Progressive Hyperspectral Image Destriping With an Adaptive Frequencial Focus

Erting Pan, *Student Member, IEEE*, Yong Ma, Xiaoguang Mei, Fan Fan, Jun Huang, and Jiayi Ma, *Senior Member, IEEE*

***Abstract*— Limited by the imaging paradigm, stripes are pervasive in remote sensing scenes, and its intensity, density, and periodicity differ dramatically among different imaging systems. Worse, it always coexists with random noises caused by unstable imaging condition. However, current destriping methods are victim to undue ideal assumptions and fail to accurately eliminate stripes against diverse practical degradation, yielding excessive or inadequate destriping results. This study proposes a progressive hyperspectral destriping method with an adaptive frequency focus for accurate destriping and delicate restoration. Specifically, a hierarchical decomposition and reconstruction framework based on progressive wavelet learning encodes the degraded input to the frequency domain with smaller scales, easing the difficulty of restoration. Then, to avoid excessive or insufficient destriping, we devote specific efforts to finely separating noise and preserving details in the high-frequency domain. First, we devise a gradient-aware frequency attention block based on the prominent unidirectional pattern of stripes, empowering to adaptively assign weights according to their sensitivity to the spatial gradient. Second, we design a focal high-frequency loss item that is dynamically scaled according to the feature distance in the high-frequency domain, profiting in identifying and preserving details. Extensive experiments conducted on data with synthetic stripes and realistic satellite scenes validate the superiority of the proposed method over the current state-of-the-art methods. The code is available at https://github.com/EtPan/PHID.**

***Index Terms*— Frequency domain, hyperspectral image (HSI) destriping, image restoration, wavelet transform.**

I. Introduction

HYPERSPECTRAL images (HSIs) are always affected by complex imaging chains, which alter the content and visual quality of HSI. Stripes and random noise are the most common degradation phenomena in spaceborne and airborne remotely sensed images, especially more ubiquitous in the hyperspectral satellite imaging system [1], [2], [3]. It is primarily caused by unstable sensors, inconsistent responses between different detectors, and charge-coupled device (CCD) dark current interference [4], [5]. As scanners are typically imaged in a whisk- or push-broom manner, noises easily manifest with stripes or defect lines sharing a fixed direction in the scene [6], [7], [8]. Worse still, the ensuing signal distortion and loss of details would lead to failure in many subsequent interpretation tasks [9], [10], [11].

Over the past decades, many scholars have devoted their endeavors to stripes removal in remotely sensed images [12], [13], [14], [15], [16]. Previous methods solve such an ill-posed problem as destriping HSIs is finding appropriate priors. Some of them deeply rely on statistical data properties, such as histogram or moment matching [17], [18]. Some others have employed a variety of fundamental models, including domain transformation [19], [20], sparse representation [21], [22], [23], [24], and total variation [25], [26]. In contrast, inspired by the rapid development of advanced deep learning theory, various neural network-based strategies are proposed for exploring solution to recover noisy HSIs in an end-to-end manner [7], [27], [28]. These methods work by analyzing the specific properties of HSIs, and most of them can handle the ideal stripes. However, some critical issues remain.

1) The degradation of stripes in practical satellite scenes is quite complex with irregular distribution and nonuniform intensity, even mixed with impulse noise or deadlines, which makes traditional methods based on the ideal stripe assumption, i.e., sparse, low rank, and inferior in applicability [20], [29].
2) The low-rank theory-based assumption always regards the stripe distributed periodically or regularly in few bands, which easily leads to harder separation of the stripes and image signals or wrong elimination of stripe-like textures [30], [31].
3) Several intrinsic attributes of stripes, including prominent geometric, unidirectional, and anisotropic properties, are neglected in methods designed for mixed noise removal [32], [33]. In consequence, destriping results are still unsatisfactory.

Regarding these limitations, we argue that developing an advanced destriping method is of utmost significance, requiring efficiently removing stripes, finely preserving the content and details, and being adaptable to various sensors. In essence, stripes arise from a combination of two factors: inconsistent response between detector units and the motion of the large-area imaging sensor, such as whisk- or push-broom imaging [34], [35]. By this, stripes present a remarkable property with its unique structures, involving smoothness along the stripe direction, discontinuity across the stripe direction, and its appearance may be regular rather than

Manuscript received 12 April 2023; revised 20 June 2023; accepted 11 July 2023. Date of publication 21 July 2023; date of current version 2 August 2023. This work was supported in part by the Natural Science Foundation of Hubei Province (Distinguished Young Scholars), China, under Grant 2022CFA073; and in part by the Natural Science Foundation of Guangdong Province, China, under Grant 2023A1515012834. *(Corresponding author: Xiaoguang Mei.)*

The authors are with the Electronic Information School, Wuhan University, Wuhan 430072, China (e-mail: panerting@whu.edu.cn; mayong@whu.edu.cn; meixiaoguang@gmail.com; fanfan@whu.edu.cn; junhwong@whu.edu.cn; jyma2010@gmail.com).

Digital Object Identifier 10.1109/TGRS.2023.3297622

strictly random. Wavelet transform is considered a typical tool for multiscale representation, lossless decomposition, and reconstruction [36]. Also, it has been verified to be effective in modeling the inherent directionality of stripes [37], [38], [39]. We also have reviewed related research in Section II-B. However, due to the complex distribution of various stripes, destriping results by existing wavelet-based techniques still suffer from excessive or insufficient destriping [20], [28].

To this end, we propose a novel method named progressive HSI destriping with adaptive frequencial focus. First, in view of the complex distribution of strip noise in real scenes, we introduce a hierarchical representation of the wavelet transform, intending to alleviate the dilemma of severe stripes through progressive decomposition and reconstruction. After the wavelet decomposition, both the content and stripes are decomposed into low- and high-frequency components on smaller scales, whereas the noise intensity is also decomposed, allowing easier removal. Then, given that components in different frequencies usually comprise different features, we implement progressive wavelet learning in two steps: 1) destriping on low-frequency components and 2) progressive destriping and restoring on high-frequency components. Notably, high-frequency components actually retain varying noise and details. Considering the unidirectional property of stripes, we design a gradient-aware frequencial attention to adaptively assign weights for the noisy component. In addition, we add a skip connection, by which the low-frequency components denoised in the first step can serve to guide the second step and replenish more information. We also propose a focal high-frequency loss to bootstrap the model toward finely distinguishing noise and details in the high-frequency components and further avoiding excessive destriping. Sufficient experiments reveal that the proposed method can reach the state-of-the-art HSI destriping performance. Our main contributions are outlined as follows.

1) We construct a progressive wavelet learning framework (PWL) based on a wavelet transform-like design for multiscale modeling. It enables input with severe stripes to be decomposed into easier and smaller components, providing the trained model with solid generalization capability.
2) We propose a gradient-aware frequency attention block for high-frequency components. It can localize the heaviest distorted component and assign customized weights according to their sensitivity to the spatial gradient.
3) We develop a novel dynamically scaled loss item named focal high frequencial loss, where the scaling factor depends on their feature distances. Such a loss enables the model to focus on the component suffering heavier noise corruption, facilitating further optimization of results.

II. Related Works

A. HSI Destriping and Denoising

Early works solving such an ill-posed problem as HSI destriping mainly focus on finding appropriate priors. Some heavily relied on statistical data properties, such as histogram matching [17], while others employed a variety of fundamental models, including low rank [29], [40], sparse representation [22], [41], and total variation [25], [42], among others. These methods have alleviated this problem to some extent. However, it cannot be ignored that most of them only concentrate on isolated research on specific stripes corruption of the given data, and the majority of them lack comprehensive and systematic analysis of different remote sensing images and various types of stripes noise.

Notably, abundant deep learning-based methods for HSI destriping and denoising have been raised, which have dramatically improved the restoration quality compared to traditional optimization-based methods. For instance, Chang et al. [43] considered the destriping task of remote sensing images as an image decomposition problem and proposed a two-stream CNN. Based on this, HSI-DeNet [44] incorporates residual learning, dilation convolution, and multichannel filtering to improve the denoising performance on top of CNN. Zhong et al. [7] adopted a spaceborne–airborne data joint learning strategy (SGIDN) to balance training samples and test samples and used deep CNN to fully describe the degradation of real HSIs in the spatial and spectral domains to restore the original radiance information better. More recently, scholars have combined model- and data-driven ideas, modeling the inherent properties of HSIs into deep networks to further facilitate restoration performance. For example, Wei et al. [33] proposed a 3-D quasi-recurrent neural network (QRNN3D) to simultaneously explore the structural spatial–spectral correlation and global correlation along the spectral. Zhuang and Ng [32] established a fast-denoising model based on low-rank regularization and sparse representation to reduce the computational burden of large matrices and patchwise iteration. Bodrito et al. [45] proposed a trainable spectral–spatial sparse coding (T3SC) model by employing sparse coding and deep learning. In addition, our prior work [46] developed a spatial–spectral quasi attention network (SQAD) to model the local spectral correlation in recovered HSIs.

