs denoising [55], destriping [56] and inpainting [57] for detailed description.

**1-D Sparse Representation:** The sparse representation methods treat the HSIs restoration as an ill-posed inverse problem by minimizing an energy functional with sparsity constraint on the HSIs. The most representative methods are dictionary learning [8]–[11]. For example, Akhtar *et al.* [58] utilized the linear correlation between the HSIs and MSIs dictionary for HSIs super-resolution. Lu *et al.* [11] presented a superpixel based spectral-spatial adaptive sparse representation method for HSIs denoising. Chang *et al.* [8] integrated the unidirectional total variation and sparse representation regularization for mixed noise removal. Numerous spatial-spectral total variational HSI restoration methods have been proposed [2]–[7]. Yuan *et al.* [4] proposed a HSIs denoising method by employing a spectral-spatial adaptive total variation model. Henrot *et al.* [5] proposed a spatial and spectral smoothness prior with a positivity constraint for HSIs deblurring. Cheng *et al.* [7] further took advantage of nonlocal total variation for remote sensing image inpainting.

**2-D Low-Rank Matrix Recovery:** The sparse representation mathematically models the image patch as a vector. Thus, the 2-D low-rank-based matrix recovery methods have been naturally proposed to better preserve the image structure [12]–[22]. For example, Fu *et al.* [15] proposed a spectral and spatial joint low-rank model for coded HSIs reconstruction. Zhang *et al.* [12] took advantage of the low-rank property along the spectral mode by lexicographically ordering the 3-D HSIs into a 2-D matrix. Rasti *et al.* [16] further exploited the Stein’s unbiased risk estimator for fully automatic parameter selection. The nonconvex low-rank regularizers, i.e., weighted Schatten  $p$ -norm [14], smooth rank approximation [59],  $L_0$  resemble normalized  $\varepsilon$ -penalty [60], have been extensively studied to offer better approximation to the original low-rank assumption. The hybrid local sparsity + global low-rank framework has also attracted great attention. He *et al.* [20] proposed a total variation-regularized low-rank matrix factorization method for HSIs restoration. Zhao *et al.* [18] jointly utilized the sparse representation and low-rank constraint in spatial and spectral domains for HSIs denoising.

**3-D Low-Rank Tensor Approximation:** Although the vector /matrix-based methods have achieved excellent restoration

Table I  
ILLUSTRATION OF EXISTING HSIS RESTORATION METHODS. LR DENOTES THE LOW-RANK FOR SHORT.

Category	Paradigm	Method	Year	Tasks	Brief Description	Information Utilization
1-D Sparse Representation	Gradient Based Method	Shen [2]	2009	Inpainting	Huber-Markov Model	Spatial Smoothness
		Bouali [3]	2011	Destriping	Unidirectional Variational Model	Spatial Smoothness
		Yuan [4]	2012	Denoising	Spectral-Spatial Adaptive Total Variation Model	Spectral-Spatial Sparsity
		Henrot [5]	2013	Deblurring	Spectral-Spatial Smoothness + Positive Constraint	Spectral-Spatial Smoothness
		Zhao [6]	2013	Deblurring	Total Variation + $L_1$ Sparsity	Spatial Smoothness + Sparsity
	Dictionary Learning	Cheng [7]	2014	Inpainting	Multichannel Nonlocal Total Variation Model	Spectral-Spatial Smoothness + Nonlocal
		Chang [8]	2014	Denoising	Unidirectional Total Variation + Dictionary Learning	Spatial Smoothness + Sparsity
		Ye [9]	2015	Denoising	Multitask Sparse Nonnegative Matrix Factorization	Spectral-Spatial Sparsity
		Fu [10]	2015	Denoising	Adaptive Spatial-spectral Dictionary Learning	Spectral-Spatial Sparsity + Nonlocal Similarity
		Lu [11]	2016	Denoising	Spectral-Spatial Adaptive Sparse Representation	Spectral-Spatial Sparsity
2-D Low-rank Matrix Recovery	Classical LR	Zhang [12]	2014	Denoising	Spectral LR Model	Spectral Correlation
		Cao [13]	2015	Denoising	LR Matrix Factorization	Spectral Correlation
		Xie [14]	2016	Denoising	Weighted Schatten-norm LR	Spectral Correlation
		Fu [15]	2016	Reconstruction	Joint Spectral-spatial LR	Nonlocal Similarity + Spectral Correlation
		Rasti [16]	2017	Denoising	Parameter-free LR	Sparsity + Spectral Correlation
	Hybrid	Lu [17]	2013	Destriping	Graph-regularized LR Representation	Spectral Smoothness + Spectral Correlation
		Zhai [18]	2015	Denoising	Sparse Representation + Low-rank Constraint	Spectral Sparsity + Spectral Correlation
		Chang [19]	2016	Destriping	TV + LR Decomposition	Spatial Smoothness
		He [20]	2016	Denoising	TV-regularized LR Matrix Factorization	Spatial Smoothness + Spectral Correlation
		Zhuang [21]	2018	Denoising	Subspace-Based LR	Spectral Correlation
3-D Low-rank Tensor Approximation	Canonical Polyadic Decomposition	Hu [22]	2020	Denoising	Nonconvex 3D TV + Low-rank	Spectral-Spatial Smoothness + Correlation
		Xie [23]	2016	Denoising	CP + Tucker LR Decomposition	Nonlocal Similarity + Spectral Correlation
		Xue [24]	2019	Denoising	CP + Tucker LR Decomposition	Nonlocal Similarity + Spectral Correlation
		Xie [25]	2019	Inpainting	Nonconvex CP/Tucker LR Decomposition	Spectral Correlation
	Tucker Decomposition	Peng [26]	2014	Denoising	Tucker Decomposition Based Dictionary Learning	Nonlocal Similarity + Spectral Correlation
		Dong [27]	2015	Denoising	Tucker Decomposition + Laplacian Scale Mixture	Nonlocal Similarity + Spectral Correlation
		Chang [28]	2017	Denoising	Hyper-laplacian + Unidirectional LR Tensor	Nonlocal Similarity + Spectral Smoothness
		Ng [29]	2017	Inpainting	Weighted Tucker LR Decomposition	Spectral Correlation
		Wang [30]	2018	Denoising	Spatial-spectral TV-Regularized Tucker LR Decomposition	Spectral Correlation + Spectral-Spatial Smoothness
		Chen [31]	2018	Destriping	Anisotropic TV + Tucker LR Decomposition	Spectral Correlation + Spectral-Spatial Smoothness
		Ji [32]	2018	Inpainting	Nonconvex 4-order Tucker LR Decomposition	Temporal-Spatial-Spectral Correlation + Nonlocal
		Zhang [33]	2019	Reconstruction	Nonconvex Tucker LR Decomposition	Nonlocal Similarity + Spectral Correlation
		He [34]	2019	Denoising	Spectral subspace-Based Unidirectional LR Tensor Recovery	Nonlocal Similarity + Spectral Correlation
		Zhang [35]	2020	Denoising	Spatial-spectral TV-Regularized Tucker LR Decomposition	Spectral Correlation + Spectral-Spatial Smoothness
	t-SVD	Chen [36]	2020	Denoising	Weighted Group Sparsity-Regularized Tucker LR Decomposition	Spectral Correlation + Spatial Sparsity
		Chang [37]	2020	Unified	Weighted LR Tensor Recovery	Nonlocal Similarity + Spectral Correlation
		Gong [38]	2020	Denoising	Tucker Decomposition-Based Shared Dictionary Learning	Nonlocal Similarity + Spectral Correlation
	Tensor-ring Decomposition	Fan [39]	2018	Denoising	Spatial-spectral TV-Regularized t-SVD LR Decomposition	Spectral Correlation + Spectral-Spatial Smoothness
		Kong [40]	2019	Denoising	Block Diagonal t-SVD LR Decomposition	Nonlocal Similarity + Spectral Correlation
		Zheng [41]	2019	Denoising	Low-Fibered-Rank t-SVD Decomposition	Spectral Correlation
	Tensor Train Decomposition	Chen [42]	2020	Super-resolution	Tensor-Ring LR Decomposition	Nonlocal Similarity + Spectral Correlation
		Xu [43]	2020	Super-resolution	Graph-Regularized High-order Tensor-Ring LR Decomposition	Spectral Correlation
	Block Term Decomposition	Dian [44]	2019	Super-resolution	4-order Low Tensor Train Rank	Nonlocal Similarity + Spectral Correlation
	Xiong [45]	2019	Denoising	Spectral-spatial $L_0$ TV-regularized Block Term Decomposition	Spectral Correlation + Spectral-Spatial Smoothness	

results, they inevitably cause damages to the spectral-spatial structural correlation of the HSIs. To alleviate this issue, the low-rank tensor approximation methods have sprung up in recent years [23]–[45]. Various tensor decompositions have been introduced for HSIs modeling, such as CP [23], Tucker [27], t-SVD [39], tensor ring [42], tensor-train [44], block term [45] etc. The spectral correlation and non-local self-similarity are two intrinsic characteristics underlying an HSI. Peng firstly *et al.* [26] modeled the spectral and non-local similarity simultaneously. Latter, there are more sophisticated methods with additional low-rank prior knowledge about constructed tensor [23], [27], [28], [34].

Most of the low-rank tensor models regularize the sum of the rank along each mode of the constructed tensor [23], [29], [32], [33], [39], [41], [44], [61], [62]. However, it is unreasonable to enforce low-rank constraint along each mode equally, since the rank along each mode has different physical meaning and different dimension. The KBR [63] and our previous work LLRT [28] have already noticed this inconspicuous but important phenomenon. In this work, we go further along this research track and figure out the optimal low-rank tensor mode for HSIs restoration. Different from the optimal rank selection of each mode [64], [65], the optimal low-rank here means the best low-rank combination of the HSIs among each mode. Moreover, we show that our model is very flexible for general HSIs restoration, such as denoising, destriping, deblurring, and inpainting.

This work is the journal extension of our conference publi-

cation [28]. The differences are three-folds:

- **Low-rank property exploration:** In our conference version, we explore the low-rank property in the hyperspectral images and difference of the low-rank property along each mode. In this work, we go further by providing a guidance how much each low-rank property contributes to the restoration and figure out the optimal combination of the low-rank property along each mode. Moreover, in this work, we explore the low-rank properties for both the image component and the sparse error component.
- **Low-rank tensor recovery model:** In our conference version, we propose a unidirectional low-rank tensor recovery model. In this work, for HSIs modeling, we propose an optimal low-rank tensor (OLRT) model by a joint spectral and non-local low-rank tensor prior. This work additionally takes the linear degradation operator into consideration, such as the blur or sampling mask. Thus, the loss function between the two papers are obviously different. Moreover, we further extend our OLRT model to the RPCA for handling the stripe noise in HSIs.
- **HSIs restoration tasks:** In our conference version, we only handle the HSIs Gaussian random noise. In this work, we further extend our tasks to the HSIs deblurring, inpainting and destriping. As far as we know, few works have been proposed for unified HSIs restoration. Moreover, in each HSIs restoration task, the proposed OLRT has consistently outperformed the state-of-the-arts approaches by a large margin, especially for the HSIs deblurring and inpainting.

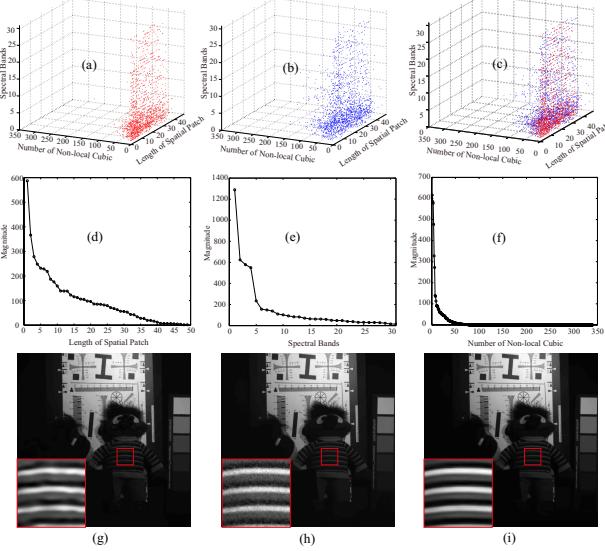


Figure 1. Low-rank property analysis of the constructed 3-order tensor along each mode via HOSVD. (a)-(c) is the visualization of singular values bigger than 1.5, bigger than 1 and smaller than 1.5, and overall bigger than 1 elements in the core tensor, respectively. (d)-(f) is the mean profile of (c) across spatial (mode-1), spectral (mode-3), and non-local self-similarity (mode-2), respectively. (g)-(i) is the corresponding denoising result via spatial, spectral, and non-local self-similarity based low-rank tensor model, respectively.

