Content-Aware Subspace Low-Rank Tensor Recovery for Hyperspectral Image Restoration

Xueyao Xiao, Wei Zhang, *Member, IEEE*, Yi Chang, *Member, IEEE*, Shuning Cao, Wei He, *Senior Member, IEEE*, Houzhang Fang, *Member, IEEE*, and Luxin Yan, *Member, IEEE*

***Abstract*— The low-rank tensor model has made great progress for hyperspectral image (HSI) restoration. Recently, the low-rank tensor methods have further been boosted with subspace learning by transforming original HSI into a low-dimensional subspace with reduced computational burden and discriminative feature representation. However, existing subspace-based methods consistently employ a fixed subspace dimension for all patches, which may violate the intrinsic dimension discrepancy of different image content, leading to information loss or redundancy. In this work, our key observation is that the intrinsic subspace of different image patches along different dimensions is different, which should be adaptively modeled for compact feature extraction. Therefore, we propose a content-aware subspace low-rank tensor recovery (CSLRTR) method by leveraging both deep network and low-rank tensor model. Specifically, we first analyze the intrinsic discrepancy of different HSI patches among both spatial and spectral dimensions and design a simple network to adaptively learn the optimal subspace dimension. The adaptive subspace learning and low-rank tensor recovery are iteratively performed and mutually promote each other. On one hand, the learned subspace would contribute to more compact low-rank representation for better restoration; on the other hand, the low-rank tensor recovery with less degradations would definitely ease the difficulty of the subspace estimation. Note that the adaptive content-aware subspace strategy has been simultaneously employed on both spectral and nonlocal dimensions, where the spectral–spatial relationship has been further strengthened with better restoration. We have performed extensive experiments on different datasets and restoration tasks and extended the content-aware subspace strategy to previous methods.**

***Index Terms*— Content-aware subspace learning, hyperspectral image (HSI) restoration, low-rank tensor recovery.**

I. Introduction

HYPERSPECTRAL image (HSI) has attracted a lot of attention in the remote sensing field recently. Compared with RGB images, HSI has abundant spectral information, which is related to the physical property of different materials. Due to this characteristic, HSI has been widely exploited in military, agriculture, medicine, and other relevant applications. Regrettably, the quality of HSI can be compromised by various factors, including circuit noise, atmospheric turbulence, and defective pixels, leading to the manifestation of undesirable damage such as noise, blurring, and other forms of degradation [1]. The degraded HSI affects greatly subsequent applications, such as target recognition [2] and classification [3], [4]. HSI restoration is necessary before the subsequent applications.

HSI restoration aims to restore clean images from various degradations such as noise, blurring, and missing data. Generally, the HSI degradation model can be formulated as follows:

$$\mathcal{Y} = \mathcal{T}(\mathcal{X}) + \mathcal{N} \tag{1}$$

where $\mathcal{Y} \in \mathbb{R}^{M\times N\times B}$ is an observed low-quality degraded image, $\mathcal{X} \in \mathbb{R}^{M\times N\times B}$ is the clean image, $\mathcal{T}(\bullet)$ is a linear degradation operator, and $\mathcal{N} \in \mathbb{R}^{M\times N\times B}$ is the noise. M, N, and B denote the number of the row, column, and spectrum of the HSI, respectively. Different $\mathcal{T}(\bullet)$'s refer to different restoration tasks. Specifically, when $\mathcal{T}(\bullet)$ is an identical matrix, it becomes an HSI denoising task. When $\mathcal{T}(\bullet)$ is a blurring operator, it becomes an HSI deblurring task. When $\mathcal{T}(\bullet)$ is a mask operator, it becomes an HSI inpainting task. Interested readers could refer to [5] for more details about HSI restoration. Essentially, restoring $\mathcal{X}$ from $\mathcal{Y}$ is a challenging ill-posed inverse problem.

The HSI priors have been designed to cope with the ill-posed nature of the HSI restoration task. Various model-based methods [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51] have been proposed, which introduces the prior information to constrain the solution space and decouples the clean image from degradation. The low-rank matrix recovery (LRMR) methods [9], [10], [11], [13], [15] have achieved great success in HSI restoration, which can well utilize the spectral redundancy. However, the low-rank matrix-based methods need to transform the 3-D HSI into the 2-D version, which would unexpectedly damage the spectral–spatial correlation. Consequently, the spotlight has shifted toward low-rank tensor-based methods [18], [22], [26], [43], [51], which emerge as a pivotal focus in HSI restoration. These methods could better utilize spectral–spatial information and preserve the spectral–spatial correlation. Moreover, nonlocal self-similarity and spectral correlation as the intrinsic

Manuscript received 11 July 2023; revised 21 August 2023; accepted 27 August 2023. Date of publication 4 September 2023; date of current version 27 September 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 62101294, Grant 62371203, and Grant 61971460. *(Xueyao Xiao and Wei Zhang contributed equally to this work.) (Corresponding author: Yi Chang.)*

Xueyao Xiao, Yi Chang, Shuning Cao, and Luxin Yan 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 430074, China (e-mail: owuchangyuo@gmail.com).

Wei Zhang is with the Artificial Intelligence Center, Peng Cheng Laboratory, Shenzhen 518066, China.

Wei He is with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430072, China.

Houzhang Fang is with the School of Computer Science and Technology, Xidian University, Xi'an 710071, China.

Digital Object Identifier 10.1109/TGRS.2023.3311482

1558-0644 © 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: XIDIAN UNIVERSITY. Downloaded on January 22,2024 at 04:18:52 UTC from IEEE Xplore. Restrictions apply.

properties of HSI have been widely applied in the low-rank tensor recovery methods.

Unfortunately, these low-rank tensor models unexpectedly suffer from heavy computational burden due to the HSI complexity, which seriously restricts its real application. Moreover, performing the low-rank analysis in the original HSI image space cannot fully exploit its redundancy. To mitigate these problems, the subspace-based methods have proposed [18], [34], [49], [52], [53], [54], [55], [56], [57]. The key idea of the subspace-based method is to project the high-dimensional HSI manifold into its low-dimensional subspace along the spectrum, so as to greatly reduce the computational complexity and enhance the compact feature representation. Thus, choosing an appropriate subspace dimension is critical for feature representation since it should be complied with the intrinsic dimension of the input HSI. Zhuang and Bioucas-Dias [18] empirically set the subspace dimension with a fixed value (eight for Washington DC Mall subscene and Pavia Center subscene). Furthermore, He et al. [34] obtained the initial value of the subspace dimension through the Hysime [58] and adjusted it with each iteration via handcrafted linear equation. These fixed or handcrafted setting do not consider the intrinsic dimension discrepancy for different HSI patches, which would inevitably introduce the structure distortion during the restoration.

In this work, our key observation is that the intrinsic dimension of different image patches is significantly different. We first illustrate the development of the low-rank models for HSI modeling. Then, we provide a detailed analysis about the intrinsic subspace dimension of different spatial and spectrum contents and discover the local similarity and global discrepancy of the HSI along both the spatial and spectral dimensions. Consequently, this motivates us to design a content-aware subspace low-rank tensor recovery (CSLRTR) model for HSI restoration. The CSLRTR contains both the deep network and low-rank tensor model. The deep network is responsible for adaptively learning the optimal subspace dimension of different content for each iteration, while the subspace-based low-rank tensor model takes the learned subspace as the guidance to adaptively extract the compact feature. Both the adaptive subspace learning and the subspace-based low-rank tensor recovery are jointly performed within an iterative manner to mutually promote each other with better restoration performance.

Compared with previous low-rank methods, the proposed CSLRTR has further perceived the high-level subspace dimension discrepancy of different HSI content, which in turn guides the low-level restoration. On the other hand, compared with the popular deep learning methods [59], [60], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], the proposed method has strictly obey the physical degradation procedure with better generalization ability, where CSLRTR does not require paired large clean-degraded training dataset. Thus, our method can be easily applied to different HSI tasks such as denoising, deblurring, and inpainting. Moreover, the joint learning and low-rank optimization framework would benefit us to avoid design very complicated network architecture. We demonstrate that a very simple network could well accommodate our subspace learning task. In addition, we have validated that the subspace dimension learning network (SDLN) can be easily applied to previous state-of-the-art (SOTA) methods to help their model the discrepancy of different HSI patches and improve their restoration performance.

Moreover, the previous methods usually perform the spatial low-rank tensor recovery in spectral dimension reduction feature subspace only. In this work, we argue that both the spatial and spectral dimensions are the two sides of the same coin. Thus, to further enhance the spectral–spatial structure correlation, the adaptive content-aware subspace strategy has been simultaneously employed on both spectral and spatial (nonlocal) dimensions. Specifically, the proposed CSLRTR contains two streams: spectral subspace-based spatial low-rank tensor model and spatial subspace-based spectral low-rank tensor model. The alternative spatial–spectral adaptive subspace learning and low-rank tensor recovery would significantly enhance the spatial–spectral HSI structural preserving. Finally, the reconstruction results of the two streams are further fused to finally obtain better clean image. The contributions are summarized as follows.

1) We discover that the intrinsic subspace of different image patches along different dimensions is different. This motivates us to replace the handcrafted subspace-based methods with content-aware deep network, which adaptively learns the intrinsic subspace dimension of different spatial and spectrum contents to perceive the fine-grained discrepancy of different HSI patches for more compact feature representation and better HSI restoration.
2) We propose a content-aware subspace low-rank tensor model for HSI restoration, in which the subspace-based low-rank tensor recovery optimization model and SDLN are jointly performed within an iterative manner, so as to mutually promote each other with better restoration performance. As far as we know, few low-level HSI restoration methods have considered the utilization of high-level information by deep network.
3) The proposed method can be applied to different HSI restoration tasks such as denoising, deblurring, and inpainting. Also, the SDLN is a very flexible module, which can be easily extended to other methods to improve their restoration effect. We demonstrate the good generalization and efficient processing of the proposed method on various tasks and datasets and outperforms SOTA methods both quantitative and qualitative.

II. Related Work

A. Low-Rank Model-Based Methods

The low-rank recovery model-based methods aim to characterize the spectral correlation and nonlocal similarity of the HSI, which commonly manifest as the low-rank property in HSI restoration. In Fig. 1, we illustrate the evolution of data representation that leverages these two properties. Initially, most methods only focus on either spectral correlation or nonlocal similarity. In Fig. 1(a), to utilize the nonlocal

Authorized licensed use limited to: XIDIAN UNIVERSITY. Downloaded on January 22,2024 at 04:18:52 UTC from IEEE Xplore. Restrictions apply.

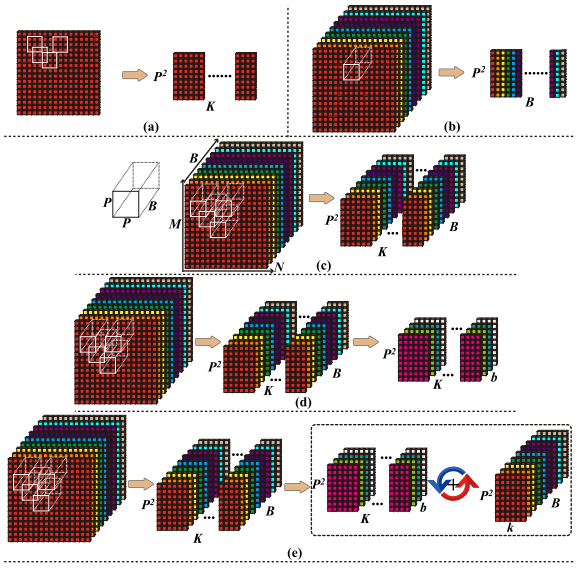


Fig. 1. Evolution of low-rank representation of HSI. (a) and (b) Matrix-based methods, which focus on either spectral correlation or nonlocal similarity [74], [75]. (c) Tensor-based methods and both spectral correlation and nonlocal similarity [25], [27]. (d) Subspace-based methods, which usually exploit the nonlocal similarity in the spectral subspace [18], [34]. (e) Proposed symmetric subspace-based methods, which utilize the nonlocal similarity in the spectral subspace and exploit the spectral correlation in the spatial (nonlocal) subspace.

similarity, these methods [74], [75] grouped K similar patches ($P \times P$) in the spatial dimension into a matrix ($P^2 \times K$) and subsequently applying a filtering operation to it. In Fig. 1(b), the 3-D HSI patch ($P \times P \times B$) is unfolded into a 2-D matrix ($P^2 \times B$) along the spectral dimension. Based on such matrix representation, Zhang et al. [10] proposed an LRMR model to exploit the spectral low-rank prior. He et al. [13] leveraged the space piecewise smoothness and proposed the total-variation-regularized low-rank matrix factorization model by introducing the total variation (TV) prior into the LRMR model. Xie et al. [14] proposed a weighted Schatten p-norm, taking into consideration the significance of different rank components, thereby making the model more closely aligned with the original low-rank assumption.

Although the low-rank matrix-based methods have produced promising restoration results, unfolding 3-D data into a 2-D matrix inevitably damages the spectral–spatial correlation of the HSI. The low-rank tensor models have been naturally proposed and extensively studied. First, the HSI was directly treated as a 3-D tensor [22], [29]. Peng et al. [22] proposed a nonlocal tensor dictionary learning (TDL) model that considered the spectral correlation and nonlocal similarity simultaneously. Wang et al. [29] introduced spatial–spectral TV (SSTV) priors and sparse priors into the low-rank tensor model, enabling the simultaneous utilization of local and global information of HSI. Next, as shown in Fig. 1(c), the new tensor ($P^2 \times K \times B$) is grouped by K similar 3-D patches ($P \times P \times B$) [25], [27], [44], and its spectral and spatial (nonlocal) dimensions would have strong low-rank property. For instance, Chang et al. [44] analyzed the discrepancy of the low-rank property along the spatial, spectral, and nonlocal modes and proposed an optimal low-rank tensor (OLRT) model. Furthermore, to alleviate computational complexity, subspace-based methods have been developed rapidly [18], [34]. These methods perform a principal component analysis on spectral dimension of the grouped tensor ($P^2 \times K \times B$) to obtain spectral subspace ($P^2 \times K \times b$) in Fig. 1(d). Zhuang and Bioucas-Dias [18] took full advantage of the low-rank structure and self-similarity and proposed a fast hyperspectral denoising (FastHyDe) based on spectral subspace.

However, the previous subspace-based methods usually perform the spatial low-rank tensor recovery in spectral subspace. We argue that both the spatial and spectral dimensions are like two sides of the same coin for HSI. To further enhance the spectral–spatial structure correlation, in Fig. 1(e), we combine the spectral subspace and spatial subspace and propose a symmetric restoration structure: spectral subspace-based spatial low-rank tensor model and spatial subspace-based spectral low-rank tensor model. The two streams explore the nonlocal similarity in spectral subspace and the spectral correlation in spatial subspace respectively. When they fuse with each other, it enables further utilization of the potentially lost information during the optimization process while simultaneously enhancing the spectral–spatial coupling.

Moreover, most of the previous low-rank recovery methods treated the HSI restoration as a low-level task without high-level understanding of different image contents. In this work, we further take the intrinsic spatial–spectral subspace dimension discrepancy of different HSI contents into consideration for compact feature representation. We introduce a deep learning network into the low-rank tensor optimization model to adaptively perceive the intrinsic dimension of the HSI along both the spatial and the spectrum. Compared with the fixed parameter of the conventional methods, the learned adaptive parameters would greatly benefit us to model different HSI contents.

