Double Low-Rank Matrix Decomposition for Hyperspectral Image Denoising and Destriping

Hongyan Zhang, *Senior Member, IEEE*, Jingyi Cai, Wei He, *Member, IEEE*, Huanfeng Shen, *Senior Member, IEEE*, and Liangpei Zhang, *Fellow, IEEE*

Abstract—Hyperspectral images (HSIs) have a wealth of applications in many areas, due to their fine spectral discrimination ability. However, in the practical imaging process, HSIs are often degraded by a mixture of various types of noise, for example, Gaussian noise, impulse noise, dead pixels, dead lines, and stripe noise. Low-rank matrix decomposition theory has been widely used in HSI denoising, and has achieved competitive results by modeling the impulse noise, dead pixels, dead lines, and stripe noise as sparse components. However, the existing low-rank-based methods for HSI denoising cannot completely remove stripe noise when the stripe noise is no longer sparse. In this article, we extend the HSI observation model and propose a double low-rank (DLR) matrix decomposition method for HSI denoising and destriping. By simultaneously exploring the low-rank characteristic of the lexicographically ordered noise-free HSI and the low-rank structure of the stripe noise on each band of the HSI, the two low-rank constraints are formulated into one unified framework, to achieve separation of the noise-free HSI, stripe noise, and other mixed noise. The proposed DLR model is then solved by the augmented Lagrange multiplier (ALM) algorithm efficiently. Both simulation and real HSI data experiments were carried out to verify the superiority of the proposed DLR method.

Index Terms—Denoising, destriping, hyperspectral imagery, low-rank constraint.

I. INTRODUCTION

HYPERSPECTRAL images (HSIs) are captured by an imaging spectrometer with hundreds of continuous, narrow spectral bands and exhibit a 3-D data structure [1]–[4]. Compared with the traditional multispectral remote sensing images, the additional spectral information in HSIs benefits a large amount of remote-sensing-related applications where fine spectral discrimination is required, for example, environmental

monitoring [5], precision agriculture [6], military surveillance [7], and face recognition [8]. However, the HSIs collected in reality are inevitably contaminated by a mixture of various types of noise, for example, Gaussian noise, impulse noise, dead pixels, dead lines, and stripe noise [9], [10]. The presence of a mixture of noise can seriously degrade the HSI quality, and it limits the subsequent processing accuracy, for example, unmixing [11], [12], anomaly detection [13], [14], classification [15], [16], and spectral curve fitting [17]. Therefore, as a preprocessing step for remote sensing image applications, the denoising of HSIs is a necessary and challenging task.

Over the past decade, research into HSI denoising has achieved significant advances. Since HSIs have a unique 3-D structure, including a 2-D spatial domain and a 1-D spectral domain, denoising research can be carried out from different perspectives. First, the most simple and straightforward method is to process the 3-D HSI data cube band by band with the 2-D gray-level image denoising algorithms, but this usually results in spectral distortion due to the neglect of the high spectral correlation among adjacent bands [18]–[20]. Second, using image transformation and the corresponding inverse transformation, the transformation-based methods denoise the HSI data in a transformed domain, which can maintain the image edges and details well [21]–[23]. However, such approaches depend on the selection of the image transformation methods and the distribution of the information present in different components. In recent years, more advanced HSI denoising approaches have been presented by taking spectral information and spatial information into consideration simultaneously. For example, considering the nonlocal spatial information, Qian and Ye [24] proposed an approach for HSI denoising based on nonlocal spectral–spatial structured sparse representation, and Wei *et al.* [25] used local and nonlocal spatial similarity simultaneously to process clusters for HSI denoising. The total variation (TV) constraint is another effective way to explore the spatial information, and it has been consequently integrated with spectral information to conduct the denoising of HSIs, for example, the spatial–spectral TV model [26], the spatial–spectral adaptive TV model [27], and the anisotropic spatial–spectral TV model [28]. In addition, treating the HSI as a 3-D structure is also a direct and comprehensive way of using the spatial and spectral information, among which BM4D is a classical method [29]. Methods based on tensor

decomposition have also achieved good restoration results, including the parallel factor analysis (PARAFAC) model [30], Tucker decomposition [31], and rank-1 decomposition [32]. However, most of the methods mentioned above are built upon prior knowledge of the HSI and can only remove one or two types of noise, and not mixed noise, due to the limitations of the used prior constraints.

Due to the powerful ability to learn data features and high processing efficiency of deep neural network, some researchers applied deep learning techniques [33], [34] to HSI denoising. For example, Xie and Li [35] extended the trainable non-linear reaction diffusion (TNRD) approach to HSI denoising. However, it still denoises each band separately, and thus does not make full use of the strong spectral correlation attributes found in HSIs. Yuan *et al.* [36] proposed an HSI denoising method based on residual convolutional neural network (HSID-CNN), which takes into account both the spatial and spectral information, and does not require preadjustment of parameters for different HSIs. However, it requires training models for different levels of noise and has poor versatility. Methods such as HSI-DeNet [37] and HSI-SDeCNN [38] have also achieved good denoising results. However, the current deep-learning-based methods are relatively simple and are not applicable for more complex real noise situations.

In recent years, low-rank matrix decomposition has played a significant role in HSI mixed noise removal, and methods for HSI denoising based on low-rank matrix decomposition have achieved excellent performances [39]–[43]. For a clean HSI, there is a high correlation between the adjacent spectral bands, which exhibit a latent low-rank structure. That is to say, the 2-D matrix lexicographically rearranged from the 3-D HSI is inherently low-rank. By the use of the noise-free HSI low-rank structure, Zhang *et al.* [44] applied the robust principal component analysis (RPCA) model to HSI denoising and proposed a method based on low-rank matrix recovery (LRMR) for HSI restoration. In this approach, the observed HSI is modeled as a summation of the noise-free HSI, Gaussian noise, and sparse component, where the sparse component includes impulse noise, dead pixels, dead lines, and stripe noise. The clean signal is then separated from the mixed noise by LRMR. The LRMR algorithm shows an excellent performance in removing HSI mixed noise, but there are still some deficiencies. One of the shortcomings is that the LRMR method exploits the low-rank characteristic from the angle of the spectrum, with the spatial information not well-utilized, and sorting the 3-D data lexicographically into 2-D data results in structural detail information loss, and a poor performance under the condition of severe Gaussian noise and complex mixed noise. To solve this problem, it is necessary to integrate spatial constraints with the low-rank matrix decomposition simultaneously. For instance, He *et al.* [45] combined low-rank regularization with a spatial TV constraint, Xue *et al.* [46] integrated spatial and spectral low-rank regularizations into a unified denoising framework, and Fan *et al.* [47] modeled the clean HSI as a 3-D low-rank tensor to handle spatial–spectral information simultaneously. What is more, Zhuang and Bioucas-Dias [48] established a fast-denoising model based on low-rank regularization and sparse representation, so as to overcome the

computational burden of large matrix singular value decomposition (SVD) and patch-wise iterative computation. In addition, Chen *et al.* [49] used nonconvex approximation instead of convex approximation when solving the low-rank matrix decomposition problem, and Rasti *et al.* [50] proposed an automatic HSI noise removal approach on the basis of sparse and low-rank modeling. All these methods have further improved the denoising performance and applicability.

Although the current HSI noise removal methods based on low-rank matrix decomposition have made significant progress, there are still some problems in stripe noise removal. Unlike impulse noise and dead pixels/lines caused by the malfunctioning of a sensor, stripe noise is a kind of artifact caused by calibration errors and sensitivity variations during line-by-line detector scanning [51]. Removing stripe noise presents many challenges due to its unique process of formation. First, according to low-rank matrix decomposition theory, if the stripes are periodically distributed or occur in the same lines/columns in several adjacent bands, it is difficult to separate the stripes and image signals as the stripes will also be taken as the low-rank component [52]. Second, the basis of these methods is that the impulse noise, dead pixels, dead lines, and stripes are deemed to be sparse compared with the entire image, and the sum of these noise items can be modeled as the sparse component under the low-rank matrix decomposition framework. However, in many cases, due to the abnormal response of the sensor detectors, the stripes span most of the image and are no longer sparse. This phenomenon leads to the failure of the low-rank-based HSI noise removal methods, and the stripes are not completely removed. Therefore, how do we conduct destriping for HSIs when the stripes are no longer sparse? Actually, the stripe noise itself has unique structures, such as smoothness along the stripe direction, discontinuity across the stripe direction, and its appearance may be regular rather than strictly random. As a result, remote sensing image destriping has been widely investigated as an individual research field [28], [52]–[54]. It is thus necessary to consider the unique properties of stripe noise separately in the denoising process for HSIs.

In this article, in view of this, we propose a unified HSI denoising and destriping framework built on double low-rank (DLR) matrix decomposition, which can remove various types of noise at the same time. The main contributions of this article can be summarized as follows.

1) The HSI observation model is first extended by modeling the observed HSI as the sum of the noise-free image, the stripes, the sparse noise, and the Gaussian noise, where the sparse noise includes impulse noise, dead pixels, dead lines, and so on. Unlike the existing observation model [44], the stripe noise is treated separately, to explore the unique characteristics of HSI stripe noise.

2) By exploiting the low-rank structure of the lexicographically ordered noise-free HSI, the low-rank property of stripe noise on each band of the HSI, and the sparsity of the other noise types, a unified HSI denoising and destriping model framework is established based on DLR decomposition, which is expected to achieve

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

ZHANG *et al.*: DLR MATRIX DECOMPOSITION FOR HSI DENOISING AND DESTRIPING 3

a better separation of the clean HSI signal from the complex HSI mixed noise.
3) The augmented Lagrange multiplier (ALM) approach is applied to solve the proposed DLR method efficiently. Simulation and real-data experiments were designed to show the validity of the proposed DLR method for HSI denoising and destriping.

The rest of this article is arranged as follows. In Section II, the DLR HSI denoising and destriping model and its corresponding optimization procedure are presented in detail. The simulation and real-data experiment results are presented and analyzed in Section III. Section IV concludes this article.

II. DLR METHOD FOR HSI DENOISING AND DESTRIPING

A. Extended HSI Observation Model

General speaking, the observed HSIs are corrupted by a mixture of various noise types during the HSI imaging procedure, including Gaussian noise, impulse noise, dead pixels, dead lines, stripe noise, and so on. In [44], impulse noise, dead pixels, dead lines, and stripes were modeled as sparse noise by the assumption that these noise items only occupy a small portion of the entire image, and the HSI observation model was built as follows:

$$\mathbf{Y} = \mathbf{L} + \mathbf{S} + \mathbf{N} \tag{1}$$

where $\mathbf{Y}$ denotes the observed HSI data; $\mathbf{L}$ stands for the clean HSI data we hope to obtain; $\mathbf{S}$ is the sparse noise; and $\mathbf{N}$ represents the Gaussian noise. The components $\mathbf{Y}$, $\mathbf{L}$, $\mathbf{S}$, and $\mathbf{N}$ are all Casorati matrices (a matrix which is rearranged by lexicographically ordering all the HSI data columns) of the same size of $IJ \times K$, where I denotes the width of the HSI, J is the height of the HSI, and K stands for the number of HSI spectral bands.

To date, most HSI denoising methods are based on the HSI observation model as shown in (1), where the stripe noise is modeled as sparse noise. However, the generation and distribution of stripe noise are actually different from other types of sparse noise, including impulse noise and dead pixels/lines [51]. In practical applications, due to the influence of the different remote sensing imaging sensors, HSIs are often corrupted by various stripes with different features [54], such as horizontal/vertical stripes with random widths, periodic/nonperiodic stripes, and very wide stripes, as in the hyperspectral digital imagery collection experiment (HYDICE) urban data set and Gaofen-5 (GF-5) data sets shown in Fig. 1. The various types of stripes in GF-5 HSIs create new challenges to the research on mixed noise removal. First, the periodic stripes and the stripes existing in the same lines/columns of some adjacent bands of the HSI will be taken as the low-rank component in the low-rank matrix decomposition process. Therefore, it is very difficult to separate such stripes from the low-rank clean HSI. Second, in some cases, due to the response of the sensor detectors, the stripes are very wide or globally distributed over the entire image and are thus no longer sparse compared with the entire image. These two points mean that the stripes no
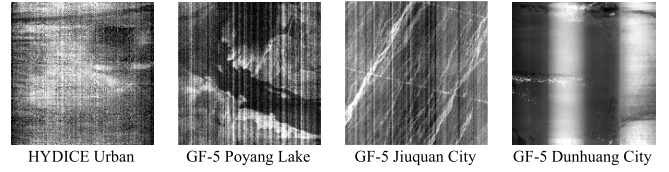

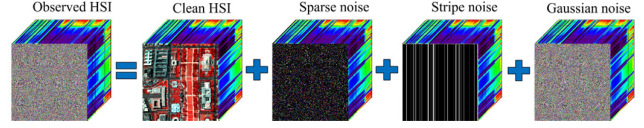Fig. 1. Real HSI data sets heavily contaminated by various types of stripe noise.

Fig. 2. Extended HSI observation model.

longer satisfy the conditions of the low-rank matrix decomposition, and thus the current HSI denoising methods fail to remove the stripes of HSIs completely. There is therefore an urgent need to extend the HSI observation model further (1).

In this article, with consideration of the unique structures of stripe noise, such as smoothness along the stripe direction and discontinuity across the stripe direction, the stripe noise is separately considered and modeled for the HSI case. Consequently, the extended HSI observation model can be written as

$$\mathbf{Y} = \mathbf{L} + \mathbf{S} + \mathbf{B} + \mathbf{N}. \tag{2}$$

Similarly, the matrices $\mathbf{Y}$, $\mathbf{L}$, $\mathbf{S}$, and $\mathbf{N}$, respectively, represent the observed HSI data, the noise-free HSI data, the sparse noise, and the Gaussian noise. Furthermore, $\mathbf{B}$ indicates the Casorati matrix of the stripe noise with the same size as $\mathbf{Y}$, $\mathbf{L}$, $\mathbf{S}$, and $\mathbf{N}$. The flowchart of the extended HSI observation model is provided in Fig. 2. The final aim of our work is to restore the clean HSI $\mathbf{L}$ from the noisy observation HSI $\mathbf{Y}$.