### III. LOW-RANK PROPERTY IN HSIS

#### A. Closer Look at Low-Rank Property of Tensor

The conventional tensor sparsity measures [23], [29], [32], [33], [39], [41], [44], [61], [62] usually extend the 2-order sparsity measure to higher-order case by simply adding up rank along each modes. However, these works neglect a fact that the different  $n$ -ranks of a higher-order tensor are not necessarily the same, indicating that the rank of each mode is closely related to its intrinsic low-rank subspace.

To understand this, we explore the low-rank property of each mode of a constructed 3-order tensor  $\mathcal{X}_i \in \mathbb{R}^{49 \times 350 \times 31}$  via HOSVD. We performed this experiment on hundreds tensor with one as a representative. In Fig. 1(a)-(c), we give a visual understanding how the singular values distribute in the core tensor. Note that, most of small singular values are trivial. We just choose the larger singular values associating with the major projection orientations. In Fig. 1(d)-(f), we show the mean profile of the core tensor across each mode. To remove the absolute dimension magnitude effect, we uniformly select the first 20 percent principal components along spatial ( $49 * 0.2 \approx 10$ ), spectral ( $31 * 0.2 \approx 7$ ), and non-local ( $350 * 0.2 \approx 70$ ) mode in a relative scale. The corresponding percentage energy value is 53.66%, 71.39%, and 99.58%, respectively. The singular values of the core tensor exhibit significant sparsity with different degrees along each mode. Along the non-local self-similarity mode (mode-2), due to the strong redundancy of the non-local cubics, the coefficients tend to be decreasing extremely fast to zeroes. While along the spatial and spectral mode, albeit still approximately decreasing along the mode, most of the coefficients are non-zeros. Consequently, the corresponding denoising results are shown in Fig. 1(g)-(i). The result obtained by unfolding along the non-

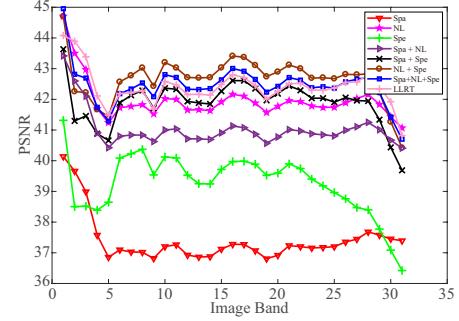


Figure 2. Effectiveness of low-rank prior along each mode and their combination. Spa, Spe, NL denotes the spatial, spectral and non-local respectively. LLRT is our conference paper [28].

local mode is much better than the others, since the sparser representation allow the most improvement [66].

The observation in Fig. 1 has truthfully reflected the intrinsic difference of structure correlation along each mode. The low-rankness of HSIs in non-local dimension is obviously stronger than that in the spectrum and spatial mode. Next is the spectral correlation. For the spatial neighborhood regions such as the spatial texture area, its low-rankness property is the weakest. Therefore, we argue non-local self-similarity is the key property contributing to HSIs restoration performance. This motivates us to treat the rank along each mode differently.

#### B. Optimal Low-Rank Tensor (OLRT) Prior for HSIs

The non-local dimension is more evidently low-rank (as shown in Fig. 1) and neglecting others can help improve efficiency. In our conference version LLRT [28], we only incorporate the low-rank property across the non-local mode and discard these weaker correlations across spatial and spectral mode. However, it might be not so rational that neglecting other useful low-rankness along other dimensions, especially in spectrum, can help improve HSIs recovery quality. To figure out the optimal low-rank modeling in HSIs, we give a comparison of the combination of low-rank prior along each mode, as shown in Fig. 2. We have the following observations:

- For single mode-based low-rank prior (red, purple, green curve), we can find that the non-local mode achieves the best performance, which further validates our conclusion: the structure correlation along the non-local mode is much stronger than that of the spatial or spectral mode.
- The spectral mode low-rank does facilitate the final recovery result (compare purple and grey). The spatial mode low-rank always bring negative influence to the final performance (compare purple and cyan, also grey and blue), which means its low-rank assumptions cannot be met, especially for the texture neighborhood areas.
- The combination of the non-local and spectral low-rank has achieved the best performance (grey curve) and works better than the combination of low-rank along all modes (blue curve), which more reasonably capture the intrinsic sparsity inside the constructed tensor.

From the above analysis, we can conclude that the non-local self-similarity is the key property contributing to HSIs

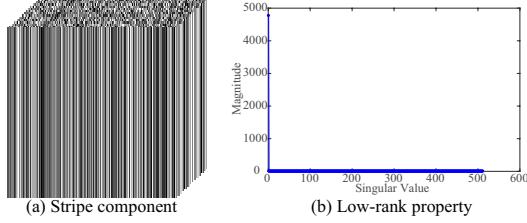


Figure 3. Low-rank property analysis of the sparse error, namely the stripe in HSIs. (a) The stripe cubic with same size as the hyperspectral images. (b) The singular value curve of (a) along the spatially vertical mode.

restoration. The spectral correlation property does facilitate the final recovery result. In this work, we propose a joint spectral and non-local low-rank tensor prior for HSIs modeling.

### C. Low-Rank Property in Sparse Error

Not only the random noise  $\mathcal{N} \in \mathbb{R}^{R \times C \times B}$  degenerates the HSIs quality, but also the structural stripe noise  $\mathcal{E} \in \mathbb{R}^{R \times C \times B}$  exists in real HSIs, which is mainly caused by the difference in the response of multi-detectors, calibration error, and so on [54]. In this work, we formulate this procedure into 3-order image decomposition framework as Eq. (1). Different from the random noise, the stripe noise exhibits a significantly directional appearance. Interested reader would refer to [19], [54] for more details about the stripe.

It is widely accepted that the high-dimensional data usually lies on an intrinsic low-dimensional manifold. The low-rank property is ubiquitous, in which exists not only in the HSIs component, but also the sparse error component. In Fig. 3, we show the low-rank property of the stripe noise. We unfold the 3-order sparse error stripe component along the spatially vertical mode, and then perform the SVD on it. From Fig. 3(b), it can be seen that the singular values rapidly decrease to zero. It is natural for us to understand this that the vertical stripe component possesses much more simple structures (similar line pattern with the same direction) than that of the image component. Our previous work mainly considered the single band case [19]. In this work, we extend this low-rank property from the single band image to 3-order cubic.

## IV. LOW-RANK TENSOR HSIS RESTORATION MODEL

### A. Notations and Preliminaries

In this paper, we denote tensors by boldface Euler script letters, e.g.,  $\mathcal{X}$ . Matrices are represented as boldface capital letters, e.g.,  $X$ ; vectors are expressed with boldface lowercase letters, e.g.,  $x$ , and scalars are denoted by lowercase letters, e.g.,  $x$ . The  $i$ -th entry of a vector  $x$  is denoted by  $x_i$ , element  $(i, j)$  of a matrix  $X$  is denoted by  $x_{ij}$ , and element  $(i, j, k)$  of a 3-order tensor  $\mathcal{X}$  is denoted by  $x_{ijk}$ . The Frobenius norm of an  $N$ -order tensor  $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$  is the square root of the sum of the squares of all its elements, i.e.,  $\|\mathcal{X}\|_F = \sqrt{\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \dots \sum_{i_N=1}^{I_N} x_{i_1 i_2 \dots i_N}^2}$ . Tensor matricization, also named as *unfolding or flattening*, is the process of reordering the elements of an  $N$ -order tensor into a matrix. The mode- $n$  matricization  $X^{(n)} \in \mathbb{R}^{I_n \times (I_1 \dots I_{n-1} I_n \dots I_N)}$  of a

tensor  $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$  is obtained by taking all the mode- $n$  fibers to be the columns of the resulting matrix. Thus, the  $n$ -rank of a given tensor can be analyzed by means of matrix techniques. The rank of the matrix unfolding  $X^{(n)}$  is equal to the  $n$ -rank of  $\mathcal{X}$ , i.e.,  $\text{rank}_n(\mathcal{X}) = \text{rank}(X^{(n)})$  [67].

### B. HSIs Restoration Model

The task of HSIs restoration is to estimate the clean image  $\mathcal{X}$  in presence of the degraded image  $\mathcal{Y}$  under different degradation situations. According to Eq. (1), the HSIs restoration problem can be formulated in the following form:

$$\{\hat{\mathcal{X}}, \hat{\mathcal{E}}\} = \arg \min_{\mathcal{X}, \mathcal{E}} \frac{1}{2} \|\mathcal{T}(\mathcal{X}) + \mathcal{E} - \mathcal{Y}\|_F^2 + \omega P_h(\mathcal{X}) + \rho P_s(\mathcal{E}), \quad (2)$$

where  $P_h(\mathcal{X})$  and  $P_s(\mathcal{E})$  denote the regularization terms to enforce the solution with desired property on the clean HSIs and sparse error respectively,  $\omega, \rho$  are tradeoff regularization parameters. In section III-B, we have analyzed the optimal low-rank property of the HSIs. Here we do not consider the sparse error  $\mathcal{E}$  and replace the  $P_h(\mathcal{X})$  with the joint spectral and non-local low-rank tensor prior as follow:

$$\begin{aligned} \{\hat{\mathcal{X}}, \hat{\mathcal{L}}_i^j\} = \arg \min_{\mathcal{X}, \mathcal{L}_i^j} & \frac{1}{2} \|\mathcal{T}(\mathcal{X}) - \mathcal{Y}\|_F^2 \\ & + \omega_j \sum_j \sum_i \left( \frac{1}{\lambda_i^2} \|\mathcal{R}_i^j \mathcal{X} - \mathcal{L}_i^j\|_F^2 + \text{rank}_j(\mathcal{L}_i^j) \right), \end{aligned} \quad (3)$$

where  $\mathcal{R}_i^j \mathcal{X}$  represents the constructed low-rank tensor for each exemplar cubic  $i$  along  $j$ -mode,  $\mathcal{L}_i^j$  is its low-rank approximation,  $i$  represents the location index of the sliding window and  $j \in \{2, 3\}$  denotes the mode along the non-local and spectral dimension.  $\omega_j$  and  $\lambda_i$  the regularization parameters, in which  $\omega_j$  is fixed for the balance between rank constraint along each mode, and  $\lambda_i$  is dynamically changed according to the noise level at different location and iteration. The basic idea of the model is that the intrinsic subspace of the non-local self-similarity and spectral correlation can be well depicted by the joint low-rank tensor prior, meanwhile the first term reflects the measurement of the linear degradation.

### C. Optimization

Due to the difficulty of estimating multiple variables directly, we adopt the alternating minimization scheme to solve the objective functional (3) with respect to the whole image  $\mathcal{X}$  and low-rank tensor  $\mathcal{L}_i^j$  per each location.

1) *Low-Rank Tensor Estimation  $\mathcal{L}_i^j$* : In this subproblem, we fix the other variable  $\mathcal{X}$  and optimize the  $\mathcal{L}_i^j$  by with its tensor unfolding formation

$$\hat{\mathcal{L}}_i^j = \arg \min_{\mathcal{L}_i^{(j)}} \frac{1}{\lambda_i^2} \|\mathcal{R}_i \mathcal{X}^{(j)} - \mathcal{L}_i^{(j)}\|_F^2 + \|\mathcal{L}_i^{(j)}\|_*, \quad (4)$$

where  $\mathcal{R}_i \mathcal{X}^{(j)}$  corresponds to the matrix of the unfolding tensor  $\mathcal{R}_i^j \mathcal{X}$  along the mode- $j$ ,  $\|\mathcal{L}_i^{(j)}\|_*$  means the matrix nuclear norm to replace  $\text{rank}(\mathcal{L}_i^{(j)})$  as its convex surrogate function. Equation (4) is a typical low-rank matrix approximation problem which has a closed-form solution and can be easily solved by the singular values thresholding algorithm [68]. In our implementation, we borrow the idea of the reweighting strategy from [69] to improve the performance. After each  $\mathcal{L}_i^{(j)}$  is obtained, the tensor folding is performed to transform them into 3-order tensors.