B. Learning-Based Methods

The learning-based methods automatically learn the HSI prior from the dataset via deep neural network and have been applied in HSI restoration [59], [60], [63], [67], [69]. Specifically, Chang et al. [59] proposed to apply the fully convolutional network for HSI restoration and designed a 2-D convolutional neural network (CNN) with dilation and residual learning model. Yuan et al. [60] introduced a 2-D spatial and 3-D spectral–spatial combined CNN to mine the character of single band and the high correction of its adjacent band. Dong et al. [63] presented a modified 3-D U-net to encode the rich multiscale information of HSI. Wei et al. [67] proposed a 3-D quasi-recurrent neural network to extract the spectral–spatial structural correlation and global correlation along spectral. Recently, the attention mechanism and transformers, which have shown success in various computer vision tasks, have also been applied in HSI restoration [73], [76], [77], [78] to explore the global spatial–spectral priors. Shi et al. [70] introduced the dual-attention denoising network to exploit the global dependency and spatial–spectral correlation. To explore the intrinsic similarity in both spatial and spectral dimensions, Li et al. [77] proposed nonlocal spatial self-attention and global spectral self-attention with

Authorized licensed use limited to: XIDIAN UNIVERSITY. Downloaded on January 22,2024 at 04:18:52 UTC from IEEE Xplore. Restrictions apply.

transformer architecture. In [78], a local–global feature-aware transformer-based residual (FATR) network was proposed for HSI denoising.

Nevertheless, the learning-based methods require a large amount of paired data to achieve satisfactory restoration results. Moreover, the deep network is always designed for one specific task, resulting in poor generalization for different tasks. In addition, the deep network directly learns an end-to-end mapping with powerful nonlinear transformation, which leads to poor interpretability and controllability. In contrast, each step of the model-based methods has a clear physical interpretation and generalization. However, the optimization methods for low-level restoration tasks are less discriminative for different spatial–spectral HSI contents, but this is exactly what deep models are good at. In this work, we leverage both the deep network and low-rank tensor recovery model and propose a CSLRTR for HSI restoration to address the adaptive subspace dimension issue, enabling the model to better adapt to various scenarios.

It is worth noting that the idea of combining model- and learning-based methods has already been proposed such as methods based on unfolding [79], [80] and plug-and-play [81], [82], [83]. The key idea of unfolding methods is to mimic the iterative optimization procedure of model-based method with an end-to-end deep network, which in essence belongs to the deep learning paradigm. They still heavily depend on the paired training datasets with less generalization ability. The plug-and-play methods resort to the half-quadratic splitting technique with separated prior and data fidelity term and replace the conventional handcrafted prior by a powerful deep denoiser. Compared with the plug-and-play methods, the proposed method does not rely on designing complicated deep network to estimate the noise-free image (tensor). Rather, the proposed method mainly utilizes the deep network to perceive the intrinsic subspace dimension (scale) of the spatial–spectral HSI patches. The single scale is significantly easier than the complex tensor to be learned.

III. CSLRTR Model

In this section, we first introduce the necessary notions and preliminaries. Then, we analyze the spatial–spectral subspace dimension discrepancy of the HSI. Motivated by the discrepancy analysis, we propose the spatial–spectral content-aware restoration framework and introduce the two-stream subspace-based low-rank tensor model and the SDLN in detail. Subsequently, we provide the detailed optimization process of the model. Finally, we analyze the time complexity of the proposed model.

A. Notions and Preliminaries

In this article, we represent tensors as the Euler script letters, e.g., $\mathcal{X}$. Matrices are denoted by boldface capital letters, e.g., $\boldsymbol{X}$. The element of an N-dimensional tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is denoted by $x_{i_1,\ldots,i_n,\ldots,i_N}$. The mode-$n$ vectors of tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ can be obtained by fixing all indices but one, and the mode-n unfolding matrix $\boldsymbol{X}^{(n)} \in \mathbb{R}^{I_n \times (I_1 \times \cdots \times I_{n-1} I_{n+1} \times \cdots \times I_N)}$ is obtained by taking the mode-n vectors of $\mathcal{X}$ to be the columns of the matrix. Therefore, the n-rank of $\mathcal{X}$ can be analyzed by means of matrix techniques and is equal to the rank of the mode-n unfolding matrix $\boldsymbol{X}^{(n)}$, e.g., $\text{rank}_n(\mathcal{X}) = \text{rank}(\boldsymbol{X}^{(n)})$. The mode-$n$ product of tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ by matrix $\boldsymbol{U} \in \mathbb{R}^{J_n \times I_n}$ is expressed as $\mathcal{Z} = \mathcal{X} \times_n \boldsymbol{U}$, where $\mathcal{Z} \in \mathbb{R}^{I_1 \times \cdots \times J_n I_{n+1} \times \cdots \times I_N}$. The Frobenius norm of a tensor $\mathcal{X}$ is defined as $\|\mathcal{X}\|_F = ((\sum_{i_1,\ldots,i_N} |x_{i_1,\ldots,i_n,\ldots,i_N}|^2))^{1/2}$.

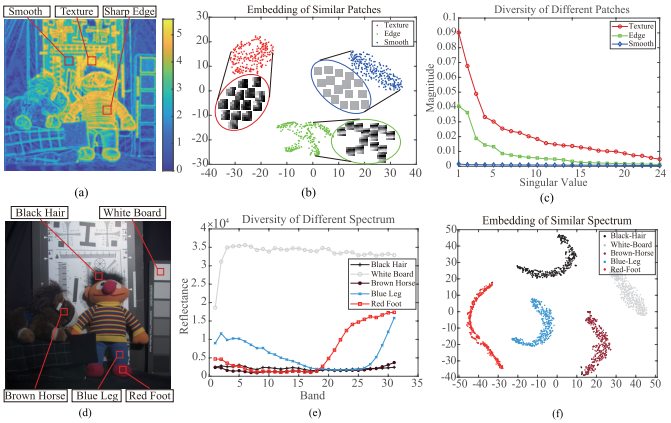


Fig. 2. (a) Different spatial image structures, including smooth, texture, and sharp edge structures (local patch entropy). (b) T-SNE visualization of different spatial image structures. (c) Corresponding singular value curves (principal component analysis). (d) Different spectral contents, including black hair, whiteboard, brown horse, blue leg, and red foot (toy). (e) Spectrum reflectance curves. (f) T-SNE visualization of different spectral color regions.

B. Spatial–Spectral Subspace Dimension Discrepancy Analysis

We analyze the intrinsic subspace discrepancy along both the spatial and spectral dimensions in Fig. 2. In Fig. 2(a), we first choose different spatial image contents such as the texture, sharp edge, and smooth patches among the whole image. Next, we search for the local similar patches of each content patch. Fig. 2(b) shows the T-SNE visual visualization results of these three typical category clusters, in which red represents the texture region, green denotes the edge region, and blue denotes the smooth region. Notably, the similar structures are grouped together within the same class, indicating that the similar contents share the similar distribution. On the contrary, these cluster centers of different contents are far apart from each other, indicating intrinsic discrepancy between different structure patches.

Furthermore, we quantitatively analyze the intrinsic dimension discrepancy of the three clusters via the singular value decomposition (SVD) in Fig. 2(c). We observe that the singular value curve in the smooth region appears relatively flat, with a majority of singular values approaching zero. In contrast, although the singular value curves for the edge region and the texture region possess the low-rank property to some extent, different structures in HSI possess different low-rank properties. Specifically, the smooth region exhibits the strongest low-rank property, followed by the edge region, while the texture region exhibits the weakest, further demonstrating the global low-rank discrepancy of the spatial dimension. Overall, both the T-SNE and SVD experiments have validated that there

Authorized licensed use limited to: XIDIAN UNIVERSITY. Downloaded on January 22,2024 at 04:18:52 UTC from IEEE Xplore. Restrictions apply.

exist both local similarity and global discrepancy in the spatial dimension.

Similarly, in Fig. 2(d)–(f), we demonstrate that the local similarity and global discrepancy observation also holds in the spectral dimension. In Fig. 2(d), we select typical different spectral signals with color contents, such as the black hair, whiteboard, brown horse, blue leg, and red foot, which naturally possess different spectral curves. Fig. 2(e) presents the spectral response curves corresponding to these distinct spectral signals. The spectral response curves of brown and black, being dark colors, exhibit similarity, while the curve for white is the most different from them, indicating the global discrepancy of the spectral dimension. Furthermore, we perform T-SNE on the spectral responses of those five color regions in Fig. 2(f). The spectrums with the same color regions are clustered together, while the cluster centers of different color regions display substantial separation, thereby emphasizing the fact of local similarity and global discrepancy in spectral dimension. Overall, this observation motivates us to model the intrinsic spatial–spectral discrepancy for better HSI restoration.

C. Spatial–Spectral Content-Aware Restoration Framework

We have discovered the local similarity and global discrepancy of HSI along the spatial and spectral dimensions. However, most model-based methods mainly focus on the similarity and ignore the discrepancy. For example, the subspace-based low-rank tensor methods often impose a uniform subspace dimension for different HSI spatial–spectral contents, neglecting the intrinsic difference would result in varying optimal subspace dimension. To address this issue, we design a CSLRTR for HSI restoration to adaptively capture the discrepancy and obtain a more compact subspace representation. The overall content-aware restoration framework is shown in Fig. 3.

In this work, we propose a CSLRTR approach for HSI restoration by integrating both deep network and low-rank tensor recovery model. Specifically, we introduce a very simple deep network to adaptively estimate the optimal subspace dimension for different spatial or spectral HSI content. The learned subspace dimension would be used to guide the subspace dictionary learning. Then, both the adaptive spectral subspace-based nonlocal low-rank approximation and nonlocal subspace-based spectral low-rank approximation are jointly performed to restore the HSI in the low-dimensional feature subspaces. Finally, we reconstruct the restored high-dimensional HSI from the low-dimensional subspaces. Note that the adaptive subspace learning and low-rank tensor recovery are performed iteratively and mutually promote each other. On one hand, the deep network would perceive different HSI content for adaptive subspace learning with more compact low-rank representation for better restoration; on the other hand, the model-driven low-rank tensor recovery with less degradations would ease the difficulty of subspace learning. Next, we describe each component in detail.

1) Spatial–Spectral Subspace-Based LRTR Model: Restoring the clean image from the degradation model is an ill-posed problem. In order to solve the problem, various prior regularization terms are used to constrain the solution range. Therefore, we can formulate the HSI restoration problem as

$$\mathcal{X} = \arg\min_{\mathcal{X}} \frac{1}{2}\|\mathcal{T}(\mathcal{X}) - \mathcal{Y}\|_F^2 + \omega_j \sum_j H_j(x) \tag{2}$$

where $\mathcal{Y} \in \mathbb{R}^{M\times N\times B}$ represents the degraded image, $\mathcal{X} \in \mathbb{R}^{M\times N\times B}$ represents the clean image, $H_j(x)$ are different regularization terms, ω_j denotes the balance parameter, and $\mathcal{T}(\cdot)$ represents different degradation operators, such as denoising, deblurring, and inpainting. The subspace-based low-rank tensor recovery has been explored with good performance [18]. The loss function is formulated as

$$\begin{aligned}\left\{\widehat{\mathcal{X}}, \widehat{\mathcal{L}}_i^3, \boldsymbol{Q}_i^3\right\} &= \arg\min_{\mathcal{X},\mathcal{L}_i^3,\boldsymbol{Q}_i^3} \frac{1}{2}\|\mathcal{T}(\mathcal{X}) - \mathcal{Y}\|_F^2 \\ &+ \omega \sum_i \left(\frac{1}{\lambda_i^2}\|\mathcal{R}_i\mathcal{X}\times_3 \boldsymbol{Q}_i^3 - \mathcal{L}_i^3\|_F^2 + rank_2\left(\mathcal{L}_i^3\right)\right) \\ &\text{s.t.} \quad (\boldsymbol{Q}_i^3)^T\boldsymbol{Q}_i^3 = \boldsymbol{I}\end{aligned} \tag{3}$$

where $\boldsymbol{Q}_i^3 \in \mathbb{R}^{B\times b}$ represents the subspace projection matrix, $\text{rank}_2(\bullet)$ denotes the nonlocal low-rank constraint, $\mathcal{L}_i^3 \in \mathbb{R}^{P^2\times K\times b}$ is the low-rank approximation, and b is the spectral subspace dimension. $\mathcal{R}_i\mathcal{X} \in \mathbb{R}^{P^2\times K\times B}$ denotes the ith grouped patches obtained by nonlocal clustering of similar patches $\mathcal{U}_i \in \mathbb{R}^{P\times P\times B}$. K is the number of the nonlocal similar patches, indicating the nonlocal dimension, and B stands for the number of bands, indicating the spectral dimension. Both the spectral correlation and the nonlocal similarity are well preserved by such representation.

However, we argue that both the spatial and spectral dimensions are the two sides of the same coin. Thus, to further enhance the spectral–spatial structure correlation, we adopt a symmetric low-rank tensor subspace restoration framework that contains two streams: spectral subspace-based spatial low-rank tensor model and spatial subspace-based spectral low-rank tensor model. These two streams aim to explore the spatial similarity in the spectral subspace and the spectral correlation in the spatial subspace. Thus, we would extend (3) spectral subspace-based LRTR model to its symmetric version, namely, the spectral and spatial subspace-based low-rank tensor model

$$\begin{aligned}\left\{\hat{\mathcal{X}}, \hat{\mathcal{L}}_i^j, \boldsymbol{Q}_i^j\right\} &= \arg\min_{\mathcal{X},\mathcal{L}_i^j,\boldsymbol{Q}_i^j} \frac{1}{2}\|\mathcal{T}(\mathcal{X}) - \mathcal{Y}\|_F^2 \\ &+ \omega_j \sum_j \sum_i \left(\frac{1}{\lambda_i^2}\|\mathcal{R}_i\mathcal{X}\times_j \boldsymbol{Q}_i^j - \mathcal{L}_i^j\|_F^2 + \text{rank}_{j'}\left(\mathcal{L}_i^j\right)\right) \\ &\text{s.t.} \quad (\boldsymbol{Q}_i^j)^T\boldsymbol{Q}_i^j = \boldsymbol{I}\end{aligned} \tag{4}$$

where $j \in \{2, 3\}$ represents the nonlocal dimension and spectral dimension and j' is the complement of j. In our previous work [44], we have proven that the combination of the spectral and nonlocal-induced low rank is most beneficial for HSI modeling. Here, we embed the subspace learning into the optimal low-rank model with more compact low-dimensional representation.

Authorized licensed use limited to: XIDIAN UNIVERSITY. Downloaded on January 22,2024 at 04:18:52 UTC from IEEE Xplore. Restrictions apply.

