

# Multigraph-Based Low-Rank Tensor Approximation for Hyperspectral Image Restoration

Na Liu<sup>1</sup>, *Member, IEEE*, Wei Li<sup>2</sup>, *Senior Member, IEEE*, Ran Tao<sup>3</sup>, *Senior Member, IEEE*, Qian Du<sup>4</sup>, *Fellow, IEEE*, and Jocelyn Chanussot<sup>5</sup>, *Fellow, IEEE*

**Abstract**—Low-rank-tensor-approximation (LRTA)-based hyperspectral image and hyperspectral imagery (HSI) restoration has drawn increasing attention. However, most of the methods construct a hidden low-rank tensor by utilizing the nonlocal self-similarity (NLSS) and global spectral correlation (GSC) inherited by HSIs. Although achieving state-of-the-art (SOTA) restoration performance, NLSS and GSC have limitations. NLSS is introduced from natural image denoising to remove spatially independent identically distributed (i.i.d.) Gaussian and impulse noise, while GSC, which is naturally possessed by HSIs, is adopted to maintain the spectral integrity and remove spectrally, i.i.d., degradations. Therefore, NLSS and GSC may not be successfully used for complex HSI restoration tasks, such as destriping, cloud removal, and recovery of atmospheric absorption bands. To solve the issue, borrowing the idea from manifold learning, the geometry information characterized by proximity relationship, is integrated with the LRTA to solve the above issue, named multigraph-based LRTA (MGLRTA). Different from most of the existing methods, the proposed MGLRTA directly models an HSI as a low-rank tensor and efficiently explores the extra proximity information on the defined graphs that are not only inherited by the low-rank constraints but also naturally possessed in HSIs. A well-posed iterative algorithm is designed to solve the restoration problem. Experimental results on different datasets that cover several severe degradation scenarios demonstrate that the proposed MGLRTA outperforms the SOTA HSI restoration methods.

**Index Terms**—Graph signal processing, hyperspectral imagery, image restoration, low-rank tensor.

Manuscript received January 1, 2022; revised April 10, 2022; accepted May 22, 2022. Date of publication May 25, 2022; date of current version June 10, 2022. This work was supported in part by the National Key Research and Development Program of China under Grant 2021YFB3900502, in part by the China Postdoctoral Science Foundation under Grant 2021M700440, in part by the National Natural Science Foundation of China under Grant 61922013, and in part by the Beijing Natural Science Foundation under Grant JQ20021. (Corresponding author: Wei Li.)

Na Liu, Wei Li, and Ran Tao are with the School of Information and Electronics and the Beijing Key Laboratory of Fractional Signals and Systems, Beijing Institute of Technology, Beijing 100081, China (e-mail: liuna@ieee.org; liwei089@ieee.org; rantao@bit.edu.cn).

Qian Du is with the Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS 39762 USA (e-mail: du@ece.msstate.edu).

Jocelyn Chanussot is with Laboratoire Jean Kuntzmann (LJK), Institut National de Recherche en Informatique et en Automatique (INRIA), Centre National de la Recherche Scientifique (CNRS), Grenoble INP, University of Grenoble Alpes, 38000 Grenoble, France, and also with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China (e-mail: jocelyn.chanussot@gipsa-lab.grenoble-inp.fr).

Digital Object Identifier 10.1109/TGRS.2022.3177719

1558-0644 © 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See <https://www.ieee.org/publications/rights/index.html> for more information.

## I. INTRODUCTION

**H**YPERSPECTRAL image and hyperspectral imagery (HSI) is one of the typical high-dimensional data, consisting of hundreds of narrow spectral bands with abundant spectral information, which has superiority in discriminating various ground targets compared with traditional multispectral remote sensing imagery [1], [2]. However, due to the passive imaging mechanism for spaceborne and airborne imaging spectrometers, HSIs are exposed to different kinds of noise sources and disturbances from atmospheric perturbation, platform turbulence, and photodetectors' miscalibration [3], [4]. Therefore, the acquired HSIs are inevitably degraded. Restoring a clean HSI from the captured degraded HSI is expected to remove the noise contained in the image and improve the quality of the image effectively.

There has been a multitude of HSI restoration algorithms proposed over the past decades [3]. At the early stage, mainly filtering-based methods (e.g., Fourier and Wavelet) were introduced from conventional optical imagery denoising approaches [5]. However, noise distribution in HSIs is often too complex (not only Gaussian or impulse noise) to be captured by a simple transform. More sophisticated HSI restoration methods are developed to pay more attention to characterize HSIs' intrinsic structures but not noise distribution. To this end, the restoration problem in an HSI is mathematically formulated as an optimization, which derives an optimal solution of the restored HSI  $\mathcal{X} \in \mathbb{R}^{M \times N \times D}$  (here,  $M$  and  $N$  are spatial sizes, and  $D$  means the number of spectral bands) from an observed degraded HSI  $\mathcal{Y} \in \mathbb{R}^{M \times N \times D}$  under the specific characterization of  $\mathcal{X}$ . These characterizations are summarized into two categories: global spectral correlation (GSC) and nonlocal self-similarity (NLSS). Most of the HSI restoration methods utilize at least one of them when formulating the restoration problem [6]–[13].