**Algorithm 1** The optimal low-rank tensor for HSIs restoration**Require:** Input image  $\mathcal{Y}$ 

- 1: **Initialize:**
- 2: • Set parameters  $\tau, \beta, \omega_j$  and the noise level  $\lambda_i^2$ ;
- 3: • Set  $\mathbf{J}^{(1)} = 0$ ,  $\mathbf{X}^{(1)} = \mathcal{Y}$ ;
- 4: • Similar cubics grouping to form the low-rank tensor;
- 5: **for**  $l=1:L$  **do**
- 6:   **Low-rank Tensor Estimation:** obtain  $\mathcal{L}_i^{(j)}$  by Eq. (4);
- 7:   **Image Restoration:** compute  $\mathbf{X}^{(l+1)}$  via Eq. (8);
- 8:   **Auxiliary Variable:** update  $\mathcal{Z}^{(l+1)}$  via Eq. (9);
- 9:   **Lagrange Multiplier:** update  $\mathcal{J}^{(l+1)}$  via Eq. (7c);
- 10:   If  $\text{mod}(l, 10) = 0$ , update cubic grouping;
- 11: **end for**

**Ensure:** Clean Image  $\mathbf{X}$ .

2) *Image Restoration*  $\mathbf{X}$ : By ignoring terms independent of  $\mathbf{X}$  in (3), we obtain following subproblem:

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X}} \frac{1}{2} \|\mathcal{T}(\mathbf{X}) - \mathcal{Y}\|_F^2 + \omega_j \sum_j \sum_i \frac{1}{\lambda_i^2} \|\mathcal{R}_i^j \mathbf{X} - \mathcal{L}_i^j\|_F^2. \quad (5)$$

Generally, Eq. (5) is a quadratic optimization and can be solved by Gauss-Seidel algorithm. However, it is time consuming and the transformation of the 3D tensor HSI into the vector would damage the spatial-spectral structural correlation.

Fortunately, the first term in Eq. (5) is an reflection of the linear measurement of different degradations  $\mathcal{T}$ . It is worth noting that the computation is efficient in frequency domain than that of the image domain [70], which is a common practice in 2D image restoration. The theoretical basis is based on the Plancherel's theorem [71], which states that the sum of the square of a function equals the sum of the square of its Fourier transform. The energy equivalence between image and Fourier domain can be built for all possible values of  $\mathbf{X}$ .

The second term in Eq. (5) is a low-rank constraint on each cubic. Most of the sparse and low-rank methods are patch/cubic based by the operator  $\mathcal{R}$ , which is an linear operator that extracts the subcubic from the HSIs. Fortunately, the operator  $\mathcal{R}$  can be also solved on a pixel-by-pixel manner with fast speed. Such a gap between the first and second term inspires us to separate the subcubic extraction operator  $\mathcal{R}$  from the linear operator  $\mathcal{T}$  so that they can be rapidly computed in Fourier transform domain and pixel-wise domain, respectively.

The main idea of the ADMM is to convert the unconstrained minimization problem on  $\mathbf{X}$  in Eq. (5) into a constrained one by introducing a auxiliary variable, such that we could decouple the Eq. (5) into two easier subproblems with closed-form solution. Thus, we introduce an auxiliary variable  $\mathcal{Z}$ , by applying ADMM to (5), we obtain

$$\begin{aligned} \{\hat{\mathbf{X}}, \hat{\mathcal{Z}}\} &= \arg \min_{\mathbf{X}, \mathcal{Z}} \frac{1}{2} \|\mathcal{T}(\mathbf{X}) - \mathcal{Y}\|_F^2 \\ &\quad + \omega_j \sum_j \sum_i \frac{1}{\lambda_i^2} \|\mathcal{R}_i^j \mathcal{Z} - \mathcal{L}_i^j\|_F^2 + \frac{\beta}{2} \|\mathcal{Z} - \mathbf{X} - \frac{\mathcal{J}}{\beta}\|_F^2, \end{aligned} \quad (6)$$

where  $\mathcal{Z} \in \mathbb{R}^{R \times C \times B}$  is an auxiliary variable,  $\mathcal{J}$  is the Lagrangian multiplier,  $\beta$  and is a positive scalar. The optimization

**Algorithm 2** The tensor RPCA model for HSIs desstriping**Require:** Input image  $\mathcal{Y}$ 

- 1: **Initialize:**
- 2: • Set parameters  $\rho, \omega_j$  and the noise level  $\lambda_i^2$ ;
- 3: • Similar cubics grouping to form the low-rank tensor;
- 4: **for**  $n=1:N$  **do**
- 5:   **Image Restoration:** compute  $\mathbf{X}$  via Eq. (11a);
- 6:   **Low-rank Approximation:** solve Eq. (11b) for  $\mathcal{L}_i^{(j)}$ ;
- 7:   **Sparse Error Estimation:** solve Eq. (11c) for  $\mathcal{E}$ ;
- 8:   If  $\text{mod}(n, 10) = 0$ , update cubic grouping;
- 9: **end for**

**Ensure:** Clean Image  $\mathbf{X}$  and stripe component  $\mathcal{E}$ .

of (6) consists of the following iterations:

$$\mathbf{X}^{(l+1)} = \arg \min_{\mathbf{X}} \frac{1}{2} \|\mathcal{T}(\mathbf{X}) - \mathcal{Y}\|_F^2 + \frac{\beta}{2} \|\mathcal{Z}^{(l)} - \mathbf{X} - \frac{\mathcal{J}^{(l)}}{\beta}\|_F^2 \quad (7a)$$

$$\mathcal{Z}^{(l+1)} = \arg \min_{\mathcal{Z}} \omega_j \sum_j \sum_i \frac{1}{\lambda_i^2} \|\mathcal{R}_i^j \mathcal{Z} - \mathcal{L}_i^j\|_F^2 + \frac{\beta}{2} \|\mathcal{Z} - \mathbf{X}^{(l+1)} - \frac{\mathcal{J}^{(l)}}{\beta}\|_F^2 \quad (7b)$$

$$\mathcal{J}^{(l+1)} = \mathcal{J}^{(l+1)} + \beta(\mathbf{X}^{(l+1)} - \mathcal{Z}^{(l+1)}) \quad (7c)$$

$$\beta^{(l+1)} = \tau \beta^{(l)}, \quad (7d)$$

where  $\tau > 1$  is a constant. Thus the variables  $\mathbf{X}$  and  $\mathcal{Z}$  can be solved with closed-form solution efficiently:

$$\mathbf{X}^{(l+1)} = \mathcal{F}^{-1} \left( \frac{\mathcal{F}(\mathcal{T}^T(\mathcal{Y}) + (\beta^{(l)} \mathcal{Z}^{(l)} - \mathcal{J}^{(l)}))}{\mathcal{F}^*(\mathcal{T}) \circ \mathcal{F}(\mathcal{T}) + \beta^{(l)} \mathcal{I}} \right), \quad (8)$$

$$\begin{aligned} \mathcal{Z}^{(l+1)} &= \left( \frac{2\omega_j}{\lambda_i^2} \sum_j \sum_i (\mathcal{R}_i^j)^T \mathcal{R}_i^j + \beta^{(l)} \mathcal{I} \right)^{-1} \\ &\quad \times \left( \frac{2\omega_j}{\lambda_i^2} \sum_j \sum_i (\mathcal{R}_i^j)^T \mathcal{L}_i^j + \beta^{(l)} \mathbf{X}^{(l+1)} + \mathcal{J}^{(l)} \right), \end{aligned} \quad (9)$$

where  $\mathcal{F}(\bullet)$  denotes the  $n$ -D fast Fourier transform,  $\mathcal{F}^*(\bullet)$  its conjugate and  $\mathcal{F}^{-1}(\bullet)$  the inverse transform,  $\mathcal{I}$  is the identity tensor, superscript  $T$  is the transpose operator,  $(\mathcal{R}_i^j)^T \mathcal{R}_i^j$  means the number of overlapping cubics that cover the pixel location  $i$  and along  $j$ -mode, and  $(\mathcal{R}_i^j)^T \mathcal{L}_i^j$  means the sum value of all overlapping reconstruction cubics that cover the pixel location  $i$  and along  $j$ -mode. Thus, Eq. (8) can be computed via 3-D fast Fourier transforms efficiently and Eq. (9) can be fast computed in pixel-to-pixel tensor division. The overall procedure is summarized in **Algorithm 1**.

**D. Extension to Tensor RPCA for HSIs Desstriping**

In HSIs, there always exists the stripe noise coexisting with the random noise [72]. In recent years, the stripe noise removal has received more and more attention. For [13], [73], this kind of methods hold the point that the stripe line is an structure noise, and introduce the mixture of Gaussians (MoG) noise assumption also its variants to adapt real mixed noise characteristics. Another research direction starts from the image decomposition perspective [17], [19], in which the stripe noise is regarded as an structural line pattern component, equally treated with the image component. Our method follows the image decomposition manner. Previous methods usually model the stripe noise via conventional  $L_1$  norm or low-rank matrix prior. In section III-C, we have analyzed the low-rank

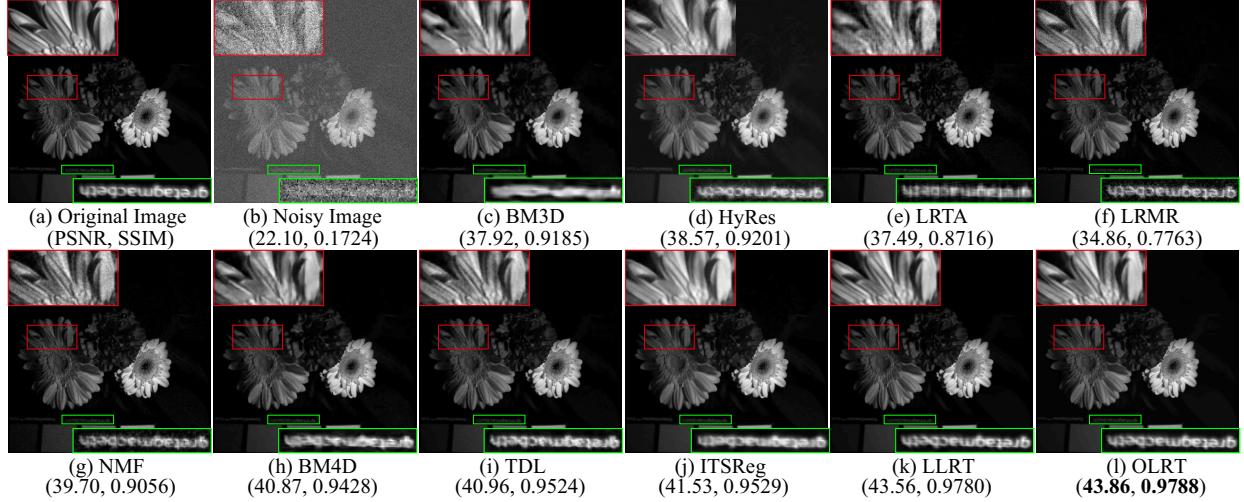


Figure 4. Simulated random noise removal results at 510nm band of image *Flower* under noise level  $\lambda^2=20$  on CAVE dataset. (a) Original image. (b) Noisy image, Denoising results by (c) BM3D, (d) HyRes, (e) LRTA, (f) LRMR, (g) NMF, (h) BM4D, (i) TDL, (j) ITSReg, (k) LLRT, (l) OLRT.