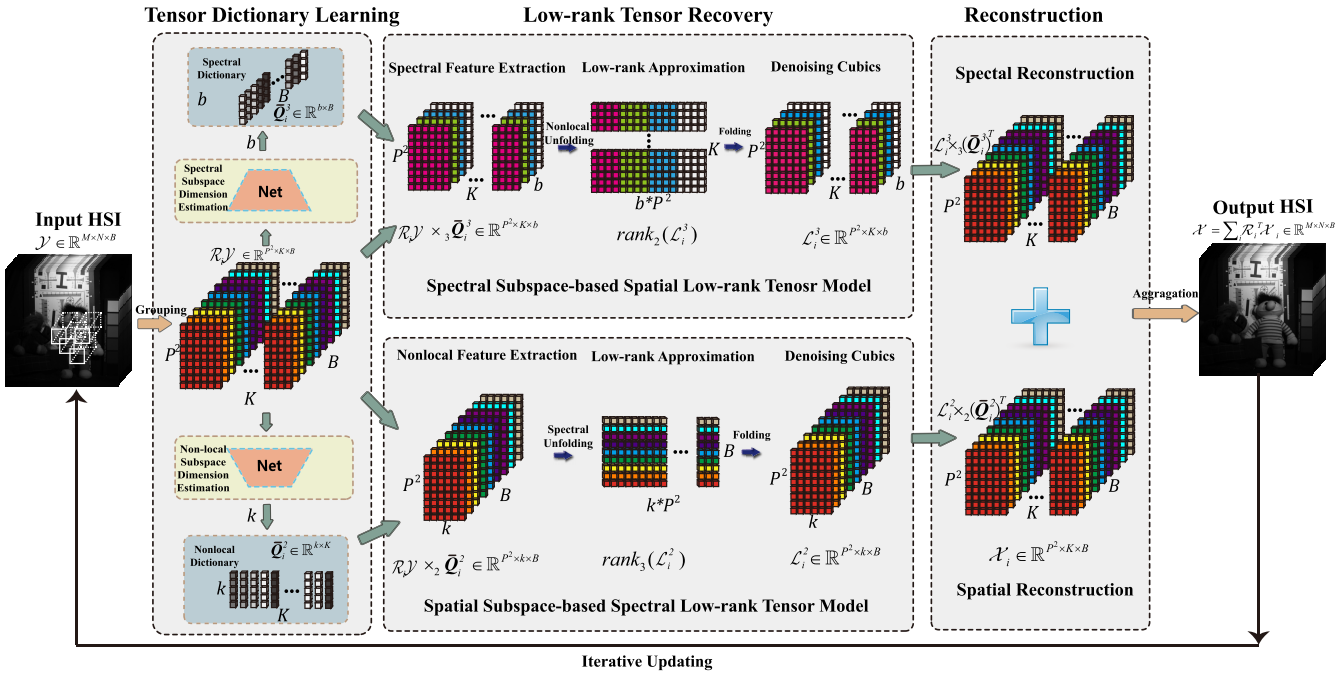


Fig. 3. Overview of the proposed CSLRTR. It consists of three parts: the first part is TDL, the second part is low-rank tensor recovery, and the third part is the reconstruction. The TDL consists of two steps, namely, learning the spectral subspace dimension and nonlocal subspace dimension by the ResNet18 classification and obtaining the dictionary through the SVD decomposition.

As we have analyzed, (3) and (4) have unexpectedly ignored the subspace dimension discrepancy for different HSI contents. The subspace transformation $\mathcal{R}_i\mathcal{X}\times_j \boldsymbol{Q}_i^j$ sets a uniform subspace dimension value across the whole image. To obtain a more compact subspace representation, we leverage the content-aware property of the deep network to adaptively learn the optimal subspace dimension of different image content. Specifically, we further extend the conventional fixed projection $\boldsymbol{Q}_i^j$ to its adaptive subspace projection matrix $\bar{\boldsymbol{Q}}_i^j$, where the learned optimal subspace dimension by the deep network is used to guide subspace projection

$$\begin{aligned}\left\{\widehat{\mathcal{X}}, \widehat{\mathcal{L}}_i^j, \bar{\boldsymbol{Q}}_i^j\right\} &= \arg\min_{\mathcal{X},\mathcal{L}_i^j,\bar{\boldsymbol{Q}}_i^j} \frac{1}{2}\|\mathcal{T}(\mathcal{X})-\mathcal{Y}\|_F^2 \\ &+ \omega_j \sum_j \sum_i \left(\frac{1}{\lambda_i^2}\|\mathcal{R}_i\mathcal{X}\times_j \bar{\boldsymbol{Q}}_i^j - \mathcal{L}_i^j\|_F^2 + \text{rank}_{j'}\left(\mathcal{L}_i^j\right)\right) \\ \text{s.t.} \quad & (\bar{\boldsymbol{Q}}_i^j)^T\bar{\boldsymbol{Q}}_i^j = \boldsymbol{I}. \end{aligned} \tag{5}$$

Equation (5) is the overall loss function. One key is to learn the optimal low-dimensional subspace for the $\bar{\boldsymbol{Q}}_i^j$. Next, we introduce details about the SDLN.

2) Subspace Dimension Learning Network: The SOTA subspace-based low-rank recovery methods usually employ a fixed subspace dimension for all patches, which may violate the intrinsic dimension discrepancy of different image content, leading to information loss or redundancy. In this work, we hypothesize that the high-level semantic understanding about the content is of great benefit for low-level restoration tasks. Therefore, we replace handcrafted subspace-based methods with content-aware deep network so as to better perceive fine-grained discrepancy for compact feature representation. In addition, this novel combination of optimization model and deep network makes the model not only perceive high-level information but also obey physical degradation procedures with better generalization ability.

The goal of SDLN is to adaptively learn the intrinsic dimension of different HSI content, where the grouped patch serves as the input and the optimal subspace dimension is the output. The subspace dimension learning task could be viewed as a classification task where each subspace dimension represents a category. The output category with the highest probability will be selected as our optimal subspace dimension. This is reasonable since learning intrinsic low dimension of a high-dimensional HSI is to discover the highest probability among different dimensions.

Compared with the conventional subspace estimate methods HySime [58], AIC [84], and MDL [85], the SDLN is robust for degradation. The HySime and AIC/MDL may work well on the clean HSI. However, with the degradation disturbance, the difficulty of the subspace dimension estimation problem has increased dramatically. Moreover, the distribution of data in deblurring and inpainting tasks does not satisfy the assumptions of HySime and AIC/MDL, which may affect the estimation of subspace dimensions. On the contrary, the SDLN benefits from the powerful nonlinear representation ability of deep network, which is more robust to the degradations.

Compared with previous deep network-based HSI restoration, the proposed deep network has three distinct advantages: free from complicated network design, relax to the fully supervised clean-degraded training dataset, and easy to extend to previous methods. For the first one, compared with previous learning-based HSI restoration network, we do not need to directly estimate the dense high-dimensional HSI in an end-to-end way. In contrast, we formulate the HSI restoration into a joint learning and optimization framework where the SDLN focuses on estimating the sparse 1-D probability value, which

Authorized licensed use limited to: XIDIAN UNIVERSITY. Downloaded on January 22,2024 at 04:18:52 UTC from IEEE Xplore. Restrictions apply.

is significantly easier to be estimated. Moreover, the deep network is only a part of CSLRTR, which serves as the guidance. These would greatly ease complicated network design, and we can employ a very simple network to achieve the goal.

For the second one, the existing fully supervised networks require the paired clean and degraded HSI. Thus, the generalization capability is severely restricted for different tasks since it is hard to prepare for different HSI degradations. Moreover, existing supervised methods all employ the synthetic models to simulate the degradation and clean pairs, which may perform unsatisfactorily on real HSI due to the domain gap between the real and synthetic datasets. For the third one, due to the novel combination of optimization model and deep network, the SDLN is a very flexible module and can be easily extended to previous methods and further improve their performances.

In this work, our SDLN does not require a paired large clean-degraded HSI training dataset. Instead, we just need to learn the intrinsic subspace dimension, which is much easier to be estimated and get rid of the specific HSI task. Here, we select the classic ResNet18 [86] classification network due to its simple architecture. Thus, the key problem is to prepare the "ground truth" of the optimal subspace dimension for training. We adopt the exhaust algorithm to obtain the pseudo-GT of subspace dimension of the grouped similar patches $\mathcal{R}_i\mathcal{X} \in \mathbb{R}^{P^2\times K\times B}$. First, we simulate continuous noise variance ranging from 1 to 50 to make the network adaptable to images with various levels of degradation. Next, taking the nonlocal dimension as an example, we enumerate all possible subspace dimensions, substitute them into the low-rank tensor approximation model (9), and manually select the best restoration result as pseudo-GT. For the spectral dimension, we can get the pseudo-GT in the same way.

After preparing training datasets, we train only one foundation model on the NTIRE2020 dataset with 800 000 different noisy level random patches ($25\times 100\times 31$, 5×5 spatial size, 100 nonlocal similar patches, and 31 bands). Note that, in this work, this model can well generalize to different HSI datasets, including simulated and real datasets, and achieved good recovery results. As for different HSI restoration tasks such as deblurring and inpainting, to achieve the optimal results, we fine-tune the foundation model with limited 32 000 patches. Note that, although the deep network may take time for dataset preparing and training, once it has been learned, the inference would be easy with very fast time.

It is worth noting that, for the method of constructing pseudo-GT, it is not necessary to enumerate all subspace dimensions. For example, it is well known that large sets of spectra consisting of natural scenes or extensive color samples can be represented compactly in just 6–8 principal components [87], [88]. As for nonlocal dimension, our previous work [44] has demonstrated that the low-rank property is even much more tight than that of the spectral dimension. According to this statistical prior knowledge, in this work, we can empirically narrow down the range of subspace dimensions, greatly reducing the searching time complexity. As for different HSI restoration tasks such as deblurring and inpainting, to achieve the optimal results, we could fine-tune the foundation model with fewer patches.

D. Optimization

Directly estimating multiple variables is difficult. To solve $\mathcal{X}, \mathcal{L}_i^j$, and $\bar{\boldsymbol{Q}}_i^j$, we adopt the variable splitting method of alternating direction method of multipliers (ADMM) [89] to decouple the variables and transform (5) into a series of subproblems.

1) Subspace Projection Matrix $\bar{\boldsymbol{Q}}_i^j$: We fix the other variables and optimize $\bar{\boldsymbol{Q}}_i^j$ with the following subproblem:

$$\min_{\bar{\boldsymbol{Q}}_i^j} \frac{1}{\lambda_i^2}\|\mathcal{R}_i^j\mathcal{X}\times_j\bar{\boldsymbol{Q}}_i^j - \mathcal{L}_i^j\|_F^2 \quad \text{s.t. } (\bar{\boldsymbol{Q}}_i^j)^T\bar{\boldsymbol{Q}}_i^j = \boldsymbol{I}. \tag{6}$$

$\bar{\boldsymbol{Q}}_i^j$ is an orthogonal matrix. According to [26], (6) has the closed-form solution. We first perform mode-j matricization on $\mathcal{R}_i\mathcal{X}$ and decompose it by SVD to get $(\boldsymbol{R}_i\boldsymbol{X})^{(j)} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T$, and then, we let $\bar{\boldsymbol{Q}}_i^j = (\boldsymbol{U}(:,1:d))^T$, where $\boldsymbol{U}(:,1:d)$ means to chose the first d left singular vectors corresponding to the d largest singular values, where d is the subspace dimension of the spectral or nonlocal. The initial value of $\mathcal{X}$ during the iterative process is equal to the degraded image $\mathcal{Y}$.

To characterize the intrinsic discrepancy of different HSI patches among both the spatial and spectral dimensions and get a compact feature representation, the subspace dimension is learned by the network

$$b = Net_{\text{spectral}}(\mathcal{R}_i\mathcal{X}) \tag{7}$$

$$k = Net_{\text{nonlocal}}(\mathcal{R}_i\mathcal{X}) \tag{8}$$

where b represents the optimal spectral subspace dimension and k represents the optimal nonlocal subspace dimension. Net$(\cdot)$ is the trained SDLN. Note that, in each iteration, the estimated subspace needs to be updated before the $\bar{\boldsymbol{Q}}_i^j$ estimation. In this article, we simply adopt the ResNet18 [86] to estimate the subspace dimension.

2) Low-Rank Approximation $\mathcal{L}_i^j$: We fix the other variables and relax the low-rank constraint term to the tensor nuclear norm term, and then, we have the following formula:

$$\hat{\mathcal{L}}_i^j = \arg\min_{\mathcal{L}_i^j} \frac{1}{\lambda_i^2}\|\mathcal{R}_i\mathcal{X}\times_j\bar{\boldsymbol{Q}}_i^j - \mathcal{L}_i^j\|_F^2 + \|\mathcal{L}_i^j\|_{tnn} \tag{9}$$

where λ_i^2 is the noise variance, $\|\mathcal{L}_i^j\|_{tnn} = \sum_p \left|\sigma_p(\boldsymbol{L}_i^{(j)})\right|_1$ is the tensor nuclear norm, $\boldsymbol{L}_i^{(j)}$ denotes the model-j matricization of the $\mathcal{L}_i^j$, and $\sigma_p(\boldsymbol{L}_i^{(j)})$ represents the pth singular value of $\boldsymbol{L}_i^{(j)}$. This minimization problem is usually solved by the singular value thresholding algorithm [27], [90], and we adopt the weighted nuclear norm minimization to boost the performance [91]. Thus, we can get the following formula:

$$\hat{\boldsymbol{L}}_i^j = \boldsymbol{U}\Psi_{\boldsymbol{W}}(\boldsymbol{\Sigma})\boldsymbol{V}^T \tag{10}$$

where $\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$ is the SVD of the mode-j matricization of the $\mathcal{R}_i\mathcal{X}\times_j\bar{\boldsymbol{Q}}_i^j$ and $\Psi_{\boldsymbol{W}}(\boldsymbol{\Sigma}) = \max(\boldsymbol{\Sigma} - \boldsymbol{W}, 0)$ is the generalized soft-thresholding operator with the weight $\boldsymbol{W}$. After obtaining $\boldsymbol{L}_i^j$, we conduct the tensor folding to transform it into the 3-D tensor $\mathcal{L}_i^j$.

Authorized licensed use limited to: XIDIAN UNIVERSITY. Downloaded on January 22,2024 at 04:18:52 UTC from IEEE Xplore. Restrictions apply.

3) Clean HSI Estimation $\mathcal{X}$: We take out the two items related to $\mathcal{X}$ and get the following formula:

$$\hat{\mathcal{X}} = \arg\min_{\mathcal{X}} \frac{1}{2}\|\mathcal{T}(\mathcal{X}) - \mathcal{Y}\|_F^2 + \omega_j \sum_j \sum_i \frac{1}{\lambda_i^2}\|\mathcal{R}_i\mathcal{X} \times_j \bar{\boldsymbol{Q}}_i^j - \mathcal{L}_i^j\|_F^2. \tag{11}$$

For the convenience of solving (11), we introduce an auxiliary variable $\mathcal{Z} \in \mathbb{R}^{M\times N\times B}$ and let $\mathcal{Z} = \mathcal{X}$. Through the augmented Lagrangian algorithm, the constrained problem is transformed into an unconstrained problem, and we have the following formula:

$$\{\hat{\mathcal{X}}, \hat{\mathcal{Z}}\} = \arg\min_{\mathcal{X},\mathcal{Z}} \frac{1}{2}\|\mathcal{T}(\mathcal{X}) - \mathcal{Y}\|_F^2 + \omega_j \sum_j \sum_i \frac{1}{\lambda_i^2}\left\|\mathcal{R}_i\mathcal{Z} \times_j \bar{\boldsymbol{Q}}_i^j - \mathcal{L}_i^j\right\|_F^2 + \frac{\beta}{2}\left\|\mathcal{Z} - \mathcal{X} - \frac{\mathcal{J}}{\beta}\right\|_F^2 \tag{12}$$

where β is penalty parameter and $\mathcal{J}$ is the Lagrange multiplier.

By ignoring variables independent $\mathcal{X}$ in (12), we get the following formula:

$$\hat{\mathcal{X}} = \arg\min_{\mathcal{X}} \frac{1}{2}\|\mathcal{T}(\mathcal{X}) - \mathcal{Y}\|_F^2 + \frac{\beta}{2}\left\|\mathcal{Z} - \mathcal{X} - \frac{\mathcal{J}}{\beta}\right\|_F^2. \tag{13}$$

The closed-form solution of (13) can be easily solved as the following iterations:

$$\mathcal{X}^{(l+1)} = \mathcal{F}^{-1}\left(\frac{\mathcal{F}\left(\mathcal{T}^T(\mathcal{Y}) + \left(\beta^{(l)}\mathcal{Z}^{(l)} - \mathcal{J}^{(l)}\right)\right)}{\mathcal{F}^*(\mathcal{T}) \circ \mathcal{F}(\mathcal{T}) + \beta^{(l)}\mathcal{I}}\right) \tag{14}$$

where $\mathcal{F}$ represents the fast Fourier transform, $\mathcal{F}^{-1}$ is the inverse transform, $\mathcal{F}^*$ is the conjugate, and $\mathcal{I}$ is the identity tensor.