GSC-based low rankness is one of the most important properties for HSIs due to its strong spectral correlation [9]. Although achieving satisfactory restoration performance, many such restoration methods [14] entail casting the HSI cubes as 2-D matrices to exploit the low-rank property. However, such a matricization of the 3-D image cube tends to obfuscate both the spatial texture structure and the spatial-spectral dependence within the image cube [15]. As an alternative to

TABLE I  
LIST OF THE REPRESENTATIVE LRTA-BASED HSI  
RESTORATION METHODS

Methods	Prior			Task	
	TV	GSC	NLSS	Denoising	Inpainting
LRTDTV [16], 2017	✓	✓		*	
3DTNN [12], 2019		✓		*	
LRTDGS [15], 2019		✓		*	
NLR-CPTD [24], 2019		✓	✓	*	
TVTR [13], 2019	✓	✓		*	
NLRR-TC [22], 2019		✓	✓		*
TV-NLRTD [25], 2020	✓	✓	✓	*	
NLTR [17], 2020		✓	✓	*	
LRTF-DFR [18], 2020		✓		*	
WLRTR [23], 2020		✓	✓	*	*
NGMeet [19], 2020		✓	✓	*	*
AATV-NN [10], 2021	✓	✓		*	
OLRT [9], 2021		✓	✓	*	*

matricization, there has been increasing interest in treating an HSI cube as a multidimensional array representation of an order-3 tensor, in which tensor theory and processes are brought to bear on the HSI restoration task. A popular strategy is to deploy a tensor decomposition, such as the well-known Tucker decomposition [15], [16], tensor ring decomposition [13], [17], and low-rank tensor factorization [12], [18]. In addition to GSC-based low-rank tensor approximation (LRTA), NLSS is also a forceful paradigm to deploy HSI restoration methods [17], [19]–[23]. NLSS-based restoration methods first construct a set of new subtensors by clustering similar patches across the image into groups. Then, restoration is performed on these newly constructed subtensors based on low-rank approximation. Table I summarizes some representative LRTA-based approaches that have been proposed recently by utilizing different prior and regularization.

By utilizing NLSS and GSC, these LRTA-based restoration methods have achieved satisfactory performance for HSI restoration. However, it is observed from Table I that most of the methods perform well when removing independent identically distributed (i.i.d.) Gaussian, impulse, and sparse noise only. However, HSIs suffer from more complex and severe degradations in a harsh environment, such as dense stripes, heavy cloud shadows, and atmospheric absorptions. To handle complex degradation, LRTA-based restoration methods considering GSC and NLSS may not be conveniently extended to guarantee the restoration performance and generalization without additional regularization terms and auxiliary information [26], [27]. One of the most frequently adopted regularization terms is total variation (TV) [10], [16], [25], [28], which effectively describes the local spatial and spectral smoothness of the restored HSIs [27], [28]. The auxiliary information may be an observation index set [9], [19], [22], [23], which is employed to indicate the complete observations and incomplete observations of the degraded HSIs. By utilizing the provided auxiliary information, the complex HSI restoration is solved by inpainting.

Recently, some deep learning methods have been proposed to treat HSI restoration [29]–[33]. They achieve the state-of-the-art (SOTA) performance in specific restoration task. However, bottlenecks still remain for convolutional-neural-network (CNN)-based deep learning methods. First, compared

with model-based methods, data-driven deep learning methods have high computational complexity in training large amounts of network parameters, which needs extra labeled data and leads to high expenditure. Second, if a deep learning model is trained for a specific task, it is impossible to adapt it to other degradation cases, which limits their generalization.

In order to remedy the deficiencies in existing strategies and better adapt to severe and complex HSI degradations, a multigraph-based LRTA (MGLRTA) is proposed for HSI restoration, which integrates geometry-based side information into LRTA to constrain a smooth solution of the restored HSIs. Specifically, the contributions are given as follows.

- 1) An LRTA framework driven by trace norm (i.e., [34]) is adopted to effectively capture the simple, yet intrinsic, structurally spatial–spectral correlation in HSI. Rather than blindly estimating the entire restored image by adopting sparse and low-rank priors as done in previous algorithms [15], we instead deploy the restoration by learning the relationship between noisy and pure pixels, as well as propagating the learned structure information to completing pixels in missing regions by imposing low-rank constraints.
- 2) Multigraph-aided LRTA restoration is developed to explore the intrinsic local proximity in HSI, which is characterized by geometry structure similarity between rows, columns, and spectral bands. Taking rows, columns, and spectral bands as a unit to define row graph, column graph, and spectral bands graph, side information is incorporated into LRTA.

As classical high-dimensional data [1], the intrinsic geometry structures of HSIs with the graph spectral analysis have been well explored in HSI classification [35] and dimensionality reduction [36], [37]. To the best of our knowledge, there are only a few works focusing on using graph-based methods in HSI restoration [28]. In [28], pixelwise graph representation is constructed in each band of the HSI and integrated into TV regularization (TVTR) as weighted coefficients. Nevertheless, the denoising performance cannot be guaranteed without a bandwise compensation scheme. In [38], a pixelwise graph regularizer is incorporated into the low-rank representation to preserve the intrinsic local structure of the recovered HSIs, but the proposed method only works on HSI destriping. Existing graph-based methods provide a promising direction to take full advantage of high GSC and local manifold structures. However, some limitations remain. On the one hand, these methods construct pixelwise graphs mainly, which makes it difficult to incorporate the spatial–spectral integration into the tensor-based reconstruction models. On the other hand, existing graph-based methods only treat single restoration tasks.

In addition, in contrast to prior NLSS-based techniques (e.g., [17]), which all use patch clustering, graph structure preservation that propagates known information of observed pixels to missing pixels is employed by learning similarity both spatially and spectrally on a graph rather than spatially only. Besides, the traditional low-rank tensor completion model requires the observation values should follow a uniform distribution, and each column/row and each spectral band should

have at least an observed entry to exactly recover the missing values [39]. However, requirements are too strict to reconstruct an HSI because noise distribution in HSI is complex and unpredictable. For example, stripes lead to the missing of an entire column, the cloud makes the pixels in an irregular patch missing across all bands, and spectral absorption loses an entire band located in a specific spectral wavelength. Thus, introducing geometry structure via a graph in the convex LRTA can effectively address the preceding problems.