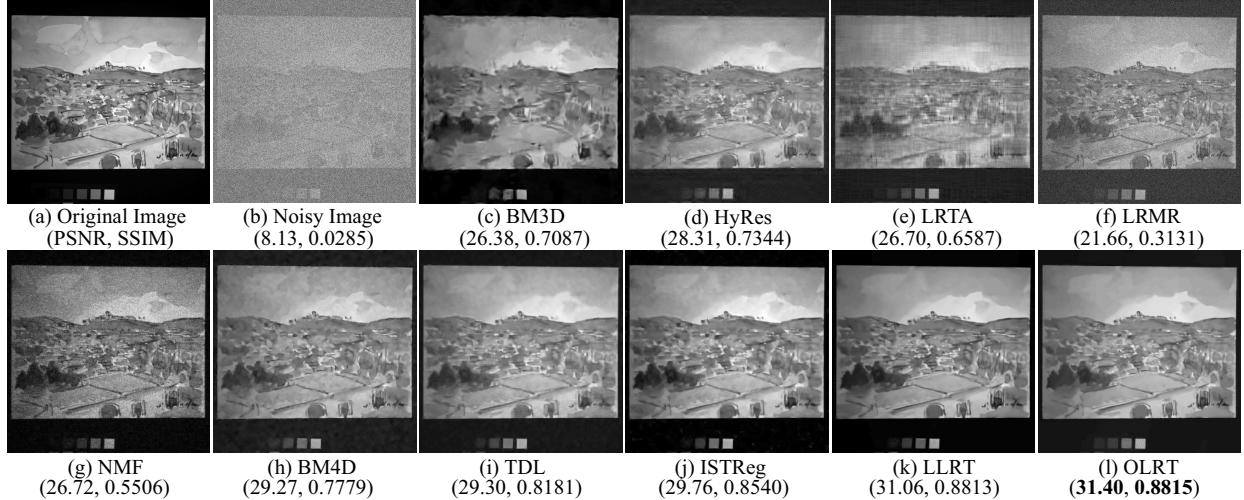


Figure 5. Simulated random noise removal results at 510nm band of image *Watercolor* under noise level  $\lambda^2=100$  on CAVE dataset. (a) Original image. (b) Noisy image, Denoising results by (c) BM3D, (d) HyRes, (e) LRTA, (f) LRMR, (g) NMF, (h) BM4D, (i) TDL, (j) ITSReg, (k) LLRT, (l) OLRT.

property of the stripe cubics. Thus, we propose a tensor RPCA for modeling both the HSIs and stripe components:

$$\left\{ \hat{\mathcal{X}}, \hat{\mathcal{L}}_i^j, \hat{\mathcal{E}} \right\} = \arg \min_{\mathcal{X}, \mathcal{L}_i^j, \mathcal{E}} \frac{1}{2} \|\mathcal{X} + \mathcal{E} - \mathcal{Y}\|_F^2 + \rho \text{rank}_1(\mathcal{E}) + \omega_j \sum_i \sum_j \left( \frac{1}{\lambda_i^2} \|\mathcal{R}_i^j \mathcal{X} - \mathcal{L}_i^j\|_F^2 + \text{rank}_j(\mathcal{L}_i^j) \right). \quad (10)$$

Here, we do not consider the linear operator  $\mathcal{T}$ . On one hand, the low-rank tensor priors effectively regularize the solution space for the HSIs and sparse error, respectively. On the other hand, better estimated HSIs will promote better sparse error estimation and vice versa. As for the optimization of (10), we still adopt the alternating minimization scheme:

$$\hat{\mathcal{X}} = \arg \min_{\mathcal{X}} \frac{1}{2} \|\mathcal{X} + \mathcal{E} - \mathcal{Y}\|_F^2 + \omega_j \sum_i \sum_j \frac{1}{\lambda_i^2} \|\mathcal{R}_i^j \mathcal{X} - \mathcal{L}_i^j\|_F^2 \quad (11a)$$

$$\hat{\mathcal{L}}_i^j = \arg \min_{\mathcal{X}, \mathcal{L}_i^j} \frac{1}{\lambda_i^2} \|\mathcal{R}_i^j \mathcal{X} - \mathcal{L}_i^j\|_F^2 + \text{rank}_j(\mathcal{L}_i^j) \quad (11b)$$

$$\hat{\mathcal{E}} = \arg \min_{\mathcal{E}} \frac{1}{2} \|\mathcal{X} + \mathcal{E} - \mathcal{Y}\|_F^2 + \rho \text{rank}_1(\mathcal{E}), \quad (11c)$$

where each subproblem admits the closed form solutions. The overall procedure is summarized in **Algorithm 2**.

## V. EXPERIMENTS AND DISCUSSION

### A. Experimental Setting

Our approach is tested on comprehensive classical HSIs restoration tasks to test the robustness and effectiveness of the proposed OLRT prior. For HSIs denoising, the competing methods include 1-D sparse representation based methods (SDS [74]), 2-D low-rank matrix recovery methods (LRMR [12], NMF [9], HyRes [16]), state-of-the-art tensor methods (BM3D [75], LRTA [76], BM4D [77], TDL [26], ITSReg [23], LLRT [28]). For HSIs deblurring, the competing methods include single image based deblurring method hyper-Laplacian (HL) [78], HSIs deblurring methods fast positive deconvolution (FPD) [5] and spectral-spatial total variation (SSTV) [79]; For HSIs inpainting, we compare with spectral-spatial

Table II  
QUANTITATIVE RESULTS OF DIFFERENT METHODS UNDER SEVERAL NOISE LEVELS ON CAVE DATASET.

$\lambda^2$	Index	Methods											
		Noisy	BM3D	SDS	LRTA	HyRes	LRMR	NMF	BM4D	TDL	ITSReg	LLRT	OLRT
5	PSNR	34.15	45.46	41.13	44.42	47.30	44.20	46.17	45.61	47.24	48.13	49.42	49.62
	SSIM	0.7496	0.9824	0.9691	0.9712	0.9875	0.9659	0.9873	0.9849	0.9887	0.9902	0.9925	0.9927
	ERGAS	112.65	30.53	54.64	34.92	24.84	36.71	28.68	30.40	24.95	21.56	19.09	18.57
	SAM	0.4823	0.1024	0.1590	0.1398	0.0970	0.2001	0.0975	0.1161	0.0845	0.0916	0.0728	0.0721
10	PSNR	28.13	42.09	39.74	41.36	41.75	39.27	43.15	44.59	44.30	45.77	46.67	47.07
	SSIM	0.4371	0.9665	0.9484	0.9499	0.9689	0.9094	0.9702	0.9784	0.9797	0.9802	0.9872	0.9877
	ERGAS	236.40	45.06	61.86	49.53	46.96	64.81	39.65	33.33	34.86	30.53	26.74	25.53
	SAM	0.7199	0.1395	0.2160	0.1719	0.1583	0.3343	0.1358	0.1295	0.1025	0.1086	0.0841	0.0840
30	PSNR	18.59	36.40	32.10	36.15	36.36	31.36	36.53	38.90	39.03	40.51	41.55	41.78
	SSIM	0.0988	0.9034	0.6709	0.8787	0.9133	0.6451	0.8565	0.9277	0.9486	0.9488	0.9683	0.9685
	ERGAS	709.29	88.29	145.88	91.40	86.32	157.65	86.25	65.38	63.54	53.05	48.20	46.70
	SAM	1.0414	0.2489	0.5050	0.2479	0.2501	0.6021	0.2465	0.2598	0.1520	0.1374	0.1192	0.1180
50	PSNR	14.15	32.66	25.32	32.44	33.85	26.67	31.98	35.96	36.42	37.75	38.93	39.13
	SSIM	0.0432	0.8320	0.3451	0.7932	0.8646	0.4000	0.7113	0.8666	0.9175	0.9271	0.9521	0.9484
	ERGAS	1181.95	115.06	280.88	118.64	114.73	264.28	123.23	91.51	85.58	70.16	65.52	63.46
	SAM	1.1741	0.2877	0.7006	0.2843	0.3189	0.7534	0.3148	0.3575	0.2000	0.1619	0.1424	0.1457
100	PSNR	8.13	29.27	17.90	29.20	30.61	20.84	26.95	30.82	32.91	33.01	35.40	35.41
	SSIM	0.0122	0.7460	0.1047	0.6945	0.7676	0.1850	0.4643	0.6956	0.8344	0.8648	0.9143	0.9022
	ERGAS	2364.05	171.94	693.94	175.91	165.68	469.26	225.55	141.18	128.22	120.77	98.91	97.82
	SAM	1.3271	0.3938	0.9690	0.3381	0.4639	0.9306	0.4321	0.5014	0.3079	0.2376	0.1895	0.2085

total variational based inpainting (SSTVI) [79], weighted nuclear norm minimization (WNNM) [80], fast hyperspectral inpainting (FastHyIn) [21] and tensor based AWTC [29]. All codes are provided by the authors and the parameters are fine-tuned by default or following the rules in their papers to achieve the best performance. The Matlab code of proposed method can be downloaded at the author's homepage<sup>1</sup>.

The spatial and spectral quality of the denoising results are very important for the subsequent processing, but is difficult to judge visually. In order to give an overall evaluation, four quantitative indices are employed: peak signal-to-noise ratio (PSNR), structure similarity (SSIM), erreur relative globale adimensionnelle de synthese (ERGAS) [81] and spectral angle map (SAM) [82]. PSNR and SSIM are two spatial-based indexes, while ERGAS and SAM are spectral-based indexes. The bigger PSNR and SSIM values are, the smaller ERGAS and SAM values are, the better the restored images are.

We evaluate the competing methods on several representative datasets: Columbia Multispectral Database (CAVE), Washington DC, Compact High Resolution Imaging Spectrometer (CHRIS), HYDICE Urban, PaviaU, Cuprite, and Airborne Visible/Infrared Imaging Spectrometer Dataset (AVIRIS). The CAVE, Washington DC, PaviaU and Cuprite are used for simulated experiments, while the Urban, CHRIS and AVIRIS are used to test the real cases.

### B. Comparison with State-of-The-Arts

1) *HSIs Denoising*: Zero mean additive white Gaussian noises with difference variance are added to generate the noisy observations. The visual results of single band in CAVE Flower under noise level 20 are shown in Figs. 4. Compared with other methods, the OLRT exhibits more clear details in texture regions or edges, meanwhile produce clean results in smooth regions with best quantitative values. Moreover, compared with our conference work LLRT [28], the OLRT is visually more pleasant with higher quantitative values.

The overall quantitative assessment results by the competing denoising methods under different noise levels are shown in Table II. The LLRT and OLRT outperforms the state-of-the-arts by a large margin. Moreover, the OLRT achieves the best performance in most of the quantitative assessments, especially for the low level noise cases.

2) *HSIs Deblurring*: In this section, we compare the proposed method with the state-of-the-art HSIs deblurring methods. We apply the 3D convolution (Matlab function *fftn* in frequency domain) to obtain the blur HSIs. To validate the robustness of our method, we test three kinds of blur cases: light Gaussian blur, heavy Gaussian blur, and uniform blur. The quantitative results on CAVE dataset are reported in Table III. Here we have three observations. First, our OLRT has overwhelming advantage over the competing methods. That is to say, our optimal low-rank tensor prior reliably reflects the intrinsic structural correlation of the HSIs. Second, compared with the single image based HL [78], our method additionally utilizes the low-rank property in the spectrum which greatly boosts to the restoration result. Third, the proposed method is robust to various blur levels and blur cases.

In Fig. 6, we give the visual deblurring results of the competing methods. We can observe that the OLRT could well recover the detailed structures such as the hair and the text in the zoomed regions, while the compared methods fail to achieve this under all circumstance. For different blurs, our OLRT consistently obtains very satisfactory deblurring results which validates the robustness of our low-rank prior.

3) *HSIs Inpainting*: HSIs inpainting refers to the problem of recovering a clean HSI from only partial observation of its entries. We simulate three kinds of missing HSIs: i.e., the random mask with 20% missing entries, the random mask with 80% missing entries, and the 50% deadlines in the HSIs. The quantitative results on CAVE dataset are reported in Table IV. The FastHyIn could only work when the missing mask is relative small, since the FastHyIn need to guarantee some pixels intact for subspace estimation. Therefore, we only report the case 20% missing entries for FastHyIn in Table IV.

