This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Toward Universal Stripe Removal via Wavelet-Based Deep Convolutional Neural Network

Yi Chang, *Student Member, IEEE*, Meiya Chen, Luxin Yan, *Member, IEEE*, Xi-Le Zhao, Yi Li, and Sheng Zhong

***Abstract*—Stripe noise from different remote sensing imaging systems varies considerably in terms of response, length, angle, and periodicity. Due to the complex distributions of different stripes, the destriping results of previous methods may be oversmoothed or contain residual stripe. To overcome this key problem, we provide a comprehensive analysis of existing destriping methods and propose a deep convolutional neural network (CNN) for handling various kinds of stripes. Moreover, previous methods individually model the stripe or the image priors, which may lose the relationship between them. In this article, a two-stream CNN is designed to simultaneously model the stripe and image, which better facilitates distinguishing them from each other. Moreover, we incorporate the wavelet into our CNN model for better directional feature representation. Therefore, the CNN learns the discriminative representation from the external data set, while the wavelet models the internal directionality of the stripe, in which both the internal and external priors are beneficial to the destriping task. In addition, the wavelet extracts the multiscale information with a larger receptive field for global contextual information modeling; thus, we can better distinguish the stripe from the similar image line pattern structures. The proposed method has been extensively evaluated on a number of data sets and outperforms the state-of-the-art methods by substantially a large margin in terms of quantitative and qualitative assessments, speed, and robustness.**

***Index Terms*—Convolutional neural network (CNN), destriping, image decomposition, wavelet.**

I. Introduction

REMOTE-SENSING image stripe noise is mainly caused by differences in the response of adjacent detectors. Numerous research studies have been proposed to boost the development of stripe removal in the past decades. In this section, we will first provide a comprehensive and systematic review of the previous destriping methods. Then, we will analyze the remaining challenges in this field. Lastly, we will provide our solution to solve these challenging issues.

A. Related Work

In Table I, we list most of the image destriping methods and their main features. We mainly consider the year of publication, input, imaging system, utilization of the direction, utilization of the image and stripe, and the speed. Next, we will provide a brief description of each kind of destriping method.

1) Statistical Matching: The statistical matching methods usually refer to histogram matching and moment matching [1]–[5] and were the dominant approaches before 2000, which mainly include two steps: the clean reference finding and statistical matching. Thus, the success of statistical matching relies heavily on finding a clean reference. In 1979, Horn and Woodham [1] proposed the first histogram matching method for Landsat images destriping. To find a suitable reference line, Wegener [2] implicitly considered the local smoothness of the image and proposed to calculate the statistics only over homogeneous regions. This approach is generally effective for specific imaging systems in which only a portion of them have fixed stripes, such as the Moderate Resolution Imaging Spectroradiometer (MODIS). However, it is difficult to find a reference for hyperspectral images with ubiquitous stripes.

2) Digital Filtering: The filtering-based methods [6]–[19], processing stripes in the transformed frequency domain instead of the original image domain, were active between 2000 and 2010. They assume that the specific frequencies caused by stripes are sparse and can be easily distinguished from the image structures in the transformed domain. The stripe is regarded as a special kind of "noise," and the conventional filters were introduced to suppress the stripes [7]. Later, the directional property of the stripe was taken into consideration via the wavelet [8], [13], [16]. We also regard the interpolation methods [14], [15] as a median filter. Hybrid approaches combining statistical matching with the filtering method have been presented [10], [12], [18]. In our opinion, it is not suitable to simply regard the stripe as "noise" that is not identically independent and contains significantly directional structure.

3) Variational Model: In the last ten years, the variational methods are the most popular destriping methods [20]–[33].

Manuscript received October 7, 2019; accepted November 28, 2019. This work was supported in part by the projects of the National Natural Science Foundation of China under Grant 61571207, Grant 61433007, and Grant 61876203, and in part by the Project of the Hubei Provincial Natural Science Foundation of China under Grant 2018CFA089. *(Corresponding author: Luxin Yan.)*

Y. Chang, M. Chen, L. Yan, Y. Li, and S. Zhong are with the National Key Laboratory of Science and Technology on Multispectral Information Processing, School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: yichang@hust.edu.cn; my_chen@hust.edu.cn; yanluxin@hust.edu.cn; li_yi@hust.edu.cn; zhongsheng@hust.edu.cn).

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

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2019.2957153

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE I

COMPARISON OF EXISTING DESTRIPING METHODS AND THEIR PROPERTIES

Method	Year	Input	Imaging	Brief Description	Direction	Estimation
Horn [1]	1979	Single Image	Landsat MSS	Histogram Matching	No	Image
Wegener [2]	1990	Single Image	Landsat MSS	Histogram Matching	No	Image
Corsini [3]	2000	Single Image	MOS-B	Moment Equalization	No	Image
Gadallah [4]	2000	Single Image	Landsat TM	Moment Matching	No	Image
Meza [5]	2016	Multispectral	Hyperspectral	Equalization	No	Image
Simpson [6]	1998	Single Image	GOES-7	Finite Impulse Filters	No	Image
Chen [7]	2003	Single Image	CMODIS	Filtering	No	Image
Chen [8]	2006	Single Image	CMODIS	Wavelet+ FFT	Yes	Image
Liu [9]	2006	Single Image	Landsat-7	FFT + Adaptive Filtering	Yes	Image
Rakwatin [10]	2007	Single Image	MODIS	Moment Matching + Facet Filtering	No	Image
Gómez [11]	2008	Multispectral	CHRIS	Filtering	Yes	Image
Rakwatin [12]	2009	Multispectral	MODIS Band6	Moment Matching + Facet Filtering	No	Image
Münch [13]	2009	Single Image	Medical	Wavelet + FFT	Yes	Image
Wang [14]	2008	Multispectral	MODIS Band6	Interpolation	No	Image
Jung [15]	2010	Multispectral	SPOT4	Detection + Interpolation	No	Image
Chhetri [16]	2011	Single Image	Hyperspectral	Wavelet + FFT	Yes	Image
Gladkova [17]	2011	Multispectral	MODIS Band6	Multivariate Regression	No	Image
Duan [18]	2014	Single Image	Hyperspectral	Reference Calibration + Filtering	No	Image
Cao [19]	2016	Single Image	Infrared	1D Guided Filtering	Yes	Image
Shen [20]	2009	Single Image	MODIS	MAP Framework with Huber-Markov Prior	No	Image
Bisceglie [21]	2009	Single Image	MODIS	Least Squares Minimization	No	Image
Carfantan [22]	2010	Single Image	SPOT3	MAP Framework with Markov field Prior	No	Stripe
Bouali [23]	2011	Single Image	MODIS	Unidirectional Variational	Yes	Image
Fehrenbach [24]	2012	Single Image	Medical	MAP Framework with Huber-Markov Prior	Yes	Stripe
Yuan [25]	2012	Multispectral	Hyperspectral	Spectral-spatial Adaptive Total Variation	No	Image
Chang [26]	2013	Single Image	General	Framelet + Unidirectional Variational	Yes	Image
Chang [27]	2014	Single Image	General	Sparse + Unidirectional Variation	Yes	Image
Chang [28]	2015	Multispectral	General	Anisotropic Spectral-Spatial Total Variation	Yes	Image
Liu [29]	2015	Single Image	General	Sparse + Unidirectional Variational	Yes	Stripe
Aggarwal [30]	2016	Multispectral	Hyperspectral	Spectral-spatial Total Variation	No	Image
Fitschen [31]	2017	Single Image	Medical	Framelet + Unidirectional Variational	Yes	Image
Liu [32]	2018	Single Image	General	Oriented Variational	Yes	Image
Liu [33]	2018	Single Image	General	Feature based Unidirectional Variational	Yes	Image
Acito [35]	2011	Multispectral	Hyperspectral	Orthogonal Subspace Learning	No	Stripe
Lu [36]	2013	Multispectral	Hyperspectral	Graph-regularized Low-rank Representation	No	Both
Zhang [37]	2014	Multispectral	Hyperspectral	Low-rank Matrix Recovery	No	Both
Zhao [38]	2015	Multispectral	Hyperspectral	Sparse + Low-rank Matrix Recovery	No	Image
Wang [39]	2016	Multispectral	Hyperspectral	Group Low-rank Representation	Yes	Both
He [40]	2016	Multispectral	Hyperspectral	Total variation based Low-rank Representation	No	Both
Chang [41]	2016	Multispectral	General	Low-rank Image Decomposition	Yes	Both
Chen [42]	2017	Single Image	General	Group Sparsity + Unidirectional Variational	Yes	Stripe
Chang [43]	2017	Single Image	General	Transformed Low-rank Matrix Recovery	Yes	Image
Chen [44]	2018	Multispectral	Hyperspectral	Low-rank Matrix Factorization	No	Image
Cao [45]	2018	Multispectral	General	Nonlocal TV + Low-rank Matrix Factorization	Yes	Both
Liu [61]	2019	Multispectral	Hyperspectral	Wavelet + Low-rank Matrix Recovery	Yes	Both
Xie [46]	2016	Multispectral	Hyperspectral	Intrinsic Tensor Sparsity	No	Image
Chang [47]	2017	Multispectral	Hyperspectral	Unidirectional Low-rank Tensor Recovery	No	Image
Fan [48]	2017	Multispectral	Hyperspectral	Low-rank Tensor Recovery	No	Both
Chen [49]	2018	Multispectral	Hyperspectral	ASSTV + Low-rank Tensor Decomposition	Yes	Both
Cao [50]	2018	Multispectral	Hyperspectral	SSTV + Low-rank Tensor Recovery	No	Both
Fan [51]	2018	Multispectral	Hyperspectral	SSTV + Low-rank Tensor Recovery	No	Both
Wang [52]	2018	Multispectral	Hyperspectral	SSTV + Low-rank Tensor Decomposition	No	Both
Kuang [53]	2017	Single Image	Infrared	CNN	No	Image
He [54]	2018	Single Image	Infrared	CNN	No	Image
Xiao [55]	2018	Single Image	Infrared	Local-global CNN	No	Image
Xie [56]	2018	Multispectral	Hyperspectral	Residual Deep CNN	No	Stripe
Zhang [57]	2018	Multispectral	Hyperspectral	Spatial-spectral Gradient CNN	No	Stripe
Chang [58]	2019	Multispectral	Hyperspectral	Residual Deep CNN	No	Stripe
Chang [60]	2019	Single Image	Infrared	Multiscale Residual Deep CNN	No	Stripe

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE II

EFFECTIVENESS COMPARISON OF REPRESENTATIVE DESTRIPING METHODS FOR DIFFERENT KINDS OF STRIPES

	Response		Intensity		Angle		Proportion		Length		Periodicity		Width		Mixed Noise	
Method	Offset	Gain	Normal	Deadline	Vertical	Oblique	Some	Entire	Broken	Full	Yes	No	Single	Broad	Light	Heavy
Gadallah [4]	√	√	√		√		√		√	√	√		√			
Münch [13]	√		√		√		√	√	√	√	√	√	√	√		
Carfantan [22]	√	√	√		√		√	√		√	√	√	√	√		
Chang [28]	√		√		√		√	√	√	√	√	√	√	√	√	√
Liu [32]	√		√		√	√	√	√	√	√	√	√	√	√		
Zhang [37]	√		√	√	√		√		√	√	√	√	√	√	√	√
Chang [41]	√	√	√		√		√	√		√	√	√	√	√	√	
Wang [52]	√		√		√		√	√	√	√	√	√	√	√	√	√
Zhang [57]	√		√		√		√	√	√	√	√	√	√	√	√	√
Proposed	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√

To explicitly utilize the sparsity in the image, the variational-based destriping methods benefit from the L_p optimization solver [34]. They treat the destriping issue as an ill-posed inverse problem and then optimize a variational model by incorporating sparsity priors about the image. In 2009, Shen and Zhang [20] first proposed the Huber–Markov-based variational model for remote sensing image destriping within a maximum-*a posteriori* framework. These variational methods take the stripe as the isotropic "noise," which fails to capture the anisotropic property of the stripe. To model the directional characteristic of the stripe, the sophisticated unidirectional variation models have been extensively studied [23], [26]–[33] and have achieved impressive performance.

The aforementioned variational methods focus on modeling the image prior. An alternate kind of method steps toward the opposite direction by modeling the stripe prior with state-of-the-art destriping performance [22], [24], [29]. The start point of the stripe estimation based methods is that, compared with the image component, the stripe has simpler structure and less unknown variable to be estimated. After they estimate the stripe component, they could obtain the clean image via the degradation model indirectly.