4) Auxiliary Variable $\mathcal{Z}$: Dropping the irrelevant variables, we have the following iterations:

$$\mathcal{Z}^{(l+1)} = \arg\min_{\mathcal{Z}} \omega_j \sum_j \sum_i \frac{1}{\lambda_i^2}\left\|\mathcal{R}_i\mathcal{Z} \times_j \bar{\boldsymbol{Q}}_i^j - \mathcal{L}_i^j\right\|_F^2 + \frac{\beta}{2}\left\|\mathcal{Z} - \mathcal{X}^{(l+1)} - \frac{\mathcal{J}^{(l)}}{\beta}\right\|_F^2. \tag{15}$$

The above formula can be transformed as follows:

$$\mathcal{Z}^{(l+1)} = \arg\min_{\mathcal{Z}} \omega_j \sum_j \sum_i \frac{1}{\lambda_i^2}\left\|\mathcal{R}_i\mathcal{Z} - \mathcal{L}_i^j \times_j \left(\bar{\boldsymbol{Q}}_i^j\right)^T\right\|_F^2 + \frac{\beta}{2}\left\|\mathcal{Z} - \mathcal{X}^{(l+1)} - \frac{\mathcal{J}^{(l)}}{\beta}\right\|_F^2. \tag{16}$$

Equation (16) has the closed-form solution

$$\mathcal{Z}^{(l+1)} = \left(\frac{2\omega_j}{\lambda_i^2}\sum_j \sum_i (\mathcal{R}_i)^T\mathcal{R}_i + \beta^{(l)}\mathcal{I}\right)^{-1} \times \left(\frac{2\omega_j}{\lambda_i^2}\sum_j \sum_i (\mathcal{R}_i)^T\mathcal{L}_i^j \times_j \left(\bar{\boldsymbol{Q}}_i^j\right)^T + \beta^{(l)}\mathcal{X}^{(l+1)} + \mathcal{J}^{(l)}\right) \tag{17}$$

where $(\mathcal{R}_i^j)^T\mathcal{L}_i^j \times_j (\bar{\boldsymbol{Q}}_i{}^j)^T$ is the sum value of all overlapping reconstruction patches that cover the pixel location and $(\mathcal{R}_i)^T\mathcal{R}_i$ is the number of the overlapping patches. Equation (17) can be solved by the pixel-to-pixel division.

5) Update $\mathcal{J}$ and β: We have the following iterations:

$$\mathcal{J}^{(l+1)} = \mathcal{J}^{(l)} + \beta\left(\mathcal{X}^{(l+1)} - \mathcal{Z}^{(l+1)}\right) \tag{18}$$

$$\beta^{(l+1)} = \tau\beta^{(l)} \tag{19}$$

where $\tau > 1$ is constant larger than 1. The overall procedure is summarized in Algorithm 1.

Algorithm 1 CSLRTR

Require: Noisy image $\mathcal{Y}$, regularization parameters ω_j, β, noise variance λ_i^2

1: **Initialize:** $\mathcal{X}^{(1)} = \mathcal{Y}$, $\mathcal{J}^{(1)} = 0$
2: **for** t=1:T **do**
3: Group similar patches: $\mathcal{R}_i\mathcal{X}$;
4: Estimate subspace dimension k and b via Eq. (7)-(8);
5: Estimate subspace projection matrix $\bar{\boldsymbol{Q}}_i^j$ via Eq. (6);
6: Low-rank approximation $\mathcal{L}_i^j$ via Eq. (10);
7: **for** l=1:L **do**
8: Estimate clean HSI $\mathcal{X}$ via Eq. (14);
9: Estimate auxiliary variable $\mathcal{Z}$ via Eq. (17);
10: Update $\mathcal{J}$ and β via Eq. (18-19);
11: **end for**
12: **end for**

Ensure: The clean image $\boldsymbol{\mathcal{X}}$.

E. Theoretical Analysis of Time Complexity

According to Algorithm 1, CSLRTR mainly contains the nonlocal patch searching (Line 3), subspace projection (Line 5), low-rank approximation (Line 6), and clean HSI estimation (Line 8). The detailed time complexity of each step is analyzed as follows. For nonlocal patch searching, it is performed in the 2-D spectral average image for acceleration. Thus, the time complexity of the nonlocal patch searching is $\mathcal{O}\left(TP^2mnMN/(sS)\right)$, where T is the number of outer loops, P is the spatial size of the HSI patch, M and N are the spatial sizes of the HSI, S is the corresponding stride of the whole image, m and n are the spatial sizes of the searching nonlocal region, s is the corresponding stride of the nonlocal region, and MN/S^2 denotes the number of the processing patches. Next, we replace MN/S^2 with G and mn/s^2 with g for short.

For the subspace projection, it contains both the spectral and spatial (nonlocal) projection: $\mathcal{O}\left(TP^2GKB^2\right)$ and $\mathcal{O}\left(TP^2GK^2B\right)$, where K is the number of the nonlocal patches and B is the number of HSI band. For the low-rank approximation, it also contains the corresponding spectral and spatial (nonlocal) ones: $\mathcal{O}\left(TP^2GkB^2\right)$ and $\mathcal{O}\left(TP^2GK^2b\right)$, where k and b are subspace dimension of the nonlocal and spectral dimension, respectively. For the clean HSI estimation, it is performed in the 3-D FFT domain with $\mathcal{O}(TL(MNB)log(MNB))$. Note that the auxiliary variable $\mathcal{Z}$ and Lagrange multiplier $\mathcal{J}$ are very easy to be computed with simple pixel-to-pixel division and summation. Thus, we omit them from the time complexity analysis. The time complexity of the deep network is highly related to the feature size and channel. Here, we report that the total optimization time complexity of CSLRTR is $\mathcal{O}\left(TP^2G\left(g + (K+k)B^2 + K^2(B+b)\right)\right) + \mathcal{O}(TL(MNB)log(MNB))$.

Authorized licensed use limited to: XIDIAN UNIVERSITY. Downloaded on January 22,2024 at 04:18:52 UTC from IEEE Xplore. Restrictions apply.

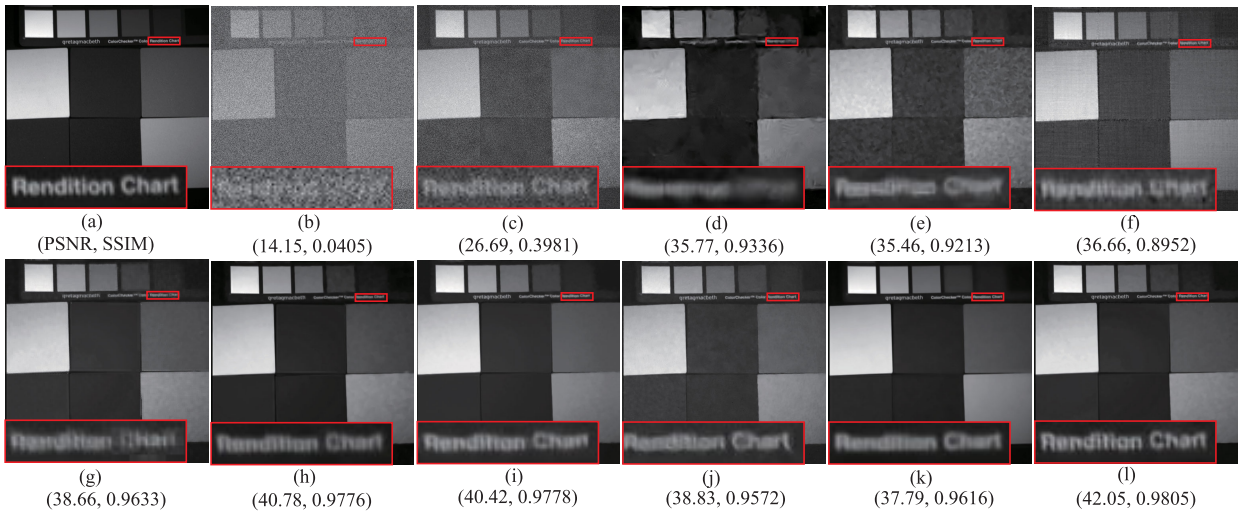


Fig. 4. Simulated noise removal results of sponges on CAVE dataset at band 7 under noise variance 50. (a) Original image (PSNR, SSIM). (b) Noisy image (14.15, 0.0405). Images denoised by (c) LRMR (26.69, 0.3981), (d) BM3D (35.77, 0.9336), (e) BM4D (35.46, 0.9213), (f) LRTDTV (36.66, 0.8952), (g) TDL (38.66, 0.9633), (h) KBRreg (40.78, 0.9776), (i) NGmeet (40.42, 0.9778), (j) QRNN3D (38.83, 0.9572), (k) HSI-DeNet (37.79, 0.9616), and (l) CSLRTR (42.05, 0.9805).

IV. Experiments and Discussion

A. Experimental Setting

Extensive experiments are conducted on multiple datasets to validate the effectiveness of the proposed CSLRTR. The experiments cover different HSI restoration tasks, including denoising, deblurring, and inpainting, and make quantitative and qualitative comparisons with several SOTA methods. Specifically, we test the proposed CSLRTR on several representative HSI datasets, including CAVE,[1] NTIRE2020,[2] Washington DC, PaviaU, Curprite, Harvard, and Indian Pines. The simulated experiments are evaluated on the NTIRE2020, CAVE, Washington DC, PaviaU, and Curprite, while the Harvard and Indian Pines datasets are used for the real experiments. For the SDLN, we select 32 images from the NTIRE2020 dataset as the training set and another 32 images as the test set. For the hyperparameters, w_j is the balance factor of the two symmetric streams, and we empirically set $w_2 = 0.5$ and $w_3 = 0.5$. K is the number of nonlocal similar patches, and we set $K = 100$. The noise level λ_i is automatically updated as introduced in [44].

The quantitative evaluation metrics are peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), spectral angle mapper (SAM), and Erreur Relative Globale Adimensionnelle de Synthese (ERGAS). PSNR and SSIM are two conventional indexes, which are used to evaluate the spatial similarity between the restored image and the reference image based on mean squared error (mse) and structural consistency, respectively. ERGAS measures the fidelity of the restored image based on the weighted sum of mse in each band. SAM is introduced to measure the spectral fidelity between the restored image and the reference image across all spatial positions. The PSNR and SSIM evaluate the spatial quality, and the ERGAS and SAM assess the spectral quality. The larger PSNR and SSIM values, the smaller ERGAS and SAM values, and the better the restored images. All experiments are implemented using MATLAB R2020a, an Intel i7 CPU at 3 GHz, 16-GB memory, and an Nvidia GeForce RTX3090 graphical processing unit (GPU). The source code would be released on the homepage of the authors.[3]

B. Evaluation on Different HSI Restoration Tasks

1) Denoising Task: We choose the representative model- and learning-based methods for comparison. The model-based methods include filtering-based methods (BM3D [74] and BM4D [7]), LRMR methods [10], and low-rank tensor recovery methods (TDL [22], KBRreg [26], LRTDTV [29], and NGmeet [34]); the learning-based methods are HSI-DeNet [59] and QRNN3D [67]. The simulated noise is the zero-mean additive white Gaussian noise and the noise variance between 10 and 50. Figs. 4–6 show the visual quality of three representative scenes denoised by competing methods. Specifically, Fig. 4 shows the denoising results of sponges image under noise variance 50 in the CAVE dataset. We could see that the proposed CSLRTR not only eliminates noise but also produces clearer and sharper text compared to other methods. Figs. 5 and 6 show the denoising results of ARAD_HS_0296 image in the NTIRE2020 and PaviaU datasets, respectively. Compared with other methods, from the zoomed areas, it could be observed that other methods lose image edge details, while the CSLRTR better preserves the structural information and gets a more visually pleasing result.

In addition, to assess the spectral restoration performance, we provide a detailed restoration comparison of each band in Fig. 7. The horizontal axis represents the band number, and the vertical axis represents the PSNR value. Here, the topmost red line corresponds to the proposed CSLRTR, indicating its superior ability to preserve the spectral–spatial structure correlation compared to other methods. To evaluate the results of image restoration more objectively, in Tables I and II, we list

[1] http://www.cs.columbia.edu/CAVE/databases/multispectral/
[2] https://data.vision.ee.ethz.ch/cvl/ntire20/
[3] https://owuchangyuo.github.io/

Authorized licensed use limited to: XIDIAN UNIVERSITY. Downloaded on January 22,2024 at 04:18:52 UTC from IEEE Xplore. Restrictions apply.

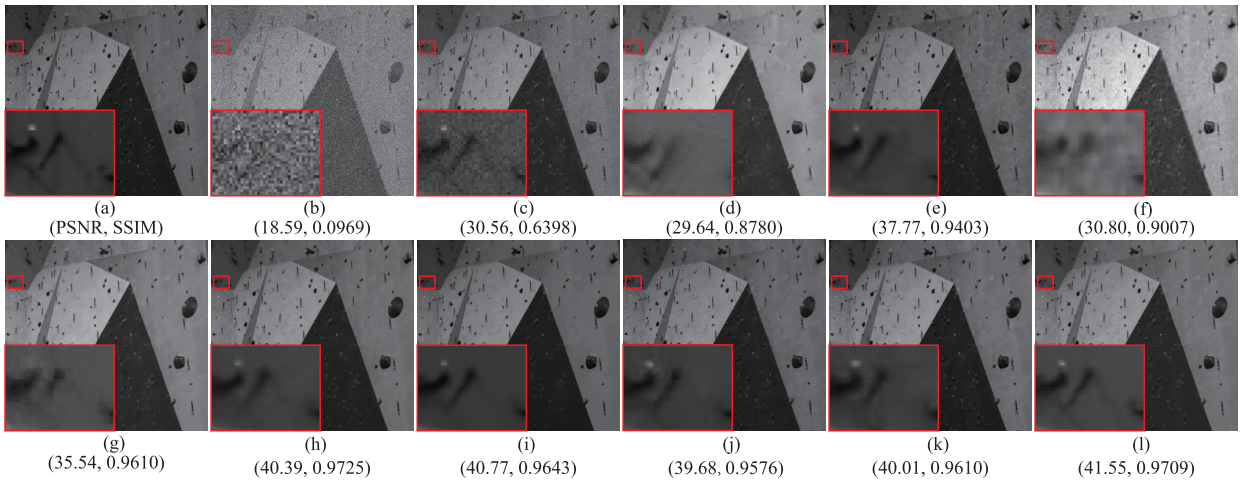


Fig. 5. Simulated noise removal results of *ARAD_HS_0296* on the NTIRE2020 dataset at band 10 under noise variance 30. (a) Original image (PSNR, SSIM). (b) Noisy image (18.59, 0.0969). Restored images of (c) LRMR (30.56, 0.6398), (d) BM3D (29.64, 0.8780), (e) BM4D (37.77, 0.9403), (f) LRTDTV (30.80, 0.9007), (g) TDL (35.54, 0.9610), (h) KBRreg (40.39, 0.9725), (i) NGmeet (40.77, 0.9643), (j) QRNN3D (39.68, 0.9576), (k) HSI-DeNet (40.01, 0.9610), and (l) CSLRTR (41.55, 0.9709).