<sup>1</sup><https://owuchangyu.github.io/>

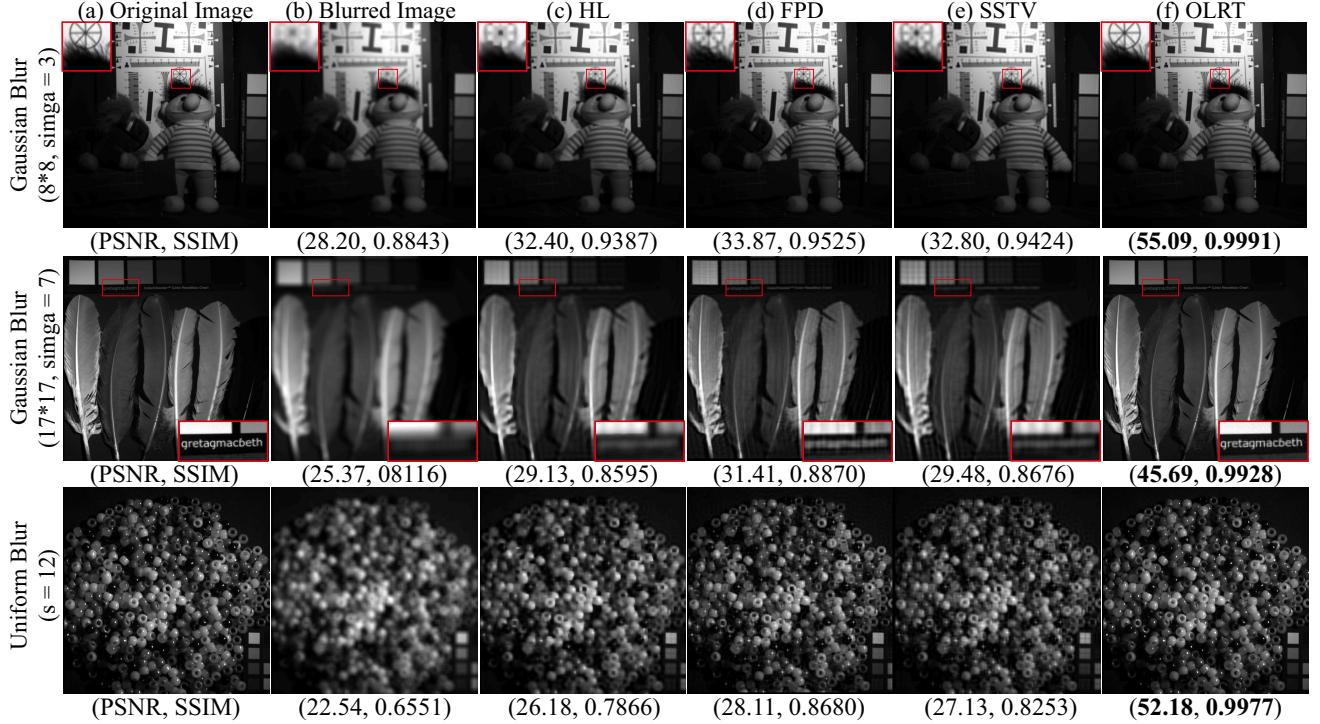


Figure 6. Simulated deblurring results under different blur level and blur case on CAVE dataset. The first row shows the light Gaussian blur on *Toy* ( $8 \times 8$ , Sigma = 3), the second row shows the heavy Gaussian blur on *feather* ( $17 \times 17$ , Sigma = 7), the third row shows the Uniform blur on *beads* ( $s = 12$ ). From the left to right columns (a) Original image at band 700nm. (b) Blurred image, Deblurring results by (c) HL, (d) FPD, (e) SSTV, (f) OLRT.

Table III  
QUANTITATIVE RESULTS OF DIFFERENT METHODS UNDER SEVERAL BLUR CASES ON CAVE DATASET.

Method	Gaussian Blur (8x8, Sigma = 3)				Gaussian Blur (17x17, Sigma = 7)				Uniform Blur (s = 12)			
	PSNR	SSIM	ERGAS	SAM	PSNR	SSIM	ERGAS	SAM	PSNR	SSIM	ERGAS	SAM
Blurred	32.61	0.9125	135.80	0.0736	28.69	0.8428	206.94	0.1020	29.62	0.8588	187.50	0.0924
HL	37.28	0.9460	83.88	0.0676	32.59	0.8819	137.14	0.1075	35.11	0.9163	104.82	0.0887
FPD	38.84	0.9617	68.48	0.0734	33.16	0.9114	125.11	0.1163	36.16	0.9467	89.65	0.0957
SSTV	37.61	0.9527	80.91	0.0658	33.08	0.8944	129.84	0.0989	35.73	0.9262	97.69	0.0844
OLRT	<b>57.02</b>	<b>0.9984</b>	<b>8.44</b>	<b>0.0224</b>	<b>50.45</b>	<b>0.9937</b>	<b>18.26</b>	<b>0.0387</b>	<b>57.84</b>	<b>0.9985</b>	<b>7.64</b>	<b>0.0228</b>

We can observe that under different missing conditions the proposed OLRT significantly outperforms the state-of-the-art HSIs inpainting methods by a large margin. It is interestingly noted that for light missing condition (20% missing) the spatially non-local based WNNM even works better than that of the spatial-spectral total variational SSTV. This partially validates that the non-local self-similarity is a key property for both the single image and hyperspectral images.

The visual comparison results are shown in Fig. 7. We can see that even for the 80% random missing entries, where the information is totally overwhelmed by the dead pixels, the proposed OLRT could still well recover clear edge and texture information with the aid of the joint spatial and spectral redundancy. Moreover, in HSIs, due to the malfunction of the sensor, there always exist the deadlines due to its push-broom imaging mechanism. We choose a remote sensing HSIs *PaviaU* and test this problem as shown in the third row. We can observe that OLRT could satisfactorily inpaint the deadlines both in terms of the quantitative and qualitative assessment.

4) *HSIs Destriping*: In HSIs, there always exists line pattern stripe noise, due to the push-broom mechanism of the multi-detector imaging systems. To model the stripe noise, we propose a tensor RPCA method for both the HSIs and stripe component. In real HSIs scenes, the stripe and random noise usually coexist which is the most common multiple degenerations. Thus, we simulate the non-periodical stripe noise along with the random noise. We choose the *Washington DC* with the size 256\*256\*31. The quantitative results are shown in Table V. We can observe that the OLRT obtains the best results under different noise levels. With the increase noise level, the advantages of proposed method over the competing methods becomes bigger in term of both the spatial and spectral assessment. That is to say, our result is robust to the stripe noise compared with the competing methods, which strongly support the effectiveness of our low-rank prior for the stripe noise. In Fig. 8, we show the destriping results of a remote sensing scene cuprite and *Washington DC*. It is easily to be seen that the competing method fail to remove the stripe

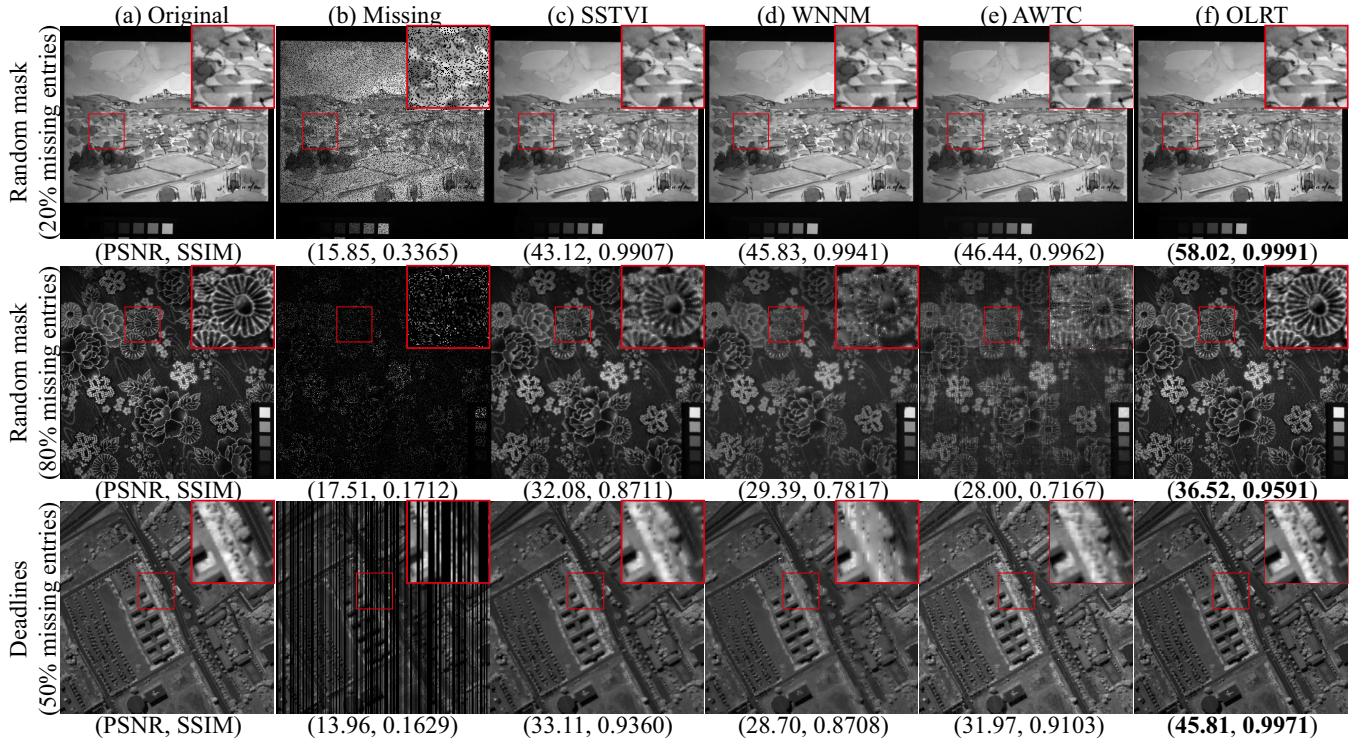


Figure 7. Simulated inpainting results under different missing level and case on CAVE and *PaviaU* dataset. The first row shows the light random mask on *watercolor* (20% missing entries), the second row shows the heavy random mask on *cloth* (80% missing entries), the third row shows the deadlines on *PaviaU* (50% missing entries). From the left to right columns (a) Original image at band 510nm. (b) Missing image, Inpainting results by (c) SSTVI, (d) WNNM, (e) AWTC, (f) OLRT.

Table IV  
QUANTITATIVE RESULTS OF DIFFERENT METHODS UNDER SEVERAL MISSING INFORMATION CASES ON CAVE DATASET.

Method	Random mask (20% missing entries)				Random mask (80% missing entries)				Deadline (50% missing entries)			
	PSNR	SSIM	ERGAS	SAM	PSNR	SSIM	ERGAS	SAM	PSNR	SSIM	ERGAS	SAM
Missing	23.24	0.6726	380.63	0.4412	17.22	0.3703	761.09	1.1261	19.28	0.5256	601.58	0.7847
WNNM	51.26	0.9974	18.37	0.0315	37.00	0.9599	88.29	0.0897	31.04	0.9705	194.76	0.0998
SSTVI	46.93	0.9947	27.64	0.0449	38.24	0.9674	73.94	0.0936	40.10	0.9706	68.24	0.0815
FastHyIn	49.62	0.9945	22.91	0.0638	—	—	—	—	—	—	—	—
AWTC	50.65	0.9962	18.83	0.0401	36.12	0.9247	100.17	0.1474	41.27	0.9741	64.37	0.0944
OLRT	<b>57.02</b>	<b>0.9992</b>	<b>9.15</b>	<b>0.0216</b>	<b>42.73</b>	<b>0.9899</b>	<b>46.25</b>	<b>0.0623</b>	<b>43.86</b>	<b>0.9940</b>	<b>61.85</b>	<b>0.0529</b>

noise, while the proposed method could well remove both the random and stripe noise simultaneously. It is worth noting that even the information has been overwhelmed by the noises, the proposed method could still recover structural edges clearly.