4) Low-Rank Matrix Recovery Model: Most of the previous single-image-based methods may lose the spectral coherence by processing each band individually. To remedy this issue, the low-rank-based matrix recovery methods [35]–[45] have been naturally proposed in recent years. They take advantage of the low-rank property along the spectral mode by lexicographically ordering the 3-D cubic into a 2-D matrix [37]. To cooperate with the global low-rank constraint, the local/nonlocal sparsity regularizer has been additionally incorporated into the low-rank model to further refine the restoration results [36], [38], [40], [42], [45]. In contrast, Chang *et al.* [41], [43] exploited a stronger low-rank property in the stripe component within an image decomposition framework. We hold the viewpoint that modeling both the image and the stripe component are useful for their decoupling.

5) Low-Rank Tensor Recovery Model: Although the vector-/matrix-based methods have achieved excellent destriping results, they inevitably cause damages to the spectral–spatial structural correlation for the 3-D inputs. To alleviate this issue, the low-rank tensor recovery methods have emerged in the last two years [46]–[52]. Some of them simply added up the ranks (or its relaxations) along each tensor mode [46], [48], [51]. To truthfully reflect the intrinsic difference of the structure correlation along each mode, we proposed a unidirectional low-rank tensor recovery model for multispectral image denoising [47]. Compared with the previous methods, the low-rank-based methods could better preserve the structure correlation and are effective for both the stripe and random noises. However, the speed of the low-rank tensor methods are extremely slow due to the large data size and complicated operations, which make them unsuitable for real-time applications.

6) Deep Convolutional Neural Network Model: Previous methods design various handcrafted priors for the image/stripe components and have achieved very great progress in destriping field. However, these handcrafted priors may not be sufficiently discriminative to distinguish the stripe from the image structures which share similar direction as the stripe. In last two years, discriminative-based deep convolutional neural network (CNN) models have been naturally proposed [53]–[60]. The CNN-based methods learn the feature of the stripe in a supervised manner and benefit us to differentiate the stripe from the image structure from a global receptive field. Several nonuniform stripe noise removal methods have been proposed for the single infrared image with a shallow CNN model [53]–[55]. Later, the residual learning strategy along with the deep CNN was introduced and achieved better performance [56]–[60]. However, these discriminative methods only consider learning from the external data set while ignoring the internally directional property of the stripe. Moreover, they neglect to explicitly model the correlation between the image and stripe.

B. Remaining Challenges

The destriping issue has been extensively studied over 40 years with very impressive results. However, there are still several challenges to be solved. In this section, we will present two key challenges with a brief analysis.

1) Robustness: We have listed the robustness of some representative destriping methods in Table II. The stripe category will be presented in Section II. The previous methods are designed for specific stripes with strong assumptions. For example, the filtering-based methods utilized the periodicity

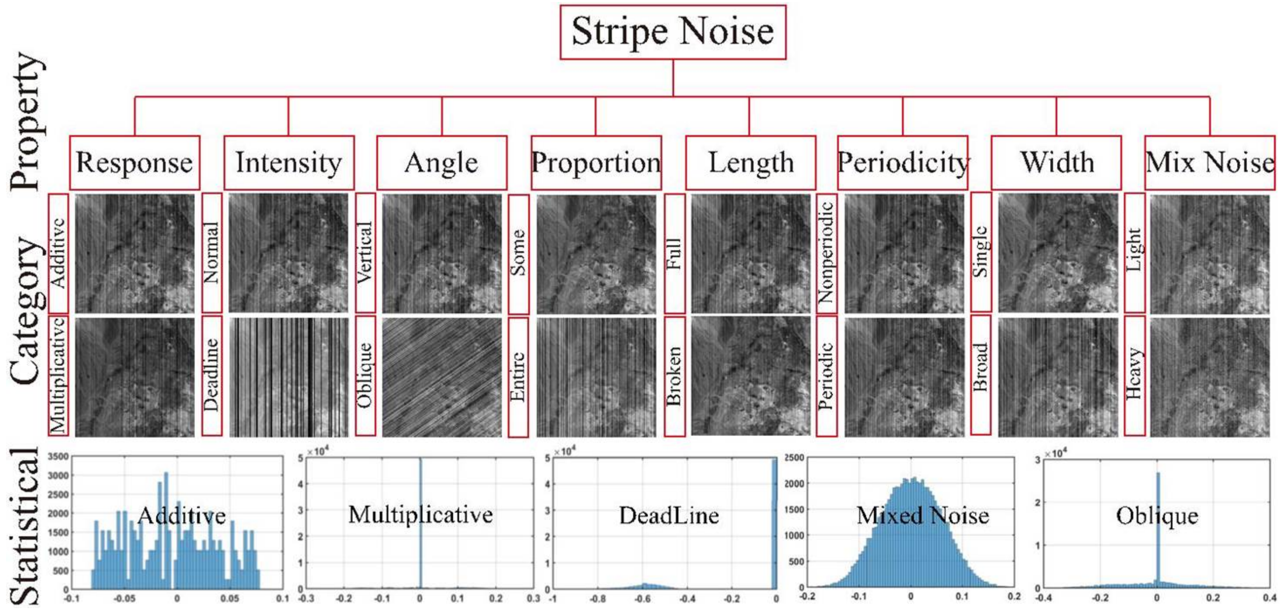

Fig. 1. Category of the stripe noise. We classify the stripe noise in both the push-broom (representative hyperspectral imaging) and whisk-broom (representative MODIS) from their appearances into sixteen classes. Their visual images are shown in the second row. Some statistical histograms of the representative stripes are shown in the last row, which is very different for each stripe.

of the stripes [6], [10]. Most of the existing destriping method are designed for the stripes, without considering the influence of the random noise [13], [22], [23], [33], [41]. In addition, the low-rank-based methods exploited the spectral correlation, where this property may be lost in a single image. Some methods assumed that the stripe is intact with full length [22]. Therefore, we can conclude that the previous methods are less effective for all kinds of stripes. The main reason is that the distributions of the various stripes are obviously different from each other, as shown in the third row in Fig. 1. Precisely, modeling these distributions of the different stripes via an exact mathematical formulation is difficult. Thus, the crucial factor for robust destriping is determining how to model various stripes with different distributions.

2) Discriminative: Even for the conventional and common vertical additive stripe, when the intensity of stripe varies a lot, especially too large, it is very hard for the previous methods to differentiate the stripe from the image structures which share the similar direction as the stripe. Consequently, the previous destriping methods may over-smooth the image structures or contain residual stripes, as shown in Fig. 2. For example, the filtering-based wavelet Fourier adaptive filter (WFAF) [13] mainly makes use of an internal handcrafted feature. There are obvious residual stripe effects in Fig. 2(left). In contrast, the deep learning-based stripe nonuniform correction (DLS-NUC) method [54] only takes advantage of the external data set. We can observe that the image structure that has the same direction as the stripe has been unexpectedly removed along with the stripe, which may be less discriminative for the stripe and sharp image edges. A similar phenomenon has also been observed in the local gradient-based unidirectional variational model [23]. In addition, most of the previous methods only extract the features of the image or the stripe component. It is naturally understood that the joint representation of

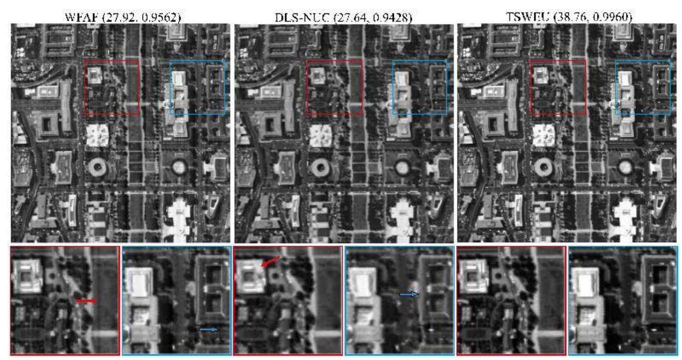

Fig. 2. Limitation of conventional methods. Residual stripe (WFAF) and over-smooth (DLS) phenomenon can be observed in previous methods.

both components is much more discriminative for separating them [41]. Therefore, discovering how to jointly utilize the internal and external features for both the image and stripe components is also a key factor for better destriping.

C. Our Solution and Contribution

To contend with the first challenge, we propose implementing the CNN for representing various stripes. As observed in Fig. 1, the distributions of various stripes are obviously different. Moreover, due to the structural correlation of the stripe "noise," the distributions are always non-Gaussian, nonidentical, and nonindependent. Therefore, it is difficult for the previous Gaussian or mixture of Gaussian (MoG)-based methods to fit these stripes precisely. We demonstrate that the U-Net [62] can fit various distributions better than the previous methods (see Section III-A1). The CNN learns the feature of the stripe from a very large receptive field, which utilizes the global contextual information and benefits us to differentiate

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

the stripe from the image structure. Moreover, to facilitate the training, we enhance the U-Net with the residual block [63] for better feature propagation and reuse (see Section III-A2).

To address the second challenge, we argue that both the learned discriminative features from the external data set and the handcrafted discriminative features extracted from the internal image are beneficial for stripe representation. We analyze the relationship between the handcrafted-based wavelet and learning-based CNN both experimentally and theoretically (see Section III-B1). The wavelet is an effective tool for modeling the intrinsical directional characteristic of the stripe [13], the multiscale representation of the image [27], and the lossless decomposition and reconstruction [64]. Thus, we propose embedding the wavelet into the end-to-end CNN network to achieve better performance (see Section III-B2).

In our previous work [41], we proposed to treat the destriping task from an image decomposition perspective, in which these two components are treated equally and decoupled iteratively. Utilizing the discriminative features from both components is beneficial for separating them. In this article, we follow the decomposition idea and implement this framework via a multitask learning-based two-stream CNN (see Section III-C1). The image stream aims at reconstructing the clear image. The stripe stream focuses on extracting the features of the stripe. The extracted features, including the intensity, location, and angle, to name a few, function as an attention map and are fed into the image stream to guide the final reconstruction (see Section III-C2).

In summary, we provide a comprehensive classification of the remote sensing stripes (see Section II) and a brief property survey of the existing destriping methods that can serve as an elementary work for beginners in this field. Moreover, we point out two major challenges in this field and propose a preliminary solution for both of them. Our contributions can be summarized as follows.

1) We formulate the destriping issue as a discriminative multitask learning problem. The two-stream CNN jointly extracts the image and stripe features interacting with each other, which makes our method more representative for various kinds of stripes with different distributions.
2) To increase the discriminative ability, apart from the external prior, we additionally utilize the internal prior via the wavelet for extracting the intrinsically directional feature in the stripe and the multiscale feature in the image.
3) We have extensively evaluated our method on various remote sensing images with state-of-the-art performance in both quantitative and qualitative assessments. Our method is effective for an arbitrary image with stripe noise.

In Section II, the category of stripe noise is analyzed. The detailed architecture is described in Section III. The experimental results and discussion are reported in Section IV. Finally, we conclude this article in Section V.

II. Category of Stripes

There are mainly two different remote sensing image systems: push-broom and whisk-broom imaging. Moreover, the striping effect has different appearances depending on the scanning mechanism of imaging instruments. The interested reader can refer to [33] for details. In this article, we provide a more comprehensive classification of the stripes in remote sensing images, as shown in Fig. 1.

According to the detector response, the stripe can be classified into additive and multiplicative types. The additive stripe is signal-independent, whereas the multiplicative stripe is signal-dependent. The intensities of the additive stripe along the stripe are ranging in a small interval, and close to a constant value. The intensity of the multiplicative stripe is highly associated with the image. That is, the stripe is much darker in the low-intensity region, whereas the stripe is brighter in the high-intensity region, which makes the multiplicative stripe more difficult to remove. Most of the previous methods focus on the additive stripe, and only the authors in [20], [22], [41] have tried to handle the multiplicative stripe. It is worth noting that the additive model can be well applied to the multiplicative case by applying the logarithm, as in [22]; however, it may fail in the presence of both the additive and multiplicative stripes.

According to the intensity, the stripe can be classified as a normal or deadline stripe. The normal stripe has a common intensity, whereas the intensity of the deadline stripe is all zeros. The deadline stripes do not convey any useful information and are caused by the failure of certain detectors. It is very difficult for conventional denoising-based methods to reconstruct the original image. Some methods resort to the spectral correlation of the multispectral images [37]. In our opinion, the removal of deadline stripes has been better treated as an image inpainting task, as in [20].

According to the angle, the stripe can be divided into vertical/horizontal and oblique ones. The stripe should be horizontal or vertical due to the imaging principle. However, for the subsequent remote sensing product, the geometric registration would cause the oblique stripe. Most of the previous methods can only handle the vertical/horizontal stripe, especially the directional-based models [23], [41]. A natural idea to process the oblique stripe is to transform it into the original domain [43]. However, this may cause an information loss due to the interpolation operator in the transformation. The recent variational model [32] can only handle the fixed angle oblique stripe, which limits its application in real settings.