The remaining of this article is organized as follows. After introducing some basic mathematical preliminaries and related work in Section II-A, the details of the MGLRTA algorithm are presented next in Section III, followed by experimental results in Section IV that compares the performance of the proposed algorithm against that of several other SOTA restoration approaches. Finally, Section V concludes this article with some concluding remarks.

## II. MATHEMATICAL PRELIMINARIES AND RELATED WORK

### A. Mathematical Preliminaries

A real-valued order- $n$  tensor is denoted as  $\mathcal{X} \in \mathbb{R}^{I_1 \times \dots \times I_k \times \dots \times I_n}$ , with each element being  $x_{i_1, \dots, i_k, \dots, i_n}$ , where  $I_k$  is the dimension of mode  $k$  for  $1 \leq k \leq n$ . A tensor may be unfolded into a matrix along a single mode; the ‘‘unfold’’ operation along mode  $k$  is defined as  $\text{unfold}_k(\mathcal{X}) = \mathcal{X}_{(k)} \in \mathbb{R}^{I_k \times (I_1 \dots I_{k-1} I_{k+1} \dots I_n)}$ . Correspondingly, the ‘‘fold’’ operation reverses the unfolding,  $\text{fold}_k(\mathcal{X}_{(k)}) = \mathcal{X}$ . The Frobenius norm of a tensor is defined as  $\|\mathcal{X}\|_F = (\sum_{i_1, \dots, i_n} |x_{i_1, \dots, i_n}|^2)^{(1/2)}$  such that  $\|\mathcal{X}\|_F = \|\mathcal{X}_{(k)}\|_F$ . The multiplication of a tensor  $\mathcal{X}$  with a matrix  $\mathbf{A} \in \mathbb{R}^{J_k \times I_k}$  on mode  $k$  is defined by  $\mathcal{Y} = \mathcal{X} \times_k \mathbf{A}$ , where  $\mathcal{Y} \in \mathbb{R}^{I_1 \times \dots \times I_{k-1} \times J_k \times I_{k+1} \times \dots \times I_n}$  and  $y_{i_1, \dots, j_k, \dots, i_n} = \sum_{i_k=1}^{I_k} x_{i_1, \dots, i_k, \dots, i_n} a_{j_k, i_k}$ . Alternatively, after an unfold operation along mode  $k$ , the multiplication is defined as  $\mathcal{Y}_{(k)} = \mathbf{A} \mathcal{X}_{(k)}$ . Trace norm is commonly used as a convex proxy for rank [34],  $\|\mathcal{X}\|_* = \sum_{k=1}^n \alpha_k \|\mathcal{X}_{(k)}\|_*$ , where  $\|\cdot\|_*$  is the matrix spectral norm, and  $\alpha_k \geq 0$  are weights such that  $\sum_{k=1}^n \alpha_k = 1$ .

### B. Graph Signal Processing and Matrix Completion on Graph

Signal processing on graph is an effective approach to describe and process high-dimensional data. The key is to construct a weighted graph incorporating the similarity information of a given dataset [40], [41]. Let an undirected and weighted graph be defined as  $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$ , where  $\mathcal{V}$  and  $\mathcal{E}$  denote the sets of vertices and edges of the graph, respectively, and  $\mathbf{W}$  denotes the weight matrix of edges. The degree of  $v_i \in \mathcal{V}$  is calculated as the sum of element values in the  $i$ th column of  $\mathbf{W}$ , i.e.,  $d_i = \sum_{j=1}^N w_{ij}$ . The degree matrix is a diagonal one  $\mathbf{D} = \text{diag}(d_1, d_2, \dots, d_i, \dots, d_N)$ . The Laplacian matrix is defined as  $\mathbf{L} = \mathbf{D} - \mathbf{W}$ . The  $k$ -nearest neighbors of  $v_i$  constitute its neighborhood set  $\mathcal{N}_i^k$ . Without loss of generality, given a graph  $\mathcal{G}$ , the graph signal is defined on the vertices of the graph  $f: \mathcal{V} \rightarrow \mathbb{R}$ .

An intuition to the low-rank constraint on a given matrix  $\mathbf{M} \in \mathbb{R}^{I_1 \times I_2}$  is that there is some linear dependence between its rows/columns [42]. By defining its rows/columns on the vertices of graphs, matrix completion on graph (MCG) is

proposed to constrain the space of solutions in a low-rank matrix completion problem [43]

$$\begin{aligned} & \min_{\mathbf{M}} \text{rank}(\mathbf{M}) \\ & \text{s.t. } \mathbf{M}_{\Omega} = \mathbf{T}_{\Omega} \end{aligned} \quad (1)$$

where  $\text{rank}(\cdot)$  indicates an operation to compute the rank of a matrix;  $\Omega \subset [I_1] \times [I_2]$  is a logical index set generated from the incomplete observation  $\mathbf{T} \in \mathbb{R}^{I_1 \times I_2}$ , wherein an element equaling to 1 indicates that the pixels are given in  $\mathbf{T}$  at the corresponding location, and vice versa, if the element equals to 0, meaning that the corresponding pixel is missing, to be smooth with respect to the geometric structure of the rows and columns of data defined on graph [44]. For example, given a recommendation system denoted by  $\mathbf{M} \in \mathbb{R}^{m \times n}$  with  $m$  users and  $n$  products, proximity information about rows and columns can be attained as usually some users share preferences, while products form clusters that receive similar ratings [45]. Not only in the recommendation system but also in other real-world applications (such as social networks), MCG has well-posed geometric analysis foundations tightly related to the low-rank matrix completion [46].