B. DLR Model for HSI Denoising and Destriping

Clearly, recovering the noise-free image $\mathbf{L}$ from the degraded observation image $\mathbf{Y}$ via (2) is an ill-posed problem. Here, maximum *a posteriori* (MAP) theory is used to estimate the noiseless image $\mathbf{L}$. The MAP formulation can be computed by

$$\mathbf{L}, \mathbf{S}, \mathbf{B} = \arg\max_{\mathbf{L}, \mathbf{S}, \mathbf{B}} \{ p(\mathbf{L}, \mathbf{S}, \mathbf{B} | \mathbf{Y}) \}. \tag{3}$$

From Bayes' rule, we have

$$\mathbf{L}, \mathbf{S}, \mathbf{B} = \arg\max_{\mathbf{L}, \mathbf{S}, \mathbf{B}} \left\{ \frac{p(\mathbf{Y} | \mathbf{L}, \mathbf{S}, \mathbf{B}) p(\mathbf{L}, \mathbf{S}, \mathbf{B})}{p(\mathbf{Y})} \right\}. \tag{4}$$

Since $p(\mathbf{Y})$ is a constant, and $\mathbf{L}$, $\mathbf{S}$, and $\mathbf{B}$ are also independent of each other, (4) can be simplified as

$$\mathbf{L}, \mathbf{S}, \mathbf{B} = \arg\max_{\mathbf{L}, \mathbf{S}, \mathbf{B}} \{ p(\mathbf{Y} | \mathbf{L}, \mathbf{S}, \mathbf{B}) p(\mathbf{L}) p(\mathbf{S}) p(\mathbf{B}) \} \tag{5}$$

With the Gaussian distribution of the Gaussian noise $\mathbf{N}$, the MAP formulation (5) can be further rewritten as the following model:

$$\mathbf{L}, \mathbf{S}, \mathbf{B} = \arg\min_{\mathbf{L}, \mathbf{S}, \mathbf{B}} \{ ||\mathbf{Y} - \mathbf{L} - \mathbf{S} - \mathbf{B}||_F^2 + C_1 Q_1(\mathbf{L}) \\ + C_2 Q_2(\mathbf{S}) + C_3 Q_3(\mathbf{B}) \} \tag{6}$$
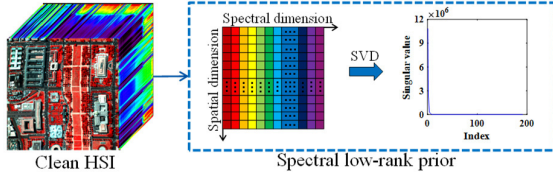
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

4

IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING

Fig. 3. Spectral low-rank prior of the clean HSI.

Fig. 4. Low-rank prior of the stripe noise for each band.

where C_1, C_2, and C_3 are the parameters, and $\| \mathbf{Y} - \mathbf{L} - \mathbf{S} - \mathbf{B} \|_F^2$ represents the data fidelity term, with $\| \cdot \|_F^2$ denoting the Frobenious norm. The central task now is to determine the formulations of the three regularization terms $Q_1(\mathbf{L})$, $Q_2(\mathbf{S})$, and $Q_3(\mathbf{B})$, respectively.

The first regularization term, $Q_1(\mathbf{L})$, is modeled as the *a priori* information of the clean HSI $\mathbf{L}$. As is well-known, from the angle of the linear spectral mixture model, the clean HSI exhibits an inherent low-rank property [44]. Specifically, the linear mixing model assumes that each spectral signature (each line of the Casorati matrix $\mathbf{L}$) can be expressed as a linear combination of all the pure spectral endmembers, indicating high correlation between the different spectral signatures [11]. For the clean HSI, the linear spectral mixture model can be represented as the following equation:

$$\mathbf{L} = \mathbf{MA} \tag{7}$$

where $\mathbf{M} \in \mathbb{R}^{IJ \times r_1}$ represents the endmember matrix, $\mathbf{A} \in \mathbb{R}^{r_1 \times K}$ stands for the abundance matrix, and r_1 denotes the number of pure spectral endmembers of the HSI.

Since the number of pure spectral endmembers r_1 is usually much less than the number of spectral bands and the spatial size of the HSI, this illustrates the HSI $\mathbf{L}$ low-rank structure

$$\text{rank}(\mathbf{L}) \leq r_1 \tag{8}$$

where $\text{rank}(\cdot)$ is the number of nonzero singular values of a matrix. Therefore, by the use of the low-rank structure of the HSI $\mathbf{L}$, as shown in Fig. 3, the regularization term $Q_1(\mathbf{L})$ can be provided as the following holds:

$$Q_1(\mathbf{L}) = \text{rank}(\mathbf{L}). \tag{9}$$

The second regularization term, $Q_2(\mathbf{S})$, models the sparsity prior of the sparse noise. The sparse noise includes impulse noise, dead pixels/lines, and so on; this noise however, can be regarded as relatively small with respect to the entire image. Therefore, the third term can be expressed as follows:

$$Q_2(\mathbf{S}) = \| \mathbf{S} \|_0 \tag{10}$$

where $\| \cdot \|_0$ represents the ℓ_0-norm, which defines the number of nonzero elements in the matrix.

The third regularization term, $Q_3(\mathbf{B})$, is related to the prior knowledge of the stripe noise. Chang *et al.* [52] found, through quantitative analysis, that the stripes in gray-level remote sensing images conform to a strict low-rank constraint. That is to say, in HSIs, the stripe noise matrix of each band can be well-expressed by the low-rank regularization

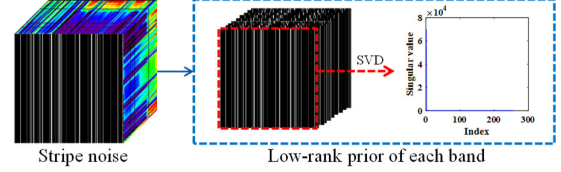$$\text{rank}(\mathbf{B}_n) \leq r_2, \quad n = 1, 2, \ldots, K \tag{11}$$

where $\mathbf{B}_n \in \mathbb{R}^{I \times J}$ represents the stripe noise matrix on the nth band of the HSI, and r_2 denotes the upper bound rank of $\mathbf{B}_n$.

Similarly, considering the low-rank constraints of stripe noise for each band, as presented in Fig. 4, the fourth regularization term $Q_3(\mathbf{B})$ can be denoted by

$$Q_3(\mathbf{B}) = \sum_{n=1}^{K} \text{rank}(\mathbf{B}_n). \tag{12}$$

By combining (9), (10), and (12), the restoration model (6) can be rewritten as

$$\mathbf{L}, \mathbf{S}, \mathbf{B} = \arg\min_{\mathbf{L}, \mathbf{S}, \mathbf{B}} \Big\{ \| \mathbf{Y} - \mathbf{L} - \mathbf{S} - \mathbf{B} \|_F^2 + C_1 \text{rank}(\mathbf{L})$$
$$+ C_2 \| \mathbf{S} \|_0 + C_3 \sum_{n=1}^{K} \text{rank}(\mathbf{B}_n) \Big\}. \tag{13}$$

In the proposed DLR model, two low-rank regularization terms are used simultaneously on the clean HSI and the stripe noise. The only difference is that the Casorati matrix for the clean HSI is low-rank, and the stripe noise matrix on each band is low-rank. For the clean HSI, even if the adjacent pixels in the local space of the image are similar, the information in the row or column is not sufficiently redundant due to the complexity of the image content, so it is not low-rank. In addition, stripe noise in each HSI band is repeatedly distributed on rows or columns due to how this noise is generated, with high information redundancy so it has a low-rank structure. The distribution pattern of stripe noise on most bands is usually different, due to the independence of the charge-coupled device elements. Therefore, if the stripe noise is lexicographically sorted and formed into a Casorati matrix, the original ordered distribution is reconfigured into an unordered distribution, no longer following the low-rank property. That is, the Casorati matrix of the noiseless HSI is low-rank, the stripe noise matrix on each band is low-rank, and the objects of the two low-rank constraints in the DLR model are different. Therefore, the two low-rank constraints in model (13) can separately induce the separation of noise-free signals and stripe noise without interfering with each other, which will be further demonstrated by our experiments.

Clearly, the optimization problem for the ℓ_0-norm and rank of the matrix is NP-hard, so it is generally relaxed to a ℓ_1-norm minimization problem and nuclear norm [55], respectively. Therefore, the proposed DLR denoising and destriping model can be derived as the equivalent formulation of (14):

$$\min_{\mathbf{L}, \mathbf{B}, \mathbf{S}} \| \mathbf{L} \|_* + \lambda_1 \| \mathbf{S} \|_1 + \lambda_2 \sum_{n=1}^{K} \| \mathbf{B}_n \|_*$$

$$\text{s.t. } \| \mathbf{Y} - \mathbf{L} - \mathbf{S} - \mathbf{B} \|_F^2 \leq \varepsilon, \ \text{rank}(\mathbf{L}) \leq r_1, \ \text{rank}(\mathbf{B}_n) \leq r_2 \tag{14}$$

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

ZHANG *et al.*: DLR MATRIX DECOMPOSITION FOR HSI DENOISING AND DESTRIPING 5

where $\lambda_1 = C_2/C_1$ and $\lambda_2 = C_3/C_1$ are the regularization parameters, ε represents the Gaussian noise variance, $\| \cdot \|_*$ denotes the nuclear norm which is defined as the sum of the singular values of a matrix, and $\| \cdot \|_1$ is the ℓ_1-norm which is the sum of the absolute values of all the elements in the matrix. Thus, in model (14), the existence of the upper bound ranks r_1 and r_2 imposes stricter rank constraints on the matrices $\mathbf{L}$ and $\mathbf{B}_n$, overcoming the problem of the nuclear norm only being an approximation of the real rank regularization, thereby further removing the noise [45]. The ALM method is used to solve the proposed DLR denoising and destriping model efficiently. Full details are given in Section II-C.

C. Optimization Procedure

The proposed DLR model (14) can be efficiently solved by the ALM optimization procedure, which minimizes the following augmented Lagrangian function:

$$\min \ell(\mathbf{L}, \mathbf{S}, \mathbf{B}, \Lambda) = \min_{\mathbf{L}, \mathbf{S}, \mathbf{B}} \| \mathbf{L} \|_* + \lambda_1 \| \mathbf{S} \|_1 + \lambda_2 \sum_{n=1}^{K} \| \mathbf{B}_n \|_*$$
$$+ <\Lambda, \mathbf{Y} - \mathbf{L} - \mathbf{S} - \mathbf{B}> + \frac{\mu}{2} \| \mathbf{Y} - \mathbf{L} - \mathbf{S} - \mathbf{B} \|_F^2$$
$$\text{s.t. } \text{rank}(\mathbf{L}) \leq r_1, \quad \text{rank}(\mathbf{B}_n) \leq r_2 \quad (15)$$

where μ stands for the penalty parameter, Λ denotes the Lagrange multiplier, and $< \cdot, \cdot >$ represents the inner product on two matrices.

A typical approach used to tackle this problem is to optimize the function (15) on one variable iteratively, while keeping the other variables fixed. Individually, in the $k + 1$th iteration, the variables are updated as follows:

$$\mathbf{L}^{k+1} = \arg \min_{\text{rank}(\mathbf{L}) \leq r_1} \ell(\mathbf{L}, \mathbf{S}^k, \mathbf{B}^k, \Lambda^k) \quad (16a)$$
$$\mathbf{S}^{k+1} = \arg \min_{\mathbf{S}} \ell(\mathbf{L}^{k+1}, \mathbf{S}, \mathbf{B}^k, \Lambda^k) \quad (16b)$$
$$\mathbf{B}^{k+1} = \arg \min_{\text{rank}(\mathbf{B}_n) \leq r_2} \ell(\mathbf{L}^{k+1}, \mathbf{S}^{k+1}, \mathbf{B}, \Lambda^k) \quad (16c)$$
$$\Lambda^{k+1} = \Lambda^k + \mu(\mathbf{Y} - \mathbf{L}^{k+1} - \mathbf{S}^{k+1} - \mathbf{B}^{k+1}). \quad (16d)$$

The optimization function (15) can then be divided into three primary subproblems, as presented in (16a)–(16c). For subproblem (16a), it can be deduced that

$$\mathbf{L}^{k+1} = \arg \min_{\text{rank}(\mathbf{L}) \leq r_1} \ell(\mathbf{L}, \mathbf{S}^k, \mathbf{B}^k, \Lambda^k)$$
$$= \arg \min_{\text{rank}(\mathbf{L}) \leq r_1} \| \mathbf{L} \|_* + <\Lambda^k, \mathbf{Y} - \mathbf{L} - \mathbf{S}^k - \mathbf{B}^k>$$
$$+ \frac{\mu}{2} \| \mathbf{Y} - \mathbf{L} - \mathbf{S} - \mathbf{B} \|_F^2$$
$$= \arg \min_{\text{rank}(\mathbf{L}) \leq r_1} \| \mathbf{L} \|_* + \frac{\mu}{2} \| \mathbf{L}$$
$$- (\mathbf{Y} - \mathbf{S}^k - \mathbf{B}^k + \Lambda^k/\mu) \|_F^2 . \quad (17)$$

The update process of $\mathbf{L}$ can be expressed by *Lemma 1* as follows.

Lemma 1 [56]: For a matrix $\mathbf{M} \in \mathbb{R}^{m \times n}$ of rank r, the SVD of it is defined as follows:

$$\mathbf{M} = \mathbf{U} \Sigma_r \mathbf{V}^* \quad (18)$$

where $\Sigma_r = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_i, \ldots, \sigma_r)$.