According to the proportion, the stripe can be divided into partial and entire proportion ones. Generally, the partial proportion stripe appears in the whisk-broom imaging system, and the entire proportion stripe occurs in the push-broom imaging system. The entire proportion stripe cannot be handled by the previous statistical matching methods [1]–[5] since no clean reference line can be found. Most of the presented destriping methods can satisfactorily remove the partial proportion stripe due to its simplicity.

According to the length, the stripe can be classified as a full or broken stripe. For the full-length stripe, it can be post-processed via the feature of its length. The broken stripe (known as random stripe in MODIS) means that each stripe may possess an arbitrary length. This would make it difficult to distinguish the stripe from the line pattern of the image

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

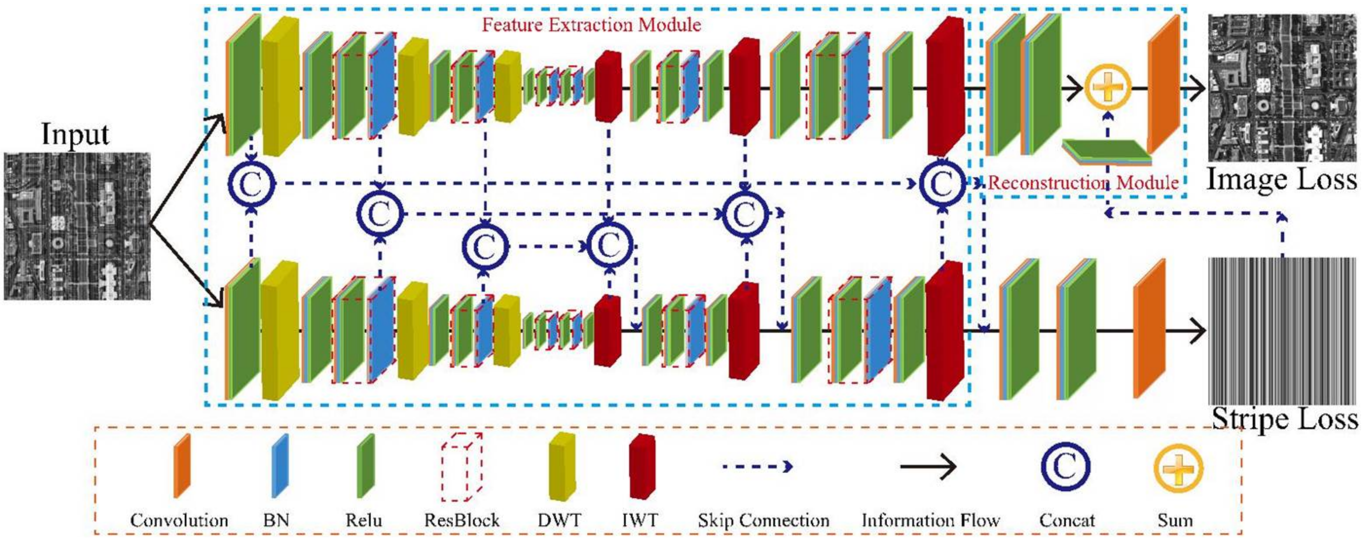

Fig. 3. Framework of the proposed network. Our two-stream wavelet enhanced U-net (TSWEU) is built on two parallel streams for stripe and image feature extraction, and a reconstruction module to restore the clean image with the guidance of the stripe. The skip connections are introduced for better information interaction among the two streams. Moreover, the wavelet is embedded into the streams for better image and stripe representation with lossless downsampling/upsampling.

structure. Moreover, the line pattern structure would be unexpectedly removed by the unsupervised methods along with the stripe. This validates that we need to extract discriminative features or utilize the contextual information to assist removing the stripe and preserving the image structure.

According to the periodicity, the stripe can be classified as a periodical or nonperiodical stripe. The periodical stripe appears only in the whisk-broom imaging system due to its imaging mechanism. The periodical stripe can be identified by the specific spectrum in the frequency domain, which has been well handled by the filtering methods [6]–[19]. The nonperiodical stripe mainly occurs only in the push-broom imaging system. Compared with the periodical stripe, the nonperiodical would inevitably damage the low-rank property or sparsity in the image. Generally, the nonperiodical stripe is much more difficult to remove [41].

According to the width, stripes can be divided into single and broad stripes. The single-width stripe can be well removed by the previous methods due to its simplicity. When a broad stripe exists, the performance of single-image-based methods would degenerate rapidly, especially smoothness-based methods [10], [19], [20]. It is worth noting that the width and the proportion are very close but are not the same. Here, a broad stripe means that several adjacent stripes have similar intensities, making the stripe extremely difficult to remove.

Normally, a stripe coexists with random noise in remote sensing images. The mixed random noise and structural stripe make the distribution of the noise complicated; therefore, simply modeling with the Gaussian or the MoG is difficult. According to the noise level, the stripe can be divided into light and heavy mixed noise. Most of the previous methods handle the mixed noise by utilizing the spectral correlation of the multispectral image. In our previous work, we proposed two elaborate models for single-image light mixed noise removal [26], [43]. For the heavy mixed noise, the useful information in the image would be overwhelmed by the noise, whereas

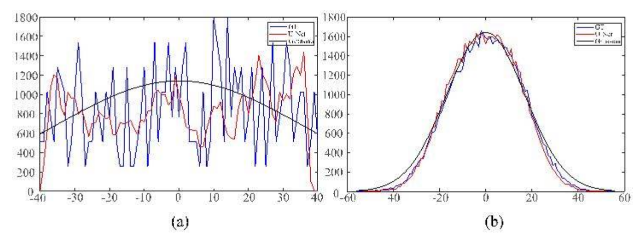

Fig. 4. Advantage of the CNN over Gaussian model for modeling the stripe. Distributions of (a) additive stripe and (b) mixed noise.

previous single-image-based methods fail to handle this case. In this article, we resort to the external clean data set for single-image heavy mixed noise removal.

III. Two-Stream-Based Wavelet Enhanced U-Net Model

As illustrated in Fig. 3, our proposed two-stream-based deep CNN is composed of two complementary components: one stream for stripe estimation and the other stream for image reconstruction. The stripe estimation stream is trained to infer the various distributions of the stripe "noise." Meanwhile, the internal directional property is extracted via the embedded wavelet. The image reconstruction stream is trained to recover the clean image with the lossless-based multiscale wavelet. Moreover, the two-stream intermediate features are further merged as an attention map for improving the discrimination.

A. External Prior: EU Model

1) Advantage Over Gaussian Model: Most of the previous methods treat the stripe as "noise" and apply the conventional Gaussian model or MoG for modeling the noise [20], [25], [44]. However, from the physical degradation and its visual appearance, we know the distribution of the stripe is obviously nonindependent. Moreover, different stripes possess distinctly different distributions in Fig. 1. It is very difficult to construct

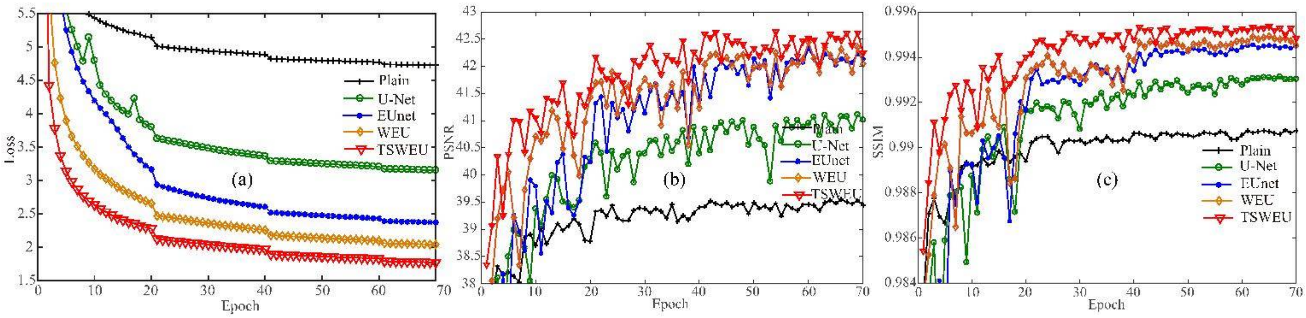

Fig. 5. Effectiveness of the U-Net, resblock, wavelet, and two-stream framework. We plot the curve of (a) training loss, (b) testing PSNR, and (c) testing SSIM. We start the CNN from the plain network, and gradually increase each term. The black curve denotes the plain CNN without upsampling and downsampling. The blue curve represents the U-Net with larger receptive field. The green curve is the EU model with short connection based residual block. The orange curve means the wavelet embedded WEU model. The red curve stands for the proposed image decomposition based TSWEU.

a precise mathematical formulation to fit the distributions of different stripes. In this article, we bypass the difficulty of constructing the handcrafted distribution function. In contrast, we start from the data-driven perspective and resort to the universal approximation ability of the CNN for an arbitrary signal [65]. We find that the CNN has a vast advantage in structural noise modeling.

To illustrate this, we plot the distributions of two kinds of stripe noise in Fig. 4, marked by the blue curve. Then, we show the distribution of estimated noises by both our CNN model (red curve) and the Gaussian model (black curve). In Fig. 4(a), the distribution of the additive stripe is very complex. The learned distribution of the CNN is much closer to the original one. In Fig. 4(b), the distribution of the mixed noise exhibits a Gaussian-like shape. Both the CNN model and Gaussian model are approximated to the ground truth. However, the CNN model can well fit the high-frequency parts, such as the range from [20, 60]. For the low-frequency parts, the CNN model can also capture the small variance. Overall, the CNN model consistently fits better than the conventional Gaussian model for different kinds of stripe noise.

2) EU: In this article, we employ the U-Net as our baseline, which has been widely used in image segmentation [62], image deblurring [66], and so on. The success of the U-Net relies heavily on the long-term skip connection between different layers for better feature reuse and information propagation. However, DenseNet [67] has demonstrated that the dense connections between both the short- and long-distance layers would be beneficial for the feature representation. Motivated by this, we additionally introduce the short-term connection-based residual blocks [63] into U-Net, as shown in the red dash blocks in Fig. 3.

To illustrate the effectiveness of the EU, we compare the EU (green curve) with the original U-Net (blue curve) and plain network (black curve), as shown in Fig. 5. We plot their training loss and destriping peak signal-to-noise ratio/structure similarity (PSNR/SSIM) curves along with the epoch. By comparing the blue curve with the black curve, we can conclude that the downsampling and upsampling operators that induced a larger receptive field is a key factor in the destriping task. This is very reasonable since the stripes always run throughout the whole image. In Fig. 5(a), the training loss of the EU is obviously lower than that of the original U-Net, which strongly supports the effectiveness of the residual blocks for better information propagation. In Fig. 5(b) and (c), the PSNR and SSIM values of the EU are consistently higher than those of the original U-Net with iterations.

B. Internal Prior: WEU Model

In this section, we first introduce the relationship between the wavelet and the CNN. Then, the advantage of the wavelet embedded in the EU is analyzed.

1) Relationship Between Wavelet and CNN: The discrete wavelet transform (DWT) is governed by the choice of filters/wavelet transform for which the wavelets are discretely sampled, such as the Haar wavelet. For the image processing task, its solution can be roughly expressed as follows:

$$\boldsymbol{X}^{(d)} = \psi(\boldsymbol{X}^{(d-1)})) \tag{1}$$

where $\boldsymbol{X}^{(d)}$ is the signal of the decomposition level d, ψ is the filtering transform operator, such as the Haar wavelet, and is the soft or hard threshold operator [68]. Similarly, the output of the dth layer of a convolutional layer can be expressed as follows:

$$\boldsymbol{X}^{(d)} = S(\boldsymbol{W}^{(d)} \otimes \boldsymbol{X}^{(d-1)} + \boldsymbol{P}^{(d)}) \in R^{R_d \times C_d \times B_d} \tag{2}$$

where $\boldsymbol{X}^{(d)}$ is the output of the dth layer, $\boldsymbol{W}^{(d)}$ is the projection matrix to be learned, $\boldsymbol{P}^{(d)}$ is the bias vector, $\otimes$ is the convolutional operator, R_d, C_d, and B_d are the spatial row, column, and channel number of the dth layer, respectively, and $S : \mathbb{R} \mapsto \mathbb{R}$ is the nonlinear activation function that handles each pixel individually, such as the sigmoid.