We define a column graph  $\mathcal{G}_c = (\mathcal{V}_c, \mathcal{E}_c, \mathbf{W}_c)$  of  $\mathbf{M}$  to represent the proximity structure between columns with each column  $\mathbf{x}_n$  being defined on  $v_n^c$ . The row graph  $\mathcal{G}_r = (\mathcal{V}_r, \mathcal{E}_r, \mathbf{W}_r)$  is defined in a similar manner to represent the similarities of rows. MCG imposes the smoothness constraint by setting  $\mathbf{m}_j^c \approx \mathbf{m}_{j'}^c$ , if  $(j, j') \in \mathcal{E}_c$  and results in the following formulation:

$$\sum_{j, j'} w_{jj'}^c \|\mathbf{m}_j^c - \mathbf{m}_{j'}^c\|_2^2 = \text{tr}(\mathbf{M} \mathbf{L}_c \mathbf{M}^T) = \|\mathbf{M}\|_{\mathcal{G}_c}^2 \quad (2)$$

where  $\text{tr}(\cdot)$  indicates an operation to compute the trace of a matrix,  $\mathbf{L}_c$  is the Laplacian of the column graph  $\mathcal{G}_c$ , and  $\|\cdot\|_{\mathcal{G}_c}^2$  is the graph Dirichlet seminorm for columns. Similarly, for row graph  $\mathcal{G}_r$ , we derive the corresponding row graph similarity relationship as  $\|\mathbf{M}\|_{\mathcal{G}_r}^2 = \text{tr}(\mathbf{M}^T \mathbf{L}_r \mathbf{M})$ .

Based on the geometric analysis and the Laplacian spectral theory from graph signal processing, MCG develops as one of the most powerful methods toward deep learning-based graph completion [45], [47]. However, for the tensor completion problems, geometric analysis has not yet found its way to well evolve from the matrix case as the nuclear norm.

## III. PROPOSED MULTIGRAPH-BASED HSI RESTORATION

### A. Motivation

The underlying low-rank nature of HSIs has been widely exploited for various application purposes. For instance, Dian *et al.* [48] have proposed an insightful work for HSI spatial resolution improvement, which effectively exploits the nonlocal similarities of the HSIs via an original nonlocal low tensor-train rank regularization. Furthermore, Chang *et al.* [23] introduced a novel weighted low-rank tensor representation for HSI restoration. Liu *et al.* [49] integrated the fibered rank constrained tensor restoration framework with an embedded plug-and-play-based regularization for HSI restoration. These works have laid a faithful foundation to

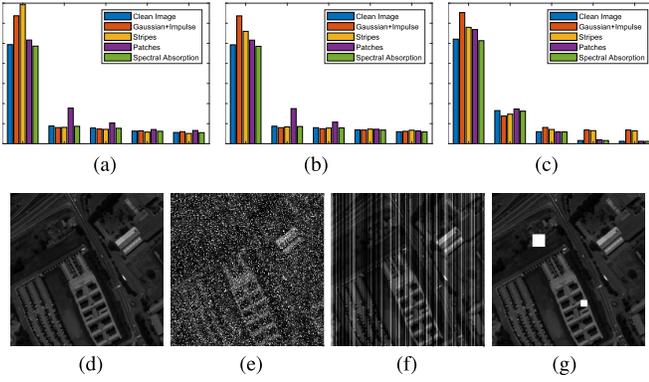


Fig. 1. Low-rankness analysis of an HSI with different degradation cases using HOSVD; above row (a)–(c): the first five singular values of HOSVD on each order; bottom row (d)–(g): the 50th band of the clean HSI and the degraded HSI with different cases as a show case. Note that noise in (e) represents Gaussian and impulse noises. (a) First order. (b) Second order. (c) Third order. (d) Clean. (e) Noise. (f) Stripes. (g) Patches.

incorporate low rankness in the proposed MGLRTA. As illustrated in Fig. 1(a)–(c), high-order singular value decomposition (HOSVD) [50] is utilized to analyze and compare the low-rank property of a given HSI with clean [i.e., Fig. 1(d)] and different degradation versions [i.e., Fig. 1(e)–(g)].

From Fig. 1(a)–(c), the singular values (in any case) decay rapidly, which explicitly indicate the low rankness of HSI. Particularly, for a clean image, the singular values on the third mode (i.e., spectral dimension), as shown in Fig. 1(c), decay much faster than that of the first two modes (spatial vertical and horizontal dimensions), as shown in Fig. 1(a) and (b), which means that a clean HSI has a stronger low rankness in spectral dimension (i.e., GSC) than spatial dimension. In addition, the presence of noise destroys the low rankness. For example, for the case of “Gaussian+Impulse” noise [i.e., Fig. 1(e)] and the case of stripes [i.e., Fig. 1(f)], the first singular value on all third orders increases obviously, and the other four singular values on the third mode decay slower than that of the clean image, which further guarantees that recovering an image by LRTA is an effective way. We expect that the real ground-target clean HSI is characterized by strong correlations in spectrally adjacent bands and high similarity between spatially neighboring pixels; both of these manifest mathematically as low rankness both spatially and spectrally. Thus, the restoration problem is formulated as an estimation of the desired  $\mathcal{X}$  with low-rank constraints on each mode as in the formulation

$$\begin{aligned} \min_{\mathcal{X}} \quad & \sum_{k=1}^3 \alpha_k \|\mathcal{X}_{(k)}\|_* \\ \text{s.t.} \quad & \mathcal{X}_{\Omega} = \mathcal{T}_{\Omega}. \end{aligned} \quad (3)$$

However, to effectively and exactly recover  $\mathcal{X}$  from the observed  $\mathcal{T}_{\Omega}$ ,  $\mathcal{T}_{\Omega}$  should follow a specific distribution with a minimal sampling rate requirement [39]. Besides, each column and each row should include at least one observation. However, for HSI, the degradation is more complex and random. Thus, the observation measurements would not satisfy the requirements of exact recovery. Thus, it is necessary to develop a reconstruction model, which does not rely much on the

strict observation requirements. To this end, we borrow ideas from graph signal processing and introduce additional side information to achieve a better restoration result by forcing smoothness regularization on the solution of  $\mathcal{X}$ .