5) *Test on Real HSIs*: To demonstrate the robustness of our method, we show in Fig. 9 various HSIs restoration results. The first row shows the degraded HSIs and the second rows present the restoration results by OLRT. From the left to the right columns, we test the proposed methods on real random noise, missing pixels, stripe noise and the mixed noise. We can observe that the proposed method has consistently obtained visually pleasure results and obtains clean image with abundant edge structures.

### C. Analysis

1) *Analysis for Different Scenes and Bands*: In Fig. 10, from a macro viewpoint, we give a detailed comparison of all

HSIs scene in CAVE. In Fig. 11, from a micro viewpoint, we give a detailed analysis about how each band is recovered for different tasks. Here we have two observations. First, the OLRT consistently obtains the best result for different scenes and different tasks of each band, especially for the deblurring and inpainting tasks. Second, compared with LLRT, the OLRT could slightly better restore the HSIs, which verifies the effectiveness of the proposed method.

2) *Effectiveness of the OLRT*: In this section, we further validate the key observation of this work: the physical meaning and dimension along each mode are markedly different and the optimal low-rank tensor model is effective for HSIs restoration. Specifically, we transfer the OLRT strategy accordingly to the state-of-the-art work ITSReg [23] and NGM [34]. For ITSReg [23], we remove the spatial mode low-rank constraint, while for NGM [34] we additionally enforce the nonlocal-subspace regularized spectral low-rank constraint. Note that,

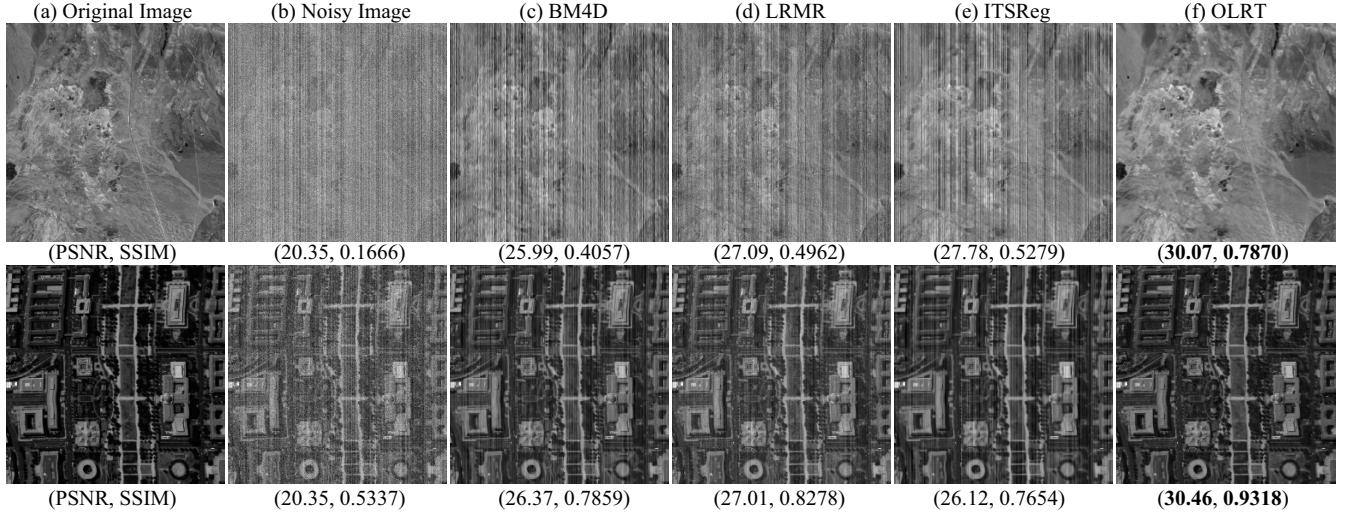


Figure 8. Simulated destriping results on *Washington DC* and *Curprite* dataset. The simulated noise level is random noise  $\lambda^2=20$  and stripe noise intensity = 20 and 50% region covered. (a) Original image. (b) Striping image, Destriping results by (c) BM4D, (d) LRMR, (e) ITSReg, (f) OLRT.

Table V  
QUANTITATIVE RESULTS OF DIFFERENT METHODS UNDER SEVERAL STRIPE LEVELS ON WASHINGTON DC.

Method	Random Noise( $\lambda^2=20$ )				Random Noise( $\lambda^2=30$ )				Random Noise( $\lambda^2=30$ )			
	Stripe Noise(Intensity = 20, 50% region)				Stripe Noise(Intensity = 30, 50% region)				Stripe Noise(Intensity = 30, 80% region)			
	PSNR	SSIM	ERGAS	SAM	PSNR	SSIM	ERGAS	SAM	PSNR	SSIM	ERGAS	SAM
Noisy	20.35	0.5347	324.83	0.4716	16.83	0.3626	487.39	0.6338	16.03	0.3252	534.05	0.6732
BM4D	26.37	0.7859	162.91	0.2443	22.93	0.6376	241.72	0.3479	20.08	0.4938	335.41	0.4693
LRMR	27.01	0.8278	151.98	0.2266	23.52	0.7030	226.58	0.3293	21.10	0.5858	298.53	0.4315
ITSReg	26.12	0.7655	167.51	0.2432	21.73	0.5688	277.64	0.3940	19.47	0.4495	359.98	0.4962
OLRT	<b>30.46</b>	<b>0.9318</b>	<b>120.51</b>	<b>0.1708</b>	<b>28.87</b>	<b>0.8945</b>	<b>136.87</b>	<b>0.1885</b>	<b>28.59</b>	<b>0.8915</b>	<b>138.98</b>	<b>0.1917</b>

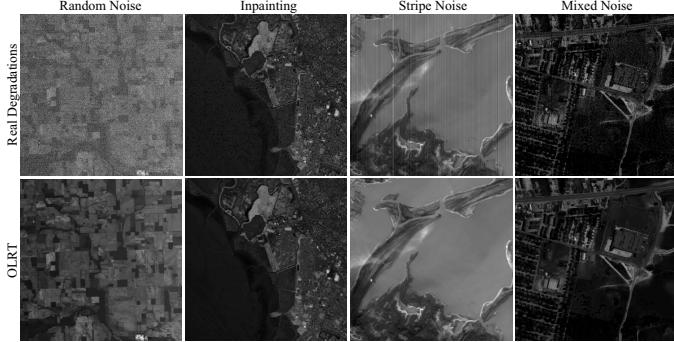


Figure 9. Real HSIs restoration results. The first to fourth column show the denoising (AVIRIS), inpainting (AVIRIS band 74), destriping (CHRIS band 5), and mixed noise removal (Urban band 204) results, respectively.

we show the comparison results on the test dataset of the released code with default noise level. Other noise levels and datasets also follow this observation. In Fig. 12, the first row shows the ITSReg denoising results with  $\lambda^2=0.1 * 255$  on *Balloons*. The second row presents the NGM denoising results with  $\lambda^2=20$  on *Stuffed\_toy*. We can observe that the improved results with OLRT strategy (third column) are better than that of the original results (second column) both quantitative and qualitative, especially for the fine texture and sharp edges. Such an interesting experiment have solidly

demonstrated the effectiveness of the proposed OLRT model.

3) *Regularization Parameter Analysis*: There are two main regularization parameters  $\omega_j$  ( $j \in \{2, 3\}$ ) enforcing the low-rank constraint along non-local mode and spectral mode, respectively. Figure 13 shows the change of the PSNR and SSIM values with the change of the parameters  $\omega_2$  and  $\omega_3$ . It is shown that the denoising results depends on the choice of  $\omega_2$ , which controls the strength of the non-local low-rank constraint. When the  $\omega_2$  is relative higher, the performance is consistently good; when  $\omega_2$  tends to be small, the performance decreases rapidly. This also validates the importance of the non-local self-similarity in HSIs restoration. The PSNR/SSIM value do not change means that the proposed method has converged to a very good minimal solution. That is to say, the choice of the regularization parameters are very robust in a certain range. In our work, we empirically find that when the  $\omega_2 \geq \omega_3$ , the restoration performance is good. In practice, we always set the  $\omega_2 = 0.5, \omega_3 = 0.5$ . It is worth noting that the noise level  $\lambda_i$  is automatically changed according to the noise level at different location and iteration. In this work, we determine the noise level by calculating the residual between each two iteration.

4) *Maximum Missing Analysis*: To demonstrate the effectiveness of the proposed method, we perform the pressure test by increasing the missing proportion, as shown in Fig. 14. For the random mask, we add two heavy missing masks on *cloth* with 90% missing entries and 99% missing entries, respec-

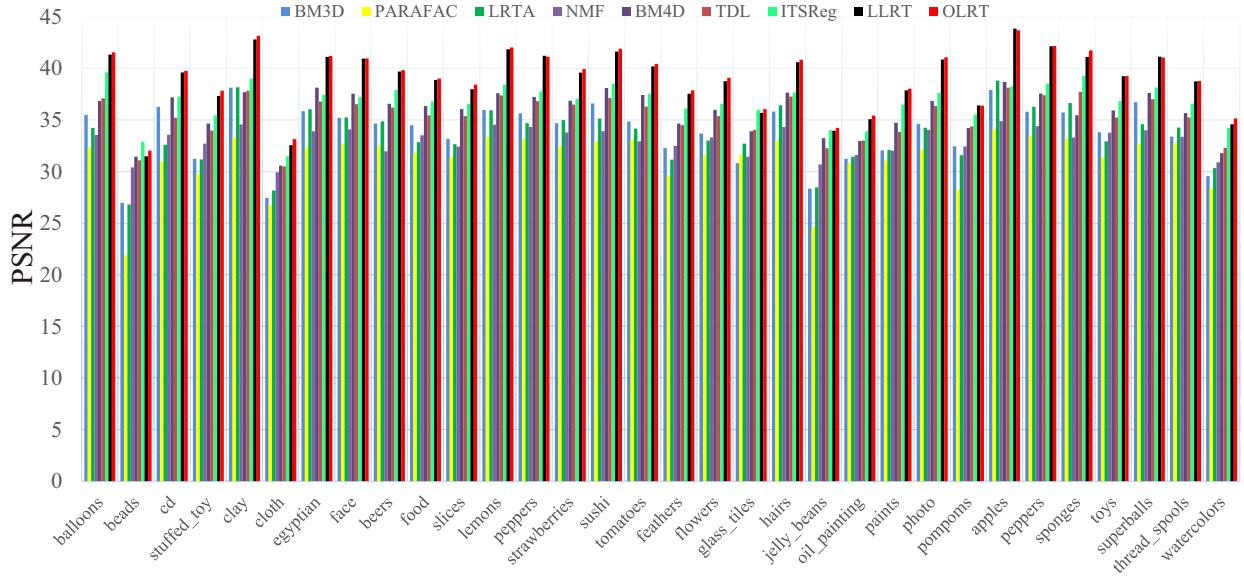


Figure 10. The quantitative PSNR value analysis of all scenes in CAVE under noise level  $\lambda^2=50$ . The horizontal axis denotes the number of HSIs scene, and the vertical axis means the PSNR values.