Although methods based on deep learning have obvious advantages over traditional methods, such algorithms neglect the unique properties of stripes and have limited generalization capabilities. We conclude that the existing methods are not accurate enough to deal with complex stripes, and their versatility is not ideal.

B. Learning in Frequency Domain

Wavelet transform-based signal decomposition techniques have been well established due to their accurate reconstruction and multiresolution analysis capabilities. Recently, different works have been proposed to incorporate it with CNNs on various tasks involving classification, image hiding, style transfer, super-resolution, and image denoising. For example, Bae et al. [47] found wavelet transform benefits data manifold simplification and proposed a wavelet residual network (WavResNet) for image denoising and single image super-resolution. Guo et al. [48] introduced a deep wavelet super-resolution (DWSR) method that combines complementary information of low- and high-frequency subbands to restore

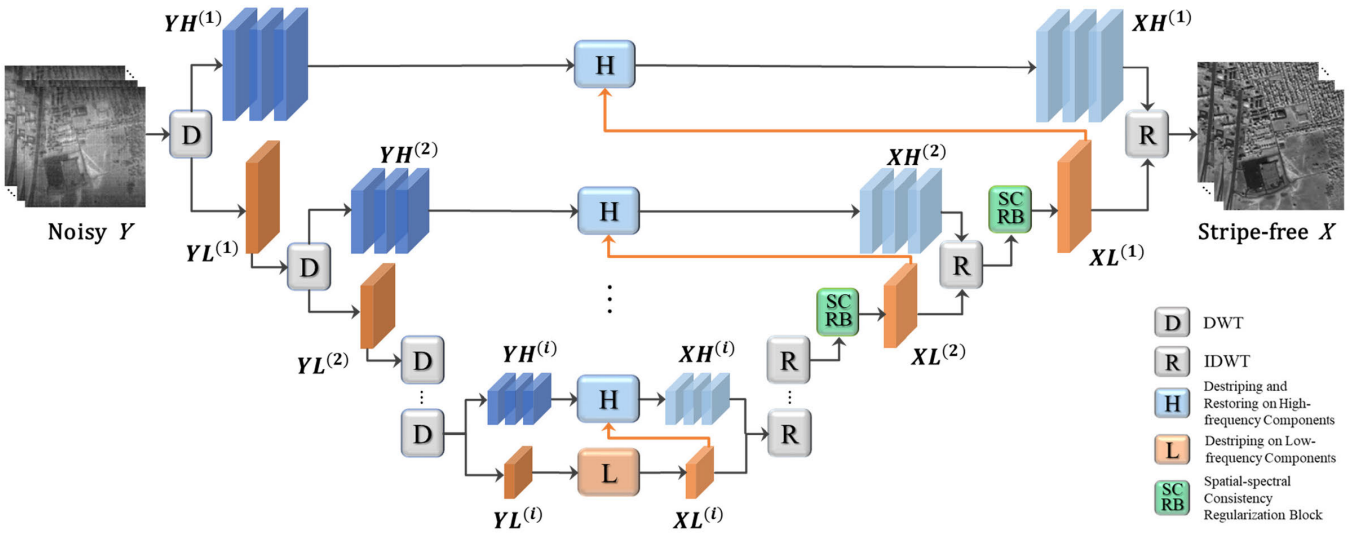


Fig. 1. PWL for HSI destriping. $\mathrm{YH}^{(i)}$ and $\mathrm{YL}^{(i)}$ denote the ith level high- and low-frequency features decomposed from noisy Y, respectively, while $\mathrm{XH}^{(i)}$ and $\mathrm{XL}^{(i)}$ are for reconstructing the clean X.

image details. Unlike the above methods only considering single-level wavelet decomposition, Liu et al. [49] proposed a multilevel wavelet convolutional neural network (MWCNN) that embeds multilevel discrete wavelet transform (DWT) and inverse DWT (IDWT) into CNN to model both context features and inner dependency between subbands. A densely self-guided wavelet network (DSWN) [50] employs the multilevel wavelet transform in a U-Net structure to recover clean real-world images.

Besides, wavelet transforms have also seen widespread use for hyperspectral destriping since the orientation of stripes from whisk- or push-broom sensors in the along-track direction [19], [39], [51]. In short, the wavelet transform concentrates the signal energy from the stripes exclusively into the relatively few wavelet subbands with vertical orientation so that destriping efforts can be easily focused there. Pande-Chhetri and Abd-Elrahman [19] applied a Fourier-domain filter to the wavelet coefficients in vertical subbands. However, previous studies have shown that Fourier-based destriping methods are not effective for nonperiodic stripes. Later, Qian and Ye [51] extended 2-D DWT and 2-D discrete cosine transform (DCT) to 3-D space and produced a dictionary from 3-D hyperspectral cubes. Liu et al. [39] proposed an iterative image decomposition method composed of a low-rank model for the stripes coupled with a group-sparse prior on the wavelet coefficients of the subbands in question. Wang et al. [28] improved destriping performance by applying cross-channel enhanced spatial and spectral attention to focus on key information. Our prior work [12] designed a denosing network in the frequency domain by embedding DWT/IDWT as lossless encoder and decoder.

In a nutshell, existing wavelet-based methods for HSI denoising focus on the advantages of multiresolution analysis provided by wavelet transform. However, they often fail to restore stripe-like textures or overlook the contribution of individual channels to stripe removal, leading to unsatisfactory results. In contrast, our proposed method considers the unique characteristics of stripes with gradient-aware frequencial attention and provides a progressive solution with adaptive frequency focus. This approach preserves high-frequency details and recovers high-fidelity clean HSIs.

III. Methodology

This work devotes to investigating a method that can effectively remove HSI stripes and noise while preserving spatial details and suppressing spectral distortion. Hence, we try to model the stripes in terms of its properties and distribution by designing sensible network architecture and a refined loss function to boost the destriping performance.

The overall architecture of our proposed method is shown in Fig. 1. First, multilevel DWT decomposes the noisy input Y into different components. Then, we conduct destriping [the process $L(\cdot)$] on the low-frequency components, in which stripes can be easier to recognize and remove by the network. Next, the process $H(\cdot)$ is performed for destriping and restoring the cleaner high-frequency components. Later, the restored low- and high-frequency components are incorporated by IDWT, and a spatial–spectral consistency regularization block (SCRB) is employed hereafter to enhance the structural fidelity. Noteworthy, the decomposition and reconstruction processes are conducted processively and recurrently. With the progressive decomposition of the noisy input, its resolution becomes smaller and smaller. It leads to the destriping process actually performing on a small-scale resolution. Also, degradation caused by stripes and noise is weakened with the decomposition process, making the removal more easily. Thus, the progressive wavelet learning architecture can achieve HSI destriping in arbitrary intensity.

A. Progressive Wavelet Learning

Our proposed method is grounded on the wavelet transform, which can decompose the noisy HSI into several same-scale subcomponents with corresponding wavelet coefficients. The Haar wavelet calculated by DWT is exploited to characterize different-frequency features in this work. As observed from

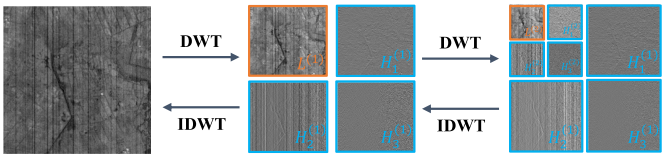


Fig. 2. Illustration of progressive DWT and IDWT process with level 2.

Fig. 2, the principal structural features of the signal are preserved in L, while H_1–H_3 comprise high-frequency noises and details. In this work, to attain the frequency subcomponents at different scales, we perform the wavelet decomposition in a stepwise manner.

Denote the wavelet decomposition process as $D(\cdot)$, and then, the decomposition of the noisy $Y \in \mathcal{R}^{W\times H\times B}$ and subsequent low-frequency components can be formulated as follows:

$$\begin{aligned}\left\{\mathrm{YL}^{(1)}, \mathrm{YH}_1^{(1)}, \mathrm{YH}_2^{(1)}, \mathrm{YH}_3^{(1)}\right\} &= D(Y)\\ \left\{\mathrm{YL}^{(2)}, \mathrm{YH}_1^{(2)}, \mathrm{YH}_2^{(2)}, \mathrm{YH}_3^{(2)}\right\} &= D(\mathrm{YL}^{(1)})\\ &\cdots\\ \left\{\mathrm{YL}^{(i)}, \mathrm{YH}_1^{(i)}, \mathrm{YH}_2^{(i)}, \mathrm{YH}_3^{(i)}\right\} &= D(\mathrm{YL}^{(i-1)})\end{aligned} \tag{1}$$

where i denotes the level of decomposition and $\mathrm{YL}^{(i)}$, $\mathrm{YH}^{(i)} \in \mathcal{R}^{(W/2i)\times(H/2i)\times B}$.