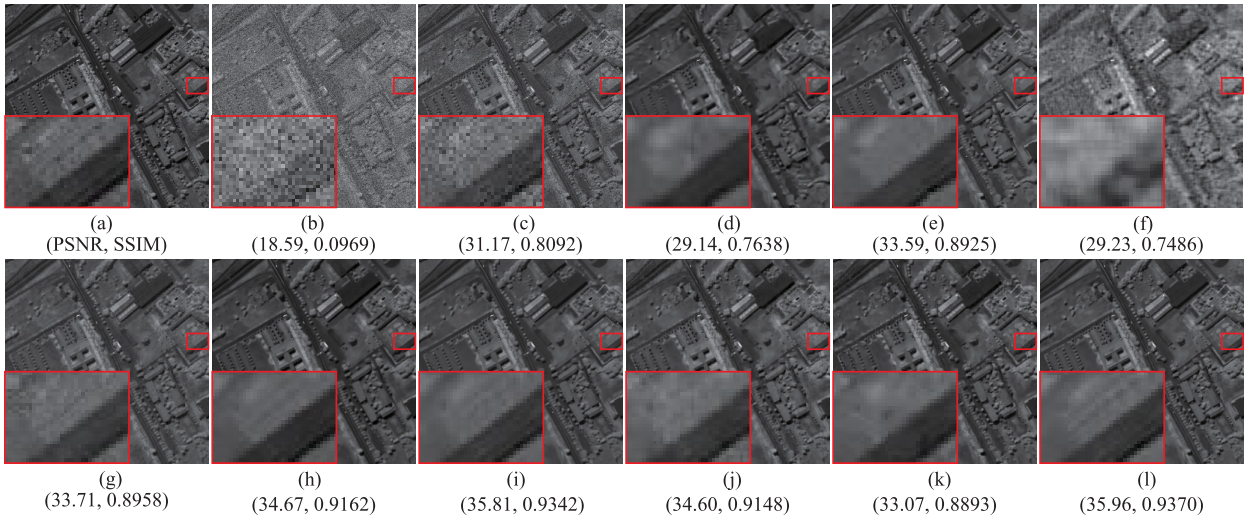


Fig. 6. Simulated noise removal results at band 80 on the Pavia University dataset under noise variance 30. (a) Original image (PSNR, SSIM). (b) Noisy image (18.59, 0.0969). Restored images of (c) LRMR (31.17, 0.8092), (d) BM3D (29.14, 0.7638), (e) BM4D (33.59, 0.8925), (f) LRTDTV (29.23, 0.7486), (g) TDL (33.71, 0.8958), (h) KBRreg (34.67, 0.9162), (i) NGmeet (35.81, 0.9342), (j) QRNN3D (34.60, 0.9148), (k) HSI-DeNet (33.07, 0.8893), and (l) CSLRTR (35.96, 0.9370).

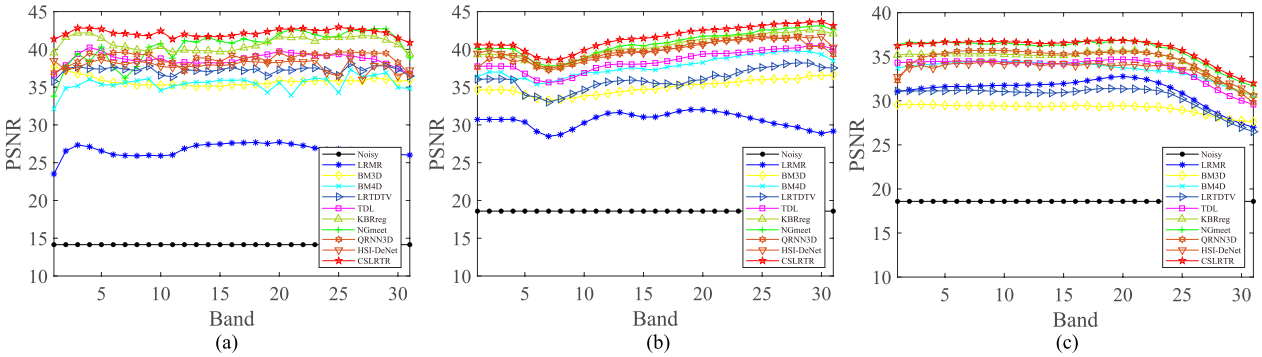


Fig. 7. PSNR values of each band. (a) Image of sponges under noise variance 50 on CAVE dataset. (b) Image of ARAD_HS_0296 under noise variance 30 on NTIRE2020 dataset. (c) Image of Pavia University under noise variance 30.

the quantitative values under five noise levels of competing methods on the CAVE and NTIRE2020 datasets. We can observe that the CSLRTR outperforms other methods in most cases, especially the value of PSNR and ERGAS.

Furthermore, to verify the generalization capability, we test CSLRTR on real measured datasets Harvard and Indian Pines. Since ground-truth clean images are not available for these datasets, quantitative metrics cannot be calculated. However,

Authorized licensed use limited to: XIDIAN UNIVERSITY. Downloaded on January 22,2024 at 04:18:52 UTC from IEEE Xplore. Restrictions apply.

TABLE I
QUANTITATIVE COMPARISON OF DIFFERENT METHODS UNDER VARIOUS NOISE LEVELS ON THE CAVE DATASET

Noise Level	Metrics	Methods										
		Noisy	LRMR	BM3D	BM4D	LRTDTV	TDL	KBRreg	NGmeet	QRNN3D	HSI-DeNet	CSLRTR
$\sigma^2 = 10$	PSNR ↑	28.13	39.28	42.14	44.60	40.91	44.28	45.15	46.69	43.50	44.47	**47.28**
	SSIM ↑	0.4410	0.9094	0.9665	0.9783	0.9510	0.9797	0.9792	0.9874	0.9728	0.9809	**0.9881**
	SAM ↓	0.1679	0.3343	0.1396	0.1295	0.1519	0.1025	0.1309	0.0852	0.1516	0.1122	**0.0800**
	ERGAS ↓	118.17	64.83	44.93	33.34	57.97	34.95	31.68	26.88	37.46	35.56	**24.61**
$\sigma^2 = 20$	PSNR ↑	22.11	34.36	38.48	41.02	39.50	41.05	42.05	43.39	42.02	41.93	**44.40**
	SSIM ↑	0.1818	0.7816	0.9339	0.9549	0.9380	0.9549	0.9619	0.9774	0.9669	0.9502	**0.9796**
	SAM ↓	0.3091	0.4945	0.1989	0.1980	0.1811	0.1285	0.1905	0.1077	0.1561	0.2592	**0.0978**
	ERGAS ↓	236.36	111.54	68.33	50.38	65.10	50.52	45.23	39.44	44.73	47.43	**34.17**
$\sigma^2 = 30$	PSNR ↑	18.59	31.38	36.28	38.79	37.79	39.03	40.38	41.62	40.54	40.07	**42.51**
	SSIM ↑	0.0969	0.6495	0.9026	0.9278	0.9116	0.9487	0.9468	0.9688	0.9569	0.9442	**0.9705**
	SAM ↓	0.4323	0.5907	0.2448	0.2593	0.2304	0.1519	0.2238	0.1219	0.1770	0.2369	**0.1139**
	ERGAS ↓	354.55	155.02	87.74	65.34	76.54	63.55	54.76	47.97	52.92	58.25	**42.52**
$\sigma^2 = 40$	PSNR ↑	16.09	29.15	34.68	37.15	36.13	37.58	39.27	39.43	39.44	38.30	**41.00**
	SSIM ↑	0.0599	0.5334	0.8734	0.8980	0.8719	0.9331	0.9432	**0.9593**	0.9452	0.9159	0.9583
	SAM ↓	0.5395	0.6680	0.2839	0.3125	0.2925	0.1972	0.1991	0.1471	0.1896	0.2559	**0.1389**
	ERGAS ↓	472.76	199.24	105.10	78.96	90.72	74.98	62.54	58.44	60.14	69.35	**50.54**
$\sigma^2 = 50$	PSNR ↑	14.15	27.40	33.78	35.87	34.62	36.42	37.90	38.87	38.51	37.13	**39.82**
	SSIM ↑	0.0405	0.4391	0.8609	0.8669	0.8219	0.9172	0.9192	0.9407	0.9204	0.8938	**0.9448**
	SAM ↓	0.6325	0.7284	0.2933	0.3575	0.3545	0.2000	0.2653	**0.1617**	0.1937	0.3178	0.1693
	ERGAS ↓	590.98	243.14	117.14	91.52	106.56	85.73	73.10	65.71	67.01	79.07	**58.04**

TABLE II
QUANTITATIVE COMPARISON OF DIFFERENT METHODS UNDER VARIOUS NOISE LEVELS ON THE NTIRE2020 DATASET

Noise Level	Metrics	Methods										
		Noisy	LRMR	BM3D	BM4D	LRTDTV	TDL	KBRreg	NGmeet	QRNN3D	HSI-DeNet	CSLRTR
$\sigma^2 = 10$	PSNR ↑	28.13	39.08	38.97	42.81	37.58	43.84	43.92	46.46	43.24	43.82	**46.57**
	SSIM ↑	0.4410	0.9362	0.9536	0.9786	0.9343	0.9834	0.9826	0.9903	0.9809	0.9842	**0.9907**
	SAM ↓	0.1679	0.0708	0.0601	0.0386	0.0493	0.0463	0.0301	0.0210	0.0381	0.0394	**0.0204**
	ERGAS ↓	118.17	41.77	40.86	26.07	51.66	24.15	23.25	17.21	24.88	35.68	**17.03**
$\sigma^2 = 20$	PSNR ↑	22.11	33.96	35.45	38.94	36.75	39.87	40.66	42.53	40.19	40.02	**42.71**
	SSIM ↑	0.1818	0.8236	0.9146	0.9545	0.9249	0.9637	0.9684	0.9780	0.9658	0.9665	**0.9799**
	SAM ↓	0.3091	0.1153	0.0884	0.0586	0.0612	0.0786	0.0418	**0.0280**	0.0502	0.0575	0.0299
	ERGAS ↓	236.36	75.32	61.23	40.68	55.52	38.02	33.84	27.13	35.39	44.65	**26.62**
$\sigma^2 = 30$	PSNR ↑	18.59	30.89	33.43	36.74	35.56	37.57	38.74	40.22	38.38	37.41	**40.48**
	SSIM ↑	0.0969	0.7059	0.8819	0.9320	0.9063	0.9445	0.9551	0.9654	0.9529	0.9389	**0.9690**
	SAM ↓	0.4323	0.1558	0.1111	0.0751	0.0785	0.0986	0.0483	0.0398	0.0589	0.0809	**0.0376**
	ERGAS ↓	354.55	106.69	77.18	52.37	62.17	49.55	42.24	35.46	43.66	49.41	**34.47**
$\sigma^2 = 40$	PSNR	16.09	28.61	31.95	35.21	34.26	36.03	37.22	38.80	37.06	36.37	**38.81**
	SSIM ↑	0.0599	0.6041	0.8520	0.9108	0.8774	0.9108	0.9416	0.9558	0.9394	0.9302	**0.9580**
	SAM ↓	0.5395	0.1985	0.1306	0.0897	0.0990	0.1143	0.0501	**0.0425**	0.0633	0.0890	0.0457
	ERGAS ↓	472.76	138.41	91.51	62.43	71.07	59.15	50.66	41.76	50.84	55.72	**41.69**
$\sigma^2 = 50$	PSNR ↑	14.15	26.84	31.09	34.03	33.00	34.03	36.12	37.54	36.09	35.15	**37.69**
	SSIM ↑	0.0405	0.5200	0.8367	0.8903	0.8396	0.9101	0.9309	0.9447	0.9270	0.9133	**0.9485**
	SAM ↓	0.6325	0.2378	0.1359	0.1034	0.1214	0.1251	0.0565	**0.0446**	0.0722	0.0976	0.0489
	ERGAS ↓	590.98	169.19	101.15	71.53	81.53	67.29	57.35	48.25	57.11	64.12	**47.44**

we can still compare the denoising performance of competing methods through visual observation. Figs. 8 and 9 show the denoising scene on the Harvard and Indian Pines datasets, respectively. From the zoomed areas, although the TDL preserves the image details, there are some noise residuals. The NGmeet method successfully removes noise but sacrifices certain image details. Compared with other methods, the CSLRTR effectively suppresses most of the noise while preserving the clearer image edges, thus achieving more visually pleasing results. These indicate that CSLRTR exhibits robustness and can adapt well to real and complex scenarios.

2) Deblurring Task: It refers to restore clear image from blurred one. We choose the fast positive deconvolution (FPD) [92], spectral–spatial TV (SSTV) [93], and weighted low-rank tensor recovery (WLRTR) method [43] as the comparison methods. In Fig. 10, we select three representative scenes under 17×17 Gaussian blur kernel to show the visual deblurring performance. We can observe that the results of FPD and SSTV remain relatively blurry, with a significant loss of image details. In contrast, both CSLRTR and WLRTR successfully restore clear image structures, making it challenging to distinguish between them visually. In terms of PSNR and SSIM values, CSLRTR still outperforms the WLRTR, which illustrates that the CSLRTR better exploits the intrinsic spectral–spatial correlation. To evaluate more objectively, we apply two Gaussian blur kernels (8×8 and 17×17) to the CAVE and NTIRE2020 datasets to obtain blur samples. The quantitative values on the CAVE dataset and NTIRE2020

Authorized licensed use limited to: XIDIAN UNIVERSITY. Downloaded on January 22,2024 at 04:18:52 UTC from IEEE Xplore. Restrictions apply.

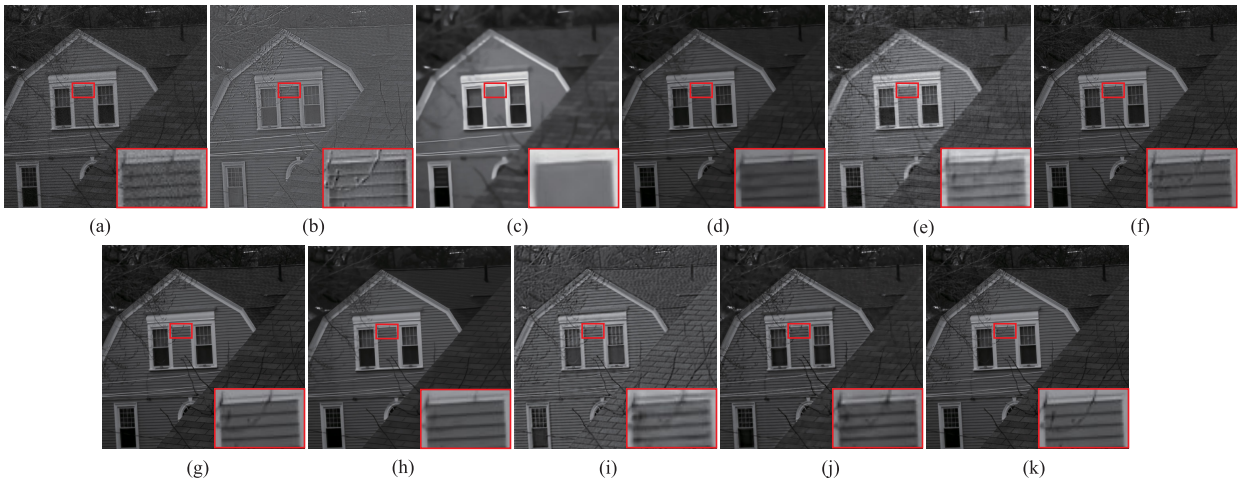


Fig. 8. Real noise removal results at band 1 on the Harvard dataset. (a) Original image. Restored images of (b) LRMR, (c) BM3D, (d) BM4D, (e) LRTDTV, (f) TDL, (g) KBRreg, (h) NGmeet, (i) QRNN3D, (j) HSI-DeNet, and (k) CSLRTR.