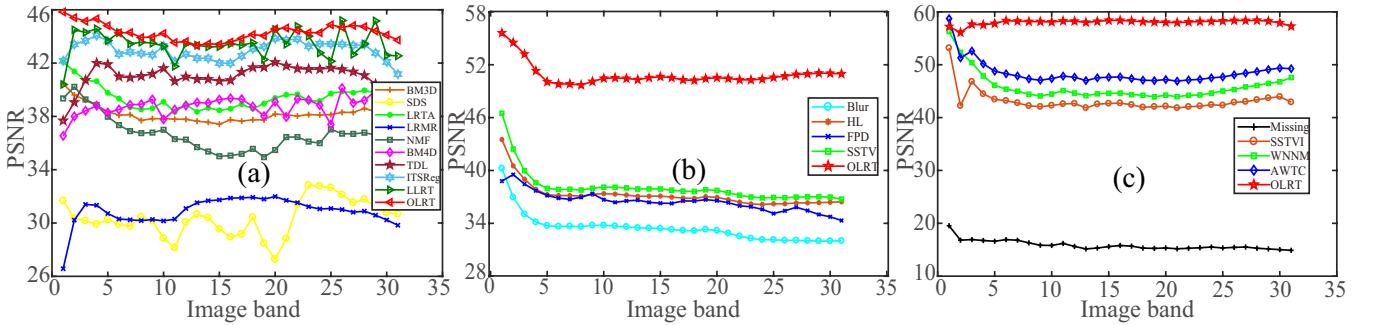


Figure 11. The spectral analysis of OLRT under different restoration tasks. The horizontal axis denotes the number of band and the vertical axis means the PSNR values. (a)  $\lambda^2=30$  denoising, (b) Gaussian blur (17\*17, Sigma = 7) deblurring, (c) 20% missing entries inpainting. The OLRT consistently obtains the best result for different tasks of each band.

tively. For the deadlines, we add two heavy missing masks on *PaviaU* with 90% missing entries and 95% missing entries, respectively. We still use the same images in Fig. 7 for better comparison. In such an extremely hard situation, the proposed OLRT could still restore the image with sharp edge yet less details. Note that, the width of most concentrated deadline is almost 70 column in the last column. The performance of OLRT is still visually acceptable with sharp edge. It is worth noting that our restoration task is based on the hyperspectral images. The amazing spatial visual appearance greatly benefits from the spectral information. Imagining that for a pixel, along the spectral axis, the pixel may be missing in some bands; while for the other bands it may be still be intact. Thus, the spectral correlation could provide redundant information for us to reconstruct the missing pixels. In summary, the non-local self-similarity plays a decisive role, and the additional spectral high correlation further boosts the restoration result.

5) *Number of the Bands*: To validate the effectiveness of the bands, we perform the experiment to analyze the relationship between the restoration performance and number of the

bands. Here, we take the denoising as examples. For denoising, we degrade the *Flower* in CAVE with  $\lambda^2 = 20$ , and gradually increase the number of bands as the input. The quantitative results are shown in Fig. 15. We have two observations. First, we can observe that the PSNR and SSIM value (the larger, the better) increase rapidly in the first 20 bands. Then, the PSNR and SSIM (mean of each band) gradually converge to a stable value. The same phenomenon has also been observed for the SAM and ERGAS. That is to say, the increase number of bands greatly benefits the OLRT for better restoration performance. Second, it is worth noting that when the image band equals to 1, the OLRT degenerates to the single image based method, and the OLRT could obtain PSNR = 38.34dB, SSIM = 0.9574, SAM = 0.0069, ERGAS = 83.46 for image denoising, in which the quantitative results are very similar to single image based denoising method BM3D (refer to Fig. 4). It is noticeable, although the restoration performance increase rapidly, the price of the growth is the increasing running time.

6) *Extension to the Color Image*: Although OLRT is proposed for HSIs which possess dozens or hundreds of contin-

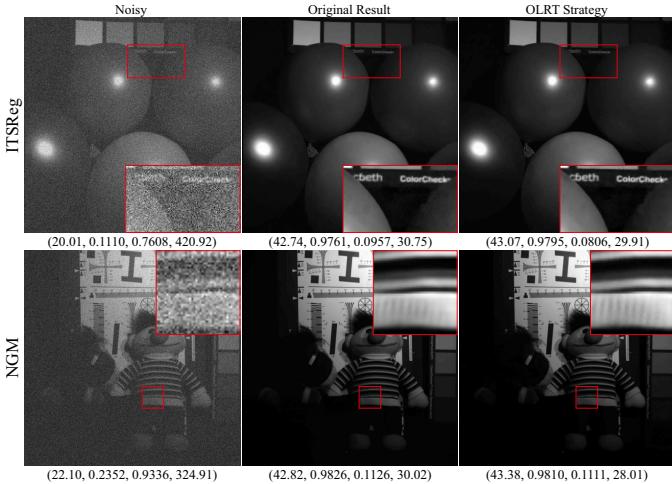


Figure 12. The effectiveness of OLRT strategy. We validate the OLRT model on the STOA HSIs restoration methods ITSReg [23] and NGM [34] by modifying their models with optimal low-rank strategy. The numbers below the images denote the (PSRN, SSIM, SAM, ERGAS) respectively.

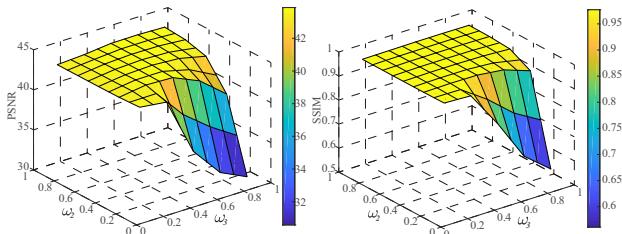


Figure 13. The regularization parameter robustness analysis. Change of the PSNR (left) and SSIM (right) values versus the parameters  $\omega_2$  and  $\omega_3$ .

uous bands, it can be extended to multispectral images with fewer bands, such as RGB color image. Here, we compare the proposed OLRT work with the SNN (LRTC) [83] and TRPCA [84] for color image processing. We choose the color image denoising task for fair comparison. We download the source code from the homepage of the author where the parameters are fine-tuned to achieve the best performance. The color image denoising results of image *castle* in BSD dataset are shown in Fig. 16. We can observe that the LRTC and TRPCA have suppressed the noise at the cost of blurry edges, while the proposed OLRT obtains smoother image with clearer texture. Overall, the OLRT has obtained better performance in terms of noise suppression and detail preserving.

7) *Convergence Analysis:* The convergence of the ADMM has been extensively discussed [85], [86]. In our work, both the two terms in Eq. (5) are always strongly convex and have Lipschitz continuous gradient. In such a situation, the convergence property of the ADMM can be well guaranteed with proper penalty parameters [86]. We also provide an empirical convergence analysis by taking the HSIs denoising as example. Figure 17 illustrates its evolutional curve of functional energy of Eq. (5) versus the iterations. We can observe that the functional energy curve monotonically decreases and the PSNR value curve monotonically increases, which verify the convergence property in practice.

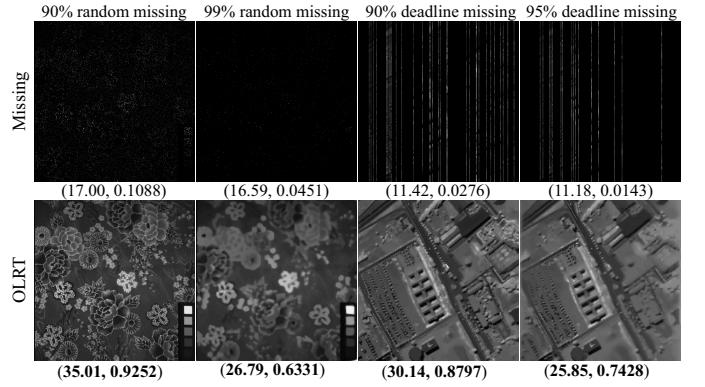


Figure 14. Robustness of the OLRT inpainting for extreme large missing proportion. Both the random mask and deadlines are tested.

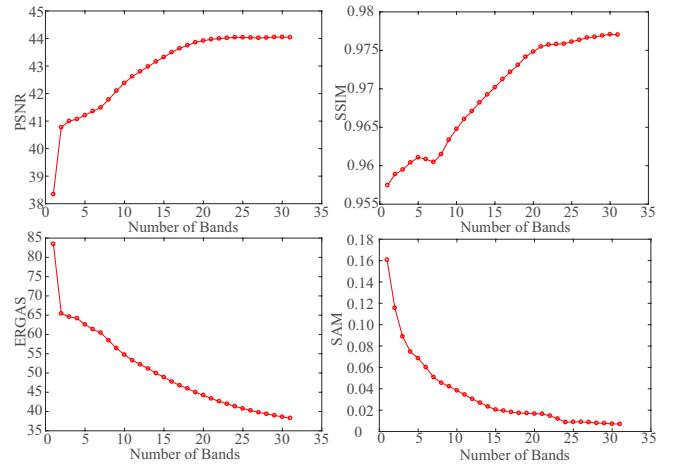


Figure 15. Influence of the numbers of bands. We show the changing curve of PSNR, SSIM, SAM, and ERGAS versus the number of the bands.

## VI. CONCLUSION

The low-rank based methods have been widely used in HSIs restoration tasks. However, it is still hard to tell which property is most beneficial for HSIs restoration, not to say how to model these properties in a more reasonable manner. In this work, we try to figure out these fundamental problems by giving a detailed analysis of the structure correlations in HSIs and an optimal combination for better restoration. Consequently, we propose an optimal low-rank tensor OLRT prior in which both the non-local and spectral low-rank property are taken into

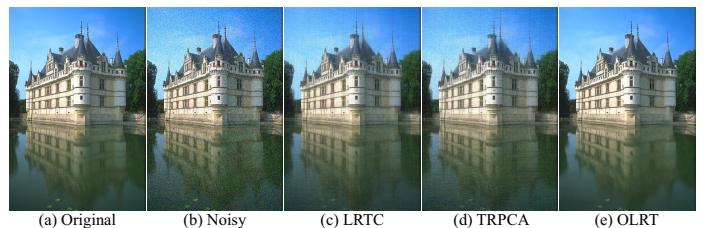


Figure 16. Color image denoising results under noise level  $\lambda^2 = 20$ . (a) original, (b) noisy, Denoising results by (c) LRTC, (d) TRPCA, (e) OLRT.

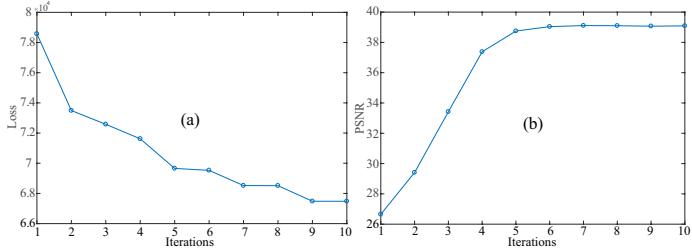


Figure 17. The empirical convergence curve of HSIs denoising result. (a) Functional energy curve, (b) PSNR value curve.

consideration simultaneously. Moreover, we further excavate the low-rank property in the sparse stripe error and extend the OLRT to the RPCA model for HSIs stripe removal. The proposed method has been extensively tested on extensive simulated and real experiment such as HSI denoising, destriping, deblurring, inpainting, and it consistently outperformed the competing state-of-the-art approaches in both quantitative assessments and visual appearance. The OLRT can also be well extended to three bands color image restoration. In the future, we would like to extend the proposed OLRT to other HSIs tasks such as reconstruction and super-resolution.