On the one hand, low-rank implies the linear dependence of rows/columns/spectral bands of  $\mathcal{X}$ , which means that proximity constraints are inherited from the low-rank approximation model. On the other hand, as illustrated in Fig. 1, additional structure information does exist in HSI, which naturally admits certain regularity or smoothness on the reconstruction of  $\mathcal{X}$ . Thus, it is reasonable to characterize structural similarity between rows/columns/spectral bands with graphs.

### B. Proposed MGLRTA Restoration

Manifold learning [51], [52] is an effective technique to find the intrinsic low-dimensional structure of nonlinear high-dimensional data. Due to inevitable imaging environmental disturbance and changing, the captured spectral profiles of HSIs have variability and shift. Therefore, the nonlinear structure exists in HSIs [53]. According to manifold learning, the similarity  $W_{ij}$  between pixel  $\mathbf{x}_i$  and  $\mathbf{x}_j$  in the original high-dimensional space is preserved after embedding the signal into low-dimensional subspace by learning an optimal projection matrix  $\tilde{\mathbf{P}}$  via solving the following problem:

$$\tilde{\mathbf{P}} = \arg \min_{\mathbf{P}^T \mathbf{B} \mathbf{X} \mathbf{P} = \mathbf{I}} \sum_{i \neq j} \|\mathbf{P}^T \mathbf{x}_i - \mathbf{P}^T \mathbf{x}_j\|^2 W_{ij} \quad (4)$$

where  $\mathbf{B}$  is a constraint matrix that avoids trivial solution. Borrowing the idea from manifold learning, the recovered HSI is capable of restoring the original clean image while preserving the similarity relationship naturally inherited by the clean image and induced by the low-rank property. To define the column graph  $\mathcal{G}_c = (\mathcal{V}_c, \mathcal{E}_c, \mathbf{W}_c)$ , each lateral slice  $\mathcal{X}(:, j, :)$  with  $1 \leq j \leq N$  is defined as an  $MD$ -dimensional signal  $\mathbf{x}_j^c$  on the vertex of the column graph  $\mathcal{G}_c$ . Then, based on different priors to the reconstructed HSI, the corresponding graph construction strategies are developed to integrate the geometry information, as shown in Fig. 2. Here, as one of the main objectives is to incorporate graph signal processing with LRTA but not to develop the graph construction method, we adopt the basic rule in (2) to construct the graph. Similarly, for the row graph  $\mathcal{G}_r$  and the spectral band graph  $\mathcal{G}_b$ , we derive the corresponding row graph similarity relationship  $\|\mathcal{X}\|_{\mathcal{G}_r}^2 = \text{tr}(\mathcal{X}_{(1)}^T \mathbf{L}_r \mathcal{X}_{(1)})$  and the band graph similarity relationship  $\|\mathcal{X}\|_{\mathcal{G}_b}^2 = \text{tr}(\mathcal{X}_{(3)}^T \mathbf{L}_b \mathcal{X}_{(3)})$ , with the Laplacian  $\mathbf{L}_r$  and  $\mathbf{L}_b$  of the row graph  $\mathcal{G}_r$  and spectral band graph  $\mathcal{G}_b$ , respectively. For the definition of affinity matrix  $\mathbf{W}_r$ ,  $\mathbf{W}_c$ , and  $\mathbf{W}_b$ , they are learned with different similarity measurements [54], [55]. For simplification, in this article, we just adopt the  $k$ -adjacent graph.

These similarity relationship terms are added to the restoration problem for regularization

$$\begin{aligned} \min_{\mathcal{X}} \quad & \sum_{k=1}^3 \alpha_k \|\mathcal{X}_{(k)}\|_* + \gamma_c \|\mathcal{X}\|_{\mathcal{G}_c}^2 + \gamma_r \|\mathcal{X}\|_{\mathcal{G}_r}^2 + \gamma_b \|\mathcal{X}\|_{\mathcal{G}_b}^2 \\ \text{s.t.} \quad & \mathcal{X}_{\Omega} = \mathcal{T}_{\Omega} \end{aligned} \quad (5)$$

where  $k$  stands for 1, 2, or 3; parameters  $\gamma_c$ ,  $\gamma_r$ , and  $\gamma_b$  control the intensity of smoothness constraints on each mode

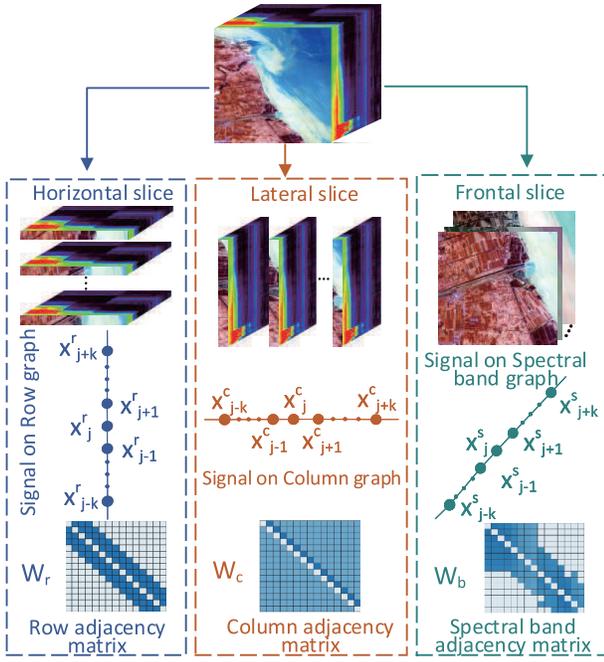


Fig. 2. Definition of the row graph, the column graph, and the spectral band graph.