The singular value shrinkage operator of matrix $\mathbf{M}$ is

$$D_\gamma(\mathbf{M}) = \arg \min_{\text{rank}(\mathbf{E}) \leq r} \delta \| \mathbf{E} \|_* + \frac{1}{2} \| \mathbf{E} - \mathbf{M} \|_F^2 \quad (19)$$

where $D_\gamma(\mathbf{M}) = \mathbf{U} D_\gamma(\Sigma_r) \mathbf{V}^*$, $D_\gamma(\Sigma_r) = \text{diag}\{\max((\sigma_i - \gamma), 0)\}$. The optimization result of (17) can be directly obtained using *Lemma 1*:

$$\mathbf{L}^{k+1} = D_{1/\mu}(\mathbf{Y} - \mathbf{S}^k - \mathbf{B}^k + \Lambda^k/\mu). \quad (20)$$

As to the $\mathbf{S}$-related subproblem (16b), we can obtain the following formulation:

$$\mathbf{S}^{k+1} = \arg \min_{\mathbf{S}} \ell(\mathbf{L}^{k+1}, \mathbf{S}, \mathbf{B}^k, \Lambda^k)$$
$$= \arg \min_{\mathbf{S}} \lambda_1 \| \mathbf{S} \|_1 + <\Lambda^k, \mathbf{Y} - \mathbf{L}^{k+1} - \mathbf{S} - \mathbf{B}^k>$$
$$+ \frac{\mu}{2} \| \mathbf{Y} - \mathbf{L}^{k+1} - \mathbf{S} - \mathbf{B}^k \|_F^2$$
$$= \arg \min_{\mathbf{S}} \lambda_1 \| \mathbf{S} \|_1$$
$$+ \frac{\mu}{2} \| \mathbf{S} - (\mathbf{Y} - \mathbf{L}^{k+1} - \mathbf{B}^k + \Lambda^k/\mu) \|_F^2 . \quad (21)$$

This can be achieved by considering the following soft-thresholding shrinkage operator [57] to solve

$$\Re_\delta(x) = \begin{cases} x - \delta, & \text{if } x > \delta \\ x + \delta, & \text{if } x < -\delta \\ 0, & \text{otherwise.} \end{cases} \quad (22)$$

The optimization result of (21) can be represented as shown below:

$$\mathbf{S}^{k+1} = \Re_{(\lambda_1/\mu)}(\mathbf{Y} - \mathbf{L}^{k+1} - \mathbf{B}^k + \Lambda^k/\mu). \quad (23)$$

For the $\mathbf{B}$-related subproblem (16c), we can obtain the following expression:

$$\mathbf{B}^{k+1} = \arg \min_{\text{rank}(\mathbf{B}_n) \leq r_2} \ell(\mathbf{L}^{k+1}, \mathbf{S}^{k+1}, \mathbf{B}, \Lambda^k)$$
$$= \arg \min_{\text{rank}(\mathbf{B}_n) \leq r_2} \lambda_2 \sum_{n=1}^{K} \| \mathbf{B}_n \|_* + <\Lambda^k, \mathbf{Y} - \mathbf{L}^{k+1}$$
$$- \mathbf{S}^{k+1} - \mathbf{B}>$$
$$+ \frac{\mu}{2} \| \mathbf{Y} - \mathbf{L}^{k+1} - \mathbf{S}^{k+1} - \mathbf{B} \|_F^2$$
$$= \arg \min_{\text{rank}(\mathbf{B}_n) \leq r_2} \lambda_2 \sum_{n=1}^{K} \| \mathbf{B}_n \|_*$$
$$+ \frac{\mu}{2} \| \mathbf{B} - (\mathbf{Y} - \mathbf{L}^{k+1} - \mathbf{S}^{k+1} + \Lambda^k/\mu) \|_F^2 . \quad (24)$$

To solve the problem (24), we can optimize the stripe noise matrix on each band separately. We can then rewrite (24) as follows:

$$\mathbf{B}^{k+1} = \arg \min_{\text{rank}(\mathbf{B}_n) \leq r_2} \sum_{n=1}^{K} (\lambda_2 \| \mathbf{B}_n \|_*$$
$$+ \frac{\mu}{2} \| \mathbf{B}_n - (\mathbf{Y}_n - \mathbf{L}_n^{k+1} - \mathbf{S}_n^{k+1} + \Lambda_n^k/\mu) \|_F^2) \quad (25)$$

where $\mathbf{Y}_n$, $\mathbf{L}_n$, $\mathbf{S}_n$, and Λ_n represent the matrices of the nth band. Equally, the optimization equation can be solved by *Lemma 1* [the same as that of (20)]:

$$\mathbf{B}_n^{k+1} = \arg \min_{\text{rank}(\mathbf{B}_n) \leq r_2} \lambda_2 \| \mathbf{B}_n \|_*$$
$$+ \frac{\mu}{2} \| \mathbf{B}_n - (\mathbf{Y}_n - \mathbf{L}_n^{k+1} - \mathbf{S}_n^{k+1} + \Lambda_n^k/\mu) \|_F^2$$
$$= D_{\lambda_2/\mu}(\mathbf{Y}_n - \mathbf{L}_n^{k+1} - \mathbf{S}_n^{k+1} + \Lambda_n^k/\mu). \quad (26)$$

By summarizing the above presentation, we arrive at the ALM approach to solve the proposed DLR model, as shown in Algorithm 1.

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

6

IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING

The inputs of the DLR algorithm are the observed HSI $\mathbf{Y}$, the upper bound ranks r_1 and r_2, the maximum number of iterations $k_{\max}$, the stopping criterion ε, and the regularization parameters λ_1 and λ_2. The output is the denoised image $\mathbf{L}$. As in [45], we initialize $\mathbf{L} = \mathbf{S} = \mathbf{B} = 0$, and $\Lambda = 0$. For variable μ shown in function (16d), it is initialized as 10^{-2} and updated as $\mu := \min(\rho\mu, \mu_{\max})$ in every iteration, in which $\rho = 1.5$. This strategy for determining the penalty parameter μ can support the convergence of the algorithm and has been widely used in the methods based on ALM [57]. In addition, based on experience, we set $\varepsilon = 10^{-6}$ and $k_{\max} = 50$ in all the experiments.

Algorithm 1 DLR Algorithm

1: **Input:** the observed HSI $\mathbf{Y}$, the upper bound ranks r_1 and r_2, the maximum number of iterations $k_{\max}$, the stopping criterion ε, and the regularization parameters λ_1 and λ_2
2: **Initialize:** $\mathbf{L} = \mathbf{S} = \mathbf{B} = 0$, $\Lambda = 0$, $\rho = 1.5$, $k = 0$, $\mu = 10^{-2}$, and $\mu_{\max} = 10^{6}$
3: **While:** $\| \mathbf{Y} - \mathbf{L}^{k+1} - \mathbf{S}^{k+1} - \mathbf{B}^{k+1} \|_{\infty} > \varepsilon$ and $k \le k_{\max}$
4: Update $\mathbf{L}^{k+1}$, $\mathbf{S}^{k+1}$, $\mathbf{B}^{k+1}$, and Λ^{k+1} via (16a)–(16d)
5: Update the penalty parameter $\mu := \min(\rho\mu, \mu_{\max})$
6: Update the iteration number $k = k + 1$
7: **End While**
8: **Output:** The restored HSI $\mathbf{L}$

D. Parameter Settings

In Algorithm 1, the parameters we need to determine are the desired ranks r_1 and r_2 and the regularization parameters λ_1 and λ_2.

The desired rank r_1 denotes the subspace dimension of the HSI [58]. In our model, clean HSIs are degraded by various types of noise, especially stripe noise, which is the case with real HSIs. This means that it is difficult to obtain the dimension of the subspace directly. Here, we resort to the HSI subspace identification algorithm, HySime [59], for a preliminary estimate of the desired rank. Moreover, the higher the noise level, the more difficult it is to separate the clean signals from the noise. That is to say, the optimal value of the desired rank r_1 is smaller [60]. The above two points support our estimation of the rank r_1. For the desired rank r_2, it represents the low-rank property of the stripe component of each band [52], and we set r_2 to 1 in all the experiments.

The regularization parameters λ_1 and λ_2 are used to control the trade-offs between the clean HSI and the sparse noise, and the clean HSI and the stripe noise, respectively. In all the experiments, we set $\lambda_2 = 1.0$. In addition, we fixed the parameter λ_1 to 0.1 in the simulation experiments and to 0.05 in the real data experiments. The analysis of parameters r_1, r_2, λ_1, and λ_2 is provided in the experimental discussion.

III. EXPERIMENTAL RESULTS AND DISCUSSION

The effectiveness of the HSI mixed noise removal method based on DLR was verified by both the simulation and real HSI data experiments.
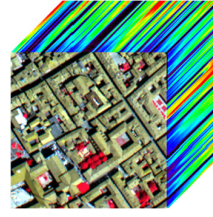

Fig. 5. Simulated data set: Pavia city center data set (R: 80, G: 34, and B: 9).

A. Experimental Setup

1) Simulated Data Experiments: The Pavia city center data set was adopted in the simulation experiments. The Pavia city center data set (available: *http://www.ehu.es/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes*) was acquired by the Reflective Optics System Imaging Spectrometer (ROSIS-03). Since the first few bands in this data set are seriously contaminated by noise, they cannot be used as references for noise removal experiments. Therefore, the first few bands of this data set were removed, and we selected a subimage size of $200 \times 200 \times 80$, which is shown in Fig. 5. Before the denoising, the gray values of the reference HSI were normalized to [0, 1] by band-by-band normalization.

To simulate a noisy image, simulated Gaussian noise, impulse noise, dead lines/pixels, and stripe noise were added to the reference HSI data set, as in the following six cases.

Case 1: Dead lines/pixels were added in band 60 to band 63. In each selected band, the number of the dead lines ranged from 3 to 5, the width of each dead line ranged from 1 to 3, and the number of the dead pixels ranged from 7 to 10. Furthermore, zero-mean Gaussian noise with the same standard deviation and the impulse noise with the same percentage were added to all bands. To demonstrate the robustness of the DLR solver under different intensities of Gaussian noise and impulse noise, the following four noise levels were set in case 1: the standard deviations of the Gaussian noise were $D = 0.001, 0.01, 0.05$, and 0.1 (the signal-to-noise ratio values were close to 50, 30, 15, and 5 dB), and the percentages of the impulse noise were $P = 0.05, 0.1, 0.15$, and 0.2, correspondingly. In addition, 40% of the bands were randomly selected to have vertical stripes added. The number of stripes was randomly selected to be between 40% and 50% of the width of the image.

Case 2: In this case, the dead lines/pixels were simulated the same as that in case 1. Besides, to more accurately simulate a real noisy HSI, zero-mean Gaussian noise and impulse noise of different intensities for each band were added to all bands [61]. Specifically, the standard deviation of the Gaussian noise was randomly selected between [0, 0.2], and the percentage of the impulse noise was randomly selected between [0, 0.2]. In addition, 20% of the bands were randomly selected to have horizontal stripes added. The number of stripes was randomly selected to be between 20% and 30% of the width of the image.

Case 3: In this case, the dead lines/pixels, Gaussian noise, and impulse noise were simulated the same way as in case 2. In addition, 40% of the bands were randomly selected to have
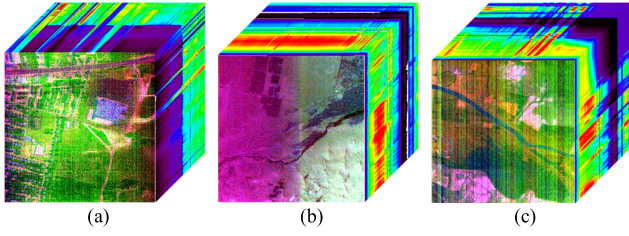
Fig. 6. Real HSI data sets. (a) HYDICE Urban data set (R: 190, G: 107, and B: 1). (b) GF-5 Dunhuang City data set (R: 194, G: 246, and B: 330). (c) GF-5 Yellow River Delta data set (R: 238, G: 305, and B: 1).

vertical stripes added. The number of stripes was randomly selected to be between 40% and 50% of the width of the image.

Case 4: In this case, the dead lines/pixels, Gaussian noise, and impulse noise were simulated the same way as in case 2. In addition, 60% of the bands were randomly selected to have vertical stripes added. The number of stripes was randomly selected to be between 60% and 70% of the width of the image.

Case 5: In this case, the dead lines/pixels, Gaussian noise, and impulse noise were simulated the same way as in case 2. In addition, 40% of the bands were randomly selected to have vertical periodic stripes added [28].

Case 6: In this case, the dead lines/pixels, Gaussian noise, and impulse noise were simulated the same way as in case 2. In addition, wide vertical stripes were added at the same position in 20 continuous bands.

2) Real HSI Data Experiments: Three real HSI data, the HYDICE Urban data set (available: *http://www.tec.army.mil/hypercube*) and the two GF-5 data sets: Dunhuang City and Yellow River Delta (available: *http://hipag.whu.edu.cn/Resource%20download.html*), were used in the real HSI data experiments. The HYDICE Urban image, as presented in Fig. 6(a), which is $307 \times 307 \times 210$ in size, is seriously corrupted by mixed noise in most bands, with especially dense stripe noise. The selected GF-5 Dunhuang City image, which is presented in Fig. 6(b), is 500×500 pixels in size and has 330 bands. The GF-5 Yellow River Delta subimage has a size of $500 \times 500 \times 305$, with some abnormal bands removed, as shown in Fig. 6(c). The two GF-5 images are both heavily contaminated by various types of stripes, including very wide stripe noise, which presents at the same position on the continuous bands, and dense stripe noise of various widths. In addition, there is a lot of mixed noise in some of the bands. Before the denoising, the gray values of real HSIs were normalized to [0, 1] by band-by-band normalization.

3) Compared Algorithms: To validate the proposed DLR method, we selected eight classical and state-of-the-art HSI denoising methods for comparison. These methods were spectral–spatial adaptive hyperspectral total variation (SSAHTV) [27], BM4D [29], LRMR [44], TV-regularized low-rank matrix factorization (LRTV) [45], local LRMR and global spatial–spectral total variation (LLRSSTV) [60], low-rank tensor decomposition with group sparse regularization (LRTDGS) [39], three-directional log-based tensor nuclear

norm (3-DLogTNN) [62], and double-factor-regularized low-rank tensor factorization (LRTFDFR) [63].

The two classical methods, SSAHTV and BM4D, were designed to remove Gaussian noise and slight sparse noise, whereas the other methods are suitable for removing various types of mixed noise. Except for BM4D, which is a parameter-free method, the experimental parameters of the competing methods were manually adjusted to the optimum according to the suggestions in the original articles. The setting details are as follows. For the SSAHTV method, we adopted the local information adaptive TV denoising mode, and to balance the data fidelity and TV constraint, we manually adjusted the regularization parameter to the optimal value of 0.6 in our experiments. For all the methods that contain low-rank constraints, the desired rank was estimated by referring to the subspace estimation technique HySime, including the rank of each submatrix of LRMR, the rank constraint of LRTV, the upper bound rank of LLRSSTV, the third-mode rank r_3 of LRTDGS, the rank of LRTFDFR, and the desired rank r_1 of the proposed DLR. In addition, in LRMR, we set the spatial size of each subcube to 20, and the step size of each subcub to 4. In LRTV, the regularization parameters were fixed as $\lambda = 10/(I \times J)^{1/2}$ and $\tau = 0.005$. For LLRSSTV, we fixed the block size as 20, the step size as 10, and the regularization parameters as $\lambda = 0.2$ and $\tau = 0.005$. For LRTDGS, the Tucker rank (r_1, r_2, r_3), r_1 and r_2, were set as 80% of the spatial size, and the parameters λ_1 and c were set to 0.1 and 500 according to degree of noise. In 3-DLogTNN, the parameters $\theta = 0.001$, $\varphi = 0.00005$, $\varpi = 0.011$, and $\omega = 10000$ and the constant ε were set to 80 for the simulated data and to 60 for the real data. For LRTFDFR, in the simulation experiment, we manually adjusted the parameters to the optimal, where regularization parameters τ, λ, and μ and the penalty parameter β were set to 0.01, 0.01, 0.04, and 15 000, respectively. In the real experiment, τ, λ, μ, and β were set to 0.2, 5, 0.04, and 15 000 as suggested in the original article. For the proposed DLR model, we fixed $\lambda_2 = 1.0$ and $r_2 = 1$ in all simulated and real data experiments; in addition, we fixed $\lambda_1 = 0.1$ in the simulation experiments, and $\lambda_1 = 0.05$ in the real data experiments, analyzed in the experimental discussion.