## REFERENCES

- [1] H. Othman and S.-E. Qian, "Noise reduction of hyperspectral imagery using hybrid spatial-spectral derivative-domain wavelet shrinkage," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp. 397–408, 2006.
- [2] H. Shen and L. Zhang, "A map-based algorithm for destriping and inpainting of remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1492–1502, 2009.
- [3] M. Bouali and S. Ladjal, "Toward optimal destriping of modis data using a unidirectional variational model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 8, pp. 2924–2935, Aug. 2011.
- [4] 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.
- [5] S. Henrot, C. Soussen, and D. Brie, "Fast positive deconvolution of hyperspectral images," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 828–833, 2013.
- [6] X.-L. Zhao, F. Wang, T.-Z. Huang, M. K. Ng, and R. J. Plemmons, "Deblurring and sparse unmixing for hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 4045–4058, 2013.
- [7] Q. Cheng, H. Shen, L. Zhang, and P. Li, "Inpainting for remotely sensed images with a multichannel nonlocal total variation model," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 175–187, 2013.
- [8] Y. Chang, L. Yan, H. Fang, and H. Liu, "Simultaneous destriping and denoising for remote sensing images with unidirectional total variation and sparse representation," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 6, pp. 1051–1055, Jun. 2014.
- [9] M. Ye, Y. Qian, and J. Zhou, "Multitask sparse nonnegative matrix factorization for joint spectral-spatial hyperspectral imagery denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2621–2639, 2015.
- [10] Y. Fu, A. Lam, I. Sato, and Y. Sato, "Adaptive spatial-spectral dictionary learning for hyperspectral image restoration," in *Proc. IEEE Conf. ICCV*, 2015, pp. 343–351.
- [11] T. Lu, S. Li, L. Fang, Y. Ma, and J. A. Benediktsson, "Spectral-spatial adaptive sparse representation for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 373–385, 2015.
- [12] 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.
- [13] X. Cao, Y. Chen, Q. Zhao, D. Meng, Y. Wang, D. Wang, and Z. Xu, "Low-rank matrix factorization under general mixture noise distributions," in *Proc. IEEE Conf. ICCV*, 2015, pp. 1493–1501.
- [14] Y. Xie, Y. Qu, D. Tao, W. Wu, Q. Yuan, and W. Zhang, "Hyperspectral image restoration via iteratively regularized weighted schatten-norm minimization," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4642–4659, 2016.
- [15] Y. Fu, Y. Zheng, I. Sato, and Y. Sato, "Exploiting spectral-spatial correlation for coded hyperspectral image restoration," in *Proc. IEEE Conf. CVPR*, 2016, pp. 3727–3736.
- [16] B. Rasti, M. O. Ulfarsson, and P. Ghamisi, "Automatic hyperspectral image restoration using sparse and low-rank modeling," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2335–2339, 2017.
- [17] X. Lu, Y. Wang, and Y. Yuan, "Graph-regularized low-rank representation for destriping of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 4009–4018, Jul. 2013.
- [18] Y.-Q. Zhao and J. Yang, "Hyperspectral image denoising via sparse representation and low-rank constraint," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 296–308, 2015.
- [19] 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, 2016.
- [20] 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.
- [21] 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 Observat. Remote Sens.*, vol. 11, no. 3, pp. 730–742, 2018.
- [22] Y. Hu, X. Li, Y. Gu, and M. Jacob, "Hyperspectral image recovery using nonconvex sparsity and low-rank regularizations," *IEEE Trans. Geosci. Remote Sens.*, 2019.
- [23] Q. Xie, Q. Zhao, D. Meng, Z. Xu, S. Gu, W. Zuo, and L. Zhang, "Multispectral images denoising by intrinsic tensor sparsity regularization," in *Proc. IEEE Conf. CVPR*, 2016, pp. 1692–1700.
- [24] J. Xue, Y. Zhao, W. Liao, and J. C.-W. Chan, "Nonlocal low-rank regularized tensor decomposition for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 5174–5189, 2019.
- [25] T. Xie, S. Li, L. Fang, and L. Liu, "Tensor completion via nonlocal low-rank regularization," *IEEE Transactions on Cybern.*, vol. 49, no. 6, pp. 2344–2354, 2018.
- [26] 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. CVPR*, 2014, pp. 2949–2956.
- [27] W. Dong, G. Li, G. Shi, X. Li, and Y. Ma, "Low-rank tensor approximation with laplacian scale mixture modeling for multiframe image denoising," in *Proc. IEEE Conf. ICCV*, 2015, pp. 442–449.
- [28] Y. Chang, L. Yan, and S. Zhong, "Hyper-laplacian regularized unidirectional low-rank tensor recovery for multispectral image denoising," in *Proc. IEEE Conf. CVPR*, 2017, pp. 5901–5909.
- [29] K. P. Ng, Q. Yuan, L. Yan, and J. Sun, "An adaptive weighted tensor completion method for the recovery of remote sensing images with missing data," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 6, pp. 3367–3381, 2017.
- [30] 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 Observat. Remote Sens.*, vol. 11, no. 4, pp. 1227–1243, 2017.
- [31] Y. Chen, T.-Z. Huang, and X.-L. Zhao, "Destriping of multispectral remote sensing image using low-rank tensor decomposition," *IEEE J. Sel. Topics Appl. Earth Observat. Remote Sens.*, vol. 11, no. 12, pp. 4950–4967, 2018.
- [32] T.-Y. Ji, N. Yokoya, X. X. Zhu, and T.-Z. Huang, "Nonlocal tensor completion for multitemporal remotely sensed images inpainting," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3047–3061, 2018.
- [33] S. Zhang, L. Wang, Y. Fu, X. Zhong, and H. Huang, "Computational hyperspectral imaging based on dimension-discriminative low-rank tensor recovery," in *Proc. IEEE Conf. ICCV*, 2019, pp. 10 183–10 192.
- [34] 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. CVPR*, 2019, pp. 6861–6870.
- [35] 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, 2020.
- [36] Y. Chen, W. He, N. Yokoya, and T.-Z. Huang, "Hyperspectral image restoration using weighted group sparsity-regularized low-rank tensor decomposition," *IEEE Transactions on Cybern.*, 2019.

- [37] Y. Chang, L. Yan, H. Fang, S. Zhong, and Z. Zhang, "Weighted low-rank tensor recovery for hyperspectral image restoration," *IEEE Transactions on Cybern.*, 2020.
- [38] X. Gong, W. Chen, and J. Chen, "A low-rank tensor dictionary learning method for hyperspectral image denoising," *IEEE Trans. Signal Process.*, vol. 68, pp. 1168–1180, 2020.
- [39] 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, 2018.
- [40] Z. Kong and X. Yang, "Color image and multispectral image denoising using block diagonal representation," *IEEE Trans. Image Process.*, vol. 28, no. 9, pp. 4247–4259, 2019.
- [41] 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, 2019.
- [42] Y. Chen, W. He, N. Yokoya, T.-Z. Huang, and X.-L. Zhao, "Nonlocal tensor-ring decomposition for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 2, pp. 1348–1362, 2020.
- [43] Y. Xu, Z. Wu, J. Chanussot, and Z. Wei, "Hyperspectral images super-resolution via learning high-order coupled tensor ring representation," *IEEE Trans. Neural Netw. Learn. Syst.*, 2020.
- [44] R. Dian, S. Li, and L. Fang, "Learning a low tensor-train rank representation for hyperspectral image super-resolution," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 9, pp. 2672–2683, 2019.
- [45] F. Xiong, J. Zhou, and Y. Qian, "Hyperspectral restoration via  $l_0$  gradient regularized low-rank tensor factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 10410–10425, 2019.
- [46] 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, 2019.
- [47] 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, 2019.
- [48] B. Lin, X. Tao, and J. Lu, "Hyperspectral image denoising via matrix factorization and deep prior regularization," *IEEE Trans. Image Process.*, vol. 29, pp. 565–578, 2019.
- [49] Q. Xie, M. Zhou, Q. Zhao, D. Meng, W. Zuo, and Z. Xu, "Multispectral and hyperspectral image fusion by ms/hs fusion net," in *Proc. IEEE Conf. CVPR*, 2019, pp. 1585–1594.
- [50] W. Dong, H. Wang, F. Wu, G. m. Shi, and X. Li, "Deep spatial-spectral representation learning for hyperspectral image denoising," *IEEE Transactions on Comput. Imag.*, vol. 5, no. 4, pp. 635–648, 2019.
- [51] A. Maffei, J. Haut, M. Paoletti, J. Plaza, L. Bruzzone, and A. Plaza, "A single model cnn for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2516–2529, 2020.
- [52] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *J. SIAM Rev.*, vol. 66, no. 4, pp. 294–310, 2009.
- [53] A. Rajwade, A. Rangarajan, and A. Banerjee, "Image denoising using the higher order singular value decomposition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 4, pp. 849–862, 2013.
- [54] Y. Chang, L. Yan, H. Fang, and C. Luo, "Anisotropic spectral-spatial total variation model for multispectral remote sensing image despeckling," *IEEE Trans. Image Process.*, vol. 24, no. 6, pp. 1852–1866, 2015.
- [55] 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.
- [56] 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, 2020.
- [57] H. Shen, X. Li, Q. Cheng, C. Zeng, G. Yang, H. Li, and L. Zhang, "Missing information reconstruction of remote sensing data: a technical review," *IEEE Geosc. Rem. Sen. M.*, vol. 3, no. 3, pp. 61–85, 2015.
- [58] N. Akhtar, F. Shafait, and A. Mian, "Sparse spatio-spectral representation for hyperspectral image super-resolution," in *Proc. of ECCV*, 2014, pp. 63–78.
- [59] H. Ye, H. Li, B. Yang, F. Cao, and Y. Tang, "A novel rank approximation method for mixture noise removal of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4457–4468, 2019.
- [60] T. Xie, S. Li, and B. Sun, "Hyperspectral images denoising via non-convex regularized low-rank and sparse matrix decomposition," *IEEE Trans. Image Process.*, vol. 29, pp. 44–56, 2019.
- [61] C. Li, Y. Ma, J. Huang, X. Mei, and J. Ma, "Hyperspectral image denoising using the robust low-rank tensor recovery," *J. Opt. Soc. Amer. A*, vol. 32, no. 9, pp. 1604–1612, 2015.
- [62] Y. Fu and W. Dong, "3d magnetic resonance image denoising using low-rank tensor approximation," *Neurocomputing*, vol. 195, pp. 30–39, 2016.
- [63] Q. Xie, Q. Zhao, D. Meng, and Z. Xu, "Kronecker-basis-representation based tensor sparsity and its application to tensor recovery," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 8, pp. 1888–1902, 2017.
- [64] J. Bioucas-Dias and J. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, 2008.
- [65] B. Rasti, M. Ulfarsson, and J. Sveinsson, "Hyperspectral subspace identification using sure," *IEEE Geosci. Remote Sens. Lett.*, vol. 46, no. 8, pp. 2481–2485, 2015.
- [66] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted  $\ell_1$  minimization," *J. Fourier Anal. Appl.*, vol. 14, no. 5–6, pp. 877–905, 2008.
- [67] L. D. Lathauwer, B. D. Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1253–1278, 2000.
- [68] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. on Optim.*, vol. 20, no. 4, pp. 1956–1982, Jan. 2010.
- [69] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *Proc. IEEE Conf. CVPR*, 2014, pp. 2862–2869.
- [70] W. Dong, L. Zhang, G. Shi, and X. Li, "Nonlocally centralized sparse representation for image restoration," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1620–1630, 2013.
- [71] R. N. Bracewell and R. N. Bracewell, *The Fourier transform and its applications*. McGraw-Hill New York, 1986.
- [72] 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, 2020.
- [73] Y. Chen, X. Cao, Q. Zhao, D. Meng, and Z. Xu, "Denoising hyperspectral image with non-i.i.d. noise structure," *IEEE Transactions on Cybern.*, vol. 48, no. 3, pp. 1054–1066, 2018.
- [74] A. Lam, I. Sato, and Y. Sato, "Denoising hyperspectral images using spectral domain statistics," in *Proc. ICPR*, 2012, pp. 477–480.
- [75] 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.
- [76] 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, 2008.
- [77] M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi, "Nonlocal transform-domain filter for volumetric data denoising and reconstruction," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 119–33, 2012.
- [78] D. Krishnan and R. Fergus, "Fast image deconvolution using hyper-laplacian priors," in *Proc. NIPS*, 2009, pp. 1033–1041.
- [79] H. Fang, C. Luo, G. Zhou, and X. Wang, "Hyperspectral image deconvolution with a spectral-spatial total variation regularization," *Can. J. Remote Sens.*, vol. 43, no. 6, 2017.
- [80] S. Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, and L. Zhang, "Weighted nuclear norm minimization and its applications to low level vision," *Int. J. Comput. Vis.*, vol. 121, no. 2, pp. 183–208, 2017.
- [81] L. Wald, *Data fusion: definitions and architectures: fusion of images of different spatial resolutions*. Presses des MINES, 2002.
- [82] R. H. Yuhas, J. W. Boardman, and A. F. Goetz, "Determination of semi-arid landscape endmembers and seasonal trends using convex geometry spectral unmixing techniques," in *Summaries of the 4th Annual JPL Airborne Geoscience Workshop*, 1993.
- [83] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208–220, 2012.
- [84] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor robust principal component analysis: Exact recovery of corrupted low-rank tensors via convex optimization," in *Proc. IEEE Conf. CVPR*, 2016, pp. 5249–5257.
- [85] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.
- [86] W. Deng and W. Yin, "On the global and linear convergence of the generalized alternating direction method of multipliers," *J. Sci. Comput.*, vol. 66, no. 3, pp. 889–916, 2016.