Fast Large-Scale Hyperspectral Image Denoising via Noniterative Low-Rank Subspace Representation

Yong Chen, Jinshan Zeng, *Member, IEEE*, Wei He, *Senior Member, IEEE*, Xi-Le Zhao, *Member, IEEE*,
Tai-Xiang Jiang, and Qing Huang

***Abstract*— Denoising of hyperspectral image (HSI) is challenging, especially when dealing with large-scale data. Model-based methods show promise in HSI denoising due to their good generalization, but they suffer from computational complexity due to complex priors [like nonlocal self-similarity (NSS)] and iterations, resulting in low efficiency for large-scale HSI processing. To address these challenges, we propose a fast large-scale HSI denoising (FallHyDe) method based on noniterative low-rank (LR) subspace representation to enjoy high denoising efficiency, effectiveness, and flexibility simultaneously. By leveraging the global spectral property of HSI, FallHyDe efficiently estimates spectral subspace and spatial representation coefficients (SRCs) from the observed noisy HSI, reducing computation complexity caused by the high spectral dimension during processing. In addition, we innovatively explore the presence of high signal-to-noise ratio bands (HSNRBs) in real HSI, enabling fast SRC estimation through a least squares problem without relying on complex priors and iterations. FallHyDe requires neither iteration nor parameter tuning, enabling our method to process large-scale HSI denoising quickly and flexibly. Experimental results on both simulated and real HSI datasets demonstrate that our proposed method not only achieves competitive results in quality but also speeds up the restoration by more than ten times than the representative fast HSI denoising methods. The code is available at https://chenyong1993.github.io/yongchen.github.io/.**

***Index Terms*— Fast denoising, hyperspectral image (HSI) denoising, large-scale HSI, low-rank (LR) subspace representation.**

Received 31 July 2024; revised 28 August 2024; accepted 9 September 2024. Date of publication 11 September 2024; date of current version 26 September 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 62101222, Grant 12471483, Grant 42271370, Grant 62376110, Grant 12371456, Grant 12171072, and Grant 62131005; in part by the Natural Science Foundation of Jiangxi, China, under Grant 20232ACB212001 and Grant 20224ACB212004; in part by the Young Elite Scientists Sponsorship Program by Jiangxi Association for Science and Technology (JXAST) under Grant 2023QT12; in part by Sichuan Science and Technology Program under Grant 2024NSFJQ0038 and Grant 2023ZYD0007; in part by the Fund of Hubei Key Laboratory of Inland Shipping Technology under Grant NHHY2023003; and in part by the National Key Research and Development Program of China under Grant 2020YFA0714001. *(Corresponding author: Wei He.)*

Yong Chen, Jinshan Zeng, and Qing Huang are with the School of Computer and Information Engineering, Jiangxi Normal University, Nanchang 330022, China (e-mail: chenyong1872008@163.com; jinshanzeng@jxnu.edu.cn; qh@jxnu.edu.cn).

Wei He is with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430072, China (e-mail: weihe1990@whu.edu.cn).

Xi-Le Zhao is with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: xlzhao122003@163.com).

Tai-Xiang Jiang is with the School of Economic Information Engineering, Southwestern University of Finance and Economics, Chengdu 611130, China (e-mail: taixiangjiang@gmail.com).

Digital Object Identifier 10.1109/TGRS.2024.3458395

I. INTRODUCTION

HYPERSPECTRAL image (HSI) is a 3-D data cube with two spatial dimensions and one spectral dimension. It is collected by a high spectral resolution sensor and has hundreds or thousands of spectral channels. HSI provides abundant spatial and spectral information, making it valuable for various applications, such as recognition [1], unmixing [2], and classification [3]. However, the physical mechanism, sensor sensitivity, and imaging conditions often introduce Gaussian noise, degrading the HSI's quality and limiting its usability in downstream applications [4]. Therefore, HSI denoising is a crucial preprocessing step to improve image quality and enable more effective use in various applications [5].

To date, numerous methods have been proposed for HSI denoising, which can be roughly divided into two categories: deep learning-based approaches and model-based approaches [6]. Deep learning methods, including supervised and unsupervised learning, leverage the powerful learning ability of deep neural networks (DNNs) for HSI denoising. Supervised learning methods learn a nonlinear mapping between noisy and clean HSIs using abundant training data [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19]. Classical supervised approaches use deep convolutional neural networks (DCNNs) [20]. For example, Chang et al. [7] introduced a DCNN with 2-D convolutions and residual learning to learn the spatial and spectral features of HSI. Yuan et al. [8] developed a combined 2-D spatial and 3-D spatial–spectral CNN to exploit the high correlation of adjacent bands. Following that, a modified 3-D U-net to handle 3-D filtering by decomposing it into 2-D spatial and 1-D spectral filtering [10], [21]. To address long-range dependencies, attention mechanisms and transformer networks have been applied to HSI denoising [13], [14], [15], [22]. Chen et al. [13] incorporated multihead global spectral attention and locally enhanced cross-spatial attention into a 3-D transformer architecture. Although these methods provide effective denoising, they lack interpretability. Model-guided network approaches aim to enhance interpretability. Bodrito et al. [16] designed a spatial–spectral network by unraveling an iterative sparse coding model. Xiong et al. [17] introduced a model-guided interpretable network by unfolding a subspace-based sparse model, capturing spectral–spatial correlation and spatial sparsity priors.

Supervised learning methods require a large amount of paired training data, which is often limited and leads to poor generalization for different types of noise. On the contrary, unsupervised learning methods offer the advantage of not needing training data while still leveraging the power of DNNs [23], [24], [25], [26], [27], [28]. Sidorov and Hardeberg [23] extended the unsupervised deep image prior (DIP) to HSI denoising, and Miao et al. [24] proposed an unsupervised denoising method based on low-rank (LR) matrix decomposition of HSI within the DIP framework. These approaches improve generalization but still suffer from limited interpretability. Moreover, unsupervised methods require many iterations for optimizing network parameters, leading to slow processing efficiency.

Model-based methods treat HSI denoising as an ill-posed problem and utilize prior information to regularize the solution space. Common priors include local smoothness (LS), LR, and nonlocal self-similarity (NSS). The LS prior exploits the high consistency of local areas and adjacent bands in HSI, resulting in similar pixel values. Total variation (TV) is a popular regularization method that captures prior LS in spatial and spectral dimensions [29], [30]. However, it ignores the strong spectral correlation in HSI, which is based on the low-dimensional subspace property of spectral signatures [31], [32], [33], [34]. To address this, model-based methods incorporate LR prior using techniques such as LR matrix approximation (LRMA) [35], [36], LR matrix factorization (LRMF) [37], [38], [39], and LR tensor factorization (LRTF) [40], [41], [42], [43]. Zhang et al. [35] first proposed an LR matrix recovery (LRMR) method, which reshapes the original 3-D HSI into a 2-D matrix and employs the matrix nuclear norm to capture the spectral correlation of HSI. Chen et al. [37] integrated the mixture of Gaussian noise assumption strategy into the LRMF. To capture the spatial-spectra correlation, Zheng et al. [42] introduced three-directional LRTF to handle different correlations along different modes of HSI. Using LR prior alone may ignore local features, leading to suboptimal denoising results. Combining LS and LR priors effectively preserves global spatial–spectral correlations and local spatial–spectral smoothness, resulting in improved denoising performance.

Combining LS and LR priors effectively preserves global spatial–spectral correlations and local spatial–spectral smoothness, resulting in improved denoising performance. He et al. [44] first incorporated band-by-band TV into the LRMA to improve the denoising results. Subsequently, various methods have been introduced by integrating TV regularization into LR prior, such as LRTDTV [45], LLRSSTV [46], SLRTFL0-TV [47], RCTV [48], and WNLRATV [49]. To address the staircase effect caused by TV regularization, enhanced TV methods were proposed to characterize the LS prior and embed it into the LR prior [50]. In addition, the NSS prior captures similarities among full-band patches in nonlocal regions, utilizing complementary information effectively [51], [52]. The LR+NSS framework, which applies the LR penalty to each aggregated nonlocal patch group, is widely used for HSI denoising and achieves state-of-the-art results [53]. Peng et al. [54] introduced the nonlocal tensor dictionary learning (TDL) method that captures NSS and spectral correlation. Similarly, NSS prior embedded in LRTF has been proposed for HSI denoising, with representative methods KBR [55], LLRT [56], and WLRTR [57]. Despite their promising results, these methods may still suffer from time-consuming computations because of the large spatial–spectral dimensions of the image.

A. Bottleneck of Large-Scale HSI Denoising

In summary, numerous HSI denoising methods have been proposed as mentioned above, but most face a bottleneck: difficulty in handling the computational burden of large-scale HSI denoising while ensuring efficiency, effectiveness, and flexibility. Therefore, the challenge of large-scale HSI denoising is how to meet all three criteria.

Efficiency: Supervised learning methods provide instantaneous results through end-to-end HSI denoising after training. However, achieving the desired performance of deep networks requires sufficient training data, leading to long training times. Moreover, supervised learning is heavily dependent on high-performance GPUs and their memory. On the other hand, unsupervised learning methods save training costs, but require thousands of iterations for parameter optimization, resulting in time-consuming processes. For example, we evaluated the QRNN3D[1] [9] and DS2DP[2] [24] methods on the Washington dc Mall data in Table I.[3] We can observe that QRNN3D performs efficiently on small-scale data, but becomes impractical for large-scale data. Similarly, DS2DP's efficiency decreases gradually as data dimensions increase. Model-based methods employ LS, LR, and NSS priors for HSI denoising, utilizing iterative optimization algorithms. However, the efficiency of these priors at each iteration is highly dependent on the dimensions of the HSI. The computational complexity of these priors increases with dimensions and iterations. Thus, model-based methods face low denoising efficiency, especially for large-scale HSI.

Effectiveness: Deep learning-based approaches can lead to satisfactory denoising results, given sufficient training data, time, suitable test data, and high-performance GPU. However, meeting all these requirements simultaneously in real scenes is challenging, leading to limited denoising performance. Model-based approaches explore the priors of the original HSI, but the effectiveness of different prior combinations varies significantly. Adding more priors may boost denoising accuracy but also increase model complexity and reduce efficiency.

Flexibility: Deep learning-based approaches are often specialized for specific data and noise conditions. They may struggle to effectively denoise test data and noise scenarios that were not present in the training data. Model-based methods combine multiple priors under a unified framework, resulting in many regularization parameters that need adjustment. The optimal parameter selection for different datasets and noise intensities is inconsistent, limiting flexibility in the denoising process across various scenes.

---

[1] https://github.com/Vandermode/QRNN3D

[2] The code is available from the authors.