4) Evaluation Indices: To assess the results of the simulation experiments quantitatively, peak SNR (PSNR), structure similarity (SSIM) [64], and mean spectral angle distance (MSAD) were used. For the HSIs, the PSNR and SSIM values between every clean band and restored band were calculated, and then the mean PSNR (MPSNR) and mean SSIM (MSSIM) values were computed.

B. Simulated Data Experiments

The quantitative assessment results for the different methods in different noise cases in the Pavia city center data set are presented in Table I. The optimal values for each evaluation indicator are marked in bold, and the suboptimal results are underlined. It can be observed that the results of the proposed DLR method were optimal in almost all cases, especially when the intensity of the stripe noise was high or there was particular

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

8   IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING

TABLE I
QUANTITATIVE ASSESSMENT RESULTS OF ALL THE COMPETING METHODS

Noise case	Noise level	Index	Noisy	SSAHTV	BM4D	LRMR	LRTV	LLRSSTV	LRTDGS	3DLogTNN	LRTFDFR	DLR
Case 1	D=0.001 P=0.05 40% stripes	MPSNR	15.82	23.00	19.93	33.76	_43.61_	35.71	37.80	36.29	39.52	**46.75**
		MSSIM	0.379	0.514	0.528	0.936	_0.984_	0.965	0.961	0.756	0.982	**0.998**
		MSAD	40.82	12.03	33.21	9.46	9.45	5.51	11.01	33.18	_4.72_	**1.96**
	D=0.01 P=0.1 40% stripes	MPSNR	13.37	22.08	20.97	33.01	36.71	34.03	36.53	34.09	_38.68_	**44.88**
		MSSIM	0.227	0.503	0.527	0.934	0.948	0.953	0.955	0.772	_0.981_	**0.996**
		MSAD	44.89	13.50	28.45	10.04	17.07	7.19	12.99	33.97	_4.91_	**2.36**
	D=0.05 P=0.15 40% stripes	MPSNR	11.78	20.90	21.03	31.57	32.97	33.20	_34.53_	28.68	34.09	**34.71**
		MSSIM	0.143	0.472	0.527	0.926	0.925	0.942	_0.953_	0.726	0.951	**0.955**
		MSAD	47.06	13.55	20.71	7.73	10.97	6.91	**5.81**	32.46	_5.87_	6.36
	D=0.1 P=0.2 40% stripes	MPSNR	10.45	19.42	20.01	28.84	29.80	29.85	**30.91**	26.16	30.35	_30.89_
		MSSIM	0.092	0.422	0.478	0.872	0.864	0.890	**0.900**	0.700	0.878	_0.895_
		MSAD	48.89	15.11	18.35	9.05	12.32	9.46	_8.10_	32.11	**7.47**	8.17
Case 2	D=0~0.2 P=0~0.2 20% stripes	MPSNR	12.72	21.34	21.45	29.86	31.38	31.04	28.59	28.11	_31.39_	**32.84**
		MSSIM	0.171	0.468	0.537	0.896	0.904	_0.912_	0.881	0.738	0.902	**0.937**
		MSAD	46.54	13.65	23.96	8.92	13.09	10.01	15.43	32.28	_7.48_	**7.47**
Case 3	D=0~0.2 P=0~0.2 40% stripes	MPSNR	12.78	21.30	21.28	29.66	31.43	31.44	_32.76_	27.12	31.28	**33.17**
		MSSIM	0.177	0.467	0.524	0.893	0.905	0.923	_0.933_	0.712	0.896	**0.942**
		MSAD	46.68	14.08	23.40	8.90	12.55	8.55	**6.82**	32.00	7.30	_7.05_
Case 4	D=0~0.2 P=0~0.2 60% stripes	MPSNR	11.85	20.50	18.81	28.03	30.35	28.95	28.19	23.70	_29.82_	**32.17**
		MSSIM	0.152	0.447	0.406	0.863	0.878	0.876	0.852	0.580	_0.880_	**0.925**
		MSAD	49.51	17.76	31.61	10.80	13.28	10.94	14.80	40.16	_8.45_	**7.82**
Case 5	D=0~0.2 P=0~0.2 periodic stripes	MPSNR	12.28	21.88	19.06	_28.28_	24.73	26.31	25.44	25.65	27.16	**32.51**
		MSSIM	0.161	0.465	0.398	0.838	0.708	0.765	0.733	0.627	_0.848_	**0.941**
		MSAD	52.26	16.48	38.66	_12.75_	28.32	20.89	26.15	45.56	17.11	**7.84**
Case 6	D=0~0.2 P=0~0.2 wide stripes	MPSNR	13.47	20.59	23.44	27.94	28.22	29.17	29.15	_29.66_	27.89	**31.27**
		MSSIM	0.184	0.408	0.637	0.818	0.820	0.836	0.833	_0.849_	0.811	**0.919**
		MSAD	45.73	15.72	17.19	15.74	15.79	18.72	16.87	_14.87_	16.25	**9.09**

stripe noise, and the superiority of the proposed DLR method was very significant. In all the cases of different stripe noise settings, the quantitative evaluation results of SSAHTV and BM4D were inferior to those of the other methods based on low-rank matrix or tensor decomposition, as the sparse noise was not taken into consideration in these two classical methods. For those typical methods based on low-rank matrix decomposition, that is, LRMR, LRTV, and LLRSSTV, relatively strong quantitative evaluation results were obtained in some cases, but the results were still inferior to those of the proposed DLR method. Considering the three more advanced competing methods, LRTDGS, 3-DLogTNN, and LRTFDFR, high evaluation results could be obtained when the intensity of stripe noise was low. However, when the intensity of the simulated stripe noise increased, or when periodic stripe noise was present, the results of these low-rank-based methods significantly decreased, which reflects the disadvantages of removing stripe noise by modeling it as sparse noise. Overall, the noise removal performance of the proposed DLR method was the best among all the methods, indicating the superiority of exploiting the unique properties of stripe noise separately.

Fig. 7 presents the noise removal results of the comparative methods in simulated noise case 4, where the images were degraded by simulated Gaussian noise, impulse noise, and vertical stripes. The average PSNR values on the three

selected bands are displayed in brackets in the figure legends. As presented in Fig. 7, the two classical methods of SSAHTV and BM4D only removed some of the Gaussian noise and impulse noise, and they could not remove the stripes. LRMR, LRTV, LLRSSTV, LRTDGS, 3-DLogTNN, and LRTFDFR delivered better performance in Gaussian noise and impulse noise removal, but with uneven stripe noise removal performances. Among the different methods, due to the patch-based processing strategy, LRMR and LLRSSTV could partially remove stripes, but when the local noise intensity was high, the stripes could not be removed. In addition, although LRTV could remove some mixed noise visually, it smoothed out the details in the images. Even though 3-DLogTNN could effectively remove Gaussian noise and impulse noise, it was not able to remove dense stripes. LRTDGS and LRTFDFR could remove most of the noise; however, the results still contained obvious residual stripes, and LRTDGS caused significant spectral distortion. As can be seen from these results, the proposed DLR method could perform HSI denoising and destriping effectively while preserving detailed information.

To compare the spectral recovery capabilities of each method, Fig. 8 shows the original spectrum (red curve) and the denoising spectrums (blue curve) of the pixel (193.68) in case 4 with the Pavia city center data set. It can be seen from Fig. 8(b) that mixed noise destroyed the spectral signature of

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

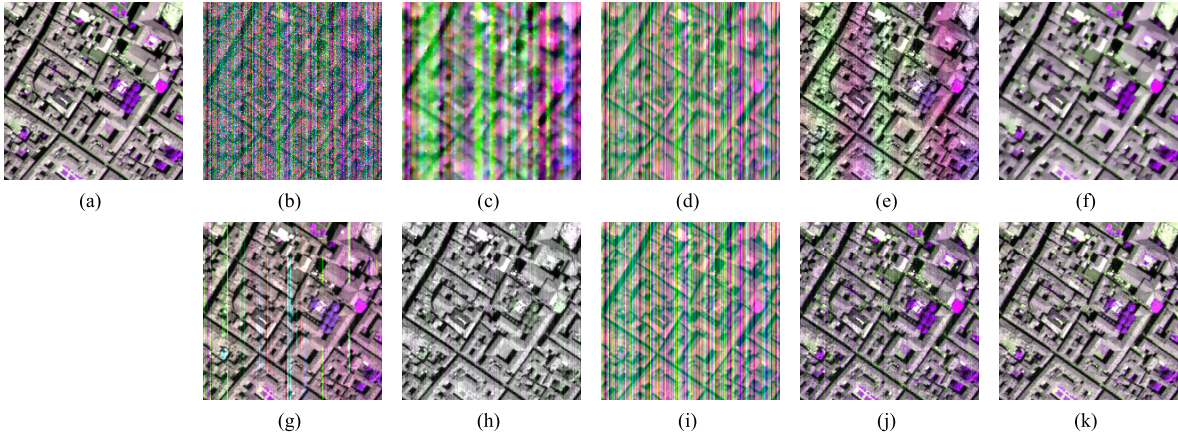ZHANG *et al.*: DLR MATRIX DECOMPOSITION FOR HSI DENOISING AND DESTRIPING

9

Fig. 7. Simulation experiment denoising false-color results in case 4. (a) Original image (R: 53, G: 26 B: 66). (b) Noisy image (11.38 dB). (c) SSAHTV (19.66 dB). (d) BM4D (16.81 dB). (e) LRMR (26.48 dB). (f) LRTV (30.39 dB). (g) LLRSSTV (27.51 dB). (h) LRTDGS (28.36 dB). (i) 3-DLogTNN (19.80 dB). (j) LRTFDFR (27.86 dB). (k) DLR (34.84 dB).
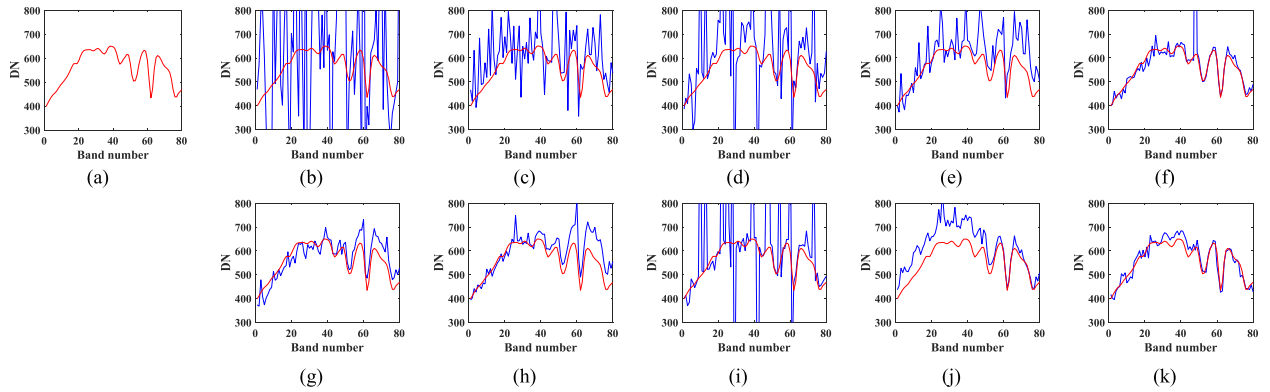

Fig. 8. Original spectrum (red curve) and the denoising results (blue curve) of pixel (193.68) in simulated case 4. (a) Original. (b) Noisy. (c) SSAHTV. (d) BM4D. (e) LRMR. (f) LRTV. (g) LLRSSTV. (h) LRTDGS. (i) 3-DLogTNN. (j) LRTFDFR. (k) DLR.
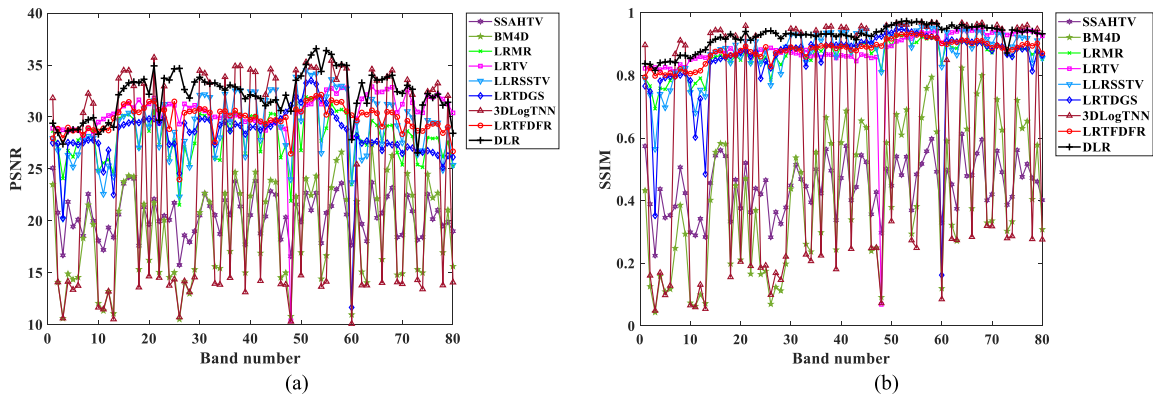

Fig. 9. (a) PSNR and (b) SSIM values on each band in simulated case 4.

the pixel, while SSAHTV, BM4D, LRMR, and 3-DLogTNN failed to restore the original spectral curve. LRTV, LLRSSTV, LRTDGS, and LRTFDFR could recover most of the spectrum, but there were still sharp noises in some bands. As shown in Fig. 8(k), the proposed DLR method could recover the spectrum better than other methods. The PSNR and SSIM values on each band in case 4 of the simulation experiments were also compared in Fig. 9. Note that 3-DLogTNN achieved high quantitative evaluation values in bands without simulated stripes, but low evaluation values in bands with stripes.