of the expected restoration. Setting the parameters  $\gamma_c$ ,  $\gamma_r$ ,  $\gamma_b$ , and  $\alpha_k$  will be discussed later in Section IV. For HSIs, the observations are often contaminated by random noise, so (5) is reformulated as

$$\min_{\mathcal{X}} \sum_{k=1}^3 \alpha_k \|\mathcal{X}_{(k)}\|_* + \frac{\tau}{2} \|\mathcal{X} - \mathcal{T}\|_F^2 + \gamma_c \|\mathcal{X}\|_{\mathcal{G}_c}^2 + \gamma_r \|\mathcal{X}\|_{\mathcal{G}_r}^2 + \gamma_b \|\mathcal{X}\|_{\mathcal{G}_b}^2. \quad (6)$$

The restoration model (6) is regarded, in essence, as the low-rank tensor completion problem incorporating the graph-based geometric side information. As discussed in [34], tensor completion builds up the relationship between the known and unknown elements to estimate the latter. The construction of the relationship facilitates the ability to capture the information from both neighbor entries and the global structure. Inspired by this core principle, the model (6) is expected to estimate the unknown  $\mathcal{X}$  by building up its relationship to the known  $\mathcal{T}_\Omega$  through observed measurements.

To introduce the TVTR-based restoration models, the graph regularization terms in (6) are replaced by a TVTR. Note that graph regularization terms developed in (6) are highly related to TVTR. In (6), if only the relationship between two adjacent columns is considered, the TV term is formulated as

$$\sum_j \|\mathbf{x}_{j+1} - \mathbf{x}_j\| = \|\mathbf{R}_c \mathcal{X}_{(2)}\|_1 = \|\mathcal{X}_{(2)}\|_{\text{TV},c} \quad (7)$$

where  $\mathbf{R}_c$  is the column differential operator (to distinguish with the degree matrix in Section III-A, and  $\mathbf{R}$  denotes the differential operator here). The TV term is capable of discouraging the solution from having oscillations and preserving piecewise smoothness of the restored  $\mathcal{X}$ . To avoid the nonlinearity and nondifferentiability, (7) is also rewritten

as  $\|\mathcal{X}_{(2)}\|_{\text{TV},c} = \|\mathbf{R}_c \mathcal{X}_{(2)}\|_F^2$ . After unfolding, the expression becomes  $\|\mathcal{X}_{(2)}\|_{\text{TV},c} = \mathcal{X}_{(2)}^\top \mathbf{R}_c^\top \mathbf{R}_c \mathcal{X}_{(2)}$ . Although enjoying similar ability in preserving the (piecewise) geometry smoothness structure in  $\mathcal{X}$ , they still have difference from the view of spectral graph theory. In fact, different from TV, which has a specific adjacency matrix denoted as a differential operator  $\mathbf{R}_c$ , the graph-based geometry term (2) learns different adjacency matrices based on the known prior, i.e.,  $j'$  can be any column as needed but not only the  $(j-1)$ th column. For example, if we want to preserve the relationship with the  $k$  nearest columns, then a  $k$ -NN-based  $\mathbf{W}_c$  is learned. In this case, the graph-based geometry term is more general and flexible.

### C. Optimization

As the trace norm is the tightest convex envelope for the rank, (6) is a convex but nonsmooth optimization because it contains nondifferential terms, belonging to the class of  $\ell_1$ -type optimization problems. Besides, entries in the tensor  $\mathcal{X}$  are interdependent, i.e., the unfolded  $\mathcal{X}$  in each order shares the same entries and cannot be optimized independently, which is the reason that lexicographical matricization is different with tensor. Equation (6) cannot be solved directly via differential calculus. Consequently, we follow the ADMM paradigm to develop a well-posed iterative scheme for the solution. The auxiliary tensor  $\mathcal{M}_k$ 's for  $\mathcal{X}$  in each mode unfolding is introduced as

$$\begin{aligned} \min_{\mathcal{X}, \mathcal{M}_k} \sum_{k=1}^3 \alpha_k \|\mathcal{M}_{k(k)}\|_* + \frac{\tau}{2} \|\mathcal{X} - \mathcal{T}\|_F^2 \\ + \gamma_c \|\mathcal{X}\|_{\mathcal{G}_c}^2 + \gamma_r \|\mathcal{X}\|_{\mathcal{G}_r}^2 + \gamma_b \|\mathcal{X}\|_{\mathcal{G}_b}^2 \\ \text{s.t. } \mathcal{X} = \mathcal{M}_k, \quad k = 1, 2, 3. \end{aligned} \quad (8)$$

In addition, we have Lagrange multipliers  $\mathcal{Y}_k$ , where, again,  $k \in \{1, 2, 3\}$ , as well as penalty parameters  $\beta$ . This renders (6) separable such that it is solved via ADMM. Specifically, the augmented Lagrangian is formed as

$$\begin{aligned} L_\beta(\mathcal{X}, \mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3, \mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Y}_3) \\ = \sum_{k=1}^3 \alpha_k \|\mathcal{M}_{k(k)}\|_* + \frac{\tau}{2} \|\mathcal{X} - \mathcal{T}\|_F^2 \\ + \gamma_c \|\mathcal{X}\|_{\mathcal{G}_c}^2 + \gamma_r \|\mathcal{X}\|_{\mathcal{G}_r}^2 + \gamma_b \|\mathcal{X}\|_{\mathcal{G}_b}^2 \\ + \langle \mathcal{X} - \mathcal{M}_k, \mathcal{Y}_k \rangle + \frac{\beta}{2} \|\mathcal{X} - \mathcal{M}_k\|_F^2. \end{aligned} \quad (9)$$

According to ADMM, the optimization problem in (6) is then easily solved via (9) with the following iterative scheme:

$$\{\mathcal{M}_{k,t+1}\} = \arg \min_{\mathcal{M}_k} L_\beta(\mathcal{X}_t, \mathcal{M}_k, \mathcal{Y}_{k,t}), \quad k = 1, 2, 3 \quad (10)$$

$$\mathcal{X}_{t+1} = \arg \min_{\mathcal{X}} L_\beta(\mathcal{X}, \mathcal{M}_{k,t+1}, \mathcal{Y}_{k,t}), \quad k = 1, 2, 3 \quad (11)$$

$$\mathcal{Y}_{k,t+1} = \mathcal{Y}_{k,t} + \beta(\mathcal{X}_{t+1} - \mathcal{M}_{k,t+1}), \quad k = 1, 2, 3. \quad (12)$$

For each subproblem, the following mathematical solutions are provided.