[3] The experiments are run on a platform with Inter i9-12900K and NVIDIA GeForce RTX 3090.

TABLE I

TIME COST (S) OF A STATE-OF-THE-ART SUPERVISED LEARNING METHOD QRNN3D AND UNSUPERVISED LEARNING METHOD DS2DP ON THE WDC WITH DIFFERENT DIMENSIONS. **N/A: OUT OF MEMORY**

Dimension ($* \times 307 \times 191$)	100	300	500	700	900	1100	1280
QRNN3D	1.80	2.52	3.70	5.33	N/A	N/A	N/A
DS2DP	499.82	1363.81	2115.77	2934.34	3796.06	4804.88	5094.02

B. Related Works

Recently, integrating LRMF and NSS prior has shown remarkable capabilities in HSI denoising. This type of method utilizes LRMF to capture global spectral correlation and reduce the spectral dimension of HSI, followed by applying NSS prior to denoise the reduced-dimensionality spatial representation coefficients (SRCs). The LRMF significantly reduces the computational burden caused by high spectral dimensions, while NSS ensures denoising accuracy. Zhuang and Bioucas-Dias [58] used LRMF to represent HSI as the product of spectral basis and SRCs, where the spectral basis is learned from the noisy HSI, and SRCs are denoised using BM3D [59]. To explore the nonlocal LR prior to SRCs, Zhuang et al. [60] further employed LRTF to regularize nonlocal similar 3-D patches. He et al. [61] developed a framework called "non-local meets global" for HSI denoising, combining spectral LR and spatial NSS prior. Despite their balanced efficiency, effectiveness, and flexibility, these methods are mainly suitable for small-scale HSI processing. The computational complexity of NSS increases exponentially with the image spatial dimension, making them unsuitable for large-scale HSI processing. In addition, the impressive results rely on the precision of the similarity block matching step.

C. Contributions

The goal of this work is to develop an efficient method for fast large-scale HSI denoising (FallHyDe) that meets these three criteria of high efficiency, effectiveness, and flexibility. The method is inspired by the fact that HSI can be well approximated by low-dimensional subspace representation, which is captured by LRMF, effectively exploring the strong spectral correlation and reducing computational costs. Specifically, the spectral basis is learned from the observed noisy HSI using singular value decomposition (SVD). Previous methods mainly relied on the observed noisy HSI for the fidelity term, but they overlooked the presence of high signal-to-noise ratio bands (HSNRBs) in the observed HSI (see Fig. 1) [62], leading to underutilization of data. This observation prompted us to investigate whether we could use the HSNRBs to guide fast and accurate restoration of SRCs. To achieve instantaneous restoration of SRCs, we propose to discard prior regularization that involves iterative solutions and has high computational complexity due to the high spatial dimension. Instead, we restore the SRCs directly from the HSNRBs. The denoised HSI is obtained through the composition of spectral basis estimation and SRCs without the need for iteration. The framework of the proposed FallHyDe is illustrated in Fig. 2. The main contributions of the proposed method compared to previous approaches are as follows.

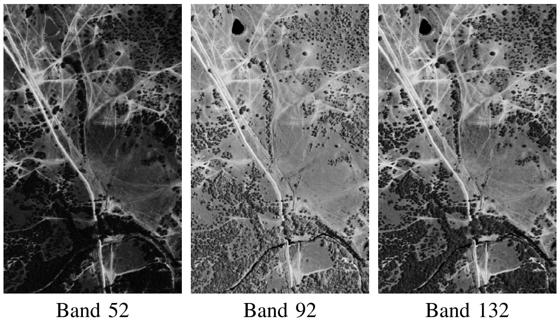


Fig. 1. Illustration of some HSNRBs in real Terrain HSI dataset.

1) *Efficiency:* Our proposed method efficiently addresses the computational complexity challenges posed by the high spatial and spectral dimensions of HSI. By fully exploiting the strong spectral correlation through LR subspace representation, we learn the spectral basis from the observed noisy HSI and quickly restore the SRCs from the HSNRBs using a noniterative least squares approach. As a result, HSI denoising can be implemented within seconds. Experimental results demonstrate that the restoration process takes less than 1 s for small-scale datasets (e.g., $200 \times 200 \times 80$) and less than 10 s for large-scale datasets (e.g., $1280 \times 307 \times 191$).
2) *Effectiveness:* Experimental results on simulated and real datasets confirm the effectiveness of the proposed method compared to other state-of-the-art HSI denoising methods.
3) *Flexibility:* The proposed method exhibits remarkable flexibility, as it does not require training data and does not rely on fine-tuning of regularization parameters. In addition, we provide theoretical proof of the method's feasibility. Consequently, it is highly flexible and can be effectively applied to large-scale HSI denoising.

The remainder of this article is organized as follows. Section II presents the problem of HSI denoising. Section III introduces the proposed FallHyDe method step by step. Section IV shows the experimental results and a discussion of the model. Section V concludes this article.

Notation: In this article, we represent scalars as lowercase letters or capital letters (e.g., b or B), vectors as lowercase bold letters (e.g., $\mathbf{x}$), matrices as capitalized boldface letters (e.g., $\mathbf{X}$), and tensors as capitalized calligraphic letters (e.g., $\mathcal{X}$). The mode-n unfolding of N-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \cdots \times I_N}$ is represented by $\mathbf{X}_{(n)} \in \mathbb{R}^{I_n \times I_1 \cdots I_{n-1} I_{n+1} \cdots I_N}$, and the inverse operator of mode-n unfolding is defined as $\text{Fold}_n(\mathbf{X}_{(n)}) = \mathcal{X}$. The Frobenius norm of $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \cdots \times I_N}$ is calculated by

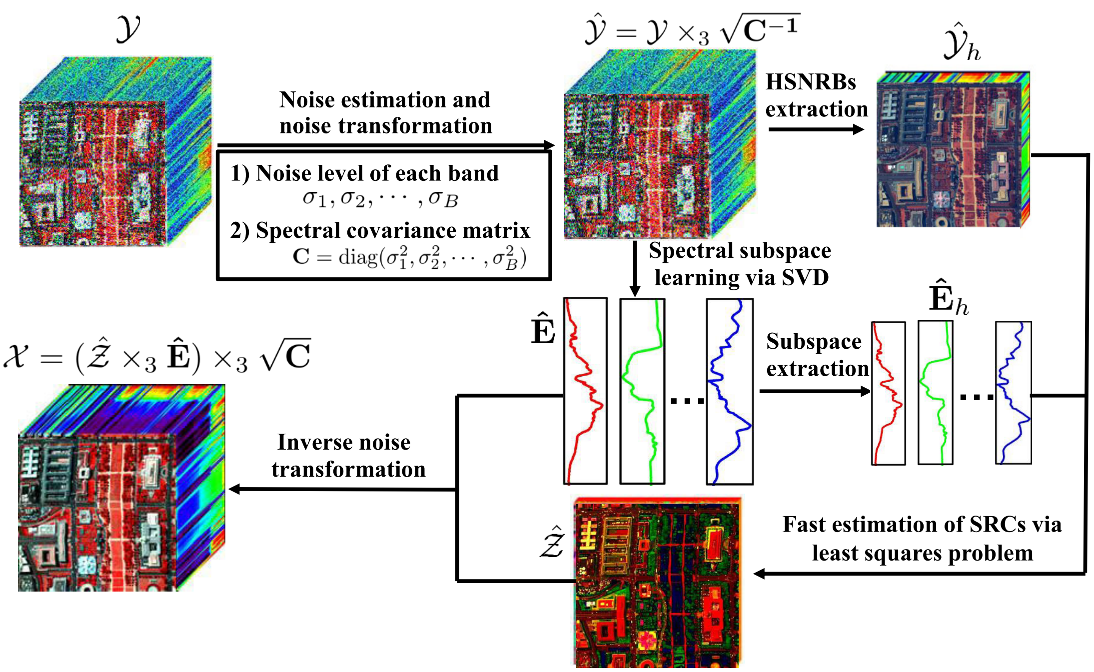


Fig. 2. Framework of the proposed FallHyDe method.

$\|\mathcal{X}\|_F = (\sum_{i_1}\sum_{i_2}\cdots\sum_{i_N} x^2_{i_1 i_2 \cdots i_N})^{(1/2)}$. The mode-$n$ product of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ and a matrix $\mathbf{A} \in \mathbb{R}^{J_n \times I_n}$ is defined as $\mathcal{X} \times_n \mathbf{A}$, where $(\mathcal{X} \times_n \mathbf{A}) \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times J_n}$ and $(\mathcal{X} \times_n \mathbf{A})_{i_i, \ldots, i_{n-1}, j_n, i_{n+1}, \ldots, i_N} = \sum_{i_n} \mathcal{X}_{i_1, \ldots, i_N} \cdot \mathbf{A}_{j_n, i_n}$.

II. FORMULATION AND LOW-RANK SUBSPACE REPRESENTATION OF HSI

In this section, we first present the HSI degradation model. Then, we introduce the LR subspace representation framework for HSI denoising.

A. Problem Formulation

Let $\mathcal{Y} \in \mathbb{R}^{M \times N \times B}$ represent observed noisy HSI with $M \times N$ pixels and B bands. The observed HSI can be described using an additive degradation model

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

where $\mathcal{X} \in \mathbb{R}^{M \times N \times B}$ and $\mathcal{N} \in \mathbb{R}^{M \times N \times B}$ represents the underlying clean HSI and additive Gaussian noise, respectively. The objective of HSI denoising is to restore the underlying clean HSI $\mathcal{X}$ from the observed HSI $\mathcal{Y}$.

B. Low-Rank Subspace Representation of HSI

The restoration of large-scale HSI faces challenges due to their high spatial and spectral dimensions, leading to extensive computational complexity to directly restore clean HSI $\mathcal{X}$. To address this, a linear spectral mixing model is used, representing each spectral signature (that is, the tube of $\mathcal{X}$) as a linear combination of several end-member basis, capturing the high spectral correlation in HSI [31]. Thus, the LR subspace representation of spectral vectors can be employed to approximate the HSI $\mathcal{X}$

$$\mathcal{X} = \mathcal{Z} \times_3 \mathbf{E} \tag{2}$$

where $\mathbf{E} \in \mathbb{R}^{B \times r}$ ($r \ll B$) represents the spectral subspace with an orthogonal constraint and $\mathcal{Z} \in \mathbb{R}^{M \times N \times r}$ represents the SRCs.