The figures show that the proposed DLR method obtained the highest PSNR and SSIM values in most of the bands, which indicated the superiority of the DLR method further.

For simulated case 5, Fig. 10 shows the denoising false-color performances of the comparison methods in the Pavia city center data set. As presented in Fig. 10(b), the original image was degraded by simulated Gaussian noise, impulse noise, and random, periodic stripes. SSAHTV, BM4D, LRTV, LRTDGS, and 3-DLogTNN could not remove the periodic stripe noise. LRMR and LLRSSTV could only remove part of

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

10   IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING
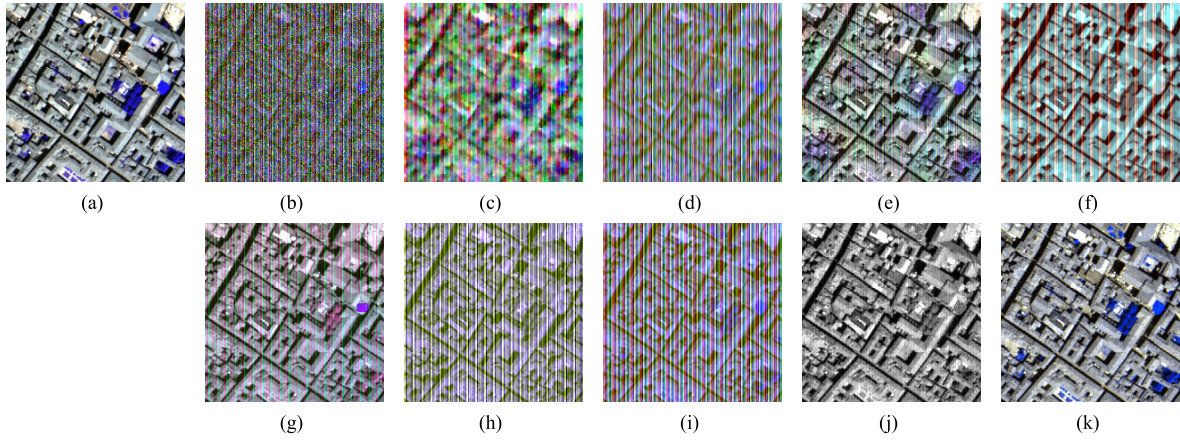
Fig. 10.   Simulation experiment denoising false-color results in case 5. (a) Original image (R: 21, G: 33, B: 55). (b) noisy image (10.01 dB). (c) SSAHTV (21.89 dB). (d) BM4D (15.06 dB). (e) LRMR (27.99 dB). (f) LRTV (24.95 dB). (g) LLRSSTV (26.18 dB). (h) LRTDGS (21.81 dB). (i) 3-DLogTNN (19.30 dB). (j) LRTFDFR (28.33 dB). (k) DLR (34.91 dB).
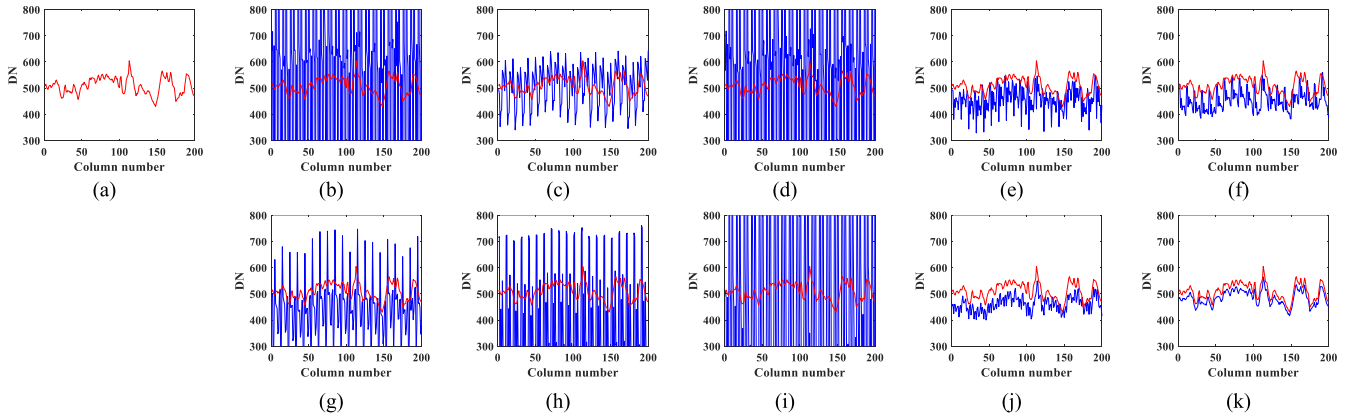

Fig. 11.   Mean vertical profiles of band 33 in simulated case 5. (a) Original profile. (b) Noisy profile. (c) SSAHTV. (d) BM4D. (e) LRMR. (f) LRTV. (g) LLRSSTV. (h) LRTDGS. (i) 3-DLogTNN. (j) LRTFDFR. (k) DLR.

the periodic stripes. Similarly, the TV-based methods also led to a loss of image detail information. From these figures, the proposed DLR method could remove mixed noise effectively, while the image details were well-preserved. The mean vertical profiles on band 33 of the simulation data set are also presented in Fig. 11 to demonstrate the destriping performance of the proposed method further. As presented in Fig. 11, the X-axis represents the number of columns, and the Y-axis is the mean digital number (DN) value of each column. From Fig. 11(a) and (b), it can be seen that at the column where the random, periodic stripe noise existed, the mean digital value changed drastically, making the original smooth mean vertical profile fluctuate abnormally. After denoising, as shown in Fig. 11(c)–(k), only the proposed DLR method could restore the original curve effectively.

Fig. 12 represents the false-color results of the different denoising methods in simulated noise case 6 with the simulation data set. The original images were contaminated by simulated mixed noise, including Gaussian noise, impulse noise, and very wide stripe noise. As can be clearly observed from Fig. 12, none of the methods could remove the wide stripe noise, except for the proposed DLR method. Some of these benchmark methods even distorted the images in the presence of stripe noise, to some extent. Fig. 12(k) is very

similar to Fig. 12(a), further showing the effectiveness of the proposed method in denoising and destriping.

C. Real HSI Data Experiments

In the following, the denoising and destriping results of the experiments performed with the three real data sets are presented.

1) HYDICE Urban Data Set: Figs. 13 and 14 show the images before and after denoising of band 107 and band 150 in the HYDICE Urban data set using the different tested methods and also show partial enlarged details in the red boxes below the images. The original images were corrupted by dense stripes and severe mixed noise, including Gaussian noise and impulse noise. As can be observed, SSAHTV and BM4D could only remove small amounts of mixed noise. Moreover, the dense noise at the edges of the images and the wide stripe noise could not be removed, and those methods even led to severe distortion of the image. LRMR and LLRSSTV were relatively successful at removing partial mixed noise, but there were still residual noises in some concentrated regions. 3-DLogTNN could hardly remove any stripes, even if it did an effective job of removing other noise. LRTV and LRTDGS were almost completely noise-free visually, but both

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

ZHANG *et al.*: DLR MATRIX DECOMPOSITION FOR HSI DENOISING AND DESTRIPING   11
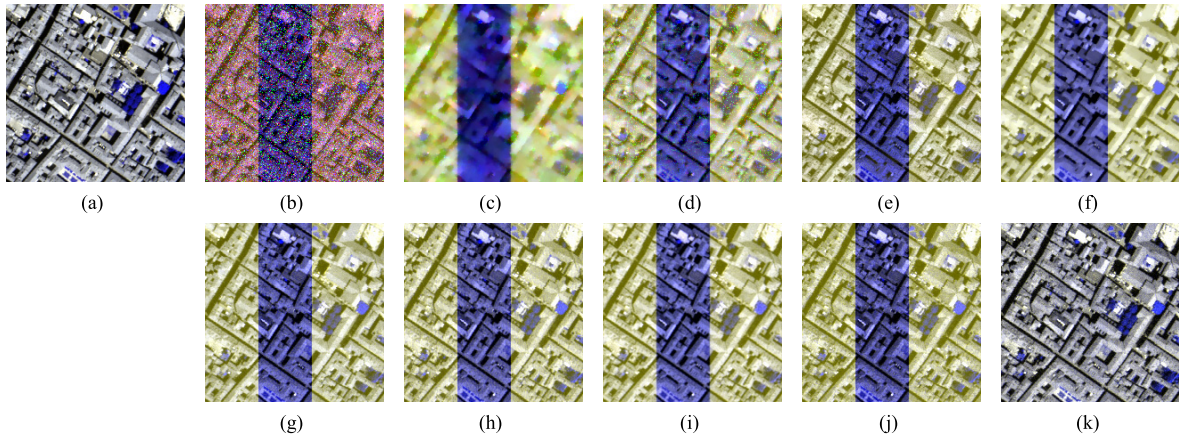

Fig. 12.   Simulation experiment denoising false-color results in case 6. (a) Original image (R: 27, G: 31, B: 53). (b) noisy image (12.32 dB). (c) SSAHTV (19.09 dB). (d) BM4D (20.38 dB). (e) LRMR (23.46 dB). (f) LRTV (22.94 dB). (g) LLRSSTV (24.02 dB). (h) LRTDGS (23.81 dB). (i) 3-DLogTNN (24.01 dB). (j) LRTFDFR (22.75 dB). (k) DLR (34.73 dB).
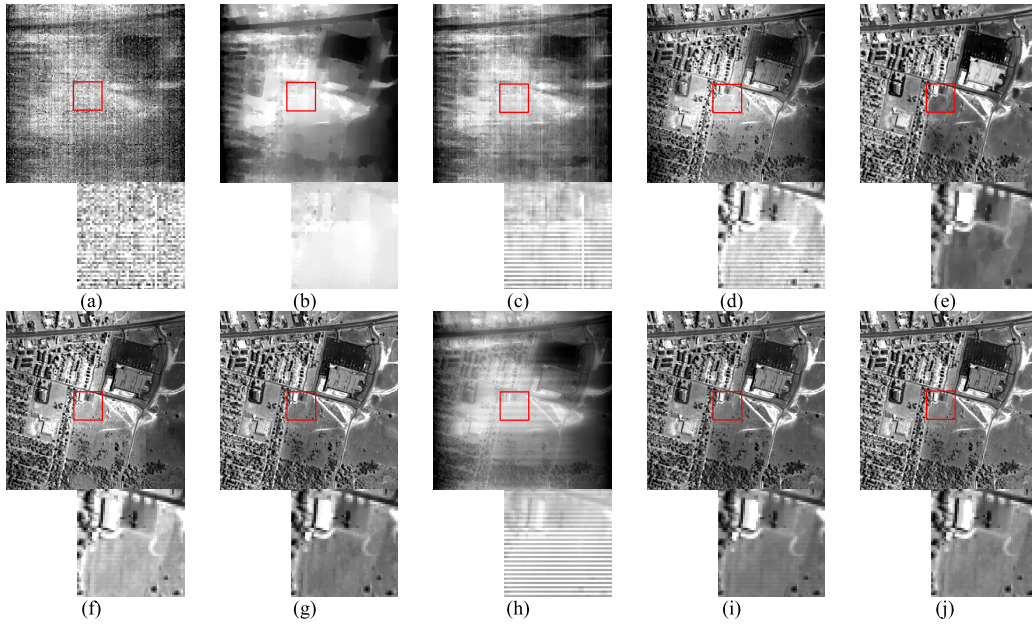

Fig. 13.   Denoising results for band 107 of the HYDICE Urban data set via the different methods. (a) Original image. The denoised images of (b) SSAHTV, (c) BM4D, (d) LRMR, (e) LRTV, (f) LLRSSTV, (g) LRTDGS, (h) 3-DLogTNN, (i) LRTFDFR, and (j) DLR.

caused spectral distortion and loss of image details. LRTFDFR achieved superior noise removal performance, but in some places, such as roofs, paths, and the enlarged area of the red box, residual noise was still present. Overall, the proposed DLR method removed most of the mixed noise while retaining details, and it achieved the best results among the different denoising methods. Fig. 15 shows the original spectrum (red curve) and denoising results (blue curve) of the pixel (23, 214). It can be seen that only the DLR method removed noise well and recovered the clean spectral curve, further verifying the practicality of the DLR method for use with real data sets.