From the mathematical formulations of (1) and (2), we can find that they are very similar to each other. Moreover, their physical meanings are the same: transform the input dth level/layer image into the feature domain via ψ or $\boldsymbol{W}^{(d)}$, activate the sparse features via the nonlinear activation function or S, and then repeat/recurse this procedure in a hierarchical manner. The main difference between wavelets and CNNs is the transformation function ψ and $\boldsymbol{W}^{(d)}$. The ψ is a fixed template for the wavelet, whereas $\boldsymbol{W}^{(d)}$ is a learnable filter for the CNN. This intrinsic similarity between them offers a theoretical basis for embedding the wavelet into the CNN.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

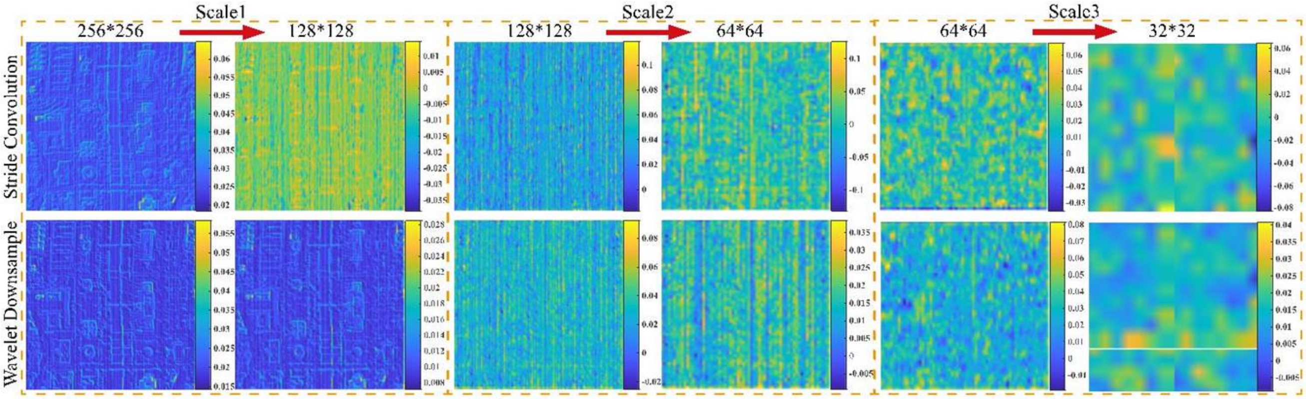

Fig. 6. Feature maps comparison between the stride convolution and wavelet. The first and second rows show the feature maps (maximum response along the channels) of the stride convolution in U-Net and wavelet, respectively. We show three different scales feature maps before (1, 3, and 5 columns) and after (2, 4, and 6 columns) the downsampling. Compared with the stride convolution, the wavelet could better preserve the structure and the line pattern of the stripe.

2) WEU: The learning-based methods have dominated computer vision since they can automatically extract abundant features from a large external data set. However, we argue that the handcrafted features from the internal prior can also be very useful when the intrinsic property can be further utilized to enhance the representative feature. In this article, we propose to embed the wavelet into the EU. The EU relies on the external data set to extract the feature of the line pattern stripe. Meanwhile, the direction-aware wavelet focuses on extracting the important feature: the directionality of the stripe. The embedded wavelet can be regarded as a feature attention reinforcement block that functions as a regularizer to extract the horizontal/vertical line pattern features of the stripe. Thus, the joint external and internal modeling makes the features more discriminative.

To differentiate the stripe from the image structure, the proposed network should capture more contextual information and possess as large a receptive field as possible. Apart from the depth and filter sizes of the U-Net, the downsampling and upsampling layers are the main means to enlarge the receptive field. However, the downsampling via the convolution with stride would inevitably introduce information loss, which is harmful to the pixel-to-pixel-level image reconstruction task. Fortunately, since DWT is invertible, it is guaranteed that all the information can be kept by such a downsampling scheme.

To illustrate the effectiveness of the wavelet, we compare the WEU with the EU on three aspects: the feature maps, the training procedure, and the testing results. Compared with the stride convolution, the wavelet can well extract the directional feature of the stripes. On the other hand, the wavelet was able to losslessly decompose and reconstruct the features, especially for the shallow features, as shown in Fig. 6. In addition, after we replaced the stride convolution with the wavelet, the training loss dropped rapidly at the early training stage and was obviously lower all the time, as shown in Fig. 5(a). Moreover, the PSNR and SSIM values of the wavelet are slightly better than that of the stride convolution.

C. Two-Stream WEU Model

1) Motivation From Decomposition: Most of the previous methods formulate the destriping as a denoising problem, in which they individually model the image or stripe prior. However, they neglect the relationship between the image and stripe as follows:

$$\boldsymbol{Y} = \boldsymbol{X} + \boldsymbol{B} + \boldsymbol{N} \tag{3}$$

where $\boldsymbol{Y} \in \mathbb{R}^{R\times C}$ is the measured image, $\boldsymbol{X}$ is the desired clear image, $\boldsymbol{B}$ is the stripe component, and $\boldsymbol{N}$ is the random noise. For the image decomposition problem, the general reconstruction model can be formulated as follows:

$$\min_{\boldsymbol{X},\boldsymbol{B}} \frac{1}{2}||\boldsymbol{X} + \boldsymbol{B} - \boldsymbol{Y}||_F^2 + \tau P(\boldsymbol{X}) + \lambda P(\boldsymbol{B}) \tag{4}$$

where the first term is the reconstruction term, and the second and third terms $P(\boldsymbol{X})$ and $P(\boldsymbol{B})$ are the priors about the image and stripe components, respectively. The proposed decomposition model aims to optimize two variables simultaneously, which can be solved via an alternative minimizing strategy. Compared with the previous "denoising" methods, the decomposition methods additionally utilize the property of the stripe and image to strengthen the representation and further build the connection between the image and stripe components, which significantly facilitates separating the two components.

2) TSWEU Module: In this article, we are motivated by our previous image decomposition-based destriping work [41], which has shown that the joint modeling of both the image and stripe is better than modeling only one of them. Our starting point is to extend the decomposition-based optimization method to the TSWEU model. The CNN model is more representative and robust than implementing handcrafted features. For example, the low rank obviously no longer holds for the oblique and mixed noise stripe, since the low rank cannot capture the angle feature automatically. Similarly, the total variational (TV) approach only extracts the horizontal and vertical first-order gradient feature of the image, whereas the WEU model can extract the multiscale feature in a hierarchical manner.

More precisely, we replace the handcrafted low-rank and total variation prior with the dual WEU, as shown in Fig. 3. For the optimization of (4), it is usually converted into three subproblems: one for optimizing the stripe, one for optimizing the image, and one for reconstruction. Analogous to iterative minimization, each WEU stream aims to extract the features of the stripe and image. Moreover, the features in the two WEU streams are merged together to influence each other. Finally, the extracted features from the two streams are imported into the reconstruction module to obtain a clear image. Thus, the final loss is defined as follows:

$$\boldsymbol{J} = \frac{\alpha}{2}\|\mathcal{F}_I([\boldsymbol{Y};\boldsymbol{W}]) - \boldsymbol{X}\|^2 + \frac{\beta}{2}\|\mathcal{F}_S([\boldsymbol{Y};\boldsymbol{W}]) - \boldsymbol{B}\|^2 \quad (5)$$

where $\mathcal{F}^I$ and $\mathcal{F}^S$ are the mapping functions about the parameter $\boldsymbol{W}$, and α and β are the balance parameters. To verify the effect of the two-stream framework, we also plot the training loss and testing values of the TSWEU model in Fig. 5. We can conclude that the two-stream approach is beneficial for the feature propagation and facilitates the destriping results.

D. Training Details

For the simulation of the stripes, we take the multiplicative stripe as an example. We generate the stripe image $\boldsymbol{Y}$ by multiplying the input image $\boldsymbol{X}$ with the coefficient matrix $\boldsymbol{A}$ via $\boldsymbol{Y} = \boldsymbol{X}. * \boldsymbol{A}$ [22]. The matrix coefficient is generated by $\boldsymbol{A} = repmat((stripe_{max} - stripe_{min}) * rand(1, size(\boldsymbol{X}, 2)) + stripe_{min}, size(\boldsymbol{X}, 1), 1)$, where $stripe_{max}$ and $stripe_{min}$ are pre-defined hyper-parameters which determine the range of the multiplicative stripe. “repmat” and “size” are the Matlab functions. The interested readers could refer to the released simulating codes for other stripe categories.

We initialize the convolutional filters with the Xavier method [63]. The learning rate is initially set as 0.0005 and decreased to a small value of 0.00005. The momentum and decay are fixed as 0.9 and 0, respectively. The ADAM solver [69] is introduced to optimize the model. We train the model with 100 epochs with a batch size of 24. The training data are normalized to [0, 1]. Compared with the image, the stripe can be regarded as the residual noise, which is much more easy to be trained with smaller training error. Moreover, the intensity and area of the stripe are usually smaller than that of the image. When computing the reconstruction error in (5), the error of the image is obviously larger than that of the stripe. Thus, we set the hyperparameter $\alpha = 0.001$ and $\beta = 1$ to balance the reconstruction error between the image and the stripe. The MatConvNet toolbox [70] is employed to train the TSWEU. It is worth noting that due to the nonuniform property of the stripe noise, a training image with a large receptive field can significantly boost the final destriping results. We randomly choose 20 000 samples from the Place2 data set with size 256*256 for training. Here, we use the natural images as the training set since the remote sensing images vary in different imaging systems.

IV. Experimental Results and Discussion

A. Experimental Setting

We have comprehensively compared the proposed TSWEU with the state-of-the-art destriping methods, including the TV [71], WFAF [13], low-rank single-image decomposition (LRSID) [41], unidirectional TV (UTV) [23], statistical linear destriping (SLD) [22], DLS-NUC [54], transformed low-rank (TLR) [43], weighted nuclear norm minimization (WNNM) [72], and the framelet [73] methods. To provide a fair comparison, we collect 15 remote sensing images as the simulated test data set and seven images as the real test data set, as shown in Fig. 7.

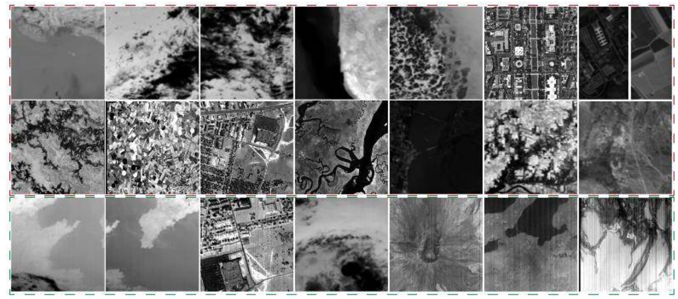

Fig. 7. Simulated and real image data set used in this article. We select 15 images as the simulated data surrounded by the red rectangle and seven images as the real data surrounded by the green rectangle.

To provide an overall evaluation of the destriping performance, several qualitative and quantitative assessments are used. The qualitative assessments include the visual inspection, the mean cross-track profile, and the power spectrum. The PSNR and SSIM [74] are employed for the quantitative assessment. All codes are provided by the authors, and the parameters are fine-tuned to achieve the best performance on average. It is worth noting that we do not adjust the parameters of competing methods for each test image but set the same parameter for all the test images. The training code and testing data sets of this article can be downloaded at the homepage of the author.[1]

B. Simulated Image Destriping

According to the different properties of the stripe, the stripe can be classified into several categories. In this section, we test the representative and difficult stripe categories.

1) Multiplicative Response: Most of the previous methods focus on the additive stripe, except for the SLD [22]. We first transform the multiplicative stripe image into the additive domain via the logarithm function, then apply these additive destriping methods, and finally re-transform the destriping results into the original domain via the exponent function. It is worth noting that this can only be used when only the multiplicative stripe exists without any additive stripe or random noise. In addition, TSWEU can remove the multiplicative stripe in any condition. From the visual results in Fig. 8, most of the compared methods remove the vertical image structure unexpectedly. In contrast, the proposed method can satisfactorily preserve the line pattern of the image structure marked by the red rectangle. Additionally, the estimated multiplicative stripe component in Fig. 8(j) is highly signal-dependent.