The LR subspace representation of HSI has several advantages: 1) it effectively preserves the strong spectral correlation in HSI, which is essential for subsequent HSI applications; 2) the computation complexity is significantly reduced since the representation works in a lower dimensional space compared to the original space of $\mathcal{X}$, resulting in faster denoising processing; and 3) the restoration results tend to be more accurate than when performed directly in the original HSI. This is because the number of unknown variables to be solved is significantly reduced [63]. As a result, the restoration of the clean HSI $\mathcal{X}$ can be achieved by estimating the spectral subspace $\mathbf{E}$ and SRCs $\mathcal{Z}$. The estimation of these components can be formulated as follows:

$$\min_{\mathbf{E}, \mathcal{Z}} \frac{1}{2} \|\mathcal{Y} - \mathcal{Z} \times_3 \mathbf{E}\|_F^2 + \lambda R(\mathcal{Z}), \quad s.t.\ \mathbf{E}^T \mathbf{E} = \mathbf{I} \tag{3}$$

where $R(\mathcal{Z})$ represents the regularization term incorporating prior information about the SRCs, and λ is a parameter that controls the tradeoff between data fidelity and regularization.

Numerous HSI denoising methods, such as FastHyDe [58], GLF [60], NGmeet [61], and RCTV [48], are based on LR subspace representation. However, these methods still lack efficiency when applied to large-scale HSI. They typically utilize LS and NSS priors to regularize SRCs $\mathcal{Z}$. The computational complexity of these priors depends mainly on the spatial dimension of $\mathcal{Z}$. As the spatial dimension increases, the denoising efficiency, particularly for NSS prior, is hindered.

In addition, the iterative solver of these algorithms further reduces efficiency. Consequently, existing methods are not suitable for large-scale HSI denoising.

III. PROPOSED FAST LARGE-SCALE HSI DENOISING METHOD

Large-scale HSI denoising faces the challenge of efficiently estimating two components from the observed HSI using LR subspace representation. The denoising efficiency bottleneck arises from the algorithm's iteration due to prior regularization and the excessive spatial dimension. In this section, we propose an FallHyDe method designed to overcome these obstacles. FallHyDe consists of five steps: 1) noise transformation; 2) spectral subspace learning; 3) extraction of HSNRBs; 4) fast estimation of SRCs; and 5) inverse noise transformation.

A. Noise Transformation

Since the ground-truth clean HSI $\mathcal{X}$ is unknown, it is a challenge to directly learn the spectral subspace $\mathbf{E}$ from it. However, when the observed HSI $\mathcal{Y}$ is degraded by additive independent and identically distributed (i.i.d.) Gaussian noise, the spectral subspace of $\mathcal{Y}$ is consistent with that of the clean HSI $\mathcal{X}$. Therefore, we can learn the spectral subspace $\mathbf{E}$ from the observed HSI $\mathcal{Y}$. In real HSI scenes, the Gaussian noise intensity in each band is typically inconsistent, resulting in non-i.i.d. noise. To transform the non-i.i.d. noise scenario into an i.i.d. noise case, Zhuang and Bioucas-Dias [58] and Zhuang et al. [60] introduced a Gaussian noise-whitening transformation strategy.

The spectral covariance matrix $\mathbf{C}$ of the Gaussian noise is assumed to be positive definite, and the noise-whitening transformation is given by

$$\hat{\mathcal{Y}} = \mathcal{Y} \times_3 \sqrt{\mathbf{C}^{-1}} \tag{4}$$

where $\hat{\mathcal{Y}}$, $(\mathbf{C}^{-1})^{1/2}$, and $\mathbf{C}^{-1}$ are the noise-whitened image, the square root of $\mathbf{C}^{-1}$, and the inverse of $\mathbf{C}$, respectively. The spectral covariance matrix $\mathbf{C}$ is defined as follows:

$$\mathbf{C} = \mathbb{E}\left(\mathbf{n}_i \mathbf{n}_i^T\right) \tag{5}$$

where $\mathbf{n}_i$ is the vectorization of the ith frontal slice of $\mathcal{N}$.

By combining the degradation model in (1) and the noise-whitening in (4), we can rewrite the degradation model for the noise-whitened image $\hat{\mathcal{Y}}$ as

$$\hat{\mathcal{Y}} = \mathcal{X} \times_3 \sqrt{\mathbf{C}^{-1}} + \mathcal{N} \times_3 \sqrt{\mathbf{C}^{-1}} = \hat{\mathcal{X}} + \hat{\mathcal{N}} \tag{6}$$

where $\hat{\mathcal{X}} = \mathcal{X} \times_3 (\mathbf{C}^{-1})^{1/2}$ and $\hat{\mathcal{N}} = \mathcal{N} \times_3 (\mathbf{C}^{-1})^{1/2}$. The noise degradation in the whitened image is standard i.i.d. Gaussian noise. To achieve the denoised HSI $\mathcal{X}$, we first restore $\hat{\mathcal{X}}$ from the noise-whitened image $\hat{\mathcal{Y}}$, and then the final clean HSI $\mathcal{X}$ is obtained using

$$\mathcal{X} = \hat{\mathcal{X}} \times_3 \sqrt{\mathbf{C}}. \tag{7}$$

After noise-whitening, the spectral vectors of $\hat{\mathcal{X}}$ still lie in a low-dimensional subspace. Therefore, the LR subspace representation in (2) can still be applied to $\hat{\mathcal{X}}$ as

$$\hat{\mathcal{X}} = \hat{\mathcal{Z}} \times_3 \hat{\mathbf{E}} \tag{8}$$

where $\hat{\mathbf{E}} \in \mathbb{R}^{B\times r}$ and $\hat{\mathcal{Z}} \in \mathbb{R}^{M\times N\times r}$ are the orthogonal spectral subspace and SRCs of $\hat{\mathcal{X}}$, respectively.

B. Spectral Subspace Learning

The spectral subspace of both $\hat{\mathcal{X}}$ and $\hat{\mathcal{Y}}$ is consistent and can be represented by the orthogonal subspace spanned by the columns of $\hat{\mathbf{E}}$. Therefore, we can learn the spectral subspace $\hat{\mathbf{E}}$ from the noise-whitened image $\hat{\mathcal{Y}}$ using the SVD as follows:

$$\hat{\mathbf{Y}}_{(3)} = \mathbf{U}\Sigma\mathbf{V}^T \tag{9}$$

where $\mathbf{U}$ and $\mathbf{V}$ are column-orthogonal matrices, and diagonal matrix Σ holds the singular values ordered by nonincreasing magnitude. When the rank of $\hat{\mathcal{X}}$ is r in the spectral dimension, the learned spectral subspace $\hat{\mathbf{E}}$ is given by

$$\hat{\mathbf{E}} = \mathbf{U}(:, 1:r). \tag{10}$$

C. Extraction of HSNRBs

The goal of the denoising process is to estimate the spectral subspace $\hat{\mathbf{E}}$ and SRCs $\hat{\mathcal{Z}}$. With the spectral subspace $\hat{\mathbf{E}}$ known, the denoising task transforms into estimating the SRCs $\hat{\mathcal{Z}}$ from the noisy HSI $\hat{\mathcal{Y}}$. Assuming that the observed HSI is of high quality ($\hat{\mathcal{Y}} \approx \hat{\mathcal{X}}$), the SRCs $\hat{\mathcal{Z}}$ can be instantly obtained by solving a least square problem. However, this assumption is not realistic. Despite the varying degrees of noise in the observed HSI, a noteworthy observation is the presence of numerous HSNRBs in the data (see Fig. 1), and it expects to retain these HSNRBs in the denoised result. By extracting these bands from the noisy HSI, we can obtain valuable information about part of the denoising result. To fully utilize these HSNRBs, we can employ these bands to guide the fast and accurate restoration of SRCs. In the following, we explain how to extract these HSNRBs and use them effectively.

The HSNRBs refer to bands in the observed HSI with small or close to zero noise intensity. To extract these bands, we estimate the noise intensity of each band in $\mathcal{Y}$. The HSNRBs of the observed HSI remain consistent before and after noise-whitening, meaning we only need to determine the noise intensity of each band in $\mathcal{Y}$. Bioucas-Dias and Nascimento [31] proposed an efficient noise estimation procedure using multiple regression theory. This method leverages the high spectral correlation and assumes that each band is approximately represented as a linear combination of the remaining $B-1$ bands

$$\mathcal{R}_i\left(\mathbf{Y}_{(3)}^T\right) = \mathcal{R}_{\partial_i}\left(\mathbf{Y}_{(3)}^T\right)\mathbf{z}_i + \mathbf{n}_i \tag{11}$$

where $\mathcal{R}_i(\mathbf{Y}_{(3)}^T)$ extracts the ith column from $\mathbf{Y}_{(3)}^T$, $\mathcal{R}_{\partial_i}(\mathbf{Y}_{(3)}^T)$ represents the matrix in $\mathbf{Y}_{(3)}^T$ excluding the ith column, $\mathbf{z}_i \in \mathbb{R}^{B-1}$ is the regression vector, and $\mathbf{n}_i \in \mathbb{R}^{MN}$ is the modeling error.

For each $i \in 1, 2, \ldots, B$, the regression vector $\mathbf{z}_i$ can be yielded by solving a least squares problem as follows:

$$\mathbf{z}_i = \arg\min_{\mathbf{z}_i} \|\mathcal{R}_i\left(\mathbf{Y}_{(3)}^T\right) - \mathcal{R}_{\partial_i}\left(\mathbf{Y}_{(3)}^T\right)\mathbf{z}_i\|_F^2. \tag{12}$$

To solve this least squares problem, we take the derivative with respect to the variable $\mathbf{z}_i$ and set the derivative to zero.

Then, optimizing (12) can be treated as solving the following linear system:

$$\left(\left(\mathcal{R}_{\partial_i}\left(\mathbf{Y}_{(3)}^T\right)\right)^T\mathcal{R}_{\partial_i}\left(\mathbf{Y}_{(3)}^T\right)\right)\mathbf{z}_i = \left(\mathcal{R}_{\partial_i}\left(\mathbf{Y}_{(3)}^T\right)\right)^T\mathcal{R}_i\left(\mathbf{Y}_{(3)}^T\right).$$

Since $\mathcal{R}_{\partial_i}(\mathbf{Y}_{(3)}^T)$ generally has the full-row rank due to the presence of noise, the above problem permits the following closed-form solution:

$$\mathbf{z}_i = \left(\left(\mathcal{R}_{\partial_i}\left(\mathbf{Y}_{(3)}^T\right)\right)^T\mathcal{R}_{\partial_i}\left(\mathbf{Y}_{(3)}^T\right)\right)^{-1}\left(\mathcal{R}_{\partial_i}\left(\mathbf{Y}_{(3)}^T\right)\right)^T\mathcal{R}_i\left(\mathbf{Y}_{(3)}^T\right).$$

With the estimate $\mathbf{z}_i$, the associated noise $\mathbf{n}_i$ in the ith band can be directly estimated as follows:

$$\hat{\mathbf{n}}_i = \mathcal{R}_i\left(\mathbf{Y}_{(3)}^T\right) - \mathcal{R}_{\partial_i}\left(\mathbf{Y}_{(3)}^T\right)\mathbf{z}_i. \tag{13}$$

We then estimate the standard deviation of Gaussian noise in the ith band as follows:

$$\sigma_i = \operatorname{std}\left(\hat{\mathbf{n}}_i\right) \tag{14}$$

where $\operatorname{std}(\cdot)$ denotes the standard deviation of all elements of the vector. In addition, by combining (5) and (13), we can obtain the spectral covariance matrix $\mathbf{C}$. Finally, we extract the HSNRBs whose noise standard deviation is less than a given threshold value, and the corresponding band numbers are denoted as $h = \{i_1, i_2, \ldots, i_k\}$. Similarly, the HSNRBs in the noise-whitened image $\hat{\mathcal{Y}}$ are represented as $\hat{\mathcal{Y}}_h$, which includes all bands in the set h extracted from $\hat{\mathcal{Y}}$.