2) GF-5 Dunhuang City Data Set: As presented in Fig. 16(a), the GF-5 Dunhuang City data set was contaminated by distinctive mixed noise, that is, very wide, periodic stripe noise at the same position on some adjacent bands. This particular stripe noise case posed a considerable challenge for the HSI denoising task. In the results of the experiments shown

in Fig. 16(b)–(i), the classical and state-of-the-art methods showed poor noise reduction effects. SSAHTV could not remove such wide stripe noise and also made the image details excessively smooth. It can also be observed that BM4D, LRMR, LLRSSTV, and 3-DLogTNN could not remove the wide stripe noise on the left side. The LRTV, LRTDGS, and LRTFDFR methods, however, showed better noise removal capabilities, but there were still some noises remaining. In addition, the denoising abilities of LRTV, LRTDGS, and LRTFDFR were partly due to their excessive smoothing of the global space, which caused inevitable distortion of the image and loss of some detail information. The denoising performance of the different methods is further illustrated by the false-color results, as shown in Fig. 17. It can be further observed in Fig. 17(g) that the global smoothing of the image by LRTV, LRTDGS, and LRTFDFR resulted in noise extending from the left half of the image to the

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

12

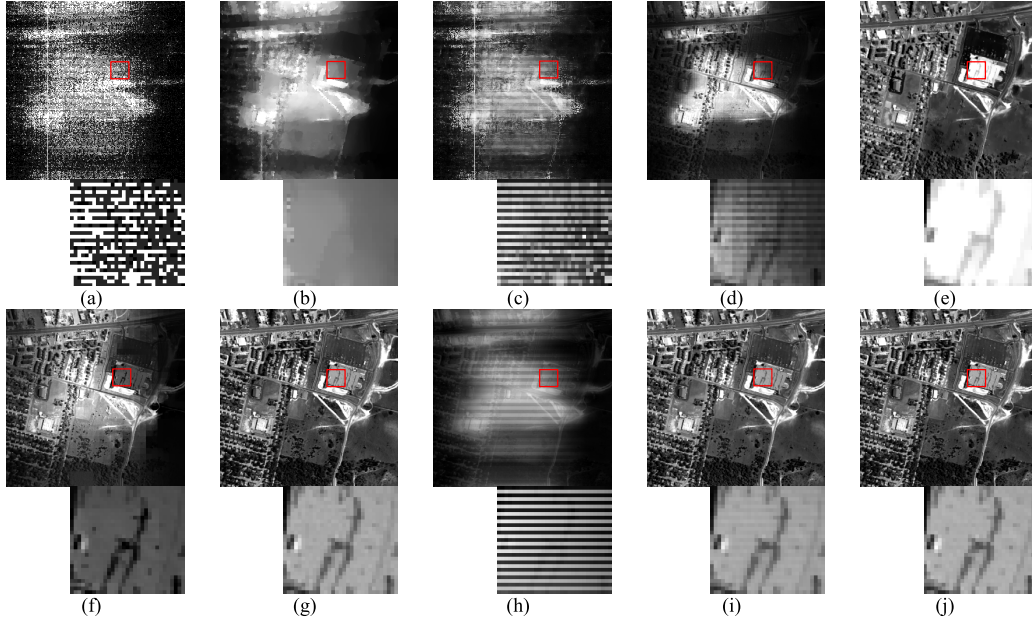IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING

Fig. 14. Denoising results for band 150 of the HYDICE Urban data set via the different methods. (a) Original image. The denoised images of (b) SSAHTV, (c) BM4D, (d) LRMR, (e) LRTV, (f) LLRSSTV, (g) LRTDGS, (h) 3-DLogTNN, (i) LRTFDFR, and (j) DLR.
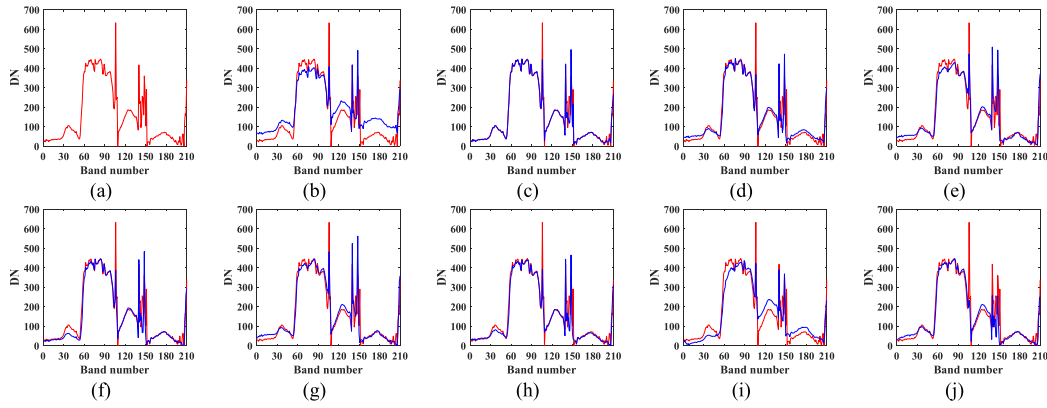

Fig. 15. Original spectrum (red curve) and denoising results (blue curve) of pixel (23.214) with the HYDICE Urban data set. (a) Original. (b) SSAHTV. (c) BM4D. (d) LRMR. (e) LRTV. (f) LLRSSTV. (g) LRTDGS. (h) 3-DLogTNN. (i) LRTFDFR. (j) DLR.

right. To sum up, the proposed DLR method, as shown in Figs. 16(j) and 17(j), performed the best of all the methods while preserving the details of the image, further illustrating the superiority of the DLR method in removing stripe noise.

3) GF-5 Yellow River Delta Data Set: Figs. 18 and 19 show the denoising images of the GF-5 Yellow River Delta data set obtained via the different methods. The images were corrupted by dense stripe noise of various widths, severe Gaussian noise, and impulse noise. The three methods, SSAHTV, BM4D, and 3-DLogTNN, could not remove this kind of mixed noise, and SSAHTV also caused significant distortion of the images. In addition, in the LRMR and LLRSSTV results, stripe noise still remained after denoising. Some methods could more or less remove some noise, such as LRTV, LRTDGS, and LRTFDFR, but some of the details in the images were also lost. It can be seen from Figs. 18(j) and 19(j) that the DLR method obtained the best noise reduction performance when dealing with stripe noise compared with the other benchmark methods.

D. Discussion

1) Parameter Analysis: In the proposed DLR method, the selection of the regularization parameters λ_1, λ_2 and ranks r_1, r_2 determines the HSI denoising results. In the following, the impacts of these parameters on the DLR method were discussed.

The influences of the changes in parameters λ_1 and λ_2 were first analyzed based on the denoising results for the images corrupted by simulated noise in cases 1–6. Notably, the rank r_1 was estimated in two ways: the HySime method and the following observation, that is, the higher the noise level, the smaller the optimal value of the desired rank. The rank r_2 was fixed as 1. The proposed DLR method was tested with different parameter λ_1 and λ_2 values, with λ_1 and λ_2 selected from [0.05, 0.1, 0.2, 0.4, 0.6, 0.8] and [0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2], respectively.

Figs. 20 and 21 show the MPSNR values of the DLR model on the simulated data set, related to parameters λ_1 and λ_2. When the intensity of Gaussian noise and impulse

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

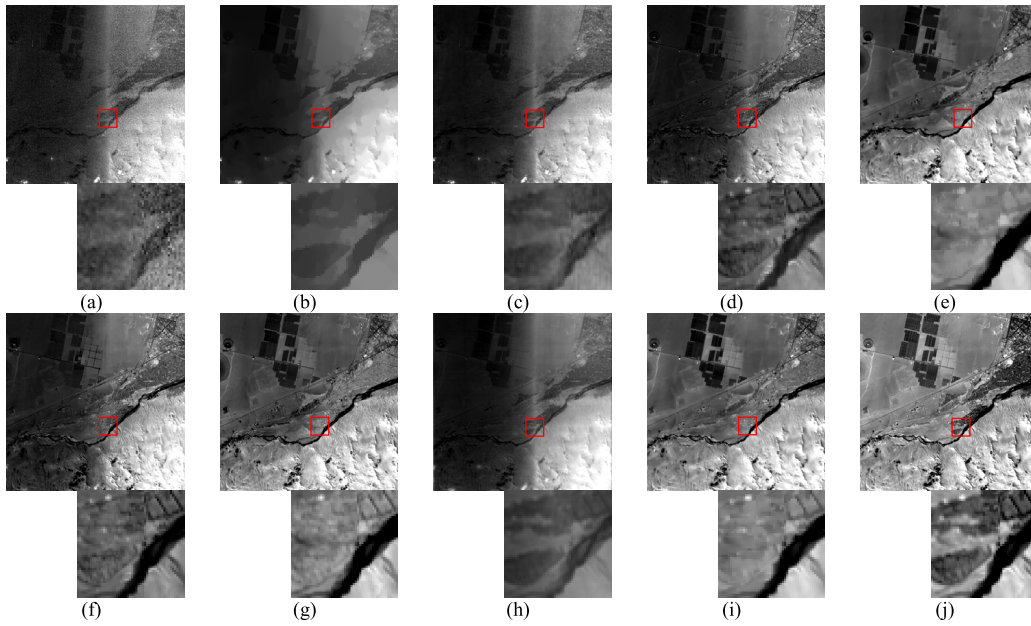ZHANG *et al.*: DLR MATRIX DECOMPOSITION FOR HSI DENOISING AND DESTRIPING   13

Fig. 16.   Denoising results for band 248 of the GF-5 Dunhuang City data set via the different methods (a) Original image. The denoised images of (b) SSAHTV, (c) BM4D, (d) LRMR, (e) LRTV, (f) LLRSSTV, (g) LRTDGS, (h) 3-DLogTNN, (i) LRTFDFR, and (j) DLR.
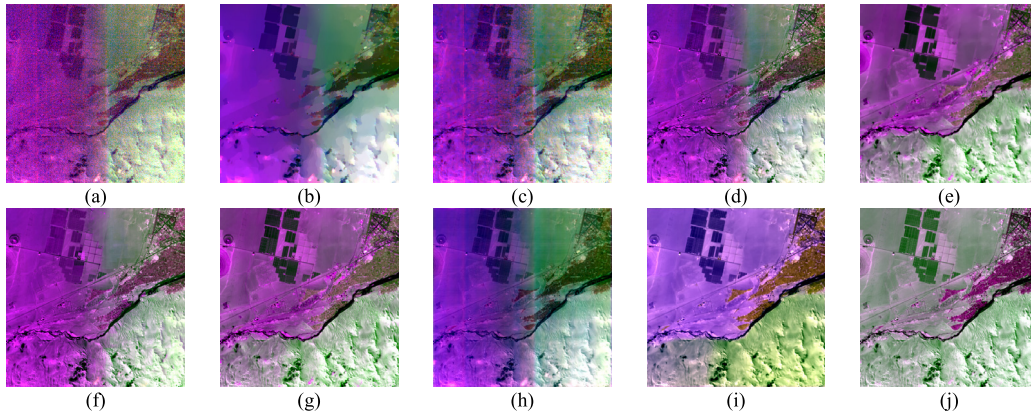

Fig. 17.   Denoising false-color results (R: 194, G: 246, B: 330) for the GF-5 Dunhuang City data set via the different methods. (a) Original image. The denoised images of (b) SSAHTV, (c) BM4D, (d) LRMR, (e) LRTV, (f) LLRSSTV, (g) LRTDGS, (h) 3-DLogTNN, (i) LRTFDFR, and (j) DLR.

noise changed in the simulated noise case 1, as shown in Fig. 20(a)–(d), the DLR method always reached optimal MPSNR values when λ_1 ranged from 0.05 to 0.6 and λ_2 ranged from 0.2 to 1.2. In cases 2–6, from Fig. 21(a)–(e), it can be seen that the denoising performance of the proposed DLR method improved only slightly when λ_2 was increased. As illustrated in Fig. 21(a)–(c), when adding random stripe noise of different intensities, a smaller value of λ_1 could bring better results. In addition, when the added stripes were special cases, as presented in Fig. 21(d)–(e), the DLR method could obtain optimal MPSNR values when λ_1 was not very low. Therefore, we fixed $\lambda_1 = 0.1$ in all the simulation experiments and $\lambda_1 = 0.05$ in the real HSI data experiments. In addition, the parameter was fixed as $\lambda_2 = 1.0$ in both the simulation and real data experiments, which ensured that the denoising effects are acceptable in all noise cases.

The rank r_1 denotes the separation boundary between the clean HSI and the mixed noise. That is to say, the first r_1 principal components of the HSI are inclined to be the clean HSI that is expected to be obtained. However, it is more difficult to distinguish the clean signals with a higher noise intensity, so only a handful of clean signals can survive among the noise, which means that the value of r_1 should be smaller [60]. Next, the effect of the upper bound rank r_1 on the DLR denoising results was investigated, and the applicability of the HySime method was tested. In the simulated noise cases 1–6, the denoising effect of the proposed DLR method was tested using different rank r_1 values for the Pavia city center data set, with r_1 varied among 2, 3, 4, 5, and 6. Similarly, the other parameters were fixed as $\lambda_1 = 0.1$ and $\lambda_2 = 1.0$, and the upper bound rank r_2 was set to 1.

Fig. 22 shows the results of the rank r_1 of the proposed DLR method when fixing the other parameters. As is shown in these
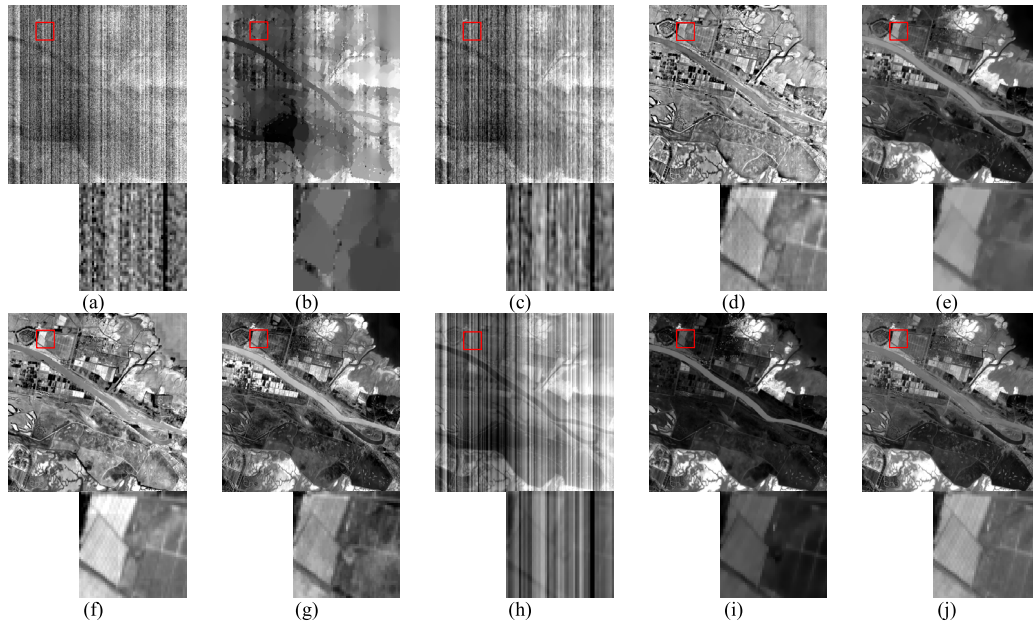
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