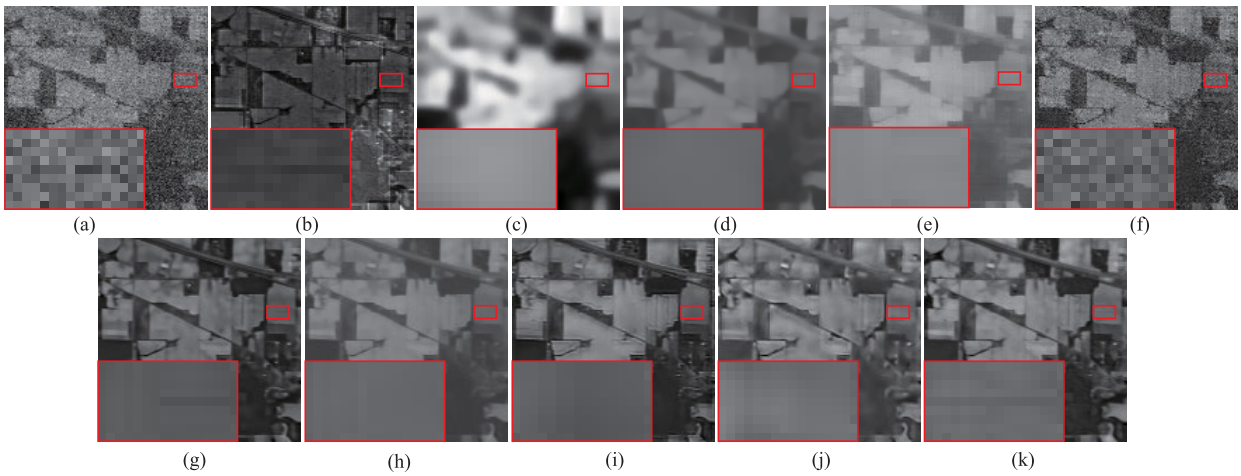


Fig. 9. Real noise removal results at band 145 on the Indian Pines dataset. (a) Original image. Restored images of (b) LRMR, (c) BM3D, (d) BM4D, (e) LRTDTV, (f) TDL, (g) KBRreg, (h) NGmeet, (i) QRNN3D, (j) HSI-DeNet, and (k) CSLRTR.

datasets are shown in Table III. It can be seen that the CSLRTR surpasses other methods over all metrics on two datasets. Besides, the CSLRTR and WLRTR outperform the FPD and SSTV by a large margin, which is because WLRTR and CSLRTR additionally utilize the spectral correlation.

3) Inpainting Task: Image inpainting is the restoration task of filling and restoring missing or damaged image regions. The competing methods include spectral–spatial total variational (SSTVI)-based inpainting [93], tensor-based AWTC [28], and WLRTR method. In Fig. 11, we select three representative scenes under 80% missing entries to show the visual inpainting performance of different methods. Both the SSTVI and AWTC result in a loss of image details. Similarly, both CSLRTR and WLRTR preserve the image details and achieve visually pleasing results, making it difficult to distinguish between them visually. However, the PSNR and SSIM values of the CSLRTR are still higher than those of WLRTR. For a more objective evaluation, we apply two random masks (20% missing entries and 80% missing entries) to the CAVE and NTIRE2020 datasets to construct missing samples. The quantitative values on the CAVE and NTIRE2020 datasets are shown in Table IV. We can observe that the CSLRTR outperforms other methods over all metrics.

C. Discussion

1) Ablation Analysis: In this article, the proposed CSLRTR contains the symmetric two-stream low-rank tensor model and the adaptive content-aware subspace strategy. To verify the impact of these components on the restoration results, we perform the ablation analysis study on the NTIRE2020 dataset with a noise level of 30. Table V lists the results for various situations. The first and second rows only exploit nonlocal similarity and spectral correlation, respectively, and the third row utilizes both nonlocal similarity and spectral correlation through the symmetric low-rank tensor model. The fourth and fifth rows combine the content-aware subspace strategy, which are one of two streams of the proposed CSLRTR, and the sixth row is the proposed CSLRTR.

Here, we have three findings. First, from the comparison results of the first three rows, we find that both nonlocal similarity and spectral correlation are indispensable, especially the spectral correlation. The first row only utilizes nonlocal similarity, and its spectral distortion is most severe based on the SAM and ERGAS metrics. Second, due to the combination of adaptive content-aware subspace strategy, the fourth and fifth rows utilize both nonlocal similarity and spectral correlation. Compared with the third rows, the restoration

Authorized licensed use limited to: XIDIAN UNIVERSITY. Downloaded on January 22,2024 at 04:18:52 UTC from IEEE Xplore. Restrictions apply.

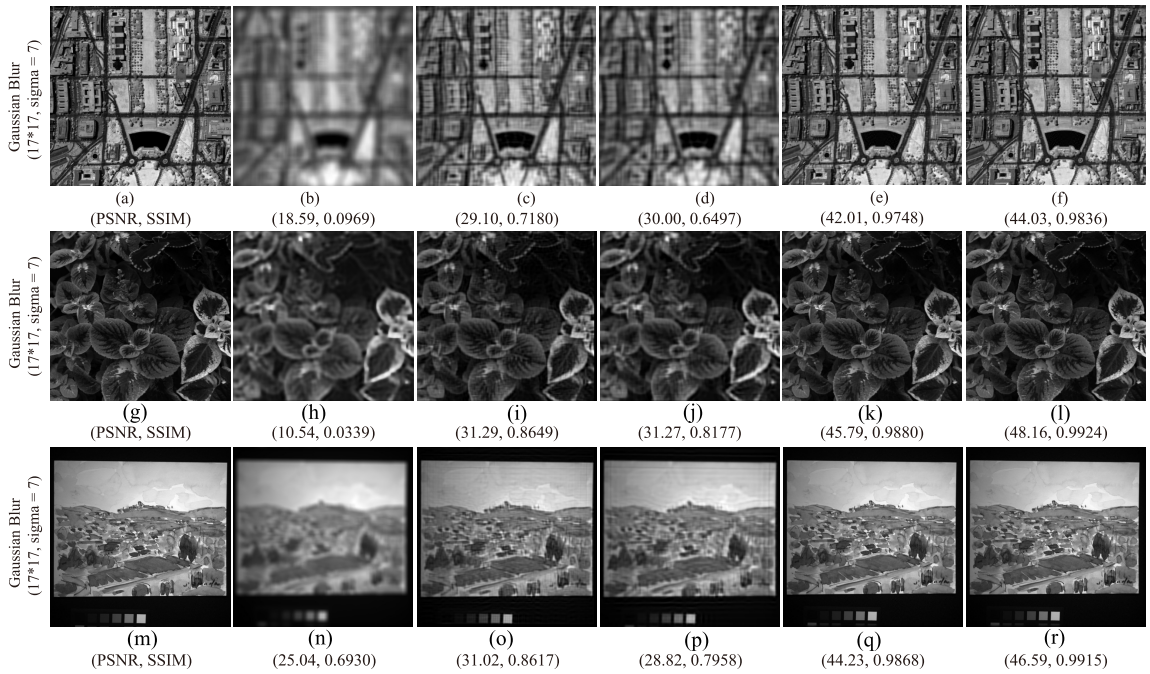


Fig. 10. Deblurring results of different datasets. (a), (g), and (m) Original image. (b) Blurred image (18.59, 0.0969). (c) FPD (29.10, 0.7180). (d) SSTV (30.00, 0.6497). (e) WLRTR (42.01, 0.9748). (f) CSLRTR (44.03, 0.9836). (h) Blurred image (10.54, 0.0339). (i) FPD (31.29, 0.8649). (j) SSTV (31.27, 0.8177). (k) WLRTR (45.79, 0.9880). (l) CSLRTR (48.16, 0.9924). (n) Blurred image (25.04, 0.6930). (o) FPD (31.02, 0.8617). (p) SSTV (28.82, 0.7958). (q) WLRTR (44.23, 0.9868). (r) CSLRTR (46.59, 0.9915).

TABLE III
QUANTITATIVE COMPARISON OF DIFFERENT METHODS UNDER SEVERAL BLUR CASES ON THE CAVE AND NTIRE2020 DATASETS

Deblur		CAVE				NTIRE2020			
		FPD	SSTV	WLRTR	CSLRTR	FPD	SSTV	WLRTR	CSLRTR
Gaussian (8×8, Sigma = 3)	PSNR ↑	38.84	37.61	55.81	**57.48**	37.65	36.31	55.67	**57.99**
	SSIM ↑	0.9617	0.9527	0.9980	**0.9986**	0.9555	0.9332	0.9982	**0.9989**
	SAM ↓	0.0734	0.0658	0.0308	**0.0259**	0.0194	0.0189	0.0045	**0.0038**
	ERGAS ↓	68.48	80.91	10.01	**8.26**	49.00	58.70	6.70	**5.17**
Gaussian (17×17, Sigma = 7)	PSNR ↑	33.16	33.08	49.69	**51.32**	28.27	31.06	48.09	**50.41**
	SSIM ↑	0.9114	0.8944	0.9930	**0.9953**	0.8886	0.8410	0.9913	**0.9946**
	SAM ↓	0.1163	0.0989	0.0518	**0.0455**	0.0381	0.0337	0.0096	**0.0087**
	ERGAS ↓	125.11	129.84	20.54	**16.96**	142.56	105.85	15.83	**12.12**

results have a significant improvement across all metrics, which demonstrates that the adaptive content-aware subspace strategy could better utilize the low-rank property and verify its effectiveness. Third, compared with the third and fourth rows, the restoration results of the sixth row have shown further improvement. This is attributed to the fusion of the two streams, which strengthens the spectral–spatial correlation. The results demonstrate that the symmetric two-stream low-rank tensor model could have a positive impact on restoration.

2) Comparison of Running Time With Different Methods: In Table VI, we report the running time of various competing methods, including tensor-based methods (LLRT [27], OLRT [44], and KBRreg [26]) and the subspace-based method (NGmeet [34]), on an image with different sizes. The time for CSLRTR does not include the time spent on MATLAB calling the network model. We observe that NGmeet has the fastest running speed. This is because NGmeet performs subspace transformation on the entire image, while CSLRTR performs subspace transformation on each image patch and includes symmetric subspace tensor recovery streams. Therefore, CSLRTR is slower than NGmeet in terms of time. However, compared to LLRT, OLRT, and KBRreg, which do not employ subspace strategies, CSLRTR demonstrates a significant improvement in running speed. It is nearly twice as fast as LLRT, three times faster than OLRT, and ten times faster than KBRreg. This clearly demonstrates that the introduction of subspace strategies effectively reduces the time complexity.

3) Effectiveness of Adaptive Subspace Learning: In Fig. 12, we show the comparison results between the conventional fixed subspace and the proposed adaptive learning subspace. We simulate noisy samples with noise level 30 on the NTIRE2020 dataset and set different fixed subspace dimensions for the spectral and nonlocal dimensions as [3, 5, 7, 9, 11, 13, 15, 17]. The red blocks represent the restoration results of adaptive content-aware subspace strategy, and the other colored blocks represent the restoration results of the fixed subspace dimension. With the increase of subspace dimension, the restoration effect shows a trend of first getting better and then getting worse. Also, our adaptive content-aware subspace strategy gets the best restoration results on all metrics, which indicates that the obtained subspace dimension by network is more accurate, contributing to more compact low-rank representation for better restoration.

Authorized licensed use limited to: XIDIAN UNIVERSITY. Downloaded on January 22,2024 at 04:18:52 UTC from IEEE Xplore. Restrictions apply.

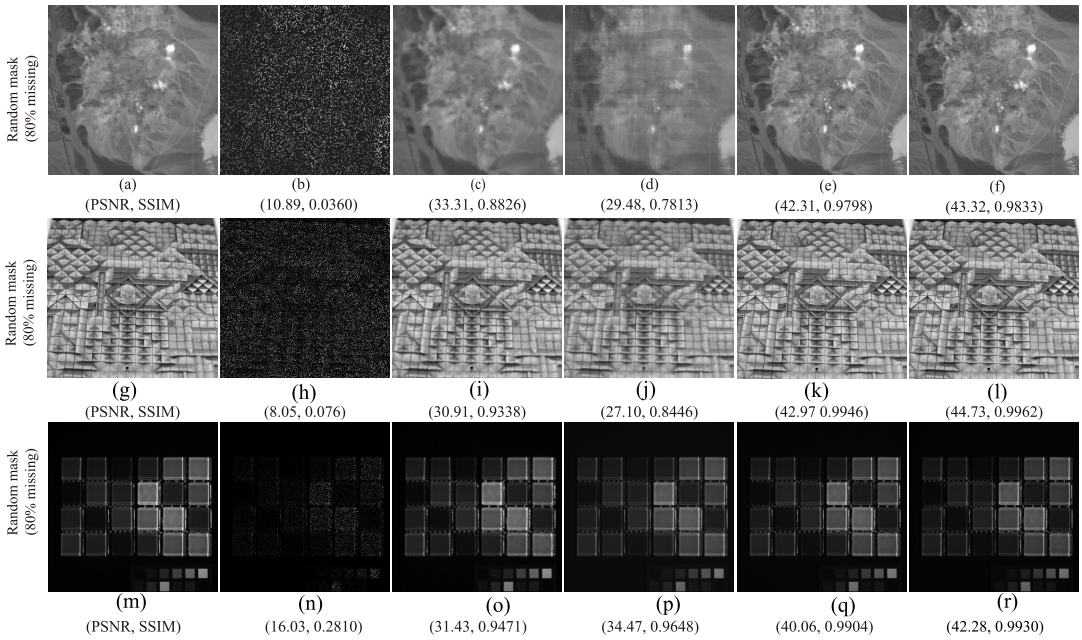


Fig. 11. Inpainting results of different datasets. (a), (g), and (m) Original Image (PSNR, SSIM). (b) Missing image (10.89, 0.0360). (c) SSTVI (33.31, 0.8826). (d) AWTC (29.48, 0.7813). (e) WLRTR (42.31, 0.9798). (f) CSLRTR (43.32, 0.9833). (h) Missing image (8.05, 0.076). (i) SSTVI (30.91, 0.9338). (j) AWTC (27.10, 0.8446). (k) WLRTR (42.97 0.9946). (l) CSLRTR (44.73, 0.9962). (n) Missing image (16.03, 0.2810). (o) SSTVI (31.43, 0.9471). (p) AWTC (34.47, 0.9648). (q) WLRTR (40.06, 0.9904). (r) CSLRTR (42.28, 0.9930).