2) Proportion: The removal of the full proportion stripe in the push-broom system is usually more difficult since the stripe covers the whole image space. In Fig. 9(c) and (h), the TV

[1] http://www.escience.cn/people/changyi/index.html.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

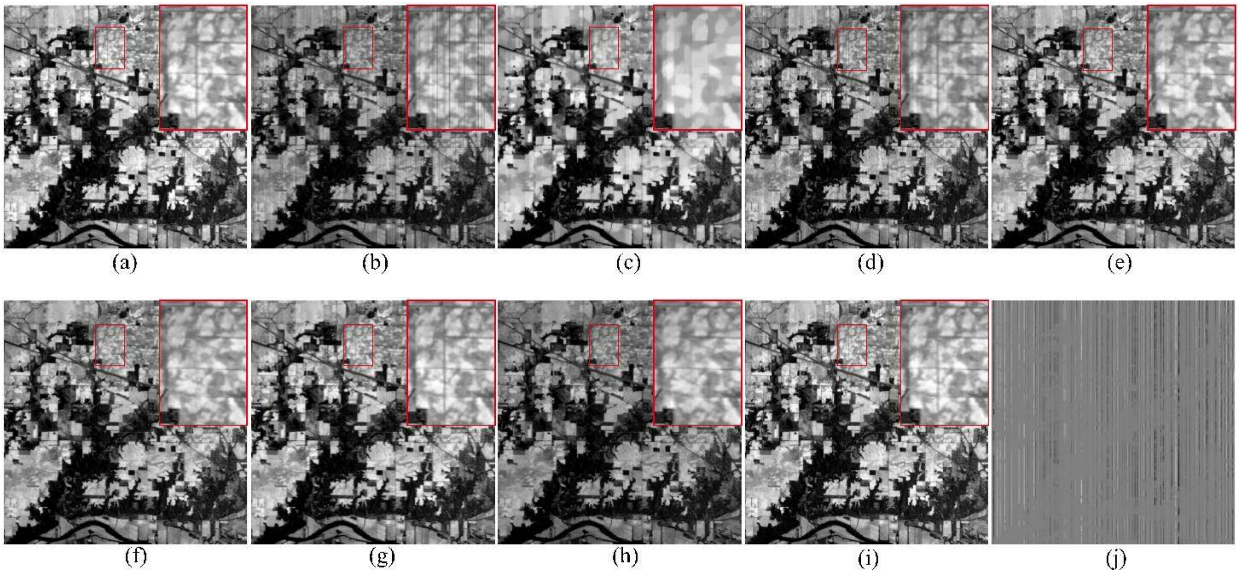

Fig. 8. Simulated destriping results for the multiplicative case. (a) Original HSI NS_line band 142. (b) Degraded with multiplicative stripes. Destriping results by (c) TV, (d) WFAF, (e) LRSID, (f) UTV, (g) SLD, (h) DLS-NUC, and (i) TSWEU. (g) Estimated stripes by TSWEU.

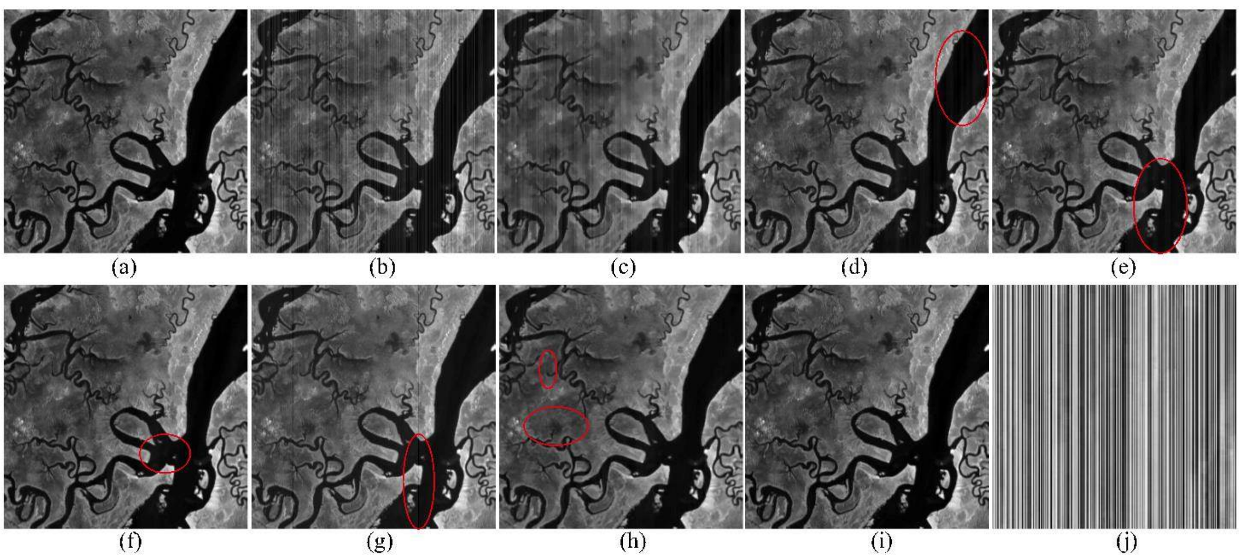

Fig. 9. Simulated destriping results for the full proportion case. (a) Original HSI Suwannee band 70. (b) Degraded with full proportion stripes. Destriping results by (c) TV, (d) WFAF, (e) LRSID, (f) UTV, (g) SLD, (h) DLS-NUC, and (i) TSWEU. (g) Estimated stripes by TSWEU.

and DLS-NUC has oversmoothed the details unexpectedly. There are residual stripes for the WFAF, LRSID, and SLD that are especially obvious in the low intensity region marked by the red ellipse. The estimated stripe and image components achieved with TSWEU are visually pleasing and quantitatively better than that of the other methods.

3) Length: Although it seems counterintuitive, it is much more difficult to remove the random stripe than the full-length stripe. On the one hand, the random stripe is more difficult to differentiate from the line pattern of the image texture. On the other hand, the cross-assumption over the whole image is no longer valid, such as the low-rank assumption for LRSID and the rank 1 assumption for SLD. In Fig. 10, the existing methods are less effective for the random-length stripes, which usually exist in the MODIS band 33. The results of TSWEU [see Fig. 10(i) and (j)] show that our method can well handle the random-length stripe with a significant advantage over the previous methods.

We also test the effectiveness of all competing methods for different stripe noise level, as shown in Table III. Our TSWEU consistently outperforms the state-of-the-art methods by a large margin of at least 5 dB, except for the SLD. It is worth noting that we simulate the full-length stripe with exactly rank 1, which perfectly fits the strong assumption of SLD. That is the main reason why SLD performs well on the full-length stripe, whereas it works poorly on the random-length stripe. Moreover, with the increasing level of the stripe noise, the advantage of the TSWEU is much larger. Due to space limitations, we do not show the quantitative results on the other kinds of stripes.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

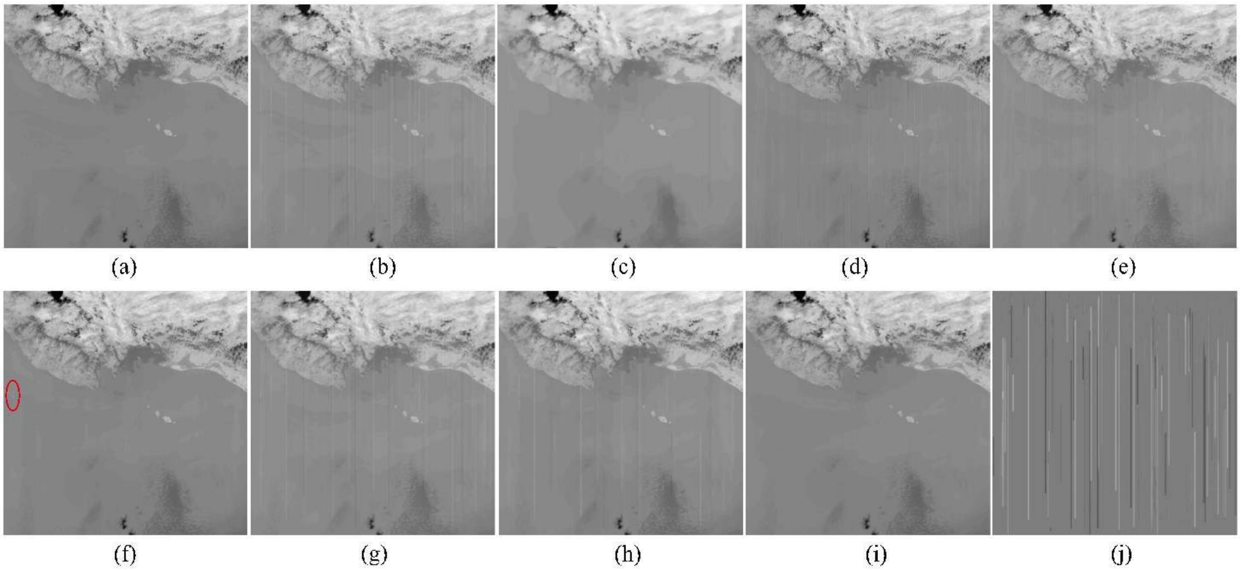

Fig. 10. Simulated destriping results for the random length case. (a) Original MODIS image Aqua band 31. (b) Degraded with random length stripes. Destriping results by (c) TV, (d) WFAF, (e) LRSID, (f) UTV, (g) SLD, (h) DLS-NUC, and (i) TSWEU. (g) Estimated stripes by TSWEU.

TABLE III
QUANTITATIVE ASSESSMENTS OF DIFFERENT METHODS UNDER DIFFERENT NOISE LEVELS

Category	Stripe Level	Index	Method							
			Noisy	TV [71]	WFAF [13]	LRSID [41]	UTV [23]	SLD [22]	DLS [54]	TSWEU
Full Length	{-10,10}	PSNR	33.04	33.99	39.19	40.19	41.57	44.79	36.38	**46.82**
		SSIM	0.8685	0.9227	0.9839	0.9878	0.9923	0.9964	0.9732	**0.9973**
	{-20,20}	PSNR	26.98	29.41	35.44	36.17	36.87	41.87	34.47	**43.63**
		SSIM	0.6879	0.8304	0.9748	0.9855	0.9842	0.9949	0.9638	**0.9953**
	{-30,30}	PSNR	23.71	26.93	33.28	33.93	34.25	39.01	32.75	**42.92**
		SSIM	0.5650	0.7563	0.9640	0.9720	0.9731	0.9495	0.9546	**0.9954**
	{-40,40}	PSNR	21.36	25.12	31.17	31.66	32.75	38.33	31.29	**41.78**
		SSIM	0.4708	0.6902	0.9507	0.9440	0.9671	0.9914	0.9362	**0.9933**
	{-50,50}	PSNR	19.51	24.05	30.23	31.13	31.66	37.13	29.18	**41.47**
		SSIM	0.4000	0.6391	0.9443	0.9386	0.9615	0.9903	0.8986	**0.9934**
	{-100,100}	PSNR	13.90	20.48	25.14	23.84	26.16	33.20	20.62	**38.13**
		SSIM	0.2064	0.4779	0.8804	0.7274	0.9033	0.9703	0.5892	**0.9900**
Random Length	{-40,40}	PSNR	32.59	27.09	34.26	33.27	39.54	34.50	33.02	**51.45**
		SSIM	0.9162	0.7523	0.9414	0.9470	0.9869	0.9519	0.9387	**0.9985**

4) Periodicity: The periodical stripe shows regular patterns and always exists in the whisk-broom imaging system and is indeed easier to be removed than the nonperiodical stripe. Here, we choose the typical MODIS Aqua band 22. The periodicity is 10. Note that the stripes not only are periodical but also have a broad width. In Fig. 11, we can observe that most of the compared methods can well remove the periodical stripe with satisfactory visual results. The estimated stripe achieved by TSWEU is composed of exactly periodical lines.

5) Deadline/Inpainting (Intensity): The deadlines are usually caused by the malfunction of certain detectors, which makes the problem difficult. It is difficult to recover the useful information for the conventional single-image-based destriping methods. Here, we consider this problem as an image inpainting issue and compare the proposed TSWEU with the WNNM [72] and Framelet [73] methods [25% and 50% missing pixels Fig. 12]. It is worth noting that the location of the deadlines should be provided in advance. The WNNM and TSWEU can well reconstruct the missing deadlines with a pleasing visual appearance. Moreover, the TSWEU obtains higher quantitative indexes.

6) Oblique Stripe (Angle): Most of the existing destriping methods are designed for the horizontal or vertical stripe only. The previous methods need to rotate the oblique stripe image into the horizontal/vertical ones, which would inevitably cause an information loss due to the interpolation operator. In contrast, our method can handle the oblique stripe with an arbitrary angle in the original image domain, mainly because we have fed TSWEU with the oblique stripe for training. Thus, the proposed TSWEU could well handle the angle variation. Here, we compare the TSWEU with the TLR [43] under different rotation angles, as shown in Fig. 13. We have three observations. First, the proposed method can better remove the oblique stripes than the TLR from both the visual and quantitative assessments. Second, the estimated oblique stripes in the last column do not contain any residual image structure.