14   IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING

Fig. 18.   Denoising results for band 305 of the GF-5 Yellow River Delta data set via the different methods (a) Original image. The denoised images of (b) SSAHTV, (c) BM4D, (d) LRMR, (e) LRTV, (f) LLRSSTV, (g) LRTDGS, (h) 3-DLogTNN, (i) LRTFDFR, and (j) DLR.
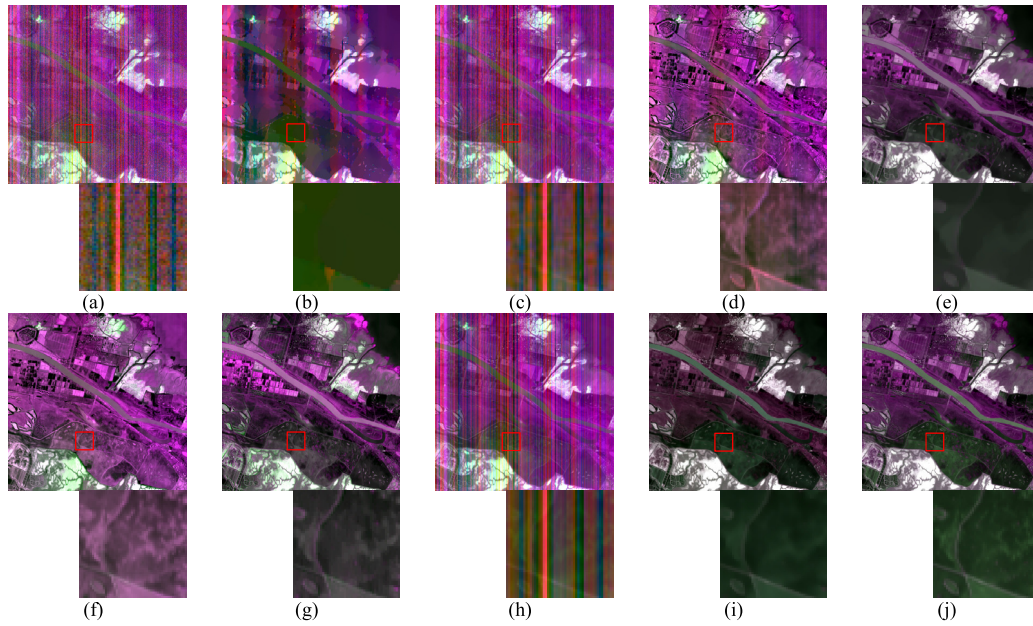

Fig. 19.   Denoising false-color results (R: 302, G: 2, B: 303) for the GF-5 Yellow River Delta data set via the different methods. (a) Original image. The denoised images of (b) SSAHTV, (c) BM4D, (d) LRMR, (e) LRTV, (f) LLRSSTV, (g) LRTDGS, (h) 3-DLogTNN, (i) LRTFDFR, and (j) DLR.
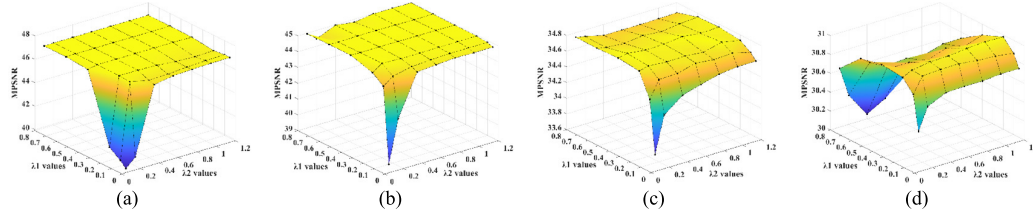

Fig. 20.   Change in the MPSNR values of DLR in simulated case 1, with respect to parameters λ_1 and λ_2. (a) $V = 0.001$, $P = 0.05$. (b) $V = 0.01$, $P = 0.1$. (c) $V = 0.05$, $P = 0.15$. (d) $V = 0.1$, $P = 0.2$.

figures, the optimal value of the rank r_1 was small and identical for each data set in all the simulated noise cases. The content in parentheses below the images represents the rank estimated by the HySime method. It can be seen that the estimated r_1 only provided a high evaluation value according to the MPSNR index in some of the simulated cases. For the presence of
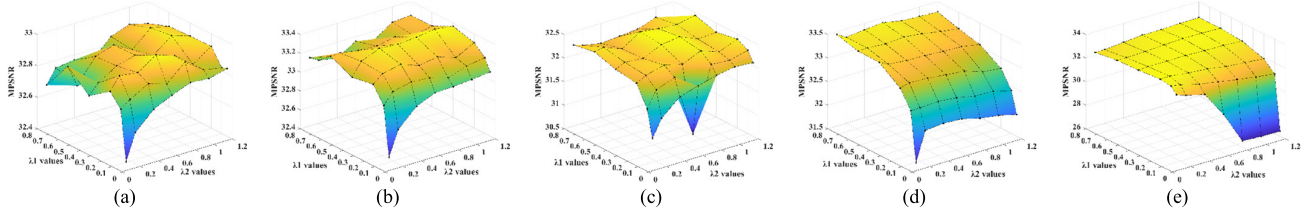
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

ZHANG *et al.*: DLR MATRIX DECOMPOSITION FOR HSI DENOISING AND DESTRIPING

15

Fig. 21. Change in the MPSNR values of DLR in the different simulated noise cases, with respect to parameters λ_1 and λ_2. (a) Case 2. (b) Case 3. (c) Case 4. (d) Case 5. (e) Case 6.
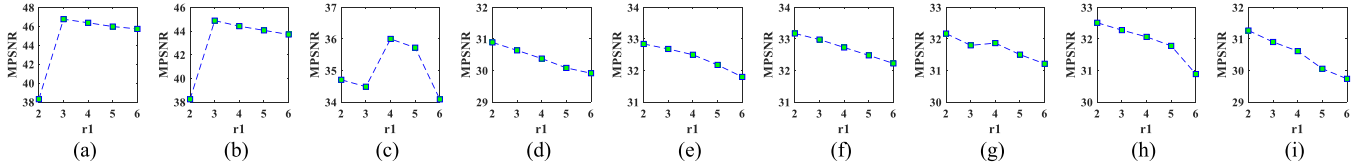

Fig. 22. Change in the MPSNR values of DLR in the different simulated noise cases, with respect to parameter r_1. Case 1: (a) $D = 0.001$, $P = 0.05$ (estimated $r_1 = 3$), (b) $D = 0.01$, $P = 0.1$ (estimated $r_1 = 3$), (c) $D = 0.05$, $P = 0.15$ (estimated $r_1 = 2$), and (d) $D = 0.1$, $P = 0.2$ (estimated $r_1 = 2$); (e) Case 2 (estimated $r_1 = 2$), (f) Case 3 (estimated $r_1 = 1$), (g) Case 4 (estimated $r_1 = 1$), (h) Case 5 (estimated $r_1 = 10$), and (i) Case 6 (estimated $r_1 = 2$).
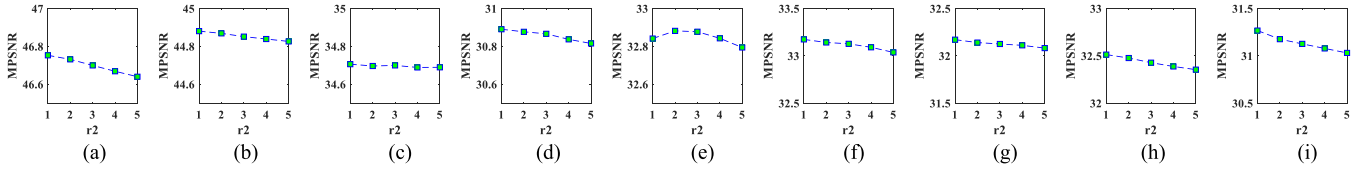

Fig. 23. Change in the MPSNR values of DLR in the different simulated noise cases, with respect to parameter r_2. (a)–(d) Case 1. (e)–(i) Cases 2–6.

periodic stripe noise, as in case 4, the upper bound rank r_1 could not be well-estimated by the HySime method. However, in the presence of stripe noise, the noise intensity was usually of the same level. As a result, the rank r_1 was usually the same. In view of this, the rank r_1 could be determined by simultaneously referring to the HySime subspace identification method and the intensity of the mixed noise.

As discussed in Section II-D, the rank r_2 of the stripe noise matrix was considered to be 1 and was verified by the simulated experiments. To make our model universal, the value of the rank r_2 was changed among 1, 2, 3, 4, and 5, to test the denoising results. Fig. 23 presents the results of the proposed DLR method with different values of rank r_2. As illustrated by these figures, an increase in r_2 value resulted in a slight deterioration in the denoising performance. To sum up, the rank r_2 was fixed as 1 in our experiments to achieve an acceptable level of denoising performance.

2) Low-Rank Analysis of Decomposed Components: In this part, we will illustrate that the two low-rank regularization terms in the DLR model, which are adopted for the Casorati matrix of HSI and the stripe noise matrix on each band, can constrain the clean HSI and the stripe noise, respectively, without interfering with each other.

Fig. 24(a)–(d) and (g)–(j) shows the decomposed components of band 42 in case 5 of the simulation experiment and band 302 in the GF-5 Yellow River Delta data set after DLR denoising. It can be seen visually that the DLR method could separate a clean image, sparse noise, and stripe noise,

which potentially showed the effectiveness of the two low-rank regularization terms in the DLR model. We also conducted SVDs on the related matrices. First, the SVDs were performed on the stripe noise and clean HSI for each band (taking 42 and 302 bands as examples). As shown in Fig. 24(e) and (k), it can be seen that the number of nonzero singular values of stripe noise (blue curve) was very small, while the nonzero singular values of the clean HSI (red curve) were large. That is, the stripe noise on each band was low rank, while the HSI on each band was not. Second, the SVDs of Casorati matrix of HSI and Casorati matrix of stripe noise were executed. As can be seen from Fig. 24(f) and (l), the singular value curve of the Casorati matrix of HSI rapidly dropped to 0, while the singular value curve of the Casorati matrix for stripe noise slowly dropped to 0, which showed that the Casorati matrix of HSI was low-rank and the Casorati matrix of stripe noise was not. Similar observations could be made in the other experiments, but due to space constraints the illustration of the decompositions was not given here.

To sum up, the objects of the two low-rank structures of the clean HSI and the stripe noise are different. That is, the two low-rank regularizations terms in the proposed DLR model do not conflict during the model optimization process and can achieve better matrix decomposition performance.

3) Run Time: Comparing the run time is also a useful way to evaluate the efficiency of different denoising methods. In the case of ensuring a good denoising result, less run time is more conducive to application processing. In the proposed DLR method, the main computational cost is the update of the three
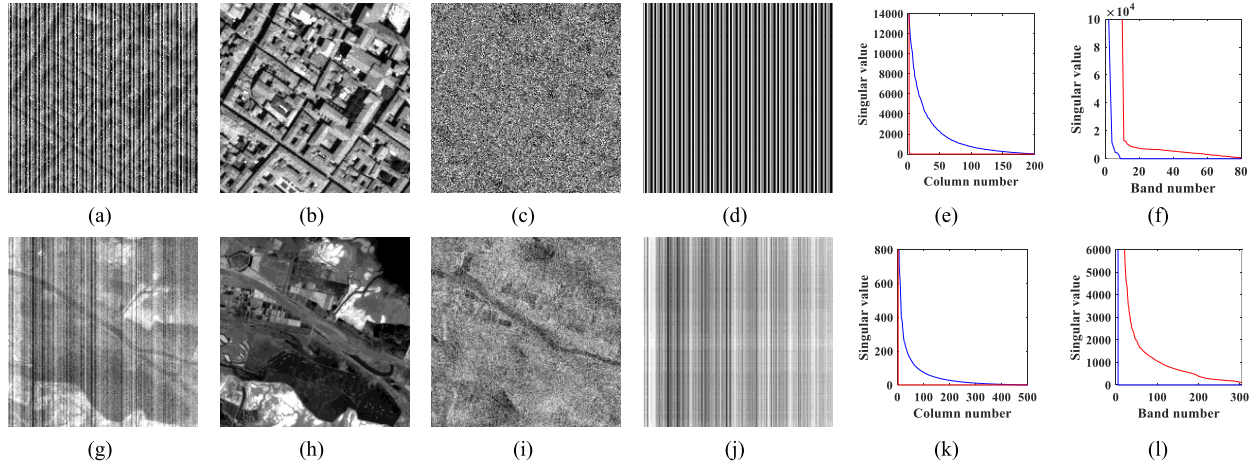
Fig. 24. Results of each component after DLR denoising and analysis of Casorati matrix and band-wise matrix. The results of band 42 in case 5 of simulation experiment. (a) Noisy. (b) Denoising. (c) Sparse term. (d) Stripe term. (e) SVDs of band-wise HSI matrix (blue) and stripe matrix (red). (f) SVDs of Casorati HSI matrix and stripe matrix. The results of band 302 in the GF-5 Yellow River Delta data set. (g) Noisy. (h) Denoising. (i) Sparse term. (j) Stripe term. (k) SVDs of band-wise HSI matrix and stripe matrix. (l) SVDs of Casorati HSI matrix and stripe matrix.
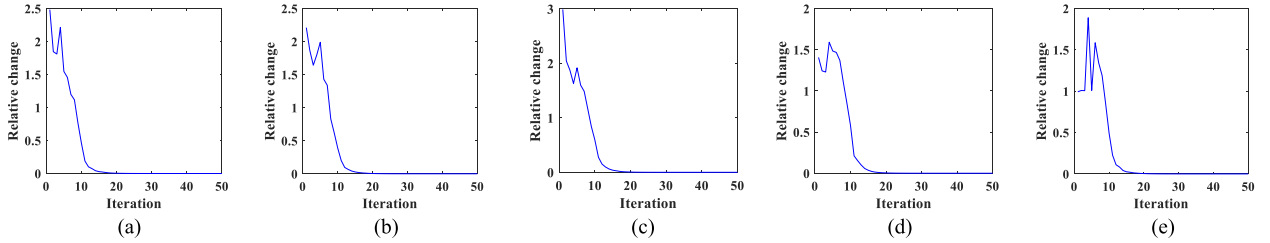

Fig. 25. Relative change values of DLR, with respect to the iteration number. (a) Case 2. (b) Case 3. (c) Case 4. (d) Case 5. (e) Case 6.

TABLE II

RUN TIME OF THE COMPETING METHODS IN THE THREE REAL HSI DATA EXPERIMENTS (IN SECONDS)

HSI Dataset	SSAHTV	BM4D	LRMR	LRTV	LLRSSTV	LRTDGS	3DLogTNN	LRTFDFR	DLR
HYDICE Urban	459	932	436	<u>378</u>	556	412	1337	523	**141**
GF-5 Dunhuang City	1538	4113	1502	1587	2150	<u>1348</u>	4308	2277	**359**
GF-5 Yellow River Delta	1608	3564	1523	1386	1897	<u>1232</u>	4046	2538	**347**

variables during each iteration, and the update calculations include the shrinkage operation and SVD. The run time of the denoising methods with the three real HSIs is listed in Table II. In Table II, we also marked the optimal values for each evaluation indicator in bold and underlined the suboptimal results. All the experiments were performed on the same server, with MATLAB 2017a, using an Intel Xeon E5-2630 CPU with 2.40 GHz and 128 GB of memory. As can be clearly observed from Table II, the proposed DLR denoising method had the least run time of all the methods. In addition, as shown in the previous experimental analysis, the denoising results of all competing methods were clearly inferior to that of the DLR method. To summarize, the proposed DLR method obtains better denoising and destriping results while consuming less run time and is thus more suitable for practical application processing than the other methods.