D. Fast Estimation of SRCs

In this section, we present how to fast restore the SRCs from the HSNRBs $\hat{\mathcal{Y}}_h$ without introducing complex prior regularizations. By combining (6) and (8), the degradation model can be formulated as follows:

$$\hat{\mathcal{Y}} = \hat{\mathcal{Z}} \times_3 \hat{\mathbf{E}} + \hat{\mathcal{N}}. \tag{15}$$

Due to the presence of noise perturbation, we cannot directly estimate the SRCs $\hat{\mathcal{Z}}$ from noisy HSI $\hat{\mathcal{Y}}$ in (15). To address this challenge and distinguish the HSNRBs, we have reformulated the degradation model as follows:

$$\begin{pmatrix}\hat{\mathcal{Y}}_{\partial_h}\\ \hat{\mathcal{Y}}_h\end{pmatrix} = \begin{pmatrix}\hat{\mathcal{Z}} \times_3 \hat{\mathbf{E}}_{\partial_h}\\ \hat{\mathcal{Z}} \times_3 \hat{\mathbf{E}}_h\end{pmatrix} + \begin{pmatrix}\hat{\mathcal{N}}_{\partial_h}\\ \hat{\mathcal{N}}_h\end{pmatrix} \tag{16}$$

where $\hat{\mathcal{Y}}_{\partial_h}$ represents the remaining band in $\hat{\mathcal{Y}}$ excluding the HSNRBs, and $h \cup \partial_h = 1, 2, \ldots, B$ with $h \cap \partial_h = \emptyset$. Using the MATLAB command, we can obtain $\hat{\mathcal{Y}}_{\partial_h} = \hat{\mathcal{Y}}(:, :, \partial_h)$, $\hat{\mathcal{Y}}_h = \hat{\mathcal{Y}}(:, :, h)$ $\hat{\mathcal{N}}_{\partial_h} = \hat{\mathcal{N}}(:, :, \partial_h)$, and $\hat{\mathcal{N}}_h = \hat{\mathcal{N}}(:, :, h)$. Similarly, we divide the spectral subspace $\mathbf{E}$ into two parts: $\hat{\mathbf{E}}_{\partial_h} = \hat{\mathbf{E}}(\partial_h, :)$ and $\hat{\mathbf{E}}_h = \hat{\mathbf{E}}(h, :)$. As the noise standard deviation of the HSNRBs is small or close to zero, we approximate the noise $\hat{\mathcal{N}}_h$ as $\mathbf{O}$.

To eliminate the noise disturbance, we restore the SRCs $\mathcal{Z}$ from the HSNRBs. The restored SRCs $\hat{\mathcal{Z}}$ satisfy the following equation:

$$\hat{\mathcal{Y}}_h = \hat{\mathcal{Z}} \times_3 \hat{\mathbf{E}}_h + \hat{\mathcal{N}}_h. \tag{17}$$

The SRCs $\hat{\mathcal{Z}}$ can be restored by solving the following least squares problem:

$$\hat{\mathcal{Z}} = \arg\min_{\hat{\mathcal{Z}}} \|\hat{\mathcal{Y}}_h - \hat{\mathcal{Z}} \times_3 \hat{\mathbf{E}}_h\|_F^2 \tag{18}$$

Algorithm 1 FallHyDe: Fast Large-Scale HSI Denoising Method

```
Require: Noisy HSI $\mathcal{Y}$ and subspace dimension $r$.
1: Noise and spectral covariance matrix estimations: $\sigma_i$ and $\mathbf{C}$.
2: Noise transformation and HSNRBs extraction: $\hat{\mathcal{Y}}$ and $\hat{\mathcal{Y}}_h$.
3: Spectral subspace $\hat{\mathbf{E}}$ learning via Eq. (10).
4: Fast restoration of SRCs $\hat{\mathcal{Z}}$ via Eq. (19).
5: Inverse noise transformation via Eq. (20).
Ensure: Denoised HSI $\mathcal{X}$.
```

which is equivalent to the following problem:

$$\hat{\mathbf{Z}}_{(3)} = \arg\min_{\hat{\mathbf{Z}}_{(3)}} \|\hat{\mathbf{Y}}_{h,(3)} - \hat{\mathbf{E}}_h\hat{\mathbf{Z}}_{(3)}\|_F^2.$$

We take the derivative with respect to the variable $\hat{\mathbf{Z}}_{(3)}$ and set the derivative to zero. Then, optimizing (18) can be treated as solving the following linear system:

$$\left(\hat{\mathbf{E}}_h^T\hat{\mathbf{E}}_h\right)\hat{\mathbf{Z}}_{(3)} = \hat{\mathbf{E}}_h^T\hat{\mathbf{Y}}_{h,(3)}.$$

In particular, when $\hat{\mathbf{E}}_h$ is full-column rank, the above problem can be efficiently solved using simple matrix multiplication

$$\hat{\mathcal{Z}} = \operatorname{Fold}_3\left(\left(\hat{\mathbf{E}}_h^T\hat{\mathbf{E}}_h\right)^{-1}\left(\hat{\mathbf{E}}_h^T\hat{\mathbf{Y}}_{h,(3)}\right)\right). \tag{19}$$

E. Inverse Noise Transformation

With the transformed spectral subspace $\hat{\mathbf{E}}$ and SRCs $\hat{\mathcal{Z}}$ known, the final denoised HSI $\mathcal{X}$ is restored as follows:

$$\mathcal{X} = \left(\hat{\mathcal{Z}} \times_3 \hat{\mathbf{E}}\right) \times_3 \sqrt{\mathbf{C}}. \tag{20}$$

To summarize the above steps, the proposed FallHyDe is presented in Algorithm 1. To show the efficiency of the proposed FallHyDe, we give the computational complexity of Algorithm 1. The computational complexity of noise estimation and spectral covariance matrix estimation in line 1 is $\mathcal{O}(MNB^3)$ and $\mathcal{O}(MNB)$, respectively. The noise transformation in line 2 and its inverse transformation in line 5 involve tensor-matrix multiplication, each with a computational complexity of $\mathcal{O}(MNB^2)$. The spectral subspace learning in line 4 and the restoration of the SRCs in line 5 require an SVD operator and a small-scale matrix multiplication, with computational complexities of $\mathcal{O}(MNB^2)$ and $\mathcal{O}(MNB_hr)$, respectively, where B_h ($B_h<B$) is the number of bands in the HSNRBs. In summary, the computational complexity of our FallHyDe is approximately $\mathcal{O}(MNB^3)$. Notably, our method does not rely on iterative solutions, making it efficient for processing large-scale HSIs without requiring high-performance computing support. In addition, the proposed method does not involve the selection of regularization parameters, and the subspace dimension r can be estimated by HySime [31], providing flexibility in HSI denoising tasks.

F. Theoretical Analysis

The feasibility of the proposed method is guaranteed by the following theorem.

Theorem 1: Suppose that the HSNRBs in the observed HSI need to be approximatively retained in the denoising results, i.e., $\mathcal{Y}_h \approx \mathcal{X}_h$, $\hat{\mathcal{Y}}_h \approx \hat{\mathcal{X}}_h$, and $\hat{\mathcal{N}}_h \approx \mathbf{O}$, and the band number of HSNRBs is greater than the subspace dimension, i.e., $B_h > r$. Then, the solution of $\hat{\mathcal{Z}}$ in (19) is also the desired solution of (8).

Proof: Matrix $\hat{\mathbf{E}}$ is a column orthogonal matrix, thus the desired solution $\hat{\mathcal{Z}}$ of (8) is unique and can be expressed as

$$\hat{\mathcal{Z}} = \hat{\mathcal{X}} \times_3 \hat{\mathbf{E}}^T. \tag{21}$$

Since (17) is a subequation extracted from (8), the desired solution $\hat{\mathcal{Z}}$ in (21) also satisfies (17). Therefore, we can conclude that (17) has a solution with respect to the variable $\hat{\mathcal{Z}}$.

From the linear mixing model, r is the number of endmember signatures, and it is equal to the rank of $\hat{\mathbf{X}}_{(3)}$. Since the equation $\hat{\mathcal{X}} = \hat{\mathcal{Z}} \times_3 \hat{\mathbf{E}}$ holds, we can deduce that $\hat{\mathbf{E}}$ and $\hat{\mathbf{Z}}_{(3)}$ are full-column rank and full-row rank, respectively. As $\hat{\mathbf{Y}}_{h,(3)}$ is approximatively retained in the whole HSI $\hat{\mathbf{X}}_{(3)}$, the number of endmember and rank of $\hat{\mathbf{Y}}_{h,(3)}$ are approximately equal to r. According to the rank relation of the matrix multiplication operation: $\text{rank}(\hat{\mathbf{Y}}_{h,(3)}) \le \min(\text{rank}(\hat{\mathbf{E}}_h), \text{rank}(\hat{\mathbf{Z}}_{(3)}))$, we have that the rank of $\hat{\mathbf{E}}_h$ is r.

In summary, equation $\hat{\mathcal{Y}}_h = \hat{\mathcal{Z}} \times_3 \hat{\mathbf{E}}_h$ in (17) with respect to the variable $\hat{\mathcal{Z}}$ satisfies two conditions: 1) the equation has a solution and 2) coefficient matrix $\hat{\mathbf{E}}_h$ is full-column rank. Therefore, we can conclude that (17) has a unique solution. Since both (19) and (21) satisfy (17), they can be approximated as the same solution. Thus, (19) is also the solution of (8). This completes the proof. □

IV. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we conduct simulated and real experiments to validate the superiority of the proposed method in HSI denoising. For comparison, we select ten state-of-the-art HSI denoising methods: BM3D [59], LRMR [35], NMoG [37], KBR [55], LLRT [56], WLRTR [57], NGmeet[4] [61], RCTV [48], FastHyDe [58], and FastHyMix [64]. Among these methods, BM3D, NGmeet, RCTV, FastHyDe, and FastHyMix are considered representative fast HSI denoising approaches, with FastHyMix being based on deep learning. To ensure a fair comparison, we downloaded the running codes of the comparison methods from the authors' homepages and set their parameters as suggested in their respective articles or used the default values provided in their code. All experiments are performed in MATLAB R2020b using a desktop with 32-GB RAM, with an Intel Core i9-10850K CPU at 3.60 GHz.