1) Update  $\mathcal{M}_k$  for  $k = 1, 2$ , and 3 via

$$\begin{aligned} \mathcal{M}_{k,t+1} = \arg \min_{\mathcal{M}_k} \alpha_k \|\mathcal{M}_{k(k)}\|_* \\ + \frac{\beta}{2} \left\| \mathcal{X} - \mathcal{M}_k + \frac{\mathcal{Y}_k}{\beta} \right\|_F^2. \end{aligned} \quad (13)$$

Because it can be easily proved that  $(\beta/2)\|\mathcal{X} - \mathcal{M}_k + (\mathcal{Y}_k/\beta)\|_F^2 = (\beta/2)\|\mathcal{X}_{(k)} - \mathcal{M}_{k(k)} + ((\mathcal{Y}_{k(k)})/\beta)\|_F^2$ , (13) has a closed-form solution

$$\mathcal{M}_{k,t+1} = \mathbf{fold}_k \left[ \mathcal{D}_{\alpha_k/\beta} \left( \mathcal{X}_{(k)} + \frac{\mathcal{Y}_{k(k)}}{\beta} \right) \right], \quad k = 1, 2, 3 \quad (14)$$

where  $\mathcal{D}_\tau(\mathbf{X})$  is singular-value thresholding operator [56]

$$\mathcal{D}_\tau(\mathbf{X}) = \mathbf{US}_\tau[\boldsymbol{\Sigma}]\mathbf{V}^H \quad (15)$$

with  $\mathbf{X} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H$  being a singular value decomposition of matrix  $\mathbf{X}$  and the matrix-valued shrinkage operator  $\mathcal{S}_\tau[x] = \text{sgn}(x) \max\{|x| - \tau, 0\}$  applied to each element of  $\mathbf{X}$ .

2) Update  $\mathcal{X}$ : We have

$$\begin{aligned} \mathcal{X}_{t+1} = \arg \min_{\mathcal{X}} & \gamma_c \|\mathcal{X}\|_{\mathcal{G}_c}^2 + \gamma_r \|\mathcal{X}\|_{\mathcal{G}_r}^2 + \gamma_b \|\mathcal{X}\|_{\mathcal{G}_b}^2 \\ & + \frac{\tau}{2} \|\mathcal{X} - \mathcal{T}\|_F^2 + \frac{\beta}{2} \left\| \mathcal{X} - \mathcal{M}_{k,t+1} + \frac{\mathcal{Y}_k}{\beta} \right\|_F^2. \end{aligned} \quad (16)$$

Here, the three graph preularization terms are also interdependent due to the unfolding operation in each term of the same tensor. However, for each term, they share the same way to derive the solutions. Thus, we compute the solution for each term separately, corresponding to the unfolding of  $\mathcal{X}$  in each mode. Then,  $\mathcal{X}$  is ultimately updated with

$$\mathcal{X}_{t+1} = \frac{1}{3} \sum_{k=1}^3 \mathbf{fold}_k(\mathcal{X}_{(k),t+1}). \quad (17)$$

Taking column graph regularization term optimization as example, we have

$$\begin{aligned} \mathcal{X}_{(2),t+1} = \arg \min_{\mathcal{X}_{(2)}} & \gamma_c \mathbf{tr}(\mathcal{X}_{(2)}^T \mathbf{L}_c \mathcal{X}_{(2)}) \\ & + \frac{\tau}{2} \|\mathcal{X}_{(2)} - \mathcal{T}_{(2)}\|_F^2 \\ & + \frac{\beta}{2} \left\| \mathcal{X}_{(2)} - \mathcal{M}_{2(2),t+1} + \frac{\mathcal{Y}_{2(2),t}}{\beta} \right\|_F^2. \end{aligned} \quad (18)$$

By setting the derivative of (18) subject to  $\mathcal{X}_{(2)}$  to be zero, the optimality condition is expressed as

$$\begin{aligned} \gamma_c \mathbf{L}_c \mathcal{X}_{(2)} + \tau(\mathcal{X}_{(2)} - \mathcal{T}_{(2)}) \\ + \beta \left( \mathcal{X}_{(2)} - \mathcal{M}_{2(2),t+1} + \frac{\mathcal{Y}_{2(2),t}}{\beta} \right) = 0 \end{aligned} \quad (19)$$

which is rewritten as

$$(\gamma_c \mathbf{L}_c + (\tau + \beta) \mathbf{I}_N) \mathcal{X}_{(2)} = \beta \mathcal{M}_{2(2),t+1} - \mathcal{Y}_{2(2),t} + \tau \mathcal{T}_{(2)}. \quad (20)$$

Through the column-stack vectorization operator  $\mathbf{vec}(\cdot)$ , it becomes

$$\begin{aligned} (\gamma_c \mathbf{L}_c \otimes \mathbf{I}_{MD} + (\tau + \beta) \mathbf{I}_N \otimes \mathbf{I}_{MD}) \mathbf{vec}(\mathcal{X}_{(2)}) \\ = \mathbf{vec}(\beta \mathcal{M}_{2(2),t+1} - \mathcal{Y}_{2(2),t} + \tau \mathcal{T}_{(2)}). \end{aligned} \quad (21)$$