TABLE IV

QUANTITATIVE COMPARISON OF DIFFERENT METHODS UNDER MISSING INFORMATION CASES ON THE CAVE AND NTIRE2020 DATASETS

Inpainting		CAVE				NTIRE2020			
		SSTVI	AWTC	WLRTR	CSLRTR	SSTVI	AWTC	WLRTR	CSLRTR
Random mask (20% missing entries)	PSNR ↑	46.93	50.65	59.68	**60.63**	46.39	53.77	65.31	**69.63**
	SSIM ↑	0.9947	0.9962	0.9994	**0.9995**	0.9955	0.9983	0.9998	**0.9999**
	SAM ↓	0.0449	0.0401	0.0219	**0.0209**	0.0147	0.0100	0.0028	**0.0018**
	ERGAS ↓	27.64	18.83	5.94	**5.28**	18.00	8.14	2.01	**1.21**
Random mask (80% missing entries)	PSNR ↑	38.24	36.12	46.51	**48.76**	37.27	33.52	47.41	**49.41**
	SSIM ↑	0.9674	0.9247	0.9928	**0.9957**	0.9618	0.9063	0.9944	**0.9971**
	SAM ↓	0.0936	0.1474	0.0619	**0.0548**	0.0321	0.0630	0.0151	**0.0109**
	ERGAS ↓	73.94	100.17	28.66	**22.21**	51.79	82.88	15.96	**12.82**

TABLE V

ABLATION OF DIFFERENT COMPONENTS

Nonlocal	Spectral	Subspace	PSNR ↑ / SSIM ↑ / SAM ↓ / ERGAS ↓
✓			32.96 / 0.7859 / 0.0899 / 80.20
	✓		38.28 / 0.9520 / 0.0526 / 44.23
✓	✓		39.58 / 0.9617 / 0.0438 / 38.23
✓		✓	40.29 / 0.9683 / 0.0391 / 35.11
	✓	✓	40.27 / 0.9682 / 0.0406 / 35.23
✓	✓	✓	**40.47 / 0.9690 / 0.0376 / 34.47**

TABLE VI

RUNNING TIME COMPARISON WITH SOTA METHODS (SECONDS)

Resolution	Methods				
	KBRreg	OLRT	LLRT	NGmeet	CSLRTR
128*128*31	85.63	26.25	17.55	4.36	11.23
256*256*31	347.21	103.54	77.34	11.82	32.47
512*512*31	1435.03	387.69	298.12	47.33	138.04

4) Comparison of Subspace Dimension With NGmeet: Here, we compare our adaptive subspace learning with the handcrafted linear subspace updating in NGmeet [34]. For fairness, we only select one stream of the spectral subspace-based spatial low-rank tensor model to compare with NGmeet. In Fig. 13, we randomly select ten images from the CAVE dataset. For CSLRTR, we average the subspace dimensions for each iteration of the different patches and images as shown in Fig. 13(a), with the error bars indicating the variance of subspace dimensions among the ten images. In addition, since NGmeet applies the same subspace dimension strategy to each image in the CAVE dataset, the variance of subspace dimensions is zero for different images. Fig. 14(b) represents the variations of PSNR values during each iteration for NGmeet and CSLRTR, with the error bars corresponding to the variance of PSNR values among the ten images.

It can be observed that, with the increase of iteration, the variation characteristics of subspace dimension is nonlinear and the overall trend is gradually increasing, which is reasonable, because the noise would be gradually removed with the iteration. The difference is that NGmeet linearly updates the subspace dimension, while the proposed methods shows a nonlinear increasing curve. In Fig. 13(b), we observe that the PSNR value of the CSLRTR is higher than that of NGmeet. This result could support the effectiveness of our adaptive subspace learning.

Authorized licensed use limited to: XIDIAN UNIVERSITY. Downloaded on January 22,2024 at 04:18:52 UTC from IEEE Xplore. Restrictions apply.

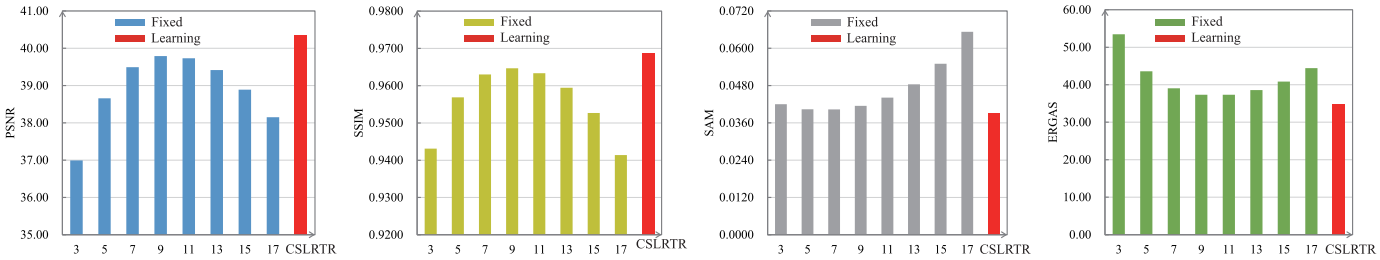


Fig. 12. Comparison between fixed and adaptive subspace strategy. The horizontal axis denotes the subspace dimension ranging from 3 to 17. The last red column means the proposed adaptive subspace dimension. The larger the PSRN and SSIM, the smaller the SAM and ERGAS, and the better the performance.

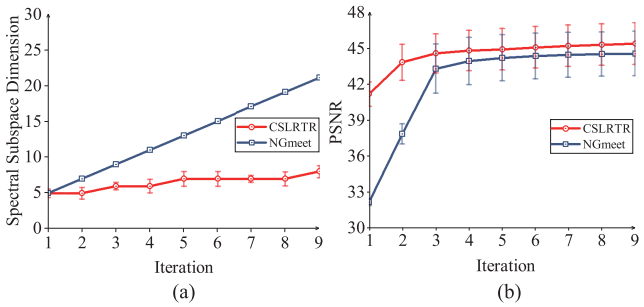


Fig. 13. Comparison of the subspace dimension between the CSLRTR and the NGmeet in the iterative denoising process on the CAVE dataset under noise variance 20. (a) Variation of spectral subspace dimension during iteration. (b) Variation of PSNR value during iteration.

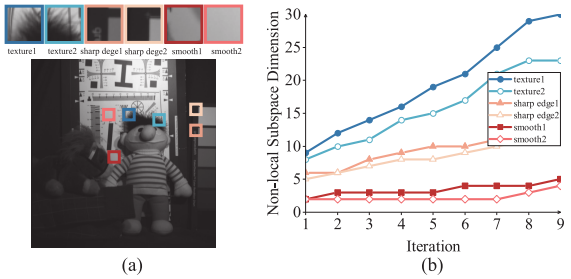


Fig. 14. Illustration of learned subspace dimension variation of different HSI patches. (a) Different spatial image structures, including smooth, texture, and sharp edge structures. (b) Variation of nonlocal subspace dimension during the iteration of each patch in (a).

5) How Does Adaptive Subspace Learning Work: To further analyze whether the proposed adaptive subspace learning network has satisfactorily characterized the subspace dimension, we specially select the typical different structure regions, including texture region, sharp edged region, and smooth region. Fig. 14 shows how the subspace dimension varies with the increase of iterations for different HSI patches. We can observe that the subspace dimensions increase with the growth of iterations for all contents. The obvious difference is the more complex of the region, the larger the subspace dimension. This is very reasonable since the complex texture would possess larger subspace than that of the smooth content in essence. This phenomenon strongly demonstrates that the proposed method has indeed learned the low-rank discrepancy of different HSI contents.

6) Extend the SDLN to SOTA Methods: To further validate the universality of our main idea of modeling the discrepancy, we embed the proposed SDLN into the previous SOTA methods: KBRreg [26], WLRTR [43], OLRT [44], and NGmeet [34]. Note that the SDLN is a very flexible module

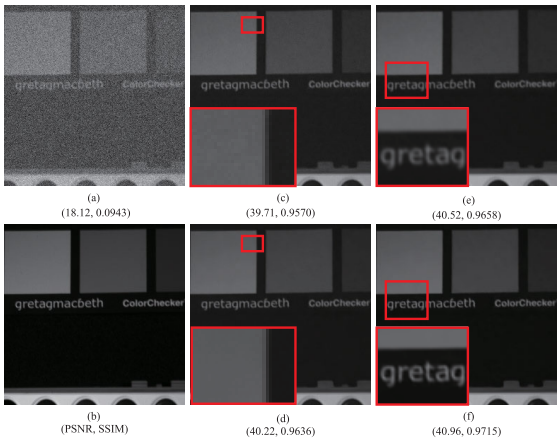


Fig. 15. Effectiveness of SDLN model on the SOTA HSI restoration methods KBRreg and NGmeet by modifying their models with SDLN. (a) Noisy image (18.12, 0.0943). (b) Original Image (PSNR, SSIM). (c) KBRreg (39.71, 0.9570). (d) KBRreg+ (40.22, 0.9636). (e) NGmeet (40.52, 0.9658). (f) NGmeet+ (40.96, 0.9715).

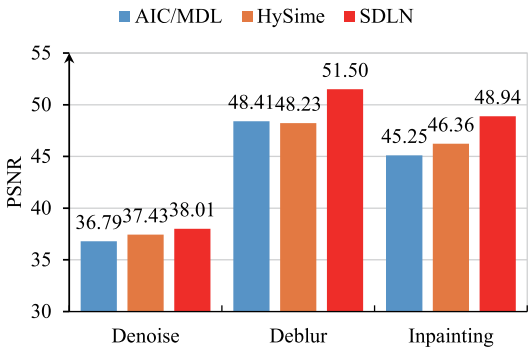


Fig. 16. Comparison results with conventional subspace dimension estimation methods on different HSI restoration tasks.

and can be very easily extended to previous methods and further improve their performances. In Table VII, we present the recovery results by applying the SDLN to other methods on the CAVE dataset ($256 \times 256 \times 31$ for fast testing). "+" denotes the improved method with SDLN. We can observe that after introducing SDLN, a noticeable improvement for all methods in the recovery metrics can be gained. Note that NGmeet is a subspace tensor-based method where the subspace strategy is applied to the entire image, which does not consider the discrepancy between different regions. After introducing SDLN, the recovery results are improved by 0.42 dB. This strongly highlights the effectiveness of SDLN to model the discrepancy of different HSI patches.

We also show the visualization results in Fig. 15. On one hand, comparing Fig. 15(c) with Fig. 15(d), the SDLN can

Authorized licensed use limited to: XIDIAN UNIVERSITY. Downloaded on January 22,2024 at 04:18:52 UTC from IEEE Xplore. Restrictions apply.

TABLE VII
UNIVERSALITY OF THE PROPOSED SDLN WHEN EMBED INTO OTHER METHODS ON THE CAVE DATASET

Noise Level	Metrics	Methods								
		Noisy	KBRreg	KBRreg+	WLRTR	WLRTR+	OLRT	OLRT+	NGmeet	NGmeet+
$\sigma^2 = 30$	PSNR ↑	18.23	39.35	39.71	39.80	40.26	40.17	40.78	40.25	40.67
	SSIM ↑	0.0961	0.9528	0.9586	0.9597	0.9645	0.9611	0.9677	0.9614	0.9667
	SAM ↓	0.4441	0.1215	0.0965	0.0871	0.0831	0.0912	0.0834	0.0913	0.0842
	ERGAS ↓	357.36	57.59	55.93	54.42	53.42	53.65	49.80	52.81	51.29

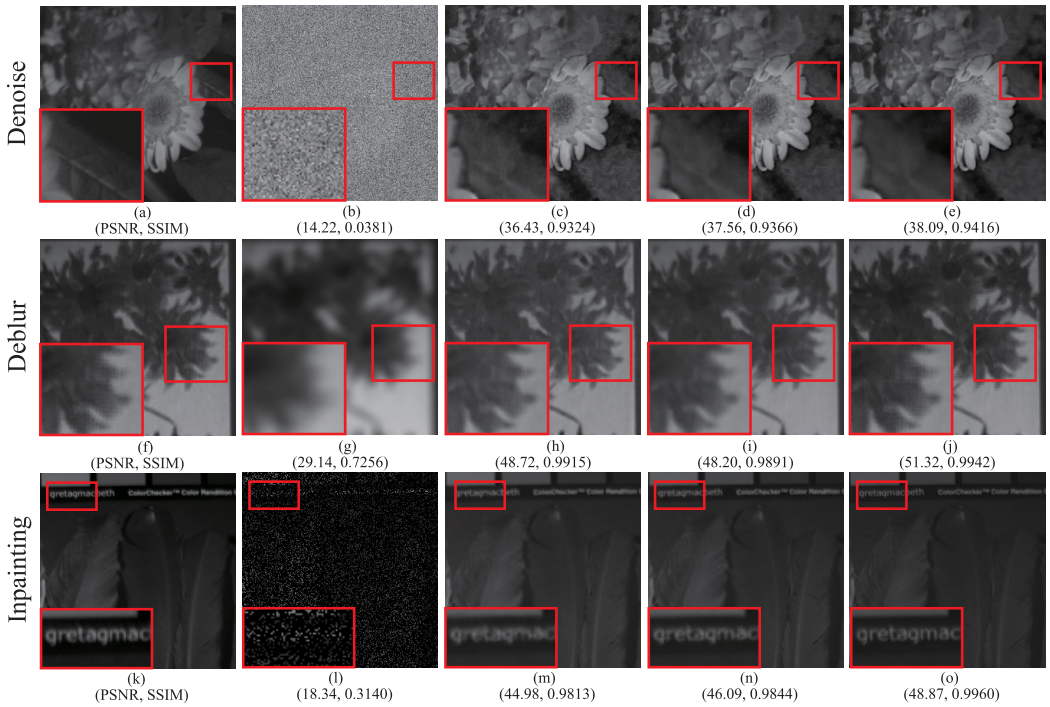


Fig. 17. Visualization comparison results of conventional subspace estimation methods AIC/MDL, HySime, and SDLN on different HSI restoration tasks. (a), (f), and (k) Original image (PSNR, SSIM). (b) Noisy image (14.22, 0.0381). (c) AIC/MDL (36.43, 0.9324). (d) HySime (37.56, 0.9366). (e) SDLN (38.09, 0.9416). (g) Blurred image (29.14, 0.7256). (h) AIC/MDL (48.72, 0.9915). (i) HySime (48.20, 0.9891). (j) SDLN (51.32, 0.9942). (l) Missing image (18.34, 0.3140). (m) AIC/MDL (44.98, 0.9813). (n) HySime (46.09, 0.9844). (o) SDLN (48.87, 0.9960).

help remove the slight residual noise with smoothness; on the other hand, comparing Fig. 15(e) with Fig. 15(f), the SDLN can facilitate the image structure preserving. Both the quantitative and qualitative results have demonstrated that SDLN could perceive fine-grained discrepancy for compact feature representation and improve their performance.

7) Comparison With Conventional Subspace Dimension Estimation Methods: To further evaluate the advantage of the proposed SDLN over the conventional methods, we present the comparative results of the most commonly used subspace dimension estimation methods (HySime [58], AIC [84], and MDL [85]) under different image restoration tasks with the same setting. In Fig. 16, for the denoising task, it can be observed that SDLN outperforms HySime and AIC/MDL (AIC for nonlocal subspace dimension and MDL for spectral subspace dimension) with a PSNR improvement of 0.58 and 1.22 dB, respectively. For the deblurring task, SDLN achieves PSNR improvements of 3.27 and 3.09 dB compared to HySime and AIC/MDL, respectively. For the inpainting task, SDLN achieves PSNR improvements of 2.58 and 3.69 dB compared to HySime and AIC/MDL, respectively. In Fig. 17, we show their visualization comparison results. The visual results of SDLN are consistently better than those of other methods in terms of noise suppression and image details

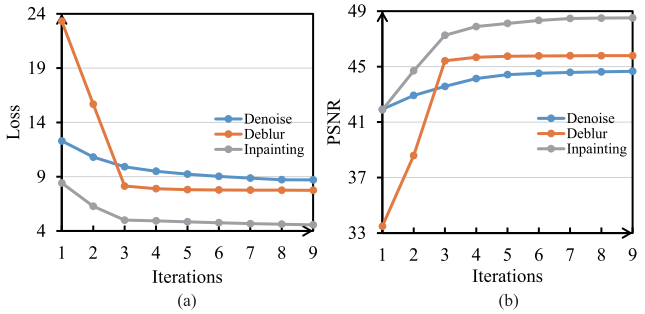


Fig. 18. Empirical convergence curve on different HSI restoration tasks. (a) Functional energy curve. (b) PSNR value curve.

recovery. These results indicate that SDLN is more robust to the degradations and provides more accurate subspace dimension estimation for better HSI restoration.