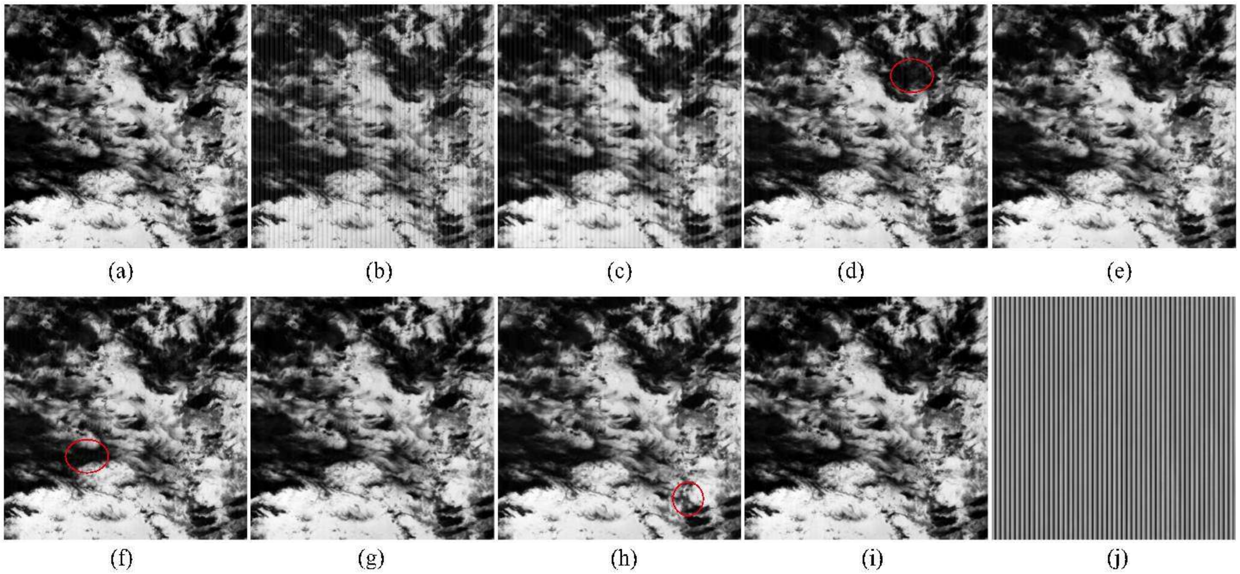

Fig. 11. Simulated destriping results for the periodical and broad case. (a) Original MODIS image Aqua band 22. (b) Degraded with periodical and broad stripes. Destriping results by (c) TV, (d) WFAF, (e) LRSID, (f) UTV, (g) SLD, (h) DLS-NUC, and (i) TSWEU. (g) Estimated stripes by TSWEU.

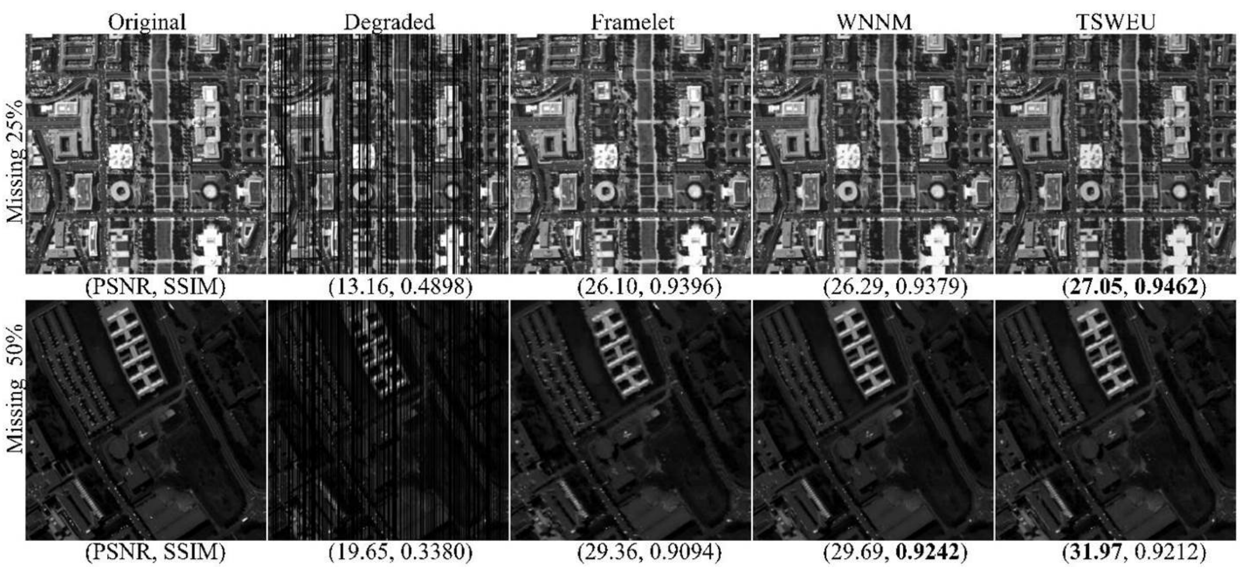

Fig. 12. Simulated destriping results for the deadline case. The first and second rows show the 25% and 50% missing of the HSI Washington DC and paviaU band 20, respectively. From the first to the last column, we show the original, the degrade, and the inpainting results of Framelet, WNNM, and TSWEU.

Last but not least, TSWEU is very robust to the angle of the stripe, where the conventional horizontal/vertical stripe can be regarded as a special case in our method.

7) Mixed Noise: Random noise usually coexists with the stripe in the remote sensing images. Previous methods always rely on the spectral correlation of the multispectral images, while few works handle this problem from a single image. We can observe that the existing destriping methods may fail unexpectedly in the presence of random noise. There are obvious residual stripes in the destriping results, such as the LRSID, UTV, and SLD in Fig. 14(e)–(g). Although the WFAF and DLS-NUC have removed the stripe well, random noise still remains in the results [see Fig. 14(d) and (h)]. For TSWEU, we have satisfactorily decoupled the clean image [see Fig. 14(i)] and the mixed noise [see Fig. 14(j)].

C. Real Image Destriping

To demonstrate the robustness of our algorithm, we test the proposed TSWEU on real stripe remote sensing images, as shown in Fig. 15. We have chosen four representative nonperiodical stripe images of the push-broom-based hyperspectral imaging system and three periodical stripe images of the cross track-based moderate resolution imaging system. It is shown that the TSWEU has completely removed the stripe and consistently achieved a visually pleasing quality for all cases, whereas other competing methods may fail for certain cases. For example, the SLD has achieved very impressive destriping results for the simulated stripes. However, for the real stripes with nonrank 1, the performance of SLD decreases rapidly. It is worth noting that all competing methods fail for the random-length stripe in MODIS Terra band 33. Our method

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

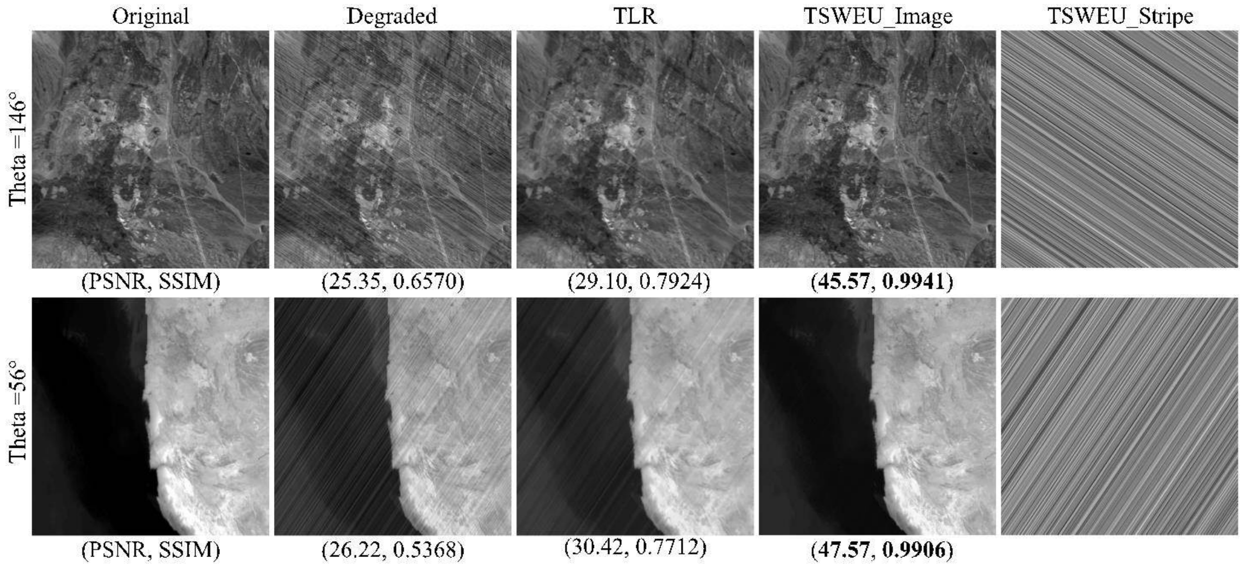

Fig. 13. Simulated destriping results for the oblique case. The first and second rows show the 146° and 56° stripe on the Cuprite band 10 and Terra band 31, respectively. From the first to the last column, we show the original, the degrade, the destriping results of TLR and TSWEU, estimated stripe by TSWEU.

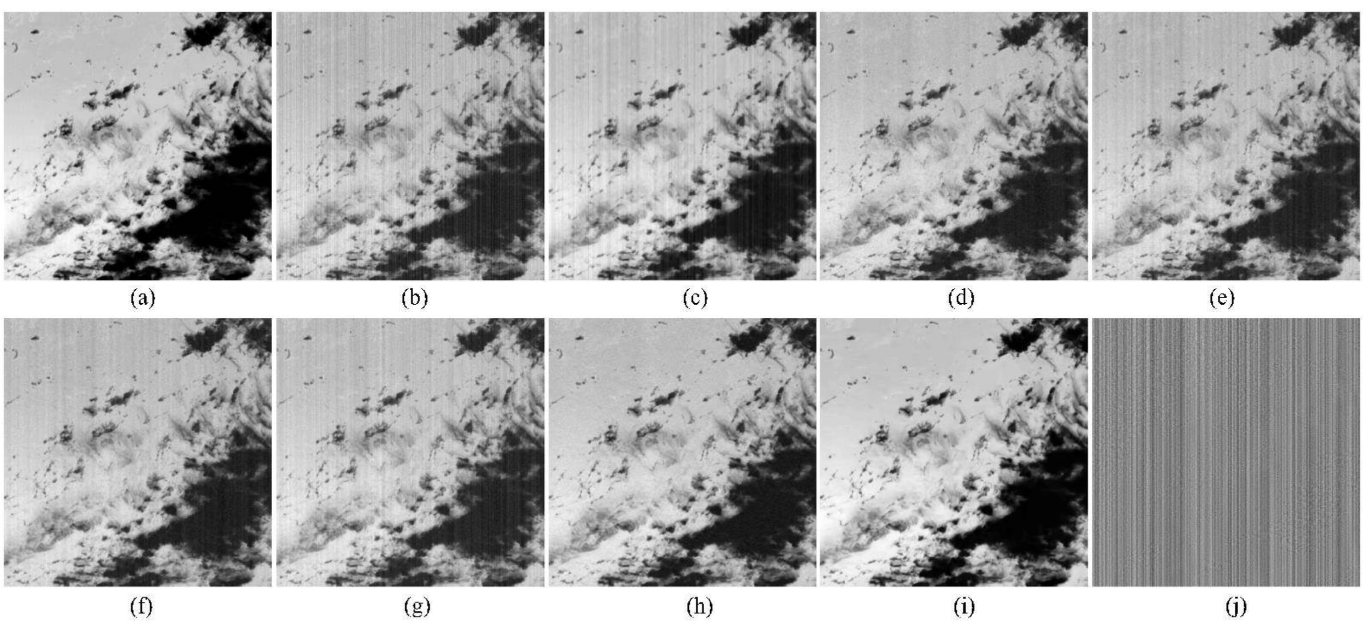

Fig. 14. Simulated destriping results for the mixed noise case. (a) Original MODIS Aqua band 22. (b) Degraded with mixed noise. Restoration results by (c) TV, (d) WFAF, (e) LRSID, (f) UTV, (g) SLD, (h) DLS-NUC, and (i) TSWEU. (g) Estimated noise by TSWEU.

still works well in this case. Overall, the results of the proposed method are consistent for all test images and exhibit good visual quality with fewer artifacts than those obtained by the compared methods.