---

### Algorithm 1 MGLRTA HSI Restoration Algorithm

---

- 1: **Input:** degraded HSI  $\mathcal{X}$  with  $\mathcal{X}_\Omega = \mathcal{T}_\Omega$ ; parameters  $\alpha_k$ ,  $k \in \{1, 2, 3\}$ ;  $\gamma_c, \gamma_r, \gamma_b$ ;  $\tau$ ;  $T_{\max}$ ;
  - 2: **Initialization:**  $\mathbf{Y}_{k,0} = \mathbf{0}$ ,  $k \in \{1, 2, 3\}$ ;  $\eta_1 = 10^{-5}$ ;  $t = 0$ ;  $\beta = 1e - 1$  and  $\mu = 1.05$ ;
  - 3: Set  $\mathcal{X}_\Omega = \mathcal{T}_\Omega$  and  $\mathcal{X}_{\bar{\Omega}} = \mathbf{0}$ ;
  - 4:  $\mathcal{M}_{k,t} = \mathbf{unfold}_k(\mathcal{X}_t)$  for  $k \in \{1, 2, 3\}$ ;
  - 5: **repeat**
  - 6:   update  $\mathcal{M}_{k,t+1}$  via (14) for  $k = 1, 2, 3$ ;
  - 7:   update  $\mathcal{X}_{t+1}$  via (17) with each mode-unfolding updated as like (22);
  - 8:   update Lagrange multiplier  $\mathcal{Y}_{k,t+1}$  via (12);
  - 9:    $t \leftarrow t + 1$ ;
  - 10: **until**  $\|\mathcal{X}_t - \mathcal{X}_{t-1}\|_F < \eta_1$  **or**  $t > T_{\max}$ ;
  - 11: **return**  $\mathcal{X} = \mathcal{X}_t$ .
- 

Because  $\mathbf{L}_c$  is symmetric positive semidefinite, the conjugate gradient (CG) algorithm [57] is applied to compute a fast approximate solution. If the size of dataset is not that large, an explicit solution can be derived directly as

$$\begin{aligned} \mathcal{X}_{(2)} = (\gamma_c \mathbf{L}_c + (\tau + \beta) \mathbf{I}_N)^{-1} \\ \times (\beta \mathcal{M}_{2(2),t+1} - \mathcal{Y}_{2(2),t} + \tau \mathcal{T}_{(2)}). \end{aligned} \quad (22)$$

The final optimization procedure for the proposed MGLRTA restoration framework is presented in Algorithm 1.

1) *Computational Complexity:* The per-iteration computational complexity of Algorithm 1 mainly includes two parts.

1) *Update  $\mathcal{M}_k$ :* It involves computation of the nuclear proximal solutions by SVD for each mode with the complexity  $O(M^2ND + MN^2D + MND^2)$ .

2) *Update  $\mathcal{X}$ :* It mainly involves the inverse of a matrix and multiplication for each mode, so the overall computation complexity is also  $O(M^2ND + MN^2D + MND^2)$ .

2) *Convergence:* As the trace norm is a convex proxy for rank, the objective function considered in the restoration problem (6) is convex, but not differentiable. However, based on the block coordinate descent theory, the introduction of variables  $\mathcal{M}_k$  in the augmented Lagrangian (9) leads to closed solutions for both  $\mathcal{X}$  [i.e., (17)] and  $\mathcal{M}_k$  [i.e., (14)] [34], [56].

## IV. EXPERIMENTS AND ANALYSIS

One of the most important advantages of the proposed MGLRTA is simple and easy to interpret. The proposed model is efficient and may be generally extended according to different purposes by developing other LRTA models and graph-based geometry preserving schemes.

The proposed MGLRTA is compared with several classic and SOTA HSI restoration algorithms under different degradation scenarios. The comparison methods are TV-regularized low-rank matrix factorization (LRTV) [58], weighted group sparsity-regularized low-rank tensor decomposition (LRTDGS) [15], HaLRTC [34], tensor ring completion with TVTR [13], weighted low-rank tensor recovery (WLRTTR) [23], CNN-based deep image prior (DIP) [30], [59], and the trainable spectral-spatial sparse (T3SC) coding

model [60]. LRTV is often adopted as a benchmark for the representative matrix-based denoising method, which integrates the nuclear norm, TVTR, and  $\ell_1$ -norm together in a unified framework. HaLRTC is a tensor completion method, which is also usually adopted as a benchmark in HSI restoration. TVTR, WLRT, and the proposed MGLRTA are developed based on HaLRTC. LRTDGS is proposed for mixed noise removal. DIP is a deep-learning-based restoration method, extended from natural image restoration. T3SC is a two-layer supervised restoration model, combining deep learning techniques with the classical sparse coding strategy. For a fair comparison, we use the codes provided by the authors and carefully tuned the parameters by default or follow the rules to achieve the best performance.

A variety of full-reference image-quality measures is employed to quantitatively gauge the performance of the reconstruction techniques under consideration. Specifically, we use the peak signal-to-noise ratio (PSNR), the spectral angle mapper (SAM), the dimensionless global relative error of synthesis (ERGAS) [61], and the structural similarity index measure (SSIM) [62]. PSNR, ERGAS, and SSIM are 2-D image-distortion measures. These measures are calculated for each spectral band, and then, the results across all bands are averaged. On the other hand, SAM is defined between two hyperspectral pixel vectors; we thus average SAM across all pixel vectors in the image. Higher PSNR and SSIM values imply better-quality fusion, while lower SAM and ERGAS values correspond to better results.

#### A. Parameter Analysis

The MGLRTA has