After the decomposition, the noisy HSI turns into multiple noisy wavelet subcomponents. Then, two mappings $L(\cdot)$ and $H(\cdot)$ are required to recover these components and obtain stripeless ones. First, for the top low-frequency noisy component, a process $L(\cdot)$ is utilized to recover $\mathrm{YL}^{(i)}$ into the clean one $\mathrm{XL}^{(i)}$, denoted as

$$\mathrm{XL}^{(i)} = \mathrm{L}\left(\mathrm{YL}^{(i)}\right). \tag{2}$$

Hereafter, the principal structure of stripeless components $\mathrm{XL}^{(i)}$ is obtained. Then, we conduct $H(\cdot)$ on the high-frequency components with the guidance of $\mathrm{XL}^{(i)}$ for further destriping and restoring, signified as

$$\left[\mathrm{XH}_1^{(i)}, \mathrm{XH}_2^{(i)}, \mathrm{XH}_3^{(i)}\right] = \mathrm{H}\left(\mathrm{YL}^{(i)}, \mathrm{YH}_1^{(i)}, \mathrm{YH}_2^{(i)}, \mathrm{YH}_3^{(i)}\right). \tag{3}$$

Then, the recovered wavelet components at the small scale can be combined into the low-frequency component at a larger scale (twice at each level) by IDWT, denoted as $R(\cdot)$. In addition, an SCRB is utilized to enhance the HSI structural fidelity as below,

$$\mathrm{XL}^{(i-1)} = \mathrm{SCRB}\left(\mathrm{R}\left(\mathrm{XL}^{(i)}, \quad \mathrm{XH}_1^{(i)}, \mathrm{XH}_2^{(i)}, \mathrm{XH}_3^{(i)}\right)\right). \tag{4}$$

With the reconstruction process denoted as (3) and (4), the stripeless components can also be recovered progressively. Noteworthy, the decomposition and reconstruction processes are conducted processively and recurrently. With the progressive decomposition of the noisy input, its resolution becomes smaller and smaller. It leads to the destriping process actually performing on a small-scale resolution. Also, degradation is relatively weakened with the decomposition process, making the removal more easily. Thus, the progressive wavelet learning architecture can cope with the complex distribution of stripe and noise and facilitate practical HSI destriping.

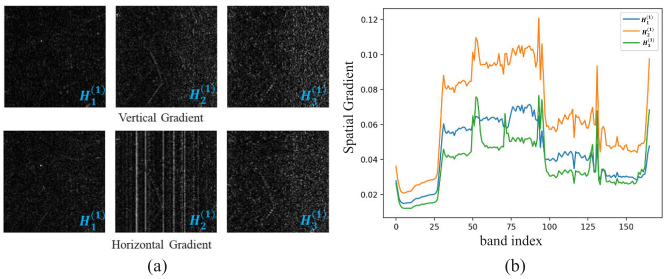


Fig. 3. (a) Illustrations of gradient saliency in different directions for high-frequency components. (b) Their spatial gradient curves.

B. Gradient-Aware Frequencial Attention

As we all know, the direction property of stripes highly depends on the mode of satellite imagery. Existing satellite imagery is essentially a horizontal whisk broom or a vertical push broom to scan a large area, resulting in stripes along the vertical or horizontal direction. Apparently, after decomposition, as shown in Fig. 2, the three high-frequency components preserve diverse details by filtering in different directions, and one of them always suffers the most. In other words, due to the unidirectional property, stripes tend to be more prominent in a certain direction. Therefore, averaging all high-frequency components would easily lead to excessive or insufficient destriping. On this account, we have to treat these components customized to achieve fine noise separation and high-fidelity details restoration in the high-frequency domain.

A natural idea is that assigned different weights to components in different directions. As shown in Fig. 3(a), when the input image has vertical strip noise, we can rapidly and accurately locate the most severely degraded component, i.e., H_2, according to the saliency of the horizontal gradient. Evidently, integrating the gradient calculation in different directions can assist us in locating the components that require higher weights. Therefore, we devised the formula for calculating the spatial gradient as follows:

$$G_{\mathrm{spa}} = \frac{1}{\mathrm{MN}}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\sqrt{G_v(m,n)^2 + G_h(m,n)^2 + \epsilon} \tag{5}$$

where M and N indicate the spatial size of the high-frequency components, respectively, G_v and G_h denote the vertical gradient and the horizontal gradient, respectively, and ϵ is an infinitesimal constant and empirically set to 10^{-6}. Fig. 3(b) shows the spatial gradient curves of these three components along the spectra dimension and evidence that it is reasonable to assign various weights according to spatial gradients.

Furthermore, integrating spatial gradient calculated as (5), we design a gradient-aware frequencial attention block, termed GFA. As shown in Fig. 4, GFA plays the role of a weight regulator, who assigns diverse weights according to the spatial gradient and average feature of each high-frequency component and generates weights value between 0 and 1 by the final sigmoid function.

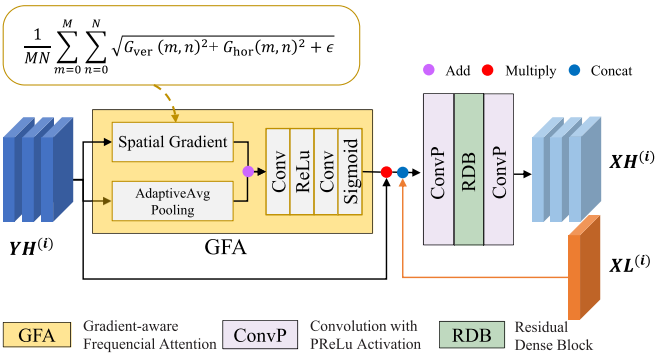


Fig. 4. Details in the process $H(\cdot)$ within gradient-aware frequencial attention.

Despite the adaptive weights generated from GFA, we also associate $\mathrm{XL}^{(i)}$ into the process $H(\cdot)$, assuring the intact high-frequency textures and details with the guidance of the corresponding destriping low-frequency component. As shown in Fig. 4, the main body of $H(\cdot)$ is a residual dense block [52], which not only preserves the feedforward nature but also fully utilizes all the layers within it via dense local connections.

C. Focal High-Frequency Loss Function

The loss function of our model involves three terms: fidelity loss, frequency domain loss, and spectral angular loss. In concrete, the fidelity loss measures the difference before and after HSI destriping from a global perspective, the frequency domain loss forces the low- and high-frequency components to retain the sufficient structure and detailed information upon destriping, and the spectral angular loss constrains the recovered spectra fidelity. The overall loss function $\mathcal{L}_{\text{total}}$ can be expressed as follows:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{fid}} + \lambda_1 \cdot \mathcal{L}_{\text{freq}} + \lambda_2 \cdot \mathcal{L}_{\text{spec}} \tag{6}$$

where $\mathcal{L}_{\text{fid}}$, $\mathcal{L}_{\text{freq}}$, and $\mathcal{L}_{\text{spec}}$ denote the fidelity loss, the frequency domain loss, and the spectral angel loss, respectively, and λ_1 and λ_2 denote the coefficients of the corresponding items. To ensure balanced convergence, they are set to 0.5 and 0.25 based on experiments.

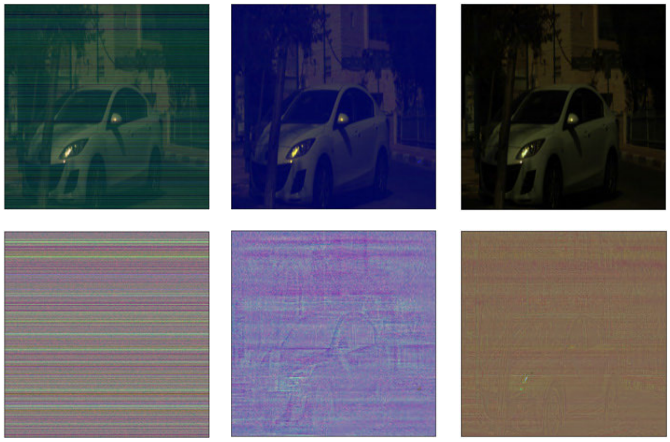
In concrete, L1loss$(\cdot)$ (i.e., mean abs error) is employed for fidelity loss, which is formulated as

$$\mathcal{L}_{\text{fid}} = \text{L1loss}(x, x') \tag{7}$$

where x and x' represent the denoising result and the noisy-free ground truth, respectively. Besides, following the decomposition idea in this work, to incorporate the unique properties of high- and low-frequency domain features, we adopt L2loss$(\cdot)$ (i.e., mean square error) for the structural and content features and design a focal high frequencial loss for the finely high-frequency details, which can be calculated as follows:

$$\mathcal{L}_{\text{freq}} = \text{L2loss}(xl, xl') + \sum i = 1^{I} \frac{1}{2^{(3-i)}} \cdot \text{FHL}_i(xh, xh') \tag{8}$$

where xl and xl' represent the low-frequency features of the denoising result and the noisy-free ground truth, respectively, xh and xh' are the corresponding high-frequency features, and i symbolizes the level of decomposition (I is configured as 3 in this work).



(a) (b) (c)

Fig. 5. Comparison of destriping results with and without focal high-frequency loss function. (a) Noisy image and its noise. (b) Result and its noise residual map produced by the network without focal high-frequency loss function. (c) Result and its noise residual map produced by the network with focal high-frequency loss function.