To further comprehensively evaluate the effectiveness of the proposed method on the real images, we introduce the no-reference assessments: inverse coefficient of variation (ICV) [23] and mean relative deviation (MRD) [20]. To reduce the bias, we compute the ICV and MRD five times in different regions and report the mean ICVs and MRDs as MICV and MMRD, respectively. The no-reference quantitative assessments are listed in Table V. We have two main observations. First, for the MICV index, the LRSID and TSWEU usually obtain the first and second best results. The ICV evaluation index could partially reflect the degree of stripe removal. ICV index is a reflection of the noise removal including both the stripe and random noise. It is worth noting that the LRSID, TSWEU, and DLS-NUC could well handle the random and stripe noise simultaneously, whereas other WFAF, UTV, and SLD could only handle the stripe noise. Unfortunately, the random noise is ubiquitous, even sometimes invisible to the naked eye. In summary, the MICV index means that the proposed TSWEU and LRSID have achieved better noise removal performance, largely due to stripe removal. Second, the WFAF has mostly obtained the best MMRD index. It is worth noting that the MRD is proportional to the

Fig. 15. Real destriping results for various remote sensing images. (From left to right) Real degrade image, the destriping results of LRSID, UTV, SLD, DLS-NUC, TSWEU. (From top to bottom) Hyperspectral image CHRIS band41, LakeMonona band105, MtStHelens band117, Urban band103, and the MODIS image Terra band27, Terra band30, Terra band33.

difference between the destriping and stripe image. That is why the MRD index of the stripe image itself is zero. In fact, this index is reasonable only when there is not any stripe on the calculating windows. Otherwise, the more stripes you remove, the worse index you obtain. Unfortunately, the stripes in the HSI are everywhere, and the periodic in MODIS is exactly ten.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

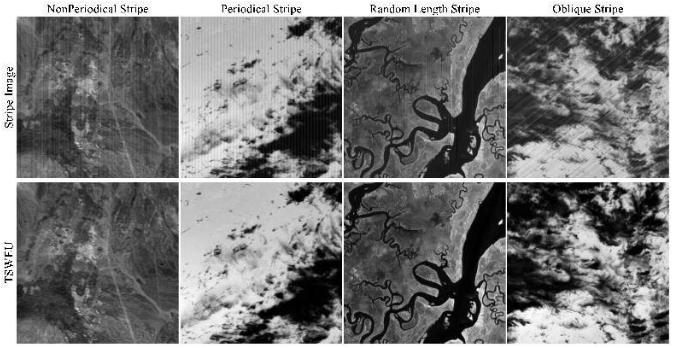

Fig. 16. Effectiveness of the proposed method with single model for different stripe noise types. (From left to right) Nonperiodical, periodical, random length, and oblique stripe, respectively.

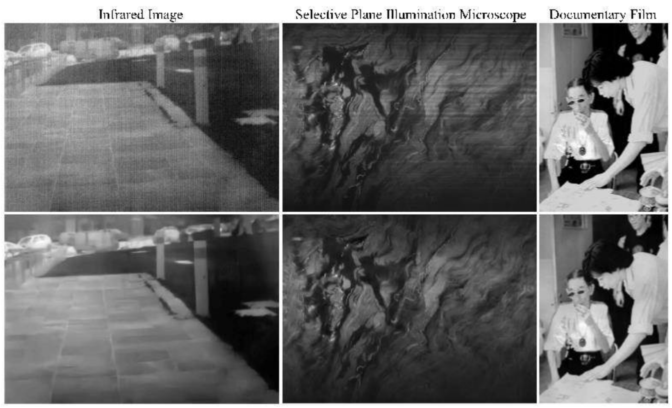

Fig. 17. Generalization of the proposed method for various imaging systems. (From left to right) IR, SPIM, and old film stripe images.

That is the main reason why WFAF and SLD obtain better MMRD, since there are obvious residual stripe in real images.

D. Discussion

1) Ablation Study: In this section, we study the effectiveness of each term in our article as shown in Table IV. We report the average PSNR/SSIM of each method on the simulated data sets of stripe noise between [−20, 20]. We can observe that the UNet obtains much better performance than that of the plain net. The skip connection promotes the information propagation over a long distance and the downsample/upsample operator that benefits enlarging the receptive field work to facilitate improving the destriping performance significantly. Moreover, the embedded residual blocks that promote the information propagation over a short distance help to improve the performance. Furthermore, the wavelet that replaces the conventional downsample/upsample operator with a lossless reconstruction slightly contributes to the final performance. Lastly, the two-stream strategy obviously improves the SSIM.

2) Single Model for Different Stripes: Explicitly modeling the distributions of the complex stripe noises with different stripe categories and stripe intensity levels is extremely hard for previous methods. Thanks to the universal approximation theorem [65] which states that a feed-forward network with a single hidden layer containing a finite number of neurons can approximate continuous functions on compact subsets of $\mathbb{R}^n$, under mild assumptions on the activation function. Here, we show that our single model is robust to different stripe noise with different scenarios. In Fig. 16, we perform an experiment to train one single model for different stripe categories with different remote sensing images. We train one single model on 60 000 images, where every 15 000 samples are simulated for each kind of stripe. Here, we just select four kinds of stripes due to limitation of the GPU memory. However, we think it is possible to use one single model for all kinds of the stripes, which requires very large data set and powerful computer. We can observe that the stripes have been completely removed by single TSWEU model, which strongly validates the effectiveness and robustness of TSWEU to any stripe noise and image scenario.

3) Generalization for Other Imaging: To validate the generalization of the proposed method, we perform several destriping experiments on various real stripe images. For each kind of

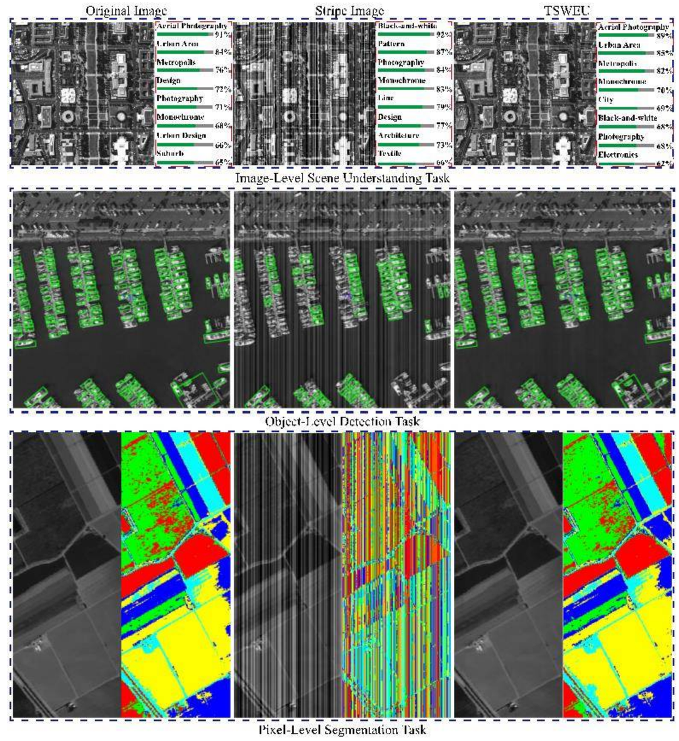

Fig. 18. Evaluation of destriping results of the proposed TSWEU on different high-level tasks. From the first to the third rows, we show the image-level, object level, and pixel-level task before and after destriping. (From left to right) Original, striped, and the destriping images, respectively.

TABLE IV
ABLATION STUDY OF EACH TERM

Model	CNN	Skip	Res	Wavelet	TS	PSNR/SSIM
Plain	√	×	×	×	×	39.39/0.9892
UNet	√	√	×	×	×	40.67/0.9923
EUNet	√	√	√	×	×	41.72/0.9932
WEU	√	√	√	√	×	42.11/0.9930
TSWEU	√	√	√	√	√	**42.27/0.9943**

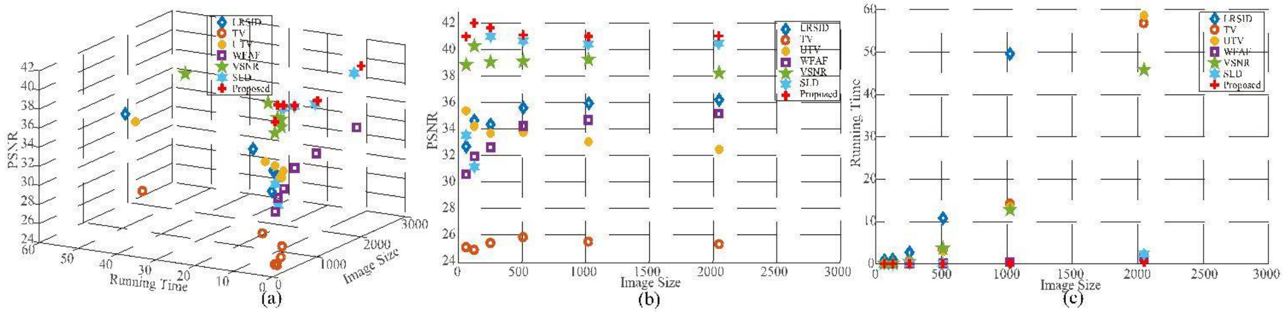

Fig. 19. (a) Relationship among image size, running time, and the performance. (b) Relationship among image size and the performance. (c) Relationship among image size and the running time.

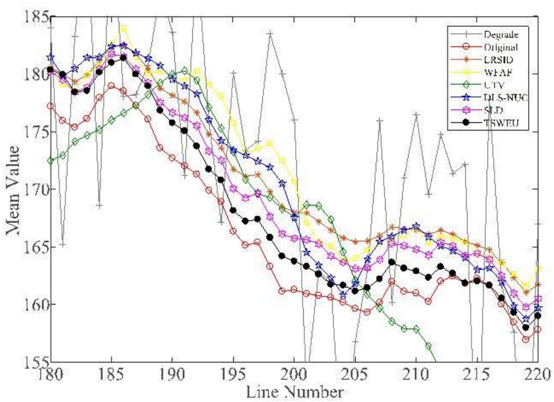

Fig. 20. Mean cross profile analysis. The horizontal axis means the column number of the image, and the vertical axis denotes the mean intensity value of the image.

TABLE V

QUANTITATIVE ASSESSMENT OF THE REAL STRIPE IMAGES. THE FIRST AND SECOND BEST RESULTS ARE MARKED BY THE RED AND BLUE

Images	Index	Methods						
		Noisy	WFAF	LRSID	UTV	SLD	DLS-NUC	TSWEU
Terra	MICV	23.39	53.36	100.94	94.97	69.18	94.34	98.89
band27	MMRD	0	3.61	4.86	5.01	4.19	5.94	5.77
Terra	MICV	19.74	72.28	153.62	88.43	83.45	94.45	105.19
band30	MMRD	0	1.68	1.85	2.02	1.70	2.43	1.91
Terra	MICV	51.36	77.53	251.58	116.09	93.87	108.63	119.84
band33	MMRD	0	0.41	0.83	0.91	0.99	0.71	0.97
CHRIS	MICV	10.68	27.19	34.55	30.62	30.07	33.05	46.78
band41	MMRD	0	8.23	9.08	8.67	8.56	9.17	8.63
Urban	MICV	11.25	14.26	16.46	14.53	13.33	16.92	17.83
band103	MMRD	0	3.90	3.49	4.70	4.94	5.13	4.41
LakeMonona	MICV	12.16	20.97	34.55	30.62	21.60	35.47	33.56
band105	MMRD	0	6.94	7.43	7.53	7.12	7.57	7.32
MtStHelens	MICV	16.94	28.85	48.34	27.98	30.30	34.15	37.20
band117	MMRD	0	3.84	4.17	4.40	4.31	4.34	5.60

the images, the cause of the stripe and the imaging technique is totally different. Those test real image scenes and stripe categories are completely "unknown" to the training data set. In Fig. 17, the stripes in different images are completely different, such as the mixed noise in infrared image, multiplicative stripe in SPIM [75], and the sparse broken stripe in old film documentary. The destriping results of TSWEU are visual pleasure without any residual stripe. Moreover, both the image edges and textures are well preserved. It is worth noting that our model is trained on the natural image, and could be directly applied to those images without any re-training or other operations. Our TSWEU has achieved very impressive destriping results on those different images with completely different stripes. The destriping results have strongly supported the generalization ability of our model. The reasons are twofold. First, the proposed model has actually learned the intrinsic line pattern property of the stripes. It is not a hard thing for CNN to extract the line pattern features with low-dimensional manifold. Second, our TSWEU learns not only the stripe but also the image. The learned natural image shares the similar point, edge, profile features with the other image, which could be well transferred to other images.

4) Effectiveness for High-Level Tasks: Considering the destriping is a pre-processing for subsequent application, we further demonstrate the effectiveness of the destriping results on several high-level vision tasks. We comprehensively evaluate the destriping results on image-level scene understanding, object-level detection, and pixel-level segmentation tasks, as shown in Fig. 18. We employ the Google Vision API https://cloud.google.com/vision/ on the images before and after destriping to perform the scene recognition. For the scene recognition, the stripes have brought negative influence on the recognition, where the API recognizes the Washington DC as "Black-and-white," "Pattern" to name a few. After the destriping by TSWEU, we can observe that the recognition labels are highly consistent with that of the original image, such as "Aerial Photography," "Urban Area," "Metropolis" with similar confidence scores. For the object detection, we take the DOTA [76] image as example. The stripe obviously reduces the detection number (79) of small ship objects. After the destriping by TSWEU, the number of the detected ship objects (147) is even slightly higher than that of the original image (145). For the segmentation, we use the unsupervised k-means clustering method as semantic segmentation for the HSI Salinas. The number of the class is five. The maximum iteration is set as ten times. The segmentation result of the stripe image is completely false without any structural information. After the destriping, the semantic segmentation result is meaningful and very similar to that of the original image. Overall, the proposed TSWEU could significantly improve the performance of the subsequent high-level vision tasks.