A. Experiments on HSI With Synthetic Noise

1) Experimental Setting: To compare the denoising performance with state-of-the-art nonlocal methods KBR, LLRT, and WLRTR, which have high computational complexity due to their NSS prior to the original HSI data, we select a small-scale dataset for testing. The Pavia City Center image (PaC[5]) is chosen as the small-scale simulated data, and a subimage of size 200 × 200 × 80 is selected as the reference. To demonstrate the efficiency of our method on large-scale HSI denoising, we utilize the entire Washington DC Mall dataset (WDC), acquired by the hyperspectral digital imagery collection experiment (HYDICE) sensor. The WDC dataset contains 1208 × 307 pixels and 191 spectral bands, which can relatively be considered a large-scale HSI dataset. Following the settings in the previous works [56], [57], [61], we add additive Gaussian noise with noise standard deviations $\sigma = 10, 30, 50, 80, 100$ to the clean HSI. In addition, we introduce additive Gaussian noise, where the noise standard deviations σ are randomly chosen from uniform distributions $U[10, 100]$ and $U[30, 80]$. In addition, mixed Gaussian and stripe noises are introduced to demonstrate the robustness of the proposed method. The standard deviations of the Gaussian noise are randomly selected from the range $U[10, 100]$, and stripes are added to 20% of the bands, with the number of stripes ranging from 5 to 12. To simulate the existence of HSNRBs in real HSI noise scenarios, we randomly select some bands without adding noise. Specifically, for data with low noise intensity, more HSNRBs are exited, while for scenes with high noise intensity, fewer HSNRBs are included. Therefore, we randomly select 5/10, 4/10, 3/10, 2/10, 1/10, 3/10, and 3/10 bands as the HSNRBs for noise cases with $\sigma = 10, 30, 50, 80, 100, U[10, 100]$, and $U[30, 80]$, respectively.

2) Quantitative Indices: We quantitatively evaluate the denoising results of different methods using three common HSI restoration indices: peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and spectral angle mapper (SAM). PSNR and SSIM are spatial information-based evaluation indices, while SAM is employed to evaluate spectral information. Higher PSNR and SSIM values and lower SAM values indicate better denoising performance. In addition, we record the running time (in seconds) to assess the denoising efficiency of all methods.

3) Experimental Results on Small-Scale PaC Dataset: Table II presents the quantitative results of various methods on the PaC dataset, with the best values highlighted in bold. FastHyDe consistently outperforms other methods in most cases. Although our FallHyDe does not achieve the absolute best results in all quantitative indices, the differences in denoising accuracy between our method and the best-performing methods are negligible across various noise levels. However, in terms of denoising efficiency, the proposed method can complete denoising in only 0.2 s, more than 19, 23, 34, and 10 times faster than that of related fast HSI denoising methods RCTV, FastHyDe, NGmeet, and FastHyMix, respectively. Thus, the proposed method significantly enhances the efficiency of HSI denoising on small-scale datasets without compromising denoising accuracy. Moreover, compared with the NSS prior to image method WLRTR, our proposed method is 4500 times faster.

Fig. 3 shows the false-color image of the denoised PaC dataset with $\sigma = 100$, and an enlarged region is provided

---

[4] It is worth noting that according to the framework of NGmeet method, we have rewritten its running code, which can greatly improve the accuracy and efficiency of HSI denoising.

[5] http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes

TABLE II

QUANTITATIVE COMPARISON OF DIFFERENT COMPARISON METHODS UNDER DIFFERENT NOISE INTENSITIES ON SMALL-SCALE PaC DATASET. THE BEST RESULT IS HIGHLIGHTED IN BOLD

Case	Index	Noisy	BM3D	LRMR	NMoG	KBR	LLRT	WLRTR	RCTV	FastHyDe	NGmeet	FastHyMix	FallHyDe
10	PSNR	28.12	36.33	45.91	52.99	40.49	40.51	42.01	41.83	53.27	53.73	50.98	**54.25**
	SSIM	0.809	0.952	0.996	**0.999**	0.985	0.977	0.990	0.989	**0.999**	**0.999**	**0.999**	**0.999**
	SAM	15.785	6.042	2.274	1.384	3.431	5.210	0.047	3.449	1.233	1.225	1.237	**1.210**
	Time (s)		26.5	11.1	54.5	781.6	514.0	956.8	4.2	10.0	8.7	3.2	**0.2**
30	PSNR	18.59	32.32	40.05	48.59	34.45	35.02	36.23	34.90	51.51	51.18	49.29	**51.53**
	SSIM	0.376	0.853	0.987	0.995	0.944	0.943	0.966	0.950	**0.998**	0.997	**0.998**	**0.998**
	SAM	34.288	9.881	4.006	2.269	6.284	7.253	6.078	6.128	**1.550**	1.853	1.648	1.703
	Time (s)		30.6	10.6	54.5	759.2	508.4	970.2	3.5	7.9	8.7	2.5	**0.2**
50	PSNR	14.15	29.44	35.15	44.68	32.47	33.82	33.21	31.73	**49.07**	49.06	47.24	48.86
	SSIM	0.190	0.742	0.965	0.993	0.920	0.944	0.937	0.904	**0.998**	**0.998**	0.997	0.997
	SAM	46.627	12.175	5.933	2.579	6.129	5.520	5.103	7.773	**1.799**	1.812	1.978	1.961
	Time (s)		34.8	12.7	54.3	831.0	496.7	969.0	4.2	4.1	7.4	2.1	**0.2**
80	PSNR	10.07	26.54	29.94	39.12	29.88	30.09	30.71	29.09	**47.07**	46.79	45.24	46.92
	SSIM	0.084	0.606	0.890	0.981	0.868	0.878	0.893	0.844	**0.997**	**0.997**	0.995	0.996
	SAM	58.311	14.031	11.122	3.958	6.236	6.333	6.131	7.198	**2.033**	2.037	2.297	2.163
	Time (s)		37.6	12.7	39.1	1097.8	486.2	979.7	4.1	3.0	6.2	1.8	**0.2**
100	PSNR	8.13	24.11	26.49	39.89	28.36	28.37	29.29	28.48	43.57	42.98	42.73	**43.73**
	SSIM	0.055	0.503	0.781	0.988	0.820	0.826	0.857	0.820	**0.994**	0.992	0.993	0.993
	SAM	63.216	14.550	15.186	3.406	6.802	6.801	6.152	8.253	**2.821**	2.915	2.837	2.970
	Time (s)		39.9	12.7	36.2	1099.6	474.9	979.5	3.3	1.8	4.3	1.3	**0.2**
[10, 100]	PSNR	15.02	29.88	33.85	43.39	28.98	31.33	28.91	30.36	48.48	48.47	47.36	**48.66**
	SSIM	0.258	0.750	0.945	0.961	0.803	0.890	0.795	0.874	**0.997**	**0.997**	**0.997**	**0.997**
	SAM	50.398	13.004	7.879	21.652	15.651	9.466	9.487	12.586	1.861	**1.840**	1.900	2.044
	Time (s)		33.5	12.6	36.1	829.7	501.6	989.3	3.6	3.1	6.2	1.4	**0.2**
[30, 80]	PSNR	14.17	29.48	33.80	42.70	30.86	33.17	30.72	31.05	48.90	48.90	47.24	**49.05**
	SSIM	0.198	0.741	0.951	0.977	0.877	0.934	0.864	0.882	**0.998**	**0.998**	0.997	0.997
	SAM	48.113	12.595	7.343	11.503	10.101	6.282	6.280	10.893	1.781	**1.758**	1.979	1.902
	Time (s)		34.0	12.7	36.2	830.3	498.4	986.9	3.5	3.1	6.8	1.8	**0.2**
Mixed	PSNR	14.33	29.34	33.12	46.38	28.48	31.31	27.71	29.88	47.99	43.30	46.83	**48.22**
	SSIM	0.247	0.729	0.935	0.985	0.789	0.893	0.755	0.854	**0.997**	0.992	**0.997**	**0.997**
	SAM	52.840	14.118	8.984	10.985	16.367	8.867	0.497	12.530	**1.974**	2.393	2.008	2.081
	Time (s)		33.5	15.1	38.5	1772.3	545.1	1172.2	3.4	3.2	6.0	1.5	**0.2**
Average	PSNR	15.32	29.68	34.79	44.72	31.75	32.95	32.35	32.17	48.73	48.05	47.11	**48.90**
	SSIM	0.277	0.735	0.931	0.985	0.876	0.911	0.882	0.890	**0.997**	0.996	**0.997**	**0.997**
	SAM	46.197	12.050	7.841	7.217	8.875	6.967	4.972	8.601	**1.882**	1.979	1.986	2.004
	Time (s)		33.8	12.5	43.7	1000.2	503.2	1000.5	3.7	4.5	6.8	2.0	**0.2**

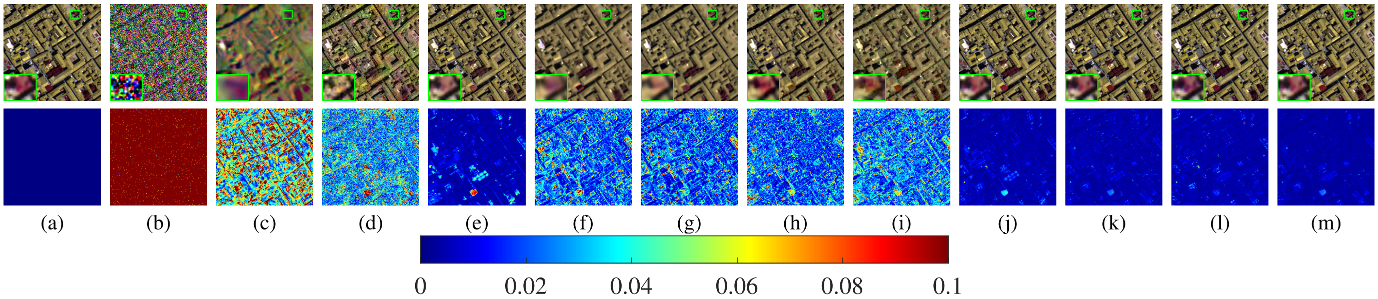


Fig. 3. Denoised results of small-scale PaC dataset under noise cases $\sigma = 100$. The first row is the visual results, and the second row is the corresponding absolute error maps between the ground truth and the denoising results. The false-color image is composed of bands (R: 49, G: 37, and B: 4). (a) Original. (b) Noisy. (c) BM3D. (d) LRMR. (e) NMoG. (f) KBR. (g) LLRT. (h) WLRTR. (i) RCTV. (j) FastHyDe. (k) NGmeet. (l) FastHyMix. (m) FallHyDe.

to visualize the differences among the denoising results. The absolute error map between the ground truth and denoised images is also presented for comparison. It is evident that all methods effectively reduce the noise in the image. However, BM3D, LRMR, NMoG, KBR, LLRT, WLRTR, and RCTV exhibit spectral distortion, leading to changes in the contrast of spatial information (circled in the green box). FastHyDe, NGmeet, FastHyMix, and our method perform similar denoising results, effectively reducing most of the noise while preserving the image details. Comparing the absolute error maps, our proposed FallHyDe demonstrates significantly higher spatial and spectral fidelity, with error maps much closer to the ground truth. These results verify that our method achieves competitive denoising performance compared to state-of-the-art methods, while also greatly improving denoising efficiency.