4) Convergence of the DLR Solver: The relative change $\| \ \mathbf{Y} - \mathbf{L}^{k+1} - \mathbf{S}^{k+1} - \mathbf{B}^{k+1} \ \|_\infty$ was used to numerically demonstrate the convergence of the DLR solver [63].

Fig. 25 presents the relative change under cases 2–6 during each iteration. It can be observed that the relative change values of the proposed method converged to 0 after the 15th iteration, which demonstrates the convergence of the DLR solver on the test data set.

5) Computational Complexity: We now analyze the computational complexity of the proposed Algorithm 1. It mainly consists of four parts as follows. The first part is updating $\mathbf{L}^{k+1} \in \mathbb{R}^{IJ \times K}$. The main computational cost of subproblem (16a) is the SVD of matrix $\mathbf{L}$ with upper bound rank r_1 which requires $O(IJKr_1)$ flops in each iteration. The second part updates $\mathbf{S}^{k+1} \in \mathbb{R}^{IJ \times K}$ using the soft-thresholding shrinkage operator, and for each iteration, $O(IJK)$ is needed. The third part updates $\mathbf{B}_n^{k+1} \in \mathbb{R}^{I \times J}, n = 1, 2, \ldots, K$ and requires the calculation of K SVDs of $I \times J$ matrices. Thus, the complexity for each iteration is $O(KIJr_2)$, where r_2 is the upper bound of the rank of the stripe noise matrix. In addition, the final part is updating the Lagrange multiplier Λ^{k+1} that requires $O(IJK)$ flops. Therefore, the overall computational

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

ZHANG *et al.*: DLR MATRIX DECOMPOSITION FOR HSI DENOISING AND DESTRIPING

17

complexity for Algorithm 1 is $O(k_{\max}(IJKr_1 + KIJr_2 + 2IJK))$, where $k_{\max}$ is the number of the iteration steps.

IV. CONCLUSION

In this article, we have proposed a DLR matrix decomposition method for HSI denoising and destriping. The extended HSI observation model is first established by introducing a new term to model the HSI stripe noise separately. Second, to promote the separation of clean HSI signals from the mixed noise further, the DLR HSI denoising and destriping model is built based on low-rank matrix decomposition theory. It simultaneously exploits the low-rank structure of the lexicographically ordered noise-free HSI, the low-rank property of the stripe noise on each band of the HSI, and the sparsity characteristics of the other sparse noise. Finally, the ALM optimization algorithm is used to solve the DLR HSI denoising and destriping model efficiently. Extensive simulation and real HSI data experiments were conducted, and the experiment results clearly showed the superior performance of the proposed DLR method, compared with the other classical and state-of-the-art denoising methods, from both visual and quantitative evaluation perspectives.

However, the proposed method still has much room for improvement. For example, the proposed method could be improved by introducing TV constraints to exploit the spatial characteristics of the stripe noise and the clean HSI. These issues will be addressed in our future work.

ACKNOWLEDGMENT

REFERENCES

[1] B. Aiazzi, L. Alparone, A. Barducci, S. Baronti, and I. Pippi, "Information-theoretic assessment of sampled hyperspectral imagers," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 7, pp. 1447–1458, Jul. 2001.

[2] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.

[3] H. Zhai, H. Zhang, L. Zhang, and P. Li, "Total variation regularized collaborative representation clustering with a locally adaptive dictionary for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 166–180, Jan. 2019.

[4] H. Zhai, H. Zhang, L. Zhang, and P. Li, "Nonlocal means regularized sketched reweighted sparse and low-rank subspace clustering for large hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, early access, Sep. 24, 2020, doi: 10.1109/TGRS.2020.3023418.

[5] A. Plaza *et al.*, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, pp. S110–S122, Sep. 2009.

[6] H. Zhai, H. Zhang, P. Li, and L. Zhang, "Hyperspectral image clustering: Current achievements and future lines," *IEEE Geosci. Remote Sens. Mag.*, early access, Jan. 18, 2020, doi: 10.1109/MGRS.2020.3032575.

[7] A. F. H. Goetz, "Three decades of hyperspectral remote sensing of the Earth: A personal view," *Remote Sens. Environ.*, vol. 113, pp. S5–S16, Sep. 2009.

[8] Z. Pan, G. Healey, M. Prasad, and B. Tromberg, "Face recognition in hyperspectral images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 12, pp. 1552–1560, Dec. 2003.

[9] 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.

[10] 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.

[11] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014–2039, Jun. 2011.

[12] L. Sun, F. Wu, C. He, T. Zhan, W. Liu, and D. Zhang, "Weighted collaborative sparse and L1/2 low-rank regularizations with superpixel segmentation for hyperspectral unmixing," *IEEE Geosci. Remote Sens. Lett.*, early access, Sep. 3, 2020, doi: 10.1109/LGRS.2020.3019427.

[13] Y. Yuan, D. Ma, and Q. Wang, "Hyperspectral anomaly detection by graph pixel selection," *IEEE Trans. Cybern.*, vol. 46, no. 12, pp. 3123–3134, Dec. 2016.

[14] H. Zhai, H. Zhang, L. Zhang, and P. Li, "Cloud/shadow detection based on spectral indices for multi/hyperspectral optical remote sensing imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 144, pp. 235–253, Oct. 2018.

[15] H. Li, G. Xiao, T. Xia, Y. Y. Tang, and L. Li, "Hyperspectral image classification using functional data analysis," *IEEE Trans. Cybern.*, vol. 44, no. 9, pp. 1544–1555, Sep. 2014.

[16] H. Zhai, H. Zhang, L. Zhang, and P. Li, "Sparsity-based clustering for large hyperspectral remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, early access, Nov. 16, 2020, doi: 10.1109/TGRS.2020.3032427.

[17] A. J. Brown, "Spectral curve fitting for automatic hyperspectral data analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1601–1608, Jun. 2006.

[18] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2, Jun. 2005, pp. 60–65.

[19] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.

[20] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[21] 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.

[22] B. Rasti, J. R. Sveinsson, M. O. Ulfarsson, and J. A. Benediktsson, "Hyperspectral image denoising using first order spectral roughness penalty in wavelet domain," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2458–2467, Jun. 2014.

[23] T. Lin and S. Bourennane, "Hyperspectral image processing by jointly filtering wavelet component tensor," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 6, pp. 3529–3541, Jun. 2013.

[24] 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.

[25] W. Wei, L. Zhang, Y. Jiao, C. Tian, C. Wang, and Y. Zhang, "Intracluster structured low-rank matrix analysis method for hyperspectral denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 866–880, Feb. 2019.

[26] 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.

[27] 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.

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

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

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

18   IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING

[31] N. Renard, S. Bourrennane, and J. Blanc-Talon, "Denoising and dimensionality reduction using multilinear tools for hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 138–142, Apr. 2008.

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

[33] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

[34] H. Zhang, Y. Liao, H. Yang, G. Yang, and L. Zhang, "A local-global dual-stream network for building extraction from very-high-resolution remote sensing images," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Dec. 16, 2020, doi: 10.1109/TNNLS.2020.3041646.

[35] W. Xie and Y. Li, "Hyperspectral imagery denoising by deep learning with trainable nonlinearity function," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 11, pp. 1963–1967, Nov. 2017.

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

[37] 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.

[38] A. Maffei, J. M. Haut, M. E. Paoletti, J. Plaza, L. Bruzzone, and A. Plaza, "A single model CNN for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2516–2529, Apr. 2020.

[39] 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.

[40] Y. Peng, J. Suo, Q. Dai, and W. Xu, "Reweighted low-rank matrix recovery and its application in image restoration," *IEEE Trans. Cybern.*, vol. 44, no. 12, pp. 1054–1066, Mar. 2014.

[41] L. Sun and B. Jeon, "A novel subspace spatial-spectral low rank learning method for hyperspectral denoising," in *Proc. IEEE Vis. Commun. Image Process. (VCIP)*, Petersburg, FL, USA, Dec. 2017, pp. 1–4.

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

[43] W. He, H. Zhang, L. Zhang, and H. Shen, "Hyperspectral image denoising via noise-adjusted iterative low-rank matrix approximation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 3050–3061, Jun. 2015.

[44] 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.

[45] 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.

[46] J. Xue, Y. Zhao, W. Liao, and S. G. Kong, "Joint spatial and spectral low-rank regularization for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 1940–1958, Apr. 2018.

[47] 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.

[48] L. Zhuang and J. M. Bioucas-Dias, "Fast hyperspectral image denoising based on low rank and sparse representations," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2016, pp. 1847–1850.

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

[50] B. Rasti, M. O. Ulfarsson, and P. Ghamisi, "Automatic hyperspectral image restoration using sparse and low-rank modeling," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2335–2339, Dec. 2017.

[51] B. Rasti, P. Scheunders, P. Ghamisi, G. Licciardi, and J. Chanussot, "Noise reduction in hyperspectral imagery: Overview and application," *Remote Sens.*, vol. 10, no. 3, p. 482, Mar. 2018.

[52] 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.

[53] 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.

[54] 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. 802–822, Feb. 2018.

[55] N. Wang, T. Yao, J. Wang, and D.-Y. Yeung, "A probabilistic approach to robust matrix factorization," in *Computer Vision—ECCV*. Berlin, Germany: Springer-Verlag, 2012, pp. 126–139.

[56] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, Mar. 2010.

[57] Z. Lin, M. Chen, and Y. Ma, "The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices," Univ. Illinois Urbana-Champaign, Champaign, IL, USA, Tech. Rep. UILU-ENG-09-2215, 2009, *arXiv:1009.5055*. [Online]. Available:

[58] H. Zhai, H. Zhang, L. Zhang, and P. Li, "Laplacian-regularized low-rank subspace clustering for hyperspectral image band selection," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 3, pp. 1723–1740, Mar. 2019.

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

[60] 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 Sen.*, vol. 11, no. 3, pp. 713–729, Mar. 2018.

[61] X. Zheng, Y. Yuan, and X. Lu, "Hyperspectral image denoising by fusing the selected related bands," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 5, pp. 2596–2609, May 2019.

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

[63] Y.-B. Zheng, T.-Z. Huang, X.-L. Zhao, Y. Chen, and W. He, "Double-factor-regularized low-rank tensor factorization for mixed noise removal in hyperspectral image," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 12, pp. 8450–8464, Dec. 2020.

[64] 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.

Hongyan Zhang (Senior Member, IEEE) received the B.S. degree in geographic information systems and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2005 and 2010, respectively.

Since 2016, he has been a Full Professor with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University. He is a Young Chang-Jiang Scholar appointed by the Ministry of Education of China. He scored First Prize in the Pairwise Semantic Stereo Challenge of the 2019 Data Fusion Contest organized by the IEEE Image Analysis and Data Fusion Technical Committee. He has authored or coauthored more than 100 research articles. His research interests include image reconstruction for quality improvement, hyperspectral information processing, and agricultural remote sensing.

Dr. Zhang was the Session Chair of several international conferences. He serves as an Associate Editor of *Photogrammetric Engineering and Remote Sensing* and *Computers and Geosciences*. He is a Reviewer for more than 30 international academic journals, including the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, and the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.

Jingyi Cai received the B.S. degree from the School of Electronic Information, Wuhan University, Wuhan, China, in 2019, where she is pursuing the M.S. degree in Surveying, Mapping and Remote Sensing (LIESMARS).

Her research interests include low-rank matrix analysis and hyperspectral image denoising and destriping.

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

ZHANG *et al.*: DLR MATRIX DECOMPOSITION FOR HSI DENOISING AND DESTRIPING

19

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

He is a Researcher with the Geoinformatics Unit, RIKEN Center for Advanced Intelligence Project, Wako,Japan. His research interests include image quality improvement, remote sensing image processing, and low-rank representation.

Huanfeng Shen (Senior Member, IEEE) received the B.S. degree in surveying and mapping engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2002 and 2007, respectively.

He is a Distinguished Professor with Wuhan University, where he serves as an Associate Dean with the School of Resource and Environmental Sciences. His research interests include remote sensing image processing, multisource data fusion, and intelligent environmental sensing. He was or is the principal investigator (PI) of two projects supported by the National Key Research and Development Program of China and six projects supported by the National Natural Science Foundation of China. He has published more than 150 peer-reviewed international journal articles, where over 60 appeared in IEEE journals, and published 4 books as a Chief Editor.

Dr. Shen is a fellow of Institution of Engineering and Technology (IET), the Education Committee Member of Chinese Society for Geodesy Photogrammetry and Cartography, and the Theory Committee Member of Chinese Society for Geospatial Information Society. He was a recipient of the First Prize in Natural Science Award of Hubei Province in 2011, the First Prize in Nature Scientific Award of China's Ministry of Education in 2015, and the First Prize in Scientific and Technological Progress Award of Chinese Society for Geodesy Photogrammetry and Cartography in 2017. He is a Senior Regional Editor of the *Journal of Applied Remote Sensing* and an Associate Editor of *Geography and Geo-Information Science* and *Journal of Remote Sensing*.

Liangpei Zhang (Fellow, IEEE) received the B.S. degree in physics from Hunan Normal University, Changsha, China, in 1982, the M.S. degree in optics from the Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, China, in 1988, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1998.

He is a "Chang-Jiang Scholar" Chair Professor appointed by the Ministry of Education of China in State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing (LIESMARS), Wuhan University. He was a Principal Scientist for the China State Key Basic Research Project from 2011 to 2016 appointed by the Ministry of National Science and Technology of China to lead the Remote Sensing Program in China. He has published more than 700 research articles and 5 books. He is the Institute for Scientific Information (ISI) Highly Cited Author. He is the holder of 30 patents. His research interests include hyperspectral remote sensing, high-resolution remote sensing, image processing, and artificial intelligence.

Dr. Zhang is a fellow of the Institution of Engineering and Technology (IET). He was a recipient of the 2010 Best Paper Boeing Award, the 2013 Best Paper ERDAS Award from the American Society of Photogrammetry and Remote Sensing (ASPRS), and the 2016 Best Paper Theoretical Innovation Award from the International Society for Optics and Photonics (SPIE). His research teams won the top three prizes of the IEEE Geoscience and Remote Sensing Society (GRSS) 2014 Data Fusion Contest, and his students have been selected as the Winners or the Finalists of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS) Student Paper Contest in recent years. He is the Founding Chair of IEEE GRSS Wuhan Chapter. He also serves as an Associate Editor or an Editor of more than ten international journals. He is serving as an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.