5) Influence of Image Size: In this section, we analyze the influence of the image size to the destriping performance and the running time. Here, we set the image size from 64×64 to 2048×2048 by making it two times larger each time. From Fig. 19(b), we can conclude that the image size has a different influence on different methods. For example, with the increasing image size, the performance of the UTV gradually decreases. From Fig. 19(c), the running time of most of the competing methods increases rapidly with the increasing image size, such as the TV, UTV, VSNR, and LRSID, whereas the running time of the TSWEU is almost constant. That is, our method is very robust to the image size, which is an important merit for large-sized remote sensing images.

6) Mean Cross-Profile Analysis: In this section, we analyze the mean cross-profile of the destriping result, as shown in Fig. 20. The mean cross-track profile of the destriping result should be closer to that of the original image, where the abrupt change (gray curve) in mean cross-track profile caused by the stripe should be smoothed. To better visualize this, we just select the row number between [180, 220]. We can observe that the destriping result of TSWEU (black curve) is much closer to the original ground truth (red curve).

V. Conclusion

In this article, we formulate the single image destriping task as an image decomposition problem, where the stripe component and image component are treated equally via a two-stream CNN. The CNN is beneficial for representing the stripe noise with more discriminative features via the external data set. Moreover, we embed the wavelet into the CNN to learn the internal directional property of the stripe better. We also provide a comprehensive category of the remote sensing stripes from their visual appearance. While previous methods may be suitable for some of them, the proposed method can well handle all of them due to the powerful representation ability of the model. The proposed method has been extensively verified on various simulated and real striped images and outperforms the state-of-the-art methods by a large margin in terms of quantitative and qualitative assessments, robustness to stripe categories, running time, and so on.

References

[1] B. K. P. Horn and R. J. Woodham, "Destriping LANDSAT MSS images by histogram modification," *Comput. Graph. Image Process.*, vol. 10, no. 1, pp. 69–83, May 1979.

[2] M. Wegener, "Destriping multiple sensor imagery by improved histogram matching," *Int. J. Remote Sens.*, vol. 11, no. 5, pp. 859–875, 1990.

[3] G. Corsini, M. Diani, and T. Walzel, "Striping removal in MOS-B data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1439–1446, May 2000.

[4] F. L. Gadallah, F. Csillag, and E. J. M. Smith, "Destriping multisensor imagery with moment matching," *Int. J. Remote Sens.*, vol. 21, pp. 2505–2511, Aug. 2000.

[5] P. Meza, J. E. Pezoa, and S. N. Torres, "Multidimensional striping noise compensation in hyperspectral imaging: Exploiting hypercubes' spatial, spectral, and temporal redundancy," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 9, pp. 4428–4441, Sep. 2016.

[6] J. J. Simpson, J. R. Stitt, and D. M. Leath, "Improved finite impulse response filters for enhanced destriping of geostationary satellite data," *Remote Sens. Environ.*, vol. 66, no. 3, pp. 235–249, Dec. 1998.

[7] J. Chen, Y. Shao, H. Guo, W. Wang, and B. Zhu, "Destriping CMODIS data by power filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 2119–2124, Sep. 2003.

[8] J. Chen, H. Lin, Y. Shao, and L. Yang, "Oblique striping removal in remote sensing imagery based on wavelet transform," *Int. J. Remote Sens.*, vol. 27, no. 8, pp. 1717–1723, Apr. 2006.

[9] J. G. Liu and G. L. K. Morgan, "FFT selective and adaptive filtering for removal of systematic noise in ETM+ imageodesy images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 12, pp. 3716–3724, Dec. 2006.

[10] 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.

[11] L. Gómez-Chova, L. Alonso, L. Guanter, G. Camps-Valls, J. Calpe, and J. Moreno, "Correction of systematic spatial noise in push-broom hyperspectral sensors: Application to CHRIS/PROBA images," *Appl. Opt.*, vol. 47, no. 28, pp. F46–F60, 2008.

[12] P. Rakwatin, W. Takeuchi, and Y. Yasuoka, "Restoration of Aqua MODIS band 6 using histogram matching and local least squares fitting," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 2, pp. 613–627, Feb. 2009.

[13] B. Münch, P. Trtik, F. Marone, and M. Stampanoni, "Stripe and ring artifact removal with combined wavelet—Fourier filtering," *Opt. Express*, vol. 17, no. 10, pp. 8567–8591, Jan. 2009.

[14] Z. Wang, L. Chen, X. Gu, and T. Yu, "Destriping MODIS data based on surface spectral correlation," in *Proc. IEEE Conf. IGRASS*, Jul. 2008, pp. 229–266.

[15] H.-S. Jung, J.-S. Won, M.-H. Kang, and Y.-W. Lee, "Detection and restoration of defective lines in the SPOT 4 SWIR band," *IEEE Trans. Image Process.*, vol. 19, no. 8, pp. 2143–2156, Aug. 2010.

[16] 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.

[17] I. Gladkova, M. D. Grossberg, F. Shahriar, G. Bonev, and P. Romanov, "Quantitative restoration for MODIS band 6 on aqua," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 6, pp. 2409–2416, Jun. 2012.

[18] Y. Duan, W. Chen, M. Wang, and L. Yan, "A relative radiometric correction method for airborne image using outdoor calibration and image statistics," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 5164–5174, Aug. 2014.

[19] Y. Cao, M. Y. Yang, and C.-L. Tisse, "Effective strip noise removal for low-textured infrared images based on 1-D guided filtering," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 12, pp. 2176–2188, Dec. 2016.

[20] H. Shen and L. Zhang, "A MAP-based algorithm for destriping and inpainting of remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1492–1502, May 2009.

[21] M. D. Bisceglie, R. Episcopo, C. Galdi, and S. L. Ullo, "Destriping MODIS data using overlapping field-of-view method," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 2, pp. 637–651, Feb. 2009.

[22] H. Carfantan and J. Idier, "Statistical linear destriping of satellite-based pushbroom-type images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 4, pp. 1860–1871, Apr. 2010.

[23] M. Bouali and S. Ladjal, "Toward optimal destriping of MODIS data using a unidirectional variational model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 8, pp. 2924–2935, Aug. 2011.

[24] J. Fehrenbach, P. Weiss, and C. Lorenzo, "Variational algorithms to remove stationary noise: Applications to microscopy imaging," *IEEE Trans. Image Process.*, vol. 21, no. 10, pp. 4420–4430, Oct. 2012.

[25] 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.

[26] Y. Chang, H. Fang, L. Yan, and H. Liu, "Robust destriping method with unidirectional total variation and framelet regularization," *Opt. Express*, vol. 21, no. 20, pp. 23307–23323, 2013.

[27] Y. Chang, L. Yan, H. Fang, and H. Liu, "Simultaneous destriping and denoising for remote sensing images with unidirectional total variation and sparse representation," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 6, pp. 1051–1055, Jun. 2014.

[28] 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.

[29] 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.

[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. H. Fitschen, J. Ma, and S. Schuff, "Removal of curtaining effects by a variational model with directional forward differences," *Comput. Vis. Image Understand.*, vol. 155, pp. 24–32, Feb. 2017.

[32] X. Liu, X. Lu, H. Shen, Q. Yuan, and L. Zhang, "Oblique stripe removal in remote sensing images via oriented variation," 2018, *arXiv:1809.02043*. [Online]. Available: https://arxiv.org/abs/1809.02043

[33] X. Liu, H. Shen, Q. Yuan, X. Lu, and C. Zhou, "A universal destriping framework combining 1-D and 2-D variational optimization methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 808–822, Feb. 2018.

[34] Z. Lin, R. Liu, and Z. Su, "Linearized alternating direction method with adaptive penalty for low-rank representation," in *Proc. NIPS*, 2011, pp. 612–620.

[35] N. Acito, M. Diani, and G. Corsini, "Subspace-based striping noise reduction in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 4, pp. 1325–1342, Apr. 2011.

[36] X. Lu, Y. Wang, and Y. Yuan, "Graph-regularized low-rank representation for destriping of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 4009–4018, Jul. 2013.

[37] 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.

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

[39] M. Wang, J. Yu, J.-H. Xue, and W. Sun, "Denoising of hyperspectral images using group low-rank representation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 9, pp. 4420–4427, Sep. 2016.

[40] 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.

[41] 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.

[42] 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.

[43] Y. Chang, L. Yan, and S. Zhong, "Transformed low-rank model for line pattern noise removal," in *Proc. IEEE Conf. ICCV*, Oct. 2017, pp. 1726–1734.

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

[45] W. Cao, Y. Chang, G. Han, and J. Li, "Destriping remote sensing image via low-rank approximation and nonlocal total variation," *IEEE Trans. Geosci. Remote Sens.*, vol. 15, no. 6, pp. 848–852, Jun. 2018.

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

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

[48] H. Fan, Y. Chen, Y. Guo, H. Zhang, and G. Kuang, "Hyperspectral image restoration using low-rank tensor recovery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 10, pp. 4589–4604, Oct. 2017.

[49] 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.

[50] W. Cao, K. Wang, G. Han, J. Yao, and A. Cichocki, "A robust PCA approach with noise structure learning and spatial–spectral low-rank modeling for hyperspectral image restoration," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 10, pp. 3863–3879, Oct. 2018.

[51] H. Fan, C. Li, Y. Guo, G. Kuang, and J. Ma, "Spatial–spectral total variation regularized low-rank tensor decomposition for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 6196–6213, Oct. 2018.

[52] Y. Wang, J. Peng, Q. Zhao, D. Meng, Y. Leung, and X.-L. Zhao, "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.

[53] X. Kuang, X. Sui, Q. Chen, and G. Gu, "Single infrared image stripe noise removal using deep convolutional networks," *IEEE Photon. J.*, vol. 9, no. 4, Aug. 2017, Art. no. 3900913.

[54] Z. He, Y. Cao, Y. Dong, J. Yang, Y. Cao, and C.-L. Tisse, "Single image based nonuniformity correction of uncooled long-wave infrared detectors: A deep learning approach," *Appl. Opt.*, vol. 57, no. 18, pp. D155–D164, 2018.

[55] P. Xiao, Y. Guo, and P. Zhuang, "Removing stripe noise from infrared cloud images via deep convolutional networks," *IEEE Photon. J.*, vol. 10, no. 4, Aug. 2018, Art. no. 7801114.

[56] W. Xie, Y. Li, and X. Jia, "Deep convolutional networks with residual learning for accurate spectral-spatial denoising," *Neurocomputing*, vol. 312, pp. 372–381, Oct. 2018.

[57] Q. Zhang, Q. Yuan, J. Li, X. Liu, H. Shen, and L. Zhang, "Hybrid noise removal in hyperspectral imagery with a spatial-spectral gradient network," 2018, *arXiv:1810.00495*. [Online]. Available: https://arxiv.org/abs/1810.00495

[58] 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.

[59] Q. Yuan, Q. Zhang, J. Li, H. Shen, and L. Zhang, "Hyperspectral image denoising employing a spatial–spectral deep residual convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 1205–1218, 2019.

[60] Y. Chang, L. Yan, L. Lu, H. Fang, and S. Zhong, "Infrared aerothermal nonuniform correction via deep multiscale residual network," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 7, pp. 1120–1124, Jul. 2019.

[61] 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.

[62] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2015, pp. 234–241.

[63] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.

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

[65] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, no. 5, pp. 359–366, 1989.

[66] X. Tao, H. Gao, X. Shen, J. Wang, and J. Jia, "Scale-recurrent network for deep image deblurring," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 8174–8182.

[67] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 4700–4708.

[68] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.

[69] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. ICLR*, 2015, pp. 1–15.

[70] A. Vedaldi and K. Lenc, "Matconvnet: Convolutional neural networks for matlab," in *Proc. ACM Int. Conf. Multimedia*, 2014, pp. 689–692.

[71] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, 1992.

[72] S. Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, and L. Zhang, "Weighted nuclear norm minimization and its applications to low level vision," *Int. J. Comput. Vis.*, vol. 121, no. 2, pp. 183–208, Jan. 2017.

[73] J.-F. Cai, R. H. Chan, and Z. Shen, "A framelet-based image inpainting algorithm," *Appl. Comput. Harmon. Anal.*, vol. 24, no. 2, pp. 131–149, Mar. 2008.

[74] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[75] P. Escande, P. Weiss, and W. Zhang, "A variational model for multiplicative structured noise removal," *J. Math. Imag. Vis.*, vol. 57, no. 1, pp. 43–55, 2017.

[76] G.-S. Xia *et al.*, "DOTA: A large-scale dataset for object detection in aerial images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3974–3983.