4) Experimental Results on Large-Scale WDC Dataset: We evaluated the denoising accuracy and efficiency of all

TABLE III

Quantitative Comparison of Different Comparison Methods Under Different Noise Intensities on Large-Scale WDC Dataset. The Best Result Is Highlighted in Bold

Case	Index	Noisy	BM3D	LRMR	NMoG	RCTV	FastHyDe	NGmeet	FastHyMix	FallHyDe
10	PSNR	28.13	37.52	47.53	52.93	44.10	54.54	54.62	50.84	**55.70**
	SSIM	0.654	0.938	0.993	**0.998**	0.987	**0.998**	**0.998**	**0.998**	**0.998**
	SAM	8.166	2.728	0.952	0.614	1.270	0.514	0.510	0.546	**0.473**
	Time (s)		740.0	217.6	1301.9	96.4	93.6	131.2	156.2	**9.7**
30	PSNR	18.59	33.98	40.57	48.69	37.68	52.87	52.92	48.59	**54.19**
	SSIM	0.252	0.849	0.971	0.993	0.952	**0.998**	**0.998**	0.997	**0.998**
	SAM	23.056	5.025	2.046	1.062	2.461	0.595	0.591	0.664	**0.514**
	Time (s)		809.8	205.7	1312.9	89.2	79.5	119.2	122.2	**9.7**
50	PSNR	14.15	31.66	36.12	45.53	34.86	52.00	52.08	45.66	**52.76**
	SSIM	0.125	0.772	0.932	0.988	0.915	0.997	0.997	0.995	**0.998**
	SAM	35.357	6.617	3.334	1.509	3.298	0.634	0.624	0.929	**0.578**
	Time (s)		985.4	200.5	945.0	91.0	78.9	126.3	91.3	**9.7**
80	PSNR	10.07	29.22	30.01	42.97	31.49	49.93	49.86	44.84	**50.14**
	SSIM	0.057	0.684	0.813	0.982	0.804	**0.996**	**0.996**	0.994	**0.996**
	SAM	48.583	8.512	6.482	2.009	5.125	0.797	0.797	1.021	**0.782**
	Time (s)		1050.8	194.3	878.2	88.7	65.2	117.9	74.2	**8.9**
100	PSNR	8.13	27.15	27.38	36.98	31.10	44.65	44.24	40.15	**46.09**
	SSIM	0.038	0.615	0.692	0.954	0.816	0.989	0.987	0.976	**0.993**
	SAM	54.794	9.778	8.879	3.963	5.058	1.537	1.578	1.698	**1.167**
	Time (s)		1101.7	194.9	503.3	82.8	37.9	95.9	49.2	**9.1**
[10, 100]	PSNR	14.53	31.87	35.69	44.37	32.90	51.79	51.86	44.42	**52.44**
	SSIM	0.164	0.778	0.918	0.970	0.876	**0.997**	**0.997**	0.993	**0.997**
	SAM	40.394	6.831	3.857	10.673	4.638	0.645	0.636	1.116	**0.605**
	Time (s)		959.0	202.7	723.4	85.6	78.6	126.1	102.5	**8.8**
[30, 80]	PSNR	13.41	31.48	35.81	47.28	33.99	51.81	51.87	44.98	**52.45**
	SSIM	0.116	0.763	0.925	0.984	0.890	**0.997**	**0.997**	0.994	**0.997**
	SAM	39.386	6.998	3.604	4.104	3.813	0.645	0.636	1.013	**0.599**
	Time (s)		982.3	204.0	1019.3	83.8	78.6	125.4	97.8	**8.4**
Mixed	PSNR	13.77	31.34	35.22	45.84	33.00	51.78	47.73	45.00	**51.86**
	SSIM	0.144	0.761	0.912	0.981	0.867	**0.997**	0.994	0.994	**0.997**
	SAM	42.084	7.552	4.087	7.596	4.426	**0.684**	0.953	1.064	0.700
	Time (s)		898.0	227.5	774.4	86.9	80.4	120.5	103.0	**8.9**
Average	PSNR	15.10	31.78	36.04	45.57	34.89	51.17	50.65	45.56	**51.95**
	SSIM	0.194	0.770	0.895	0.981	0.888	0.996	0.996	0.993	**0.997**
	SAM	36.478	6.755	4.155	3.941	3.761	0.756	0.791	1.006	**0.677**
	Time (s)		940.9	205.9	932.3	88.1	74.1	120.3	99.6	**9.2**

methods on the large-scale WDC dataset. Since KBR, LLRT, and WLRTR were unable to perform denoising in our experimental environment, we excluded them from the comparison in this experiment. The quantitative results for different noise intensities are presented in Table III. Remarkably, the proposed FallHyDe method achieved the best denoising results on the large-scale dataset in all cases. FastHyDe and NGmeet still obtained competitive denoising results, but FastHyMix showed significantly lower performance on the large-scale dataset compared to the small-scale dataset. The running time results in Table III demonstrate that our proposed FallHyDe method efficiently processes large-scale HSIs within 10 s, which is the lowest among all evaluated methods. Particularly, compared to BM3D, our method achieved an average running time gain of 100 times, and over LRMR by 20 times, over NMoG by 100 times, over RCTV by nine times, over FastHyDe by seven times, over NGmeet by 12 times, and over FastHyMix by ten times.

To further demonstrate the denoising capabilities of FallHyDe, Fig. 4 presents the false-color images of the large-scale WDC dataset before and after denoising. The denoising results of FastHyDe, NGmeet, FastHyMix, and FallHyDe outperform the other comparison methods significantly. However, the absolute error map reveals that FastHyDe, NGmeet, and FastHyMix lose some spatial information, while our FallHyDe method produces the lowest error results. It is noteworthy that most learning-based methods struggle to process large-scale HSI due to limited computational resources or extensive processing time (see Table I). In contrast, the proposed FallHyDe method not only exhibits superior denoising capability but also boasts remarkable processing speeds, making it highly efficient for large-scale HSI denoising.

B. Experiments on HSI With Real Noise

To facilitate comparisons with different methods and demonstrate that the proposed method is applicable to real-world scenarios, we further evaluate our method on real-world HSI using the HYDICE Terrain dataset with dimensions of 500 × 307 × 210. Compared to previous works that used experimental data with spatial dimensions of 200 × 200 or 256 × 256, this dataset is also relatively larger. This dataset contains bands that are seriously degraded by complex noise,

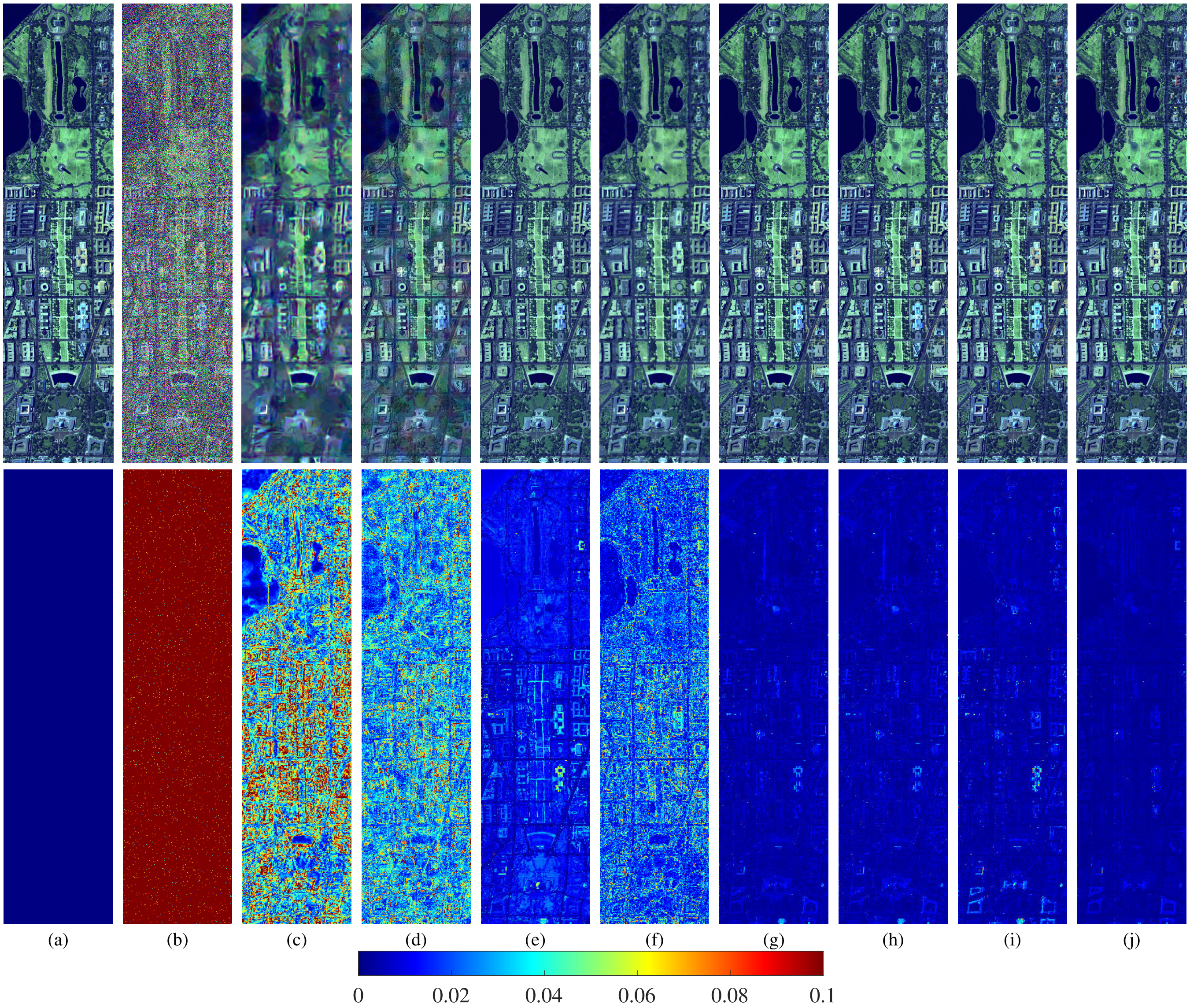


Fig. 4. Denoised results of large-scale WDC dataset under noise cases $\sigma = 100$. The first row is the visual results, and the second row is the corresponding absolute error maps between the ground truth and denoising results. The false-color image is composed of bands (R: 157, G: 109, and B: 47). (a) Original. (b) Noisy. (c) BM3D. (d) LRMR. (e) NMoG. (f) RCTV. (g) FastHyDe. (h) NGmeet. (i) FastHyMix. (j) FallHyDe.

but it also includes many HSNRBs (see Fig. 1). We empirically select the top one-fourth of bands with the lowest noise levels as the HSNRBs, which is typically feasible in real datasets. However, due to the unavailability of clean HSI in real experiments, quantitative evaluation metrics cannot be computed to illustrate the denoising results of all methods. Instead, we compare the denoising results on two representative bands with different noise levels and use the mean curve to qualitatively evaluate the denoising results. The KBR method still shows insufficient memory on our experimental device, so we exclude it from the comparison.

Fig. 5 presents the false-color results of all methods on the real dataset, where both bands are heavily degraded by noise. BM3D, LRMR, LLRT, WLRTR, and RCTV fail to restore the original image and contain stripes. FastHyMix effectively suppresses the noise but loses spatial details like edges and textures. On the other hand, NMoG, FastHyDe, NGmeet, and FallHyDe achieve similar denoising results, effectively suppressing noise and preserving structural information, resulting in visually pleasing images. Fig. 6 shows the row mean profiles of band 106, indicating that NMoG, FastHyDe, NGmeet, and FallHyDe approximately maintain the trend of the original image and produce smoother curves. Table IV lists the running times of different methods, and our proposed FallHyDe remains the fastest among all the comparison methods. In summary, although the proposed method is designed with the Gaussian noise assumption, it can also achieve superior results for real datasets.

C. Discussion

1) Comparison With Supervised Learning-Based Method:

Recent studies have shown that well-trained learning-based approaches can achieve denoising efficiently. To further illustrate the efficiency of our method, we compare the denoising results with the state-of-the-art supervised learning method

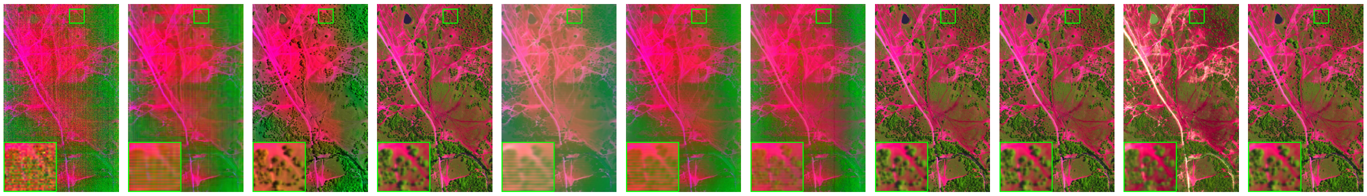

(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k)

Fig. 5. Denoised results of the real Terrain dataset. The false-color image is composed of bands (R: 151, G: 106, and B: 1). (a) Noisy. (b) BM3D. (c) LRMR. (d) NMoG. (e) LLRT. (f) WLRTR. (g) RCTV. (h) FastHyDe. (i) NGmeet. (j) FastHyMix. (k) FallHyDe.

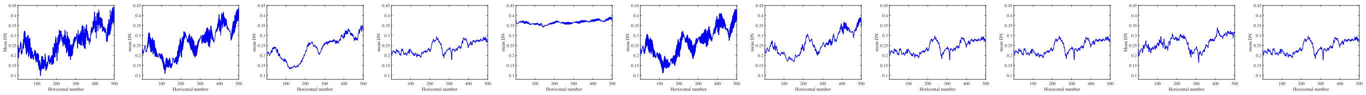

(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k)

Fig. 6. Row mean profiles of band 106 for the real Terrain dataset. (a) Noisy. (b) BM3D. (c) LRMR. (d) NMoG. (e) LLRT. (f) WLRTR. (g) RCTV. (h) FastHyDe. (i) NGmeet. (j) FastHyMix. (k) FallHyDe.

TABLE IV

RUNNING TIME OF DIFFERENT METHODS ON THE REAL TERRIAN DATASET

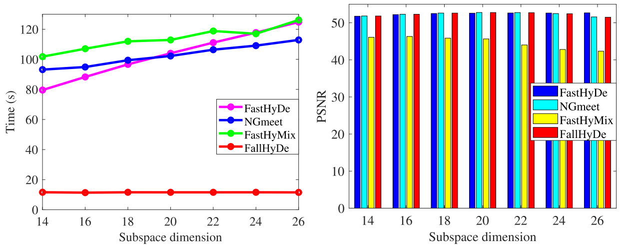
Method	BM3D	LRMR	NMoG	LLRT	WLRTR	RCTV	FastHyDe	NGmeet	FastHyMix	FallHyDe
Time (s)	292.7	91.2	230.6	2387.7	5063.7	37.2	12.4	38.0	43.5	4.5

TABLE V

QUANTITATIVE COMPARISON WITH SUPERVISED LEARNING-BASED METHOD UNDER DIFFERENT NOISE INTENSITIES ON SMALL-SCALE PaC DATASET

Method	Index	10	30	50	80	100	[10, 100]	[30, 80]
QRNN3D	PSNR	36.06	34.06	31.71	29.03	27.03	30.94	31.41
	SSIM	0.971	0.950	0.909	0.833	0.761	0.893	0.903
	Time (s)	2.1	2.1	2.1	2.1	2.1	2.1	2.1
FallHyDe	PSNR	**54.25**	**51.53**	**48.86**	**46.92**	**43.73**	**48.66**	**49.05**
	SSIM	**0.999**	**0.998**	**0.997**	**0.996**	**0.993**	**0.997**	**0.997**
	Time (s)	**0.2**	**0.2**	**0.2**	**0.2**	**0.2**	**0.2**	**0.2**

QRNN3D[6] [9]. Since QRNN3D is unable to process the large-scale WDC dataset on a single NVIDIA GeForce RTX 3090 (see Table I), we use the small-scale PaC dataset for testing. Table V presents the quantitative results and running time of the PaC dataset obtained by QRNN3D and our FallHyDe method. QRNN3D achieves relatively poor quantitative results due to the absence of these noise types in its training set, indicating its limited denoising flexibility. Even without considering the comparison of quantitative results, our method still demonstrates faster denoising efficiency compared to QRNN3D.



(a) (b)

Fig. 7. Computational time and PSNR values with different subspace dimensions. (a) Time versus subspace dimension. (b) PSNR versus subspace dimension.

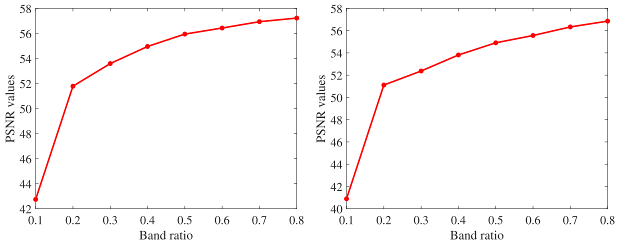
2) Robustness of Computational Efficiency and Denoising Effectiveness: To demonstrate the robustness of our method's subspace dimension selection, we compare the computational time and PSNR values of fast denoising methods such as FastHyDe, NGmeet, FastHyMix, and FallHyDe with respect to different subspace dimensions in Fig. 7. The results show that the PSNR values of FastHyDe, NGmeet, and FastHyMix remain stable as the subspace dimension increases, but their computational time increases linearly. In contrast, our method not only achieves stable denoising effects but also maintains nearly unchanged computational time with the increase of subspace dimension.

3) Impact of the Number of HSNRBs: Fig. 8 illustrates the PSNR values with varying band ratios of HSBRBs at two noise levels. As the number of HSNRBs increases, the PSNR value gradually improves due to the availability of more guidance information. In our experiments, we set the number of HSNRBs based on empirical observations and realistic scenarios to ensure practical applicability.

4) Limitations: When most of the bands in the HSI are contaminated by complex mixed noise, the estimation of the

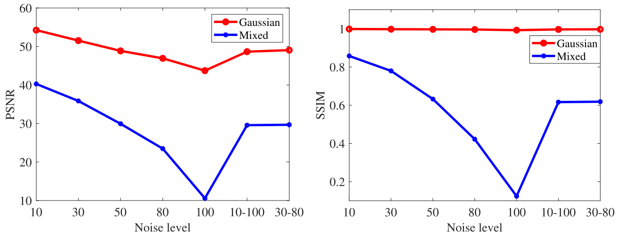
---

[6] The result is achieved by using a model published by the authors on their homepage.



(a) (b)

Fig. 8. Impact of the number of HSNRBs under two different noise levels. (a) $\sigma = 30$. (b) $\sigma = [10, 100]$.



(a) (b)

Fig. 9. Comparison of the results when 20% impulse noise is added to the Gaussian noise degradation across all bands. (a) PSNR. (b) SSIM.

spectral subspace may become inaccurate, leading to reduced denoising effectiveness. Fig. 9 shows a comparison of the results when 20% impulse noise is added to the Gaussian noise degradation across all bands. We can observe that the denoising performance quickly deteriorates after the addition of impulse noise. In fact, this is also a limitation of other Gaussian noise removal methods, including comparison methods NGmeet, FastHyDe, FastHyMix, and others. In future work, we will focus on improving the accuracy of learning the spectral subspace from HSI contaminated with complex mixed noise to address this limitation and further enhance the performance of our method.

V. CONCLUSION

In this article, we have introduced a novel noniterative LR subspace representation method to efficiently tackle the large-scale HSI denoising problem. The method decomposes the HSI into the spectral subspace and SRCs, capturing strong spectral correlation and reducing computational complexity. The spectral subspace is learned from the observed HSI, and the SRCs are efficiently estimated from guided HSNRBs. Compared to the previous works that try to denoise the HSI with different complex priors, our FallHyDe method does not require regularizers or iterations, resulting in a massive reduction in computational time. The theoretical analysis further validates the feasibility of our denoising framework. Experimental results demonstrate that FallHyDe achieves competitive denoising performance with far less computational time. In summary, our methods restore the HSI in an efficient, effective, and flexible manner, which is expected to be extended to other HSI fields.

REFERENCES

[1] M. Uzair, A. Mahmood, and A. Mian, "Hyperspectral face recognition with spatiospectral information fusion and PLS regression," *IEEE Trans. Image Process.*, vol. 24, no. 3, pp. 1127–1137, Mar. 2015.

[2] D. Hong, N. Yokoya, J. Chanussot, and X. X. Zhu, "An augmented linear mixing model to address spectral variability for hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1923–1938, Apr. 2018.

[3] J. Yang, B. Du, and L. Zhang, "From center to surrounding: An interactive learning framework for hyperspectral image classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 197, pp. 145–166, Mar. 2023.

[4] 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.

[5] Z. H. Huang et al., "Joint analysis and weighted synthesis sparsity priors for simultaneous denoising and destriping optical remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 10, pp. 6958–6982, Oct. 2020.

[6] Q. Zhang, Y. Zheng, Q. Yuan, M. Song, H. Yu, and Y. Xiao, "Hyperspectral image denoising: From model-driven, data-driven, to model-data-driven," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jun. 6, 2023, doi: 10.1109/TNNLS.2023.3278866.

[7] Y. Chang, L. Yan, 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. 2018.

[8] 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, Sep. 2018.

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

[10] 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.

[11] F. Xiong, J. Zhou, Q. Zhao, J. Lu, and Y. Qian, "MAC-Net: Model-aided nonlocal neural network for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5519414.

[12] 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.

[13] 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.*, vol. 35, no. 7, pp. 8797–8811, Jul. 2024.

[14] 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.

[15] 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.

[16] T. Bodrito, A. Zouaoui, J. Chanussot, and J. Mairal, "A trainable spectral–spatial sparse coding model for hyperspectral image restoration," in *Proc. Adv. Neural. Inf. Process. Syst.*, vol. 34, 2021, pp. 5430–5442.

[17] 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.

[18] G. Fu, F. Xiong, J. Lu, J. Zhou, J. Zhou, and Y. Qian, "Hyperspectral image denoising via spatial–spectral recurrent transformer," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5511214.

[19] Z. Huang, Z. Zhu, Z. Wang, Y. Shi, H. Fang, and Y. Zhang, "DGDNet: Deep gradient descent network for remotely sensed image denoising," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, pp. 1–5, 2023.

[20] J. Li, W. He, W. Cao, L. Zhang, and H. Zhang, "UANet: An uncertainty-aware network for building extraction from remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5608513.

[21] W. He, Q. Yao, N. Yokoya, T. Uezato, H. Zhang, and L. Zhang, "Spectrum-aware and transferable architecture search for hyperspectral image restoration," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2022, pp. 19–37.

[22] M. Li, J. Liu, Y. Fu, Y. Zhang, and D. Dou, "Spectral enhanced rectangle transformer for hyperspectral image denoising," in *Proc. CVPR*, Jun. 2023, pp. 5805–5814.

[23] O. Sidorov and J. Y. Hardeberg, "Deep hyperspectral prior: Single-image denoising, inpainting, super-resolution," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019, pp. 3844–3851.

[24] Y.-C. Miao, X.-L. Zhao, X. Fu, J.-L. Wang, and Y.-B. Zheng, "Hyperspectral denoising using unsupervised disentangled spatiospectral deep priors," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5513916.

[25] L. Zhuang, M. K. Ng, L. Gao, J. Michalski, and Z. Wang, "Eigenimage2Eigenimage (E2E): A self-supervised deep learning network for hyperspectral image denoising," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jul. 19, 2023, doi: 10.1109/TNNLS.2023.3293328.

[26] L. Zhuang, M. K. Ng, L. Gao, and Z. Wang, "Eigen-CNN: Eigenimages plus eigennoise level maps guided network for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5512018.

[27] Y. Chen, M. Chen, W. He, J. Zeng, M. Huang, and Y.-B. Zheng, "Thick cloud removal in multitemporal remote sensing images via low-rank regularized self-supervised network," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5506613.

[28] Z. Huang et al., "MD3: Model-driven deep remotely sensed image denoising," *Remote Sens.*, vol. 15, no. 2, p. 445, Jan. 2023.

[29] 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.

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

[31] 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.

[32] W. He, N. Yokoya, and X. Yuan, "Fast hyperspectral image recovery of dual-camera compressive hyperspectral imaging via non-iterative subspace-based fusion," *IEEE Trans. Image Process.*, vol. 30, pp. 7170–7183, 2021.

[33] Y. Chen, W. Lai, W. He, X.-L. Zhao, and J. Zeng, "Hyperspectral compressive snapshot reconstruction via coupled low-rank subspace representation and self-supervised deep network," *IEEE Trans. Image Process.*, vol. 33, pp. 926–941, 2024.

[34] W. He, Z. Wu, N. Yokoya, and X. Yuan, "An interpretable and flexible fusion prior to boost hyperspectral imaging reconstruction," *Inf. Fusion*, vol. 111, Nov. 2024, Art. no. 102528.

[35] 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.

[36] 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.

[37] Y. Chen, X. Cao, Q. Zhao, D. Meng, and Z. Xu, "Denoising hyperspectral image with non-IID noise structure," *IEEE Trans. Cybern.*, vol. 48, no. 3, pp. 1054–1066, Mar. 2018.

[38] T.-X. Jiang, L. Zhuang, T.-Z. Huang, X.-L. Zhao, and J. M. Bioucas-Dias, "Adaptive hyperspectral mixed noise removal," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5511413.

[39] S. Xu, X. Cao, J. Peng, Q. Ke, C. Ma, and D. Meng, "Hyperspectral image denoising by asymmetric noise modeling," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5545214.

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

[41] X. Guo, X. Huang, L. Zhang, and L. Zhang, "Hyperspectral image noise reduction based on rank-1 tensor decomposition," *ISPRS J. Photogramm. Remote Sens.*, vol. 83, pp. 50–63, Sep. 2013.

[42] 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.

[43] L. Sun, Q. Cao, Y. Chen, Y. Zheng, and Z. Wu, "Mixed noise removal for hyperspectral images based on global tensor low-rankness and nonlocal SVD-aided group sparsity," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5506617.

[44] 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.

[45] 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.

[46] W. He, H. Zhang, H. Shen, and L. Zhang, "Hyperspectral image denoising using local low-rank matrix recovery and global spatial–spectral total variation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 713–729, Mar. 2018.

[47] 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.

[48] J. Peng et al., "Fast noise removal in hyperspectral images via representative coefficient total variation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5546017.

[49] Y. Chen et al., "Hyperspectral image denoising with weighted nonlocal low-rank model and adaptive total variation regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5544115.

[50] Y. Chen, W. He, N. Yokoya, and T. 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.

[51] 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.

[52] Z. Zha et al., "Nonlocal structured sparsity regularization modeling for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5510316.

[53] Z. Huang et al., "DLRP: Learning deep low-rank prior for remotely sensed image denoising," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[54] 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.

[55] 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.

[56] 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.

[57] 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.

[58] 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.

[59] 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.

[60] 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.

[61] 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. 2020.

[62] Y. Chen, J. Zhang, J. Zeng, W. Lai, X. Gui, and T.-X. Jiang, "A guidable nonlocal low-rank approximation model for hyperspectral image denoising," *Signal Process.*, vol. 215, Feb. 2024, Art. no. 109266.

[63] M. Simões, J. Bioucas-Dias, L. B. Almeida, and J. Chanussot, "A convex formulation for hyperspectral image superresolution via subspace-based regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 6, pp. 3373–3388, Jun. 2015.

[64] L. Zhuang and M. K. Ng, "FastHyMix: Fast and parameter-free hyperspectral image mixed noise removal," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 8, pp. 4702–4716, Aug. 2023.

Yong Chen received the B.S. degree from the School of Science, East China University of Technology, Nanchang, China, in 2015, and the Ph.D. degree from the School of Mathematical Sciences, University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2020.

From 2018 to 2019, he was a Research Intern with the Geoinformatics Unit, RIKEN Center for Advanced Intelligence Project, Tokyo, Japan. He is currently with the School of Computer and Information Engineering, Jiangxi Normal University, Nanchang. His research interests include hyperspectral image processing, low-rank matrix/tensor representation, and model-driven deep learning. More information can be found on his homepage at: https://chenyong1993.github.io/yongchen.github.io/.

Jinshan Zeng (Member, IEEE) received the Ph.D. degree in mathematics from Xi'an Jiaotong University, Xi'an, China, in 2015.

He is currently a Professor with the School of Computer and Information Engineering, Jiangxi Normal University, Nanchang, China, where he has been serving as the Associate Dean since July 2022. He has authored over 60 papers in high-impact journals and conferences, such as *Journal of Machine Learning Research* (JMLR), IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE (TPAMI), IEEE TRANSACTIONS ON SIGNAL PROCESSING (TSP), IEEE TRANSACTIONS ON IMAGE PROCESSING (TIP), IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS (TNNLS), IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS), International Conference on Machine Learning (ICML), and AAAI. He has had two papers co-authored with collaborators that received the International Consortium of Chinese Mathematicians (ICCM) Best Paper Award in 2018 and 2020. His research interests include nonconvex optimization, machine learning, remote sensing, and computer vision.

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 in surveying, mapping, and remote sensing 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 representation, and deep learning.

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

He was a Post-Doctoral Researcher with Prof. Michael K. Ng with Hong Kong Baptist University, Hong Kong, from 2013 to 2014. He was a Visiting Scholar with Prof. Jose Bioucas Dias with the University of Lisbon, Lisbon, Portugal, from 2016 to 2017. He is currently a Professor with the School of Mathematical Sciences, UESTC. His research interests include image processing, machine learning, and scientific computing. More information can be found on his homepage at: https://zhaoxile.github.io/.

Tai-Xiang Jiang received the Ph.D. degree in mathematics from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2019.

He is currently a Professor with the School of Computing and Artificial Intelligence, Southwestern University of Finance and Economics, Chengdu. His research interests include sparse and low-rank modeling, tensor decomposition, and multidimensional image processing.

Qing Huang received the Ph.D. degree in computer software and theory from Wuhan University, Wuhan, China, in 2018.

He is currently an Associate Professor with the School of Computer and Information Engineering, Jiangxi Normal University, Nanchang, China. His research interests include information security, software engineering, and knowledge graph.