8) Convergence Analysis: In Fig. 18, we provide the energy function energy curve and the variation of PSNR during the iteration of the CSLRTR on different HSI restoration tasks. It can be observed that the loss value curve monotonically decreases and the PSNR value curve monotonically increases in less than nine steps. The results consistently follow the similar trends, which demonstrates the reliable convergence property of the proposed method in practice.

Authorized licensed use limited to: XIDIAN UNIVERSITY. Downloaded on January 22,2024 at 04:18:52 UTC from IEEE Xplore. Restrictions apply.

V. Conclusion

The model-based methods have achieved great effect in HSI restoration. However, the discrepancy of different HSI patches is easily ignored. In this work, we hypothesize that the high-level semantic understanding about the content is of great benefit for low-level restoration tasks and propose a content-aware subspace low-rank tensor restoration model by leveraging both deep network and low-rank tensor model. Also, the analysis experiments have suggested that the high-level information could help the model perceive fine-grained discrepancy of different HSI patches for more compact feature representation with better restoration performance. Moreover, this novel combination of optimization model and deep network makes the model strictly obey the physical degradation procedure, thereby endowing the model with a strong generalization ability. In addition, the proposed symmetric two restoration streams could further enhance the spectral–spatial coupling and preserve the spectral–spatial structural correlation. Numerous experiments on different restoration tasks verify the effectiveness of the proposed CSLRTR. The proposed CSLRTR could be easily extended to other methods and would bring a new perspective on how to jointly utilize deep learning and optimization paradigm.

References

[1] 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, Mar. 2018.

[2] X. Wu, D. Hong, J. Tian, J. Chanussot, W. Li, and R. Tao, "ORSIm detector: A novel object detection framework in optical remote sensing imagery using spatial-frequency channel features," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 5146–5158, Jul. 2019.

[3] X. Cao, L. Xu, D. Meng, Q. Zhao, and Z. Xu, "Integration of 3-dimensional discrete wavelet transform and Markov random field for hyperspectral image classification," *Neurocomputing*, vol. 226, pp. 90–100, Feb. 2017.

[4] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, and J. Chanussot, "Graph convolutional networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 5966–5978, Jul. 2020.

[5] 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.

[6] 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.

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

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

[9] 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.

[10] 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.

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

[12] X. Cao et al., "Low-rank matrix factorization under general mixture noise distributions," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1493–1501.

[13] 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.

[14] Y. Xie, Y. Qu, D. Tao, W. Wu, Q. Yuan, and W. Zhang, "Hyperspectral image restoration via iteratively regularized weighted Schatten p-norm minimization," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4642–4659, Aug. 2016.

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

[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, Dec. 2017.

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

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

[19] Y. Hu, X. Li, Y. Gu, and M. Jacob, "Hyperspectral image recovery using nonconvex sparsity and low-rank regularizations," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 532–545, Jan. 2020.

[20] 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. 4457–4469, Jul. 2019.

[21] T. Xie, S. Li, and B. Sun, "Hyperspectral images denoising via nonconvex regularized low-rank and sparse matrix decomposition," *IEEE Trans. Image Process.*, vol. 29, pp. 44–56, 2020.

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

[23] 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 Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 442–449.

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

[25] Q. Xie et al., "Multispectral images denoising by intrinsic tensor sparsity regularization," in *Proc. IEEE Conf. CVPR*, Jun. 2016, pp. 1692–1700.

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

[27] Y. Chang, L. Yan, and S. Zhong, "Hyper-Laplacian regularized unidirectional low-rank tensor recovery for multispectral image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4260–4268.

[28] M. K. 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, Jun. 2017.

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

[30] T. Xie, S. Li, L. Fang, and L. Liu, "Tensor completion via nonlocal low-rank regularization," *IEEE Trans. Cybern.*, vol. 49, no. 6, pp. 2344–2354, Jun. 2019.

[31] Y. Chen, T. Huang, and X. Zhao, "Destriping of multispectral remote sensing image using low-rank tensor decomposition," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 12, pp. 4950–4967, Dec. 2018.

[32] T. Ji, N. Yokoya, X. X. Zhu, and T. Huang, "Nonlocal tensor completion for multitemporal remotely sensed images' inpainting," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3047–3061, Jun. 2018.

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

[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/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 6868–6877.

Authorized licensed use limited to: XIDIAN UNIVERSITY. Downloaded on January 22,2024 at 04:18:52 UTC from IEEE Xplore. Restrictions apply.

[35] J. Xue, Y. Zhao, W. Liao, and J. C. Chan, "Nonlocal low-rank regularized tensor decomposition for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 5174–5189, Jul. 2019.

[36] 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/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 10182–10191.

[37] 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, Sep. 2019.

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

[39] 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, Oct. 2019.

[40] 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, Sep. 2019.

[41] 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, Dec. 2019.

[42] L. Zhuang and M. K. Ng, "Hyperspectral mixed noise removal by ℓ_1-norm-based subspace representation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 1143–1157, 2020.

[43] Y. Chang, L. Yan, X.-L. Zhao, H. Fang, Z. Zhang, and S. Zhong, "Weighted low-rank tensor recovery for hyperspectral image restoration," *IEEE Trans. Cybern.*, vol. 50, no. 11, pp. 4558–4572, Nov. 2020.

[44] Y. Chang, L. Yan, B. Chen, S. Zhong, and Y. Tian, "Hyperspectral image restoration: Where does the low-rank property exist," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 8, pp. 6869–6884, Aug. 2021.

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

[46] 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.

[47] 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.*, vol. 31, no. 11, pp. 4747–4760, Nov. 2020.

[48] Y. Chen et al., "Hyperspectral image denoising using factor group sparsity-regularized nonconvex low-rank approximation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5515916.

[49] L. Zhuang, X. Fu, M. K. Ng, and J. M. Bioucas-Dias, "Hyperspectral image denoising based on global and nonlocal low-rank factorizations," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 12, pp. 10438–10454, Dec. 2021.

[50] M. Wang, Q. Wang, and J. Chanussot, "Tensor low-rank constraint and l_0 total variation for hyperspectral image mixed noise removal," *IEEE J. Sel. Top. Sign. Proces.*, vol. 15, no. 3, pp. 718–733, Apr. 2021.

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

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

[53] J. Lin, T.-Z. Huang, X.-L. Zhao, T.-X. Jiang, and L. Zhuang, "A tensor subspace representation-based method for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 9, pp. 7739–7757, Sep. 2021.

[54] C. He, L. Sun, W. Huang, J. Zhang, Y. Zheng, and B. Jeon, "TSLRLN: Tensor subspace low-rank learning with non-local prior for hyperspectral image mixed denoising," *Signal Process.*, vol. 184, Jul. 2021, Art. no. 108060.

[55] X. Shen, W. Bao, and K. Qu, "Subspace-based preprocessing module for fast hyperspectral endmember selection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 3386–3402, 2021.

[56] S. Huang, H. Zhang, and A. Pižurica, "Subspace clustering for hyperspectral images via dictionary learning with adaptive regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5524017.

[57] Y. Yang, Y. Feng, J. Zhang, and S. Chen, "Hyperspectral image restoration via subspace-based nonlocal low-rank tensor approximation," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[58] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.

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

[60] 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.

[61] 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, 2020.

[62] 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/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 1585–1594.

[63] W. Dong, H. Wang, F. Wu, G. Shi, and X. Li, "Deep spatial–spectral representation learning for hyperspectral image denoising," *IEEE Trans. Comput. Imag.*, vol. 5, no. 4, pp. 635–648, Dec. 2019.

[64] 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.

[65] W. Liu and J. Lee, "A 3-D atrous convolution neural network for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5701–5715, Aug. 2019.

[66] A. Maffei, J. M. Haut, M. E. 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, Apr. 2020.

[67] 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.

[68] 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.

[69] S. Zhang, L. Wang, L. Zhang, and H. Huang, "Learning tensor low-rank prior for hyperspectral image reconstruction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 12001–12010.

[70] Q. Shi, X. Tang, T. Yang, R. Liu, and L. Zhang, "Hyperspectral image denoising using a 3-D attention denoising network," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 12, pp. 10348–10363, Dec. 2021.

[71] X. Cao, X. Fu, C. Xu, and D. Meng, "Deep spatial–spectral global reasoning network for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5504714.

[72] H. Zhang, H. Chen, G. Yang, and L. Zhang, "LR-Net: Low-rank spatial–spectral network for hyperspectral image denoising," *IEEE Trans. Image Process.*, vol. 30, pp. 8743–8758, 2021.

[73] H. Chen, G. Yang, and H. Zhang, "Hider: A hyperspectral image denoising transformer with spatial–spectral constraints for hybrid noise removal," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Oct. 31, 2022, doi: 10.1109/TNNLS.2022.3215751.

[74] 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.

[75] A. Buades, B. Coll, and J.-M. Morel, "Non-local means denoising," *Image Process. Line*, vol. 1, pp. 208–212, Sep. 2011.

[76] L. Pang, W. Gu, and X. Cao, "TRQ3DNet: A 3D quasi-recurrent and transformer based network for hyperspectral image denoising," *Remote Sens.*, vol. 14, no. 18, p. 4598, Sep. 2022.

[77] M. Li, Y. Fu, and Y. Zhang, "Spatial–spectral transformer for hyperspectral image denoising," 2022, *arXiv:2211.14090*.

[78] F. Wang, J. Li, Q. Yuan, and L. Zhang, "Local–global feature-aware transformer based residual network for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5546119.

[79] F. Xiong, J. Zhou, S. Tao, J. Lu, J. Zhou, and Y. Qian, "SMDS-Net: Model guided spectral–spatial network for hyperspectral image denoising," *IEEE Trans. Image Process.*, vol. 31, pp. 5469–5483, 2022.

Authorized licensed use limited to: XIDIAN UNIVERSITY. Downloaded on January 22,2024 at 04:18:52 UTC from IEEE Xplore. Restrictions apply.

[80] D. Liu et al., "An efficient unfolding network with disentangled spatial–spectral representation for hyperspectral image super-resolution," *Inf. Fusion*, vol. 94, pp. 92–111, Jun. 2023.
[81] X.-L. Zhao, W.-H. Xu, T.-X. Jiang, Y. Wang, and M. K. Ng, "Deep plug-and-play prior for low-rank tensor completion," *Neurocomputing*, vol. 400, pp. 137–149, Aug. 2020.
[82] L. Zhuang, M. K. Ng, and X. Fu, "Hyperspectral image mixed noise removal using subspace representation and deep CNN image prior," *Remote Sens.*, vol. 13, no. 20, p. 4098, Oct. 2021.
[83] Z. Lai, K. Wei, and Y. Fu, "Deep plug-and-play prior for hyperspectral image restoration," *Neurocomputing*, vol. 481, pp. 281–293, Apr. 2022.
[84] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. AC-19, no. 6, pp. 716–723, Dec. 1974, doi: 10.1109/TAC.1974.1100705.
[85] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.
[86] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[87] T.-W. Lee, T. Wachtler, and T. J. Sejnowski, "The spectral independent components of natural scenes," in *Proc. Int. Workshop Biologically Motivated Comput. Vis.* Cham, Switzerland: Springer, 2000, pp. 527–534.
[88] A. Lam, I. Sato, and Y. Sato, "Denoising hyperspectral images using spectral domain statistics," in *Proc. 21st Int. Conf. Pattern Recognit. (ICPR)*, Nov. 2012, pp. 477–480.
[89] Z. Lin, R. Liu, and Z. Su, "Linearized alternating direction method with adaptive penalty for low-rank representation," in *Proc. NIPS*, vol. 24, 2011, pp. 1–9.
[90] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, Jan. 2010.
[91] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2862–2869.
[92] S. Henrot, C. Soussen, and D. Brie, "Fast positive deconvolution of hyperspectral images," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 828–833, Feb. 2013.
[93] 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. 4, pp. 384–395, Jul. 2017.



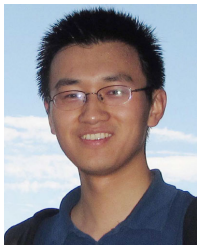
Xueyao Xiao received the B.S. degree in electronic and information engineering from Hunan University, Changsha, China, in 2021. He is currently pursuing the M.S. degree with the School of Artificial Intelligence and Automation, Huazhong University of Science and Technology (HUST), Wuhan, China.

His research interests include hyperspectral image processing, structural noise removal, and image enhancement.



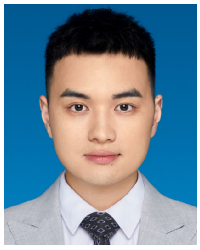
Wei Zhang (Member, IEEE) received the B.S. degree in measurement technology and instrumentation and the Ph.D. degree in instrument science and technology from Chongqing University, Chongqing, China, in 2012 and 2016, respectively.

From 2017 to 2020, he was a Post-Doctoral Fellow with the College of Optoelectronic Engineering, Chongqing University. He is currently an Assistant Researcher with the Artificial Intelligence Center, Peng Cheng Laboratory, Shenzhen, China. His research interests include optical fiber sensing, radar/lidar, image processing, neuromorphic vision, and optimization.



Yi Chang (Member, IEEE) received the B.S. degree in automation from the University of Electronic Science and Technology of China, Chengdu, China, in 2011, and the M.S. degree in pattern recognition and intelligent systems and the Ph.D. degree in control science and engineering from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2014 and 2019, respectively.

From 2014 to 2015, he was a Research Assistant with Peking University, Beijing, China. From 2019 to 2021, he was a Post-Doctoral Researcher with Peng Cheng Laboratory, Shenzhen, China. He is currently a Lecturer with the School of Artificial Intelligence and Automation, HUST. His research interests include image processing, computer vision, and machine learning.



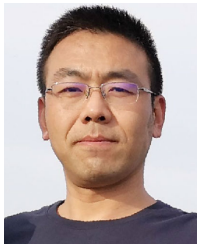
Shuning Cao received the B.S. degree in automation from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2018, and the M.S. degree in control engineering from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2020, where he is currently pursuing the Ph.D. degree with the School of Artificial Intelligence and Automation.

His research interests include image deblurring, structural noise removal, and multispectral image processing.



Wei He (Senior Member, IEEE) received the B.S. degree from the School of Mathematics and Statistics, Wuhan University, Wuhan, China, in 2012, and the Ph.D. degree from the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University, in 2017.

From 2018 to 2020, he was a Researcher with the Geoinformatics Unit, RIKEN Center for Advanced Intelligence Project, Tokyo, Japan, where he was a Research Scientist from 2020 to 2021. He is currently a Full Professor with LIESMARS, Wuhan University. His research interests include image quality improvement, remote sensing image processing, low-rank (LR) representation, and deep learning.



Houzhang Fang (Member, IEEE) received the Ph.D. degree in control science and engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2014.

He is currently an Associate Professor with the School of Computer Science and Technology, Xidian University, Xi'an, China. His research interests include target recognition and image restoration.



Luxin Yan (Member, IEEE) received the B.S. degree in electronic communication engineering and the Ph.D. degree in pattern recognition and intelligence system from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2001 and 2007, respectively.

He is currently a Professor with the School of Artificial Intelligence and Automation, HUST. His research interests include multispectral image processing, pattern recognition, and real-time embedded systems.

Authorized licensed use limited to: XIDIAN UNIVERSITY. Downloaded on January 22,2024 at 04:18:52 UTC from IEEE Xplore. Restrictions apply.