In particular, $\text{FHL}_i(\cdot)$ signifies the focal high frequencial loss. Its design likewise follows the previous idea, i.e., the high-frequency components in different directions are corrupted by noise to varying degrees, making their recovery difficult. Therefore, we design a feature distance-based loss function for the high-frequency components, allowing us to focus on recovering the heavier noise-corrupted one, which is calculated as follows:

$$\text{FHL}_i(xh, xh') = \frac{1}{c \times n \times m} \sum^{C} \sum^{M} \sum^{N} \alpha \cdot |xh_i - xh'_i| 1 \tag{9}$$

where $\alpha = \left|xh - xh'\right| \left(\alpha \in \mathcal{R}^{1\times 3}\right)$ and C, M, and N are the sizes of high-frequency components.

As represented in Fig. 5, with the focal high-frequency loss function, the proposed network could restore stripe-free results with higher spatial fidelity and are capable of preserving local details. In particular, comparing the noise residual maps in the bottom of Fig. 5, it proves that our method discriminates well between noisy and effective features in high-frequency details, avoiding the loss of critical information.

IV. Experiments

A. Experimental Settings

1) Data: We conduct destriping experiments on data with synthetic degradation and real satellite HSIs. For training, datasets are from publicly available HSI dataset ICVL[1] [53], where 120 out of 201 practically noise-free HSIs are sampled as the training set in our experiment, and along with 5 for validation, the rest constitute the testing set. Each HSI cube in ICVL has a spatial resolution of 1392 × 1300 and a spectral dimension of 31. To further demonstrate the generalization

[1] http://icvl.cs.bgu.ac.il/hyperspectral/

ability of our model, we also test real HSIs collected from satellites, involving the GF-5 Baoqing dataset and the GF-5 Wuhan dataset[2] with 155 bands acquired by the Advanced Hyperspectral Imager in GF-5 satellite [54].

2) Synthetic Noise Settings: We mimic a series of synthetic stripe noisy cases involving: 1) vertical stripes; 2) horizontal stripes; 3) a mixture of vertical stripes and non-independent and identically distributed (i.i.d) Gaussian noise; and 4) a mixture of horizontal stripes and non-i.i.d Gaussian noise. We configure some associated settings to maximize the imitation of practical noise contamination. Specifically, two-thirds of bands, 5%–45% of columns or rows in the first two cases, and one-third of bands, 5%–35% of columns or rows for the last two cases, are randomly selected to add stripes or noises. Also, the non-i.i.d. Gaussian noise in the last two cases is added to one-third of the bands with a random intensity ranging in the interval [10, 70].

3) Implementation Details: Noteworthy, considering that HSIa collected by different sensors that usually have varying spectral resolution, we expand a new dimension for both input and output, extending the 4-D tensor (batch_size, C, H, W) to a 5-D tensor (batch_size, 1, C, H, W), where C indexes the spectral dimension and H and W represent the spatial domain. It enables the transformation and calculation of feature maps to occur primarily on the other three dimensions, except "C, making the model's parameters independent of the spectral dimension. The training patches have a spatial size of 64 $\times$ 64 and the complete spectral dimension 31. To improve the robustness of training, we scale the patches with a rate in $\{1, 0.5, 0.25\}$ and randomly rotate them to further diversify the training set. Besides, 80 epochs are set up throughout the training process. To be concrete, during the first 40 epochs, we added mimic stripe noise in different directions to the training data and validated the model on cases 1 and 2. We added mixed noise for the latter 40 epochs and validated the model on cases 3 and 4. The batch size (i.e., 16) is employed to stabilize the training. Parameters are initialized with Kaiming initialization and updated by the Adam optimizer. The learning rate is set to 0.001 and decreases exponentially with epochs until the validation performance does not increase anymore. Noteworthy, all deep learning-based methods run on the same GPU RTX 3090Ti, while other methods run on the same CPU i9-10940H.

4) Evaluation Metrics: We employ four quantitative quality indices to evaluate the denoising performance, inclusive of mean peak signal-to-noise ratio (MPSNR), mean structure similarity (MSSIM), and mean feature similarity (MFSIM) for spatial-based image quality evaluation, as well as mean relative dimensionless global error synthesis (MERGAS), and mean spectral angle mapper (MSAM) for quantitative evaluation on the spectral dimension. In addition, we report the average running time, which is measured by processing a noisy HSI input with the size of 512 $\times$ 512 $\times$ 31.

[2]http://hipag.whu.edu.cn/dataset/

B. Results on Data With Synthetic Noises

To verify the effectiveness of our method in stripes and noises removal, we compare it with several existing comparative HSI destriping methods and state-of-the-art methods for mixed HSI noises removal: ASSTV [25], LRTD [29], LRTDTV [26], LRTDGS [41], QRNN3D [33], SGIDN [7], FastHyMix [32], T3SC [45], SQAD [46], MAC-Net [13], GRUNet [14], En-OctNet [28], and D2Net [12]. For most algorithms, we adopt the publicly available source code and the given parameters or trained model from the authors, except for SGIDN and T3SC, which we reproduced in the PyTorch-based framework and retrained with the synthetic data in this work.

1) Spatial Quality Comparisons: The visual results of some representative scenes are shown in Fig. 6. From the results of the first two cases, destriping methods ASSTV and LRTD effectively remove most of the stripes but with minimal image quality improvement due to the remain of random noises. Regarding the results of the mixed noise removal algorithms, the residual noise maps of LRTDTV, T3SC, QRNN3D, SQAD, and MAC-Net reveal that some essential information and texture details are excessively denoised. A large amount of strip noise still remains in the results recovered by LRTDGS, SGIDN, and T3SC, which implies that they cannot cope with such cases of individual stripes corruption. Compared with the advanced SGIDN, our proposed method is more delicate in retaining details, and the recovered image obtains the optimal quantization value. Similar conclusions can also be obtained by observing the results of the last cases. Algorithms designed only for stripes fail to remove mixed noise and show poor performance. QRNN3D, SQAD, MAC-Net, and GRUNet still suffer from excessive denoising and present oversmoothed and blurred results. FastHyMix and T3SC exhibit stripes leftovers due to the inadequate consideration of stripes. Compared with other methods, the methods based on wavelet transform, i.e., En-OctNet, D2Net, and the proposed method, can significantly reduce the loss of texture features, as can be seen from the noise residual map in Fig. 6. Our proposed method obtains the best denoising results with a superior model, achieves high-fidelity recovery, and shows better results with fewer artifacts and sharper edges.

2) Spectral Fidelity Comparisons: To verify the superiority of our method in spectral fidelity, we select one typical pixel in all four cases and draw its spectral curves recovered by comparing methods in Fig. 7. Apparently, methods, such as FastHyMix, T3SC, and MAC-Net, show unstable performance for restoring different bands, and the spectral distortion of the recovered pixel even gets worsen. Compared to the other results, the spectra curve recovered by our proposed method is much closer to the reference under all simulated degradation cases. It indicates that our method accurately eliminates the negative effect of noise in the spectral domain and further confirms the advantage of our model in maintaining high spectral fidelity.

3) Quantitative Comparisons: Table I shows the mean quantitative statistical results of the testing set for different striping noise pollution cases. The results indicate that our proposed method performs robustly even when severely cor-

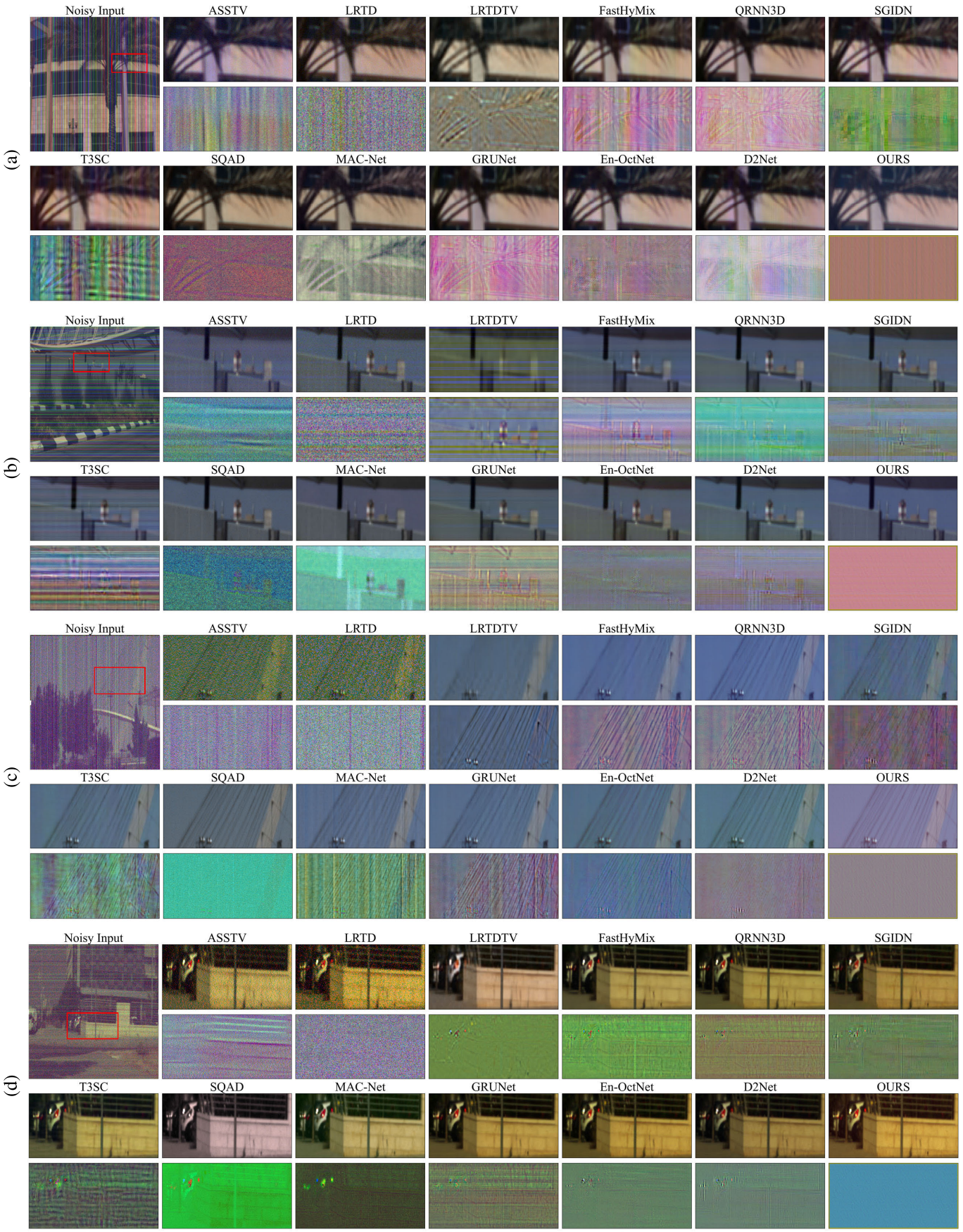


Fig. 6. Denoising results on the ICVL dataset illustrated with selected RGB bands. Four rows from top to bottom correspond to (a)–(d) Cases 1–4, respectively. In each row, we zoom in on a region of interest and show the residual noise map. We show the results from comparative destriping methods and also represent the results from state-of-the-art methods for HSI mixed noise removal.

TABLE I

QUANTITATIVE PERFORMANCE OF ALL COMPETING METHODS UNDER CASES WITH DIFFERENT SYNTHETIC STRIPES NOISES ON THE ICVL DATASET. RED: BEST RESULTS. PURPLE: SECOND

Metricx	Noisy	ASSTV [25]	LRTD [29]	LRTDTV [26]	LRTDGS [41]	FastHyMix [32]	QRNN3D [33]	SGIDN [7]	T3SC [45]	SQAD [46]	MAC-Net [13]	GRU-Net [14]	En-OctNet [28]	D2Net [12]	Ours
Case 1: Synthetic vertical stripes															
MPSNR ↑	25.698	29.313	38.188	37.698	31.954	38.656	40.918	42.064	34.867	39.663	36.089	37.145	36.021	40.417	42.664
MSSIM ↑	0.5452	0.7833	0.8962	0.9331	0.7592	0.7570	0.9695	0.9698	0.8714	0.9412	0.9274	0.9411	0.9530	0.9716	0.9810
MFSIM ↑	0.8046	0.9662	0.9880	0.9694	0.8870	0.8997	0.9907	0.9967	0.9548	0.9865	0.9863	0.9832	0.9884	0.9889	0.9933
MERGAS ↓	35.1730	28.2145	9.4970	7.0989	20.6069	20.6927	3.6517	3.6388	6.1827	7.3055	10.6023	5.7172	7.1121	3.5404	2.3742
MSAM ↓	0.5715	0.2942	0.1760	0.1193	0.3848	0.3740	0.1010	0.1412	0.1921	0.1452	0.1621	0.1316	0.1244	0.0851	0.0636
Case 2: Synthetic horizontal stripes															
MPSNR↑	25.683	29.501	37.950	37.817	36.290	38.369	38.936	41.576	31.026	39.492	36.041	37.588	38.032	40.648	42.733
MSSIM ↑	0.5442	0.7918	0.8968	0.9350	0.9246	0.7566	0.9564	0.9671	0.8220	0.9413	0.9291	0.9356	0.9581	0.9732	0.9830
MFSIM ↑	0.8034	0.9706	0.9882	0.9707	0.9620	0.8992	0.9857	0.9958	0.9454	0.9767	0.9867	0.9828	0.9919	0.9898	0.9945
MERGAS ↓	34.3328	28.1467	10.1943	5.8923	6.1719	19.6647	4.5015	3.7137	12.1287	6.6349	10.2374	5.2820	5.7485	3.2167	2.3032
MSAM↓	0.5720	0.2967	0.1755	0.1171	0.1443	0.3719	0.1177	0.1407	0.2188	0.1318	0.1674	0.1247	0.1199	0.0854	0.0664
Case 3: Synthetic vertical stripes + non-i.i.d. Gaussian noise															
MPSNR ↑	20.613	24.614	26.008	37.455	34.635	35.273	40.795	40.375	41.344	40.3260	39.677	41.322	38.957	41.685	41.748
MSSIM ↑	0.2938	0.4332	0.4418	0.9347	0.8859	0.8052	0.9719	0.9635	0.9676	0.9660	0.9583	0.9671	0.9516	0.9710	0.9718
MFSIM ↑	0.6906	0.8276	0.7911	0.9718	0.9396	0.9399	0.9837	0.9964	0.9903	0.9865	0.9841	0.9820	0.9782	0.9935	0.9894
MERGAS↓	50.6600	36.6471	37.8023	4.9425	6.1755	8.0674	2.9635	4.0369	2.8433	2.8924	3.2326	2.9446	3.5769	2.5487	2.5106
MSAM↓	0.8294	0.6084	0.6383	0.0940	0.1367	0.2206	0.0897	0.1515	0.0852	0.0682	0.0888	0.0755	0.0936	0.0673	0.0662
Case 4: Synthetic horizontal stripes + non-i.i.d. Gaussian noise															
MPSNR	27.231	28.420	32.019	38.644	37.064	41.195	44.056	42.457	37.954	40.9730	44.559	43.886	40.813	42.717	45.041
MSSIM	0.5335	0.6467	0.6692	0.9471	0.9270	0.8722	0.9835	0.9691	0.9147	0.9720	0.9815	0.9820	0.9719	0.9818	0.9865
MFSIM	0.8084	0.9063	0.8876	0.9756	0.9565	0.9576	0.9944	0.9972	0.9598	0.9890	0.9928	0.9942	0.9887	0.9958	0.9964
MERGAS↓	38.2690	30.3285	28.5305	4.1489	4.4541	7.2124	1.9044	3.5220	3.9728	2.6910	1.7988	1.9944	2.8414	2.4047	1.7234
MSAM ↓	0.7074	0.5180	0.5361	0.0759	0.0809	0.1890	0.0514	0.1407	0.1034	0.0649	0.0552	0.0527	0.0722	0.0578	0.0493
TIMES ↓	-	28.13	273.65	357.31	75.23	2.68	0.80	1.85	1.11	0.68	3.21	0.85	0.99	0.91	0.60

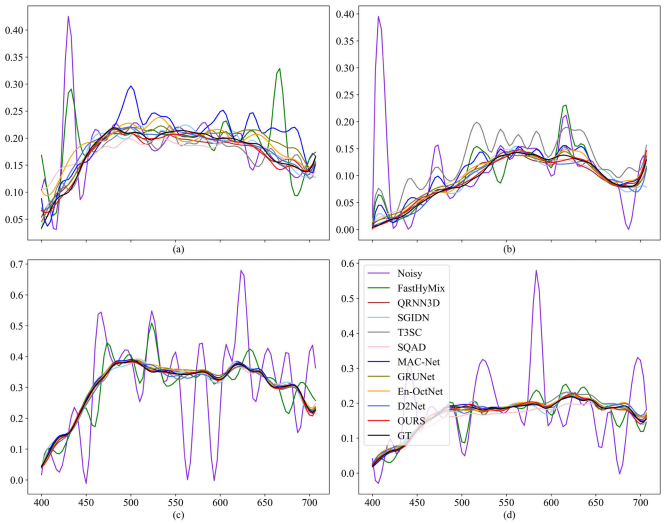


Fig. 7. Spectral fidelity of comparative methods under four synthetic noisy cases on the ICVL dataset are presented, respectively. (a) Case 1: synthetic vertical stripes. (b) Case 2: synthetic horizontal stripes. (c) Case 3: synthetic vertical stripes + non-i.i.d. Gaussian noise. (d) Case 4: synthetic horizontal stripes + non-i.i.d. Gaussian noise.

rupted by mixed noise. This demonstrates the effectiveness of our PWL, which progressively decomposes severe noise and releases the difficulty in restoration. In contrast, traditional methods, such as ASSTV and LRTD designed for destriping noise, as well as LRTDTV and LRTDGS for removing mixed noise, require time-consuming customized parameter settings and only offer moderate restoration performance. FastHyMix, T3SC, MAC-Net, and GRUNet exhibit less stable image restoration capabilities and only have relatively competitive denoising effects in some cases. Although SGIDN performs better in the MFSIM index and achieves stable and optimal values, it is still inferior to our proposed method in other indices such as MPSNR, MSSIM, MERGAS, and MSAM. Overall, our method significantly improves the spatial visual quality and spectral fidelity indicators in all synthetic noise cases, especially in the MPSNR index, which increases by an average of approximately 17 dB. However, our method sacrifices a little in the MFSIM index, which is on average 0.004 lower than SGIDN. This indicates the superior flexibility of our proposed method. To provide a more detailed quantitative comparison, we also present Fig. 8, which includes the PSNR and SSIM metrics evaluated along the entire spectral dimension for a typical HSI in each case. Although our proposed algorithm did not achieve the best performance in terms of PSNR in some bands, it still achieved a stable improvement in PSNR values compared to other advanced algorithms in all bands. This also demonstrates the robustness of our proposed method for image restoration in the spectral dimension.

C. Results on Real Satellite HSIs

We have selected ASSTV [25], LRTD [29], LRTDTV [26], FastHyMix [32], SGIDN [7], SQAD [46], MAC-Net [13], and D2Net [12] as to validate and compare the performance of HSI restoration on real satellite datasets. For the proposed method in this article, as the degradation in cases 3 and 4 with complex noise has mimic actual degradation to the greatest extent possible, we directly use the models trained under these cases to evaluate the real satellite HSIs. We have used publicly available parameter settings and models pretrained in the aforementioned synthetic experiments for most algorithms.

Figs. 9 and 10 present some representative visualization results from the GF5 Baoqing dataset and the GF5 Wuhan dataset, respectively. Apparently, complex degradation with stripes and other noises exists in these data among different bands to varying degrees. Despite this, our method exhibits stable performance with the slightest noise residual (remarkable improvements in comparison with traditional destriping methods ASSTV, LRTD, LRTDTV, and FastHyMix) and the highest

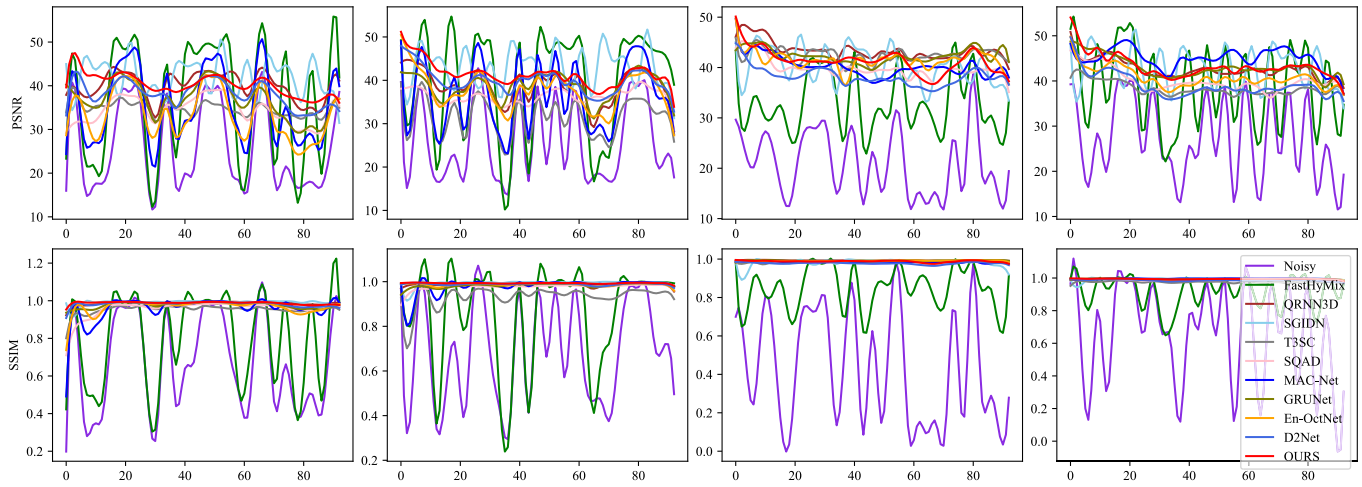


Fig. 8. Quantitative performance on PSNR and SSIM along the spectral dimension of advanced competing methods on the ICVL dataset. Four columns from left to right correspond to Cases 1–4.

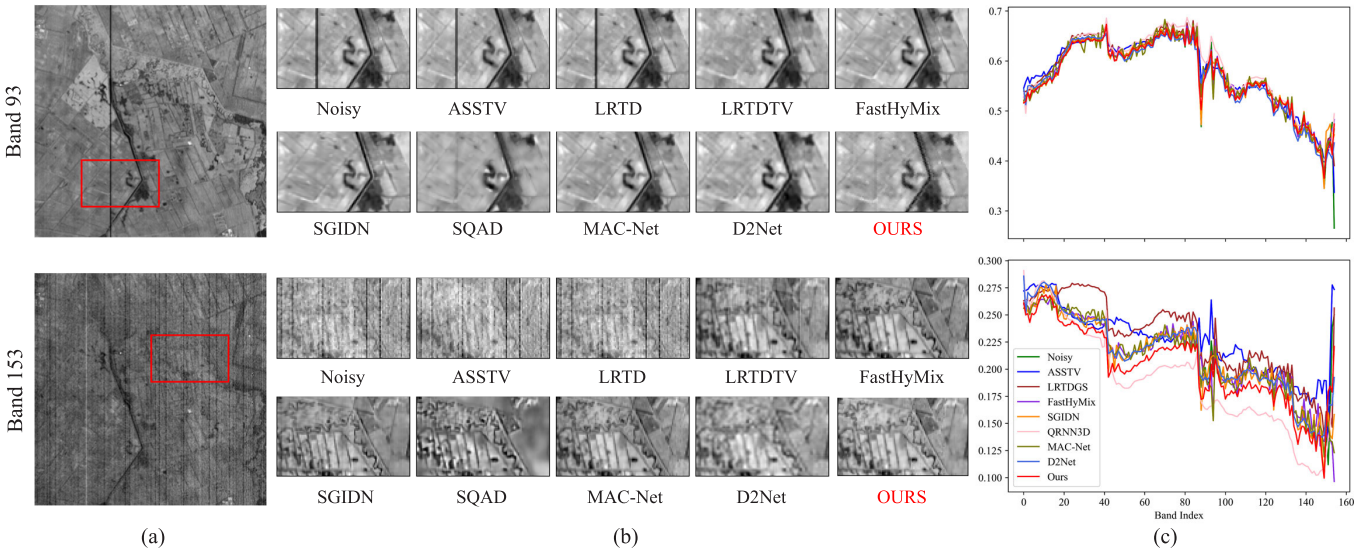


Fig. 9. Destriping results on the GF5 Baoqing dataset from some comparative methods. (a) Images of noisy bands. (b) ROI of images restored by different methods. (c) Spectral of two typical pixels recovered by different methods.

TABLE II

ABLATION STUDIES (INVOLVING THE INVESTIGATION OF SUBCOMPONENTS AND LOSS ITEMS) ON THE ICVL DATASET. WE EVALUATE THE RESULTS BY PSNR (dB), SSIM, SAM, AND THE NUMBER OF PARAMETERS (PARAMS) OF THESE NETWORKS. OUR BENCHMARK NETWORK IS INDICATED BY BOLDFACE. (NOTES: D/R DENOTES DWT/IDWT AND LFG SIGNIFIES THE GUIDANCE OF DENOISED LOW-FREQUENCY COMPONENTS)

No.	Subcomponents						Loss				Metrics			
	PWL		$L(\cdot)$	$H(\cdot)$			$\mathcal{L}_{\text{fid}}$	$\mathcal{L}_{\text{freq}}$		$\mathcal{L}_{\text{sam}}$	MPSNR	MSSIM	MSAM	Params
	D/R	SCRB		RDB	LFG	GFA		w/o FHL	with FHL					
1	✓		✓	✓			✓				35.47	0.943	0.134	1.10M
2	✓		✓	✓	✓		✓				37.75	0.965	0.112	1.11M
3	✓		✓	✓	✓	✓	✓				38.64	0.973	0.114	1.11M
4	✓	✓	✓	✓	✓	✓	✓				38.51	0.971	0.107	1.12M
5	✓	✓	✓	✓	✓	✓	✓	✓			38.86	0.975	0.109	1.12M
6	✓	✓	✓	✓	✓	✓	✓		✓		39.12	0.979	0.105	1.12M
OURS	✓	✓	✓	✓	✓	✓	✓		✓	✓	**39.28**	**0.982**	**0.096**	**1.12M**

detail recovery fidelity [especially obvious in comparing with SGIDN, SQAD, MAC-Net, and D2Net in Fig. 10(b)]. It also confirms the ability of our model to generalize in real satellite scenes. On the other hand, Figs. 9 and 10 also present the spectral curves of two typical pixels recovered by some leading methods. The original spectral curves befall some localized

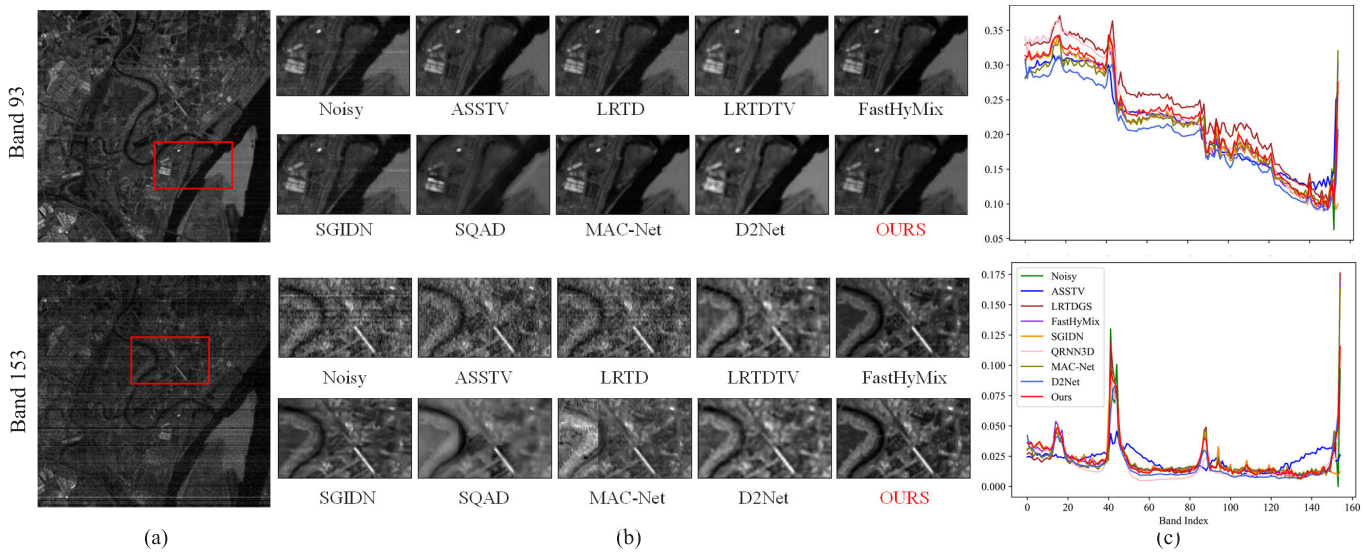


Fig. 10. Destriping results on the GF5 Wuhan dataset from some comparative methods. (a) Images of noisy bands. (b) ROI of images restored by different methods. (c) Spectral of two typical pixels recovered by different methods.

TABLE III
RESULTS OF ABLATION STUDIES ON GFA

No.	GFA		Metrics	
	Spatial Gradient	AdaptiveAvg Pooling	MPSNR	MSSIM
1	-	-	38.14	0.973
2	✓	-	38.55	0.976
3	-	✓	38.52	0.980
4	✓	✓	**39.28**	**0.982**

dense fluctuations and a few spikes because of noises or stripes corruption. It is pretty clear that the proposed destriping method alleviates this problem to a certain extent, and the overall shape of the recovered spectral curves shows more consistency and retains more meaningful intrinsic features. It verifies the effectiveness and flexibility of our model.

D. Ablation Studies

To downright verify the effectiveness of each endeavor for HSI destriping in this work, we conduct comprehensive ablation studies on the ICVL dataset, involving the investigation of subcomponents and loss items. In contrast to the settings of previous comparative experiments, the ablation experiments adjust settings in two main aspects. Specifically, we reduce the training data size to only 2/3 of the original and evaluate the model of each ablation only on the simulation of case 1. Such a setting is efficient and also can comprehensively reflect the positive effects of all modules on model training. Table II lists the relevant indicators with various ablation experiments, and a detailed analysis is carried out as follows.

The top four rows investigate the validity of each subcomponent of the backbone in the proposed method, which consists of PWL, the process $L(\cdot)$, and the process $H(\cdot)$. Obviously, the designed GFA imposes no additional computational burden and brings considerable gains in other metrics (like 2.28 dB in MPSNR, 0.22 increase in MSSIM, and 0.022 decreases in MSAM), while adding denoised low-frequency components' guidance (LFG) and SCRB increments the number of parameters by almost 0.1 M. It reveals the advantages of GFA. SCRB significantly optimizes the performance on MSAM metrics by sacrificing a little bit of spatial quality. On the other hand, the optimal model of our method is trained under the joint constraint of multiple loss items. The ablation experiments in the last four rows exemplify the irreplaceable role of each loss item, especially focal high-frequency loss function (FHL) in $\mathcal{L}_{\text{freq}}$ (brings over 0.6-dB increments in MPSNR). In a nutshell, the incremental designs in this work can be summarized into two categories: 1) both GFA in the process $H(\cdot)$ and FHL in the loss function term $\mathcal{L}_{\text{freq}}$ serve to assign adaptively frequencial focus in high-frequency domain, which significantly improves the quality of the recovered HSI, and 2) SCRB in PWL and LFG in the process $H(\cdot)$ and $\mathcal{L}_{\text{spec}}$ guarantee their high spatial–spectral fidelity.

V. CONCLUSION

In this article, we propose a progressive HSI destriping method with an adaptive frequency focus. It has a PWL, which eases the difficulty of restoration by decomposing the degraded input over multiple levels. Incorporating the unidirectional property of stripes, we develop a gradient-aware frequency attention block, which can localize the high-frequency component heaviest distorted by stripes via a spatial gradient calculator and assign customized weights according to their sensitivity to the spatial gradient. At the same time, we design a focal high-frequency loss item that is dynamically scaled according to feature distance in the high-frequency domain. Such designs facilitate the delicate recovery of high-frequency components and can effectively avoid excessive or insufficient denoising. Qualitative and quantitative experiments demonstrate the superiority of the proposed method in complex and varying HSI restoration task. In addition, we believe that the idea of this work can also be flexibly transferred or promoted

to other low-level vision tasks for future insightful research, such as HSI restoration, enhancement, unmixing, and super-resolution.

REFERENCES

[1] B. Datt, T. R. McVicar, T. G. Van Niel, D. L. B. Jupp, and J. S. Pearlman, "Preprocessing EO-1 hyperion hyperspectral data to support the application of agricultural indexes," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 6, pp. 1246–1259, Jun. 2003.

[2] F. Tsai and W. W. Chen, "Striping noise detection and correction of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 12, pp. 4122–4131, Dec. 2008.

[3] Y. Zhong, X. Wang, S. Wang, and L. Zhang, "Advances in spaceborne hyperspectral remote sensing in China," *Geo-spatial Inf. Sci.*, vol. 24, no. 1, pp. 95–120, Jan. 2021.

[4] K. C. Lawrence, B. Park, W. R. Windham, and C. Mao, "Calibration of a pushbroom hyperspectral imaging system for agricultural inspection," *Trans. ASAE*, vol. 46, no. 2, p. 513, 2003.

[5] H. Cui et al., "Multiscale intensity propagation to remove multiplicative stripe noise from remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2308–2323, Apr. 2020.

[6] X. Liu, X. Lu, H. Shen, Q. Yuan, Y. Jiao, and L. Zhang, "Stripe noise separation and removal in remote sensing images by consideration of the global sparsity and local variational properties," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 3049–3060, May 2016.

[7] Y. Zhong, W. Li, X. Wang, S. Jin, and L. Zhang, "Satellite-ground integrated destriping network: A new perspective for EO-1 hyperion and Chinese hyperspectral satellite datasets," *Remote Sens. Environ.*, vol. 237, Feb. 2020, Art. no. 111416.

[8] Y. Xu et al., "Luojia-HSSR: A high spatial–spectral resolution remote sensing dataset for land-cover classification with a new 3D-HRNet," *Geo-spatial Inf. Sci.*, pp. 1–13, Jun. 2022.

[9] L. Fang, D. Zhu, J. Yue, B. Zhang, and M. He, "Geometric-spectral reconstruction learning for multi-source open-set classification with hyperspectral and LiDAR data," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 10, pp. 1892–1895, Oct. 2022.

[10] Z. Shao, W. Wu, and D. Li, "Spatio-temporal-spectral observation model for urban remote sensing," *Geo-Spatial Inf. Sci.*, vol. 24, no. 3, pp. 372–386, Jul. 2021.

[11] L. Tang, Y. Deng, Y. Ma, J. Huang, and J. Ma, "SuperFusion: A versatile image registration and fusion network with semantic awareness," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 12, pp. 2121–2137, Dec. 2022.

[12] E. Pan, Y. Ma, X. Mei, J. Huang, F. Fan, and J. Ma, "D2Net: Deep denoising network in frequency domain for hyperspectral image," *IEEE/CAA J. Autom. Sinica*, vol. 10, no. 3, pp. 813–815, Mar. 2023.

[13] 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.

[14] Z. Lai, K. Wei, and Y. Fu, "Deep plug-and-play prior for hyperspectral image restoration," *Neurocomputing*, vol. 481, pp. 281–293, Apr. 2022.

[15] C. Wang et al., "Toward real hyperspectral image stripe removal via direction constraint hierarchical feature cascade networks," *Remote Sens.*, vol. 14, no. 3, p. 467, Jan. 2022. [Online]. Available: https://www.mdpi.com/2072-4292/14/3/467

[16] X. Rui, X. Cao, Q. Xie, Z. Yue, Q. Zhao, and D. Meng, "Learning an explicit weighting scheme for adapting complex HSI noise," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 6735–6744.

[17] P. Rakwatin, W. Takeuchi, and Y. Yasuoka, "Stripe noise reduction in MODIS data by combining histogram matching with facet filter," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1844–1856, Jun. 2007.

[18] B. Nie, L. Yang, J. Jing, and J. Zhou, "Hyperspectral image destriping method based on time-frequency joint processing method," *Optik*, vol. 172, pp. 317–327, Nov. 2018.

[19] R. Pande-Chhetri and A. Abd-Elrahman, "De-striping hyperspectral imagery using wavelet transform and adaptive frequency domain filtering," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 5, pp. 620–636, Sep. 2011.

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

[21] Q. Xie et al., "Multispectral images denoising by intrinsic tensor sparsity regularization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1692–1700.

[22] Y. Chen, T.-Z. Huang, L.-J. Deng, X.-L. Zhao, and M. Wang, "Group sparsity based regularization model for remote sensing image stripe noise removal," *Neurocomputing*, vol. 267, pp. 95–106, Dec. 2017.

[23] 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.

[24] H. Zhang, J. Cai, W. He, H. Shen, and L. Zhang, "Double low-rank matrix decomposition for hyperspectral image denoising and destriping," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5502619.

[25] Y. Chang, L. Yan, H. Fang, and C. Luo, "Anisotropic spectral–spatial total variation model for multispectral remote sensing image destriping," *IEEE Trans. Image Process.*, vol. 24, no. 6, pp. 1852–1866, Jun. 2015.

[26] 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.

[27] C. Wang et al., "Translution-SNet: A semisupervised hyperspectral image stripe noise removal based on transformer and CNN," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5533114.

[28] C. Wang et al., "Hyperspectral image stripe removal network with cross-frequency feature interaction," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5521515.

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

[30] S. Zhao, Q. Yuan, J. Li, Y. Hu, X. Liu, and L. Zhang, "A fast and effective irregular stripe removal method for moon mineralogy mapper (M^3)," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4600119.

[31] 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.

[32] L. Zhuang and M. K. Ng, "FastHyMix: Fast and parameter-free hyperspectral image mixed noise removal," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Sep. 23, 2021, doi: 10.1109/TNNLS.2021.3112577.

[33] 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.

[34] J. E. Fowler, "Compressive pushbroom and whiskbroom sensing for hyperspectral remote-sensing imaging," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 684–688.

[35] B.-L. Hu, S.-J. Hao, D.-X. Sun, and Y.-N. Liu, "A novel scene-based non-uniformity correction method for SWIR push-broom hyperspectral sensors," *ISPRS J. Photogramm. Remote Sens.*, vol. 131, pp. 160–169, Sep. 2017.

[36] C. K. Chui, *An Introduction to Wavelets*. Amsterdam, The Netherlands: Elsevier, 2016.

[37] B. Münch, P. Trtik, F. Marone, and M. Stampanoni, "Stripe and ring artifact removal with combined wavelet-Fourier filtering," *Optics Exp.*, vol. 17, no. 10, pp. 8567–8591, 2009.

[38] L. Ye and Z. Hou, "Memory efficient multilevel discrete wavelet transform schemes for JPEG2000," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 11, pp. 1773–1785, Nov. 2015.

[39] N. Liu, W. Li, R. Tao, and J. E. Fowler, "Wavelet-domain low-rank/group-sparse destriping for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 10310–10321, Dec. 2019.

[40] 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.

[41] 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.

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

[43] 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.

[44] 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.

[45] 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.

[46] E. Pan, Y. Ma, X. Mei, F. Fan, J. Huang, and J. Ma, "SQAD: Spatial–spectral quasi-attention recurrent network for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5524814.

[47] W. Bae, J. Yoo, and J. C. Ye, "Beyond deep residual learning for image restoration: Persistent homology-guided manifold simplification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 1141–1149.

[48] T. Guo, H. S. Mousavi, T. H. Vu, and V. Monga, "Deep wavelet prediction for image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 1100–1109.

[49] P. Liu, H. Zhang, W. Lian, and W. Zuo, "Multi-level wavelet convolutional neural networks," *IEEE Access*, vol. 7, pp. 74973–74985, 2019.

[50] W. Liu, Q. Yan, and Y. Zhao, "Densely self-guided wavelet network for image denoising," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2020, pp. 1742–1750.

[51] 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.

[52] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image super-resolution," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 2472–2481.

[53] B. Arad and O. Ben-Shahar, "Sparse recovery of hyperspectral signal from natural RGB images," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2016, pp. 19–34.

[54] 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.

Erting Pan (Student Member, IEEE) received the B.S. degree in electrical engineering and its automation from Northeast Normal University, Changchun, China, in 2018, and the M.E. degree in electronic and communication engineering from Wuhan University, Wuhan, China, in 2020, where she is currently pursuing the Ph.D. degree with the Multi-Spectral Vision Processing Laboratory, Electronic Information School.

Her research interests include remote sensing image processing, computer vision, and pattern recognition.

Yong Ma graduated from the Department of Automatic Control, Beijing Institute of Technology, Beijing, China, in 1997. He received the Ph.D. degree from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2003.

From 2004 to 2006, he was a Lecturer at the University of the West of England, Bristol, U.K. From 2006 to 2014, he was with the Wuhan National Laboratory for Optoelectronics, HUST, where he was a Professor of electronics. He is currently a Professor with the Electronic Information School, Wuhan University. His research projects include remote sensing of light detection and ranging (LiDAR) and infrared, as well as infrared image processing, pattern recognition, and interface circuits to sensors and actuators. His general field of research is in signal and systems.

Xiaoguang Mei received the B.S. degree in communication engineering from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2007, the M.S. degree in communications and information systems from Huazhong Normal University, Wuhan, in 2011, and the Ph.D. degree in circuits and systems from HUST in 2016.

From 2010 to 2012, he was a Software Engineer with the 722 Research Institute, China Shipbuilding Industry Corporation, Wuhan. From 2016 to 2019, he was a Post-Doctoral Fellow with the Electronic Information School, Wuhan University, Wuhan, where he is currently an Associate Professor. His research interests include hyperspectral imagery, machine learning, and pattern recognition.

Fan Fan received the B.S. degree in communication engineering and the Ph.D. degree in electronic circuit and system from the Huazhong University of Science and Technology, Wuhan, China, in 2009 and 2015, respectively.

He was a Post-Doctoral Researcher with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan. He is currently an Associate Professor with the Electronic Information School, Wuhan University. His research interests include infrared thermal imaging, machine learning, and computer vision.

Jun Huang received the B.S. and Ph.D. degrees from the Department of Electronic and Information Engineering, Huazhong University of Science and Technology, Wuhan, China, in 2008 and 2014, respectively.

He is currently working as an Associate Professor with the Electronic Information School, Wuhan University, Wuhan. His main research interests include infrared image processing and infrared spectrum processing. pattern recognition.

Jiayi Ma (Senior Member, IEEE) received the B.S. degree in information and computing science and the Ph.D. degree in control science and engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2008 and 2014, respectively.

From 2012 to 2013, he was an Exchange Student with the Department of Statistics, University of California at Los Angeles, Los Angeles, CA, USA. He held a post-doctoral position at the Electronic Information School, Wuhan University, Wuhan, from August 2014 to November 2015, and received an accelerated promotion to an Associate Professor and a Full Professor in December 2015 and December 2018, respectively. He has authored or coauthored more than 120 refereed journal articles and conference papers, including IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE TRANSACTIONS ON CYBERNETICS, IEEE TRANSACTIONS ON MULTIMEDIA, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, International Journal of Computer Vision (IJCV), Computer Vision and Pattern Recognition Conference (CVPR), The International Joint Conferences on Artificial Intelligence (IJCAI), The Association for the Advancement of Artificial Intelligence (AAAI), The International Conference on Robotics and Automation (ICRA), International Conference on Intelligent Robots and Systems (IROS), and The Association for Computing Machinery International Conference on Multimedia (ACM MM). His research interests include computer vision, machine learning, and pattern recognition.

Dr. Ma is an Editorial Board Member of *Information Fusion* and *Neurocomputing* and a Guest Editor of *Remote Sensing*. He has been identified in the 2019 Highly Cited Researchers list from the Web of Science Group. He was a recipient of the Natural Science Award of Hubei Province (first class), the Chinese Association for Artificial Intelligence (CAAI) Excellent Doctoral Dissertation Award (a total of eight winners in China), and the Chinese Association of Automation (CAA) Excellent Doctoral Dissertation Award (a total of ten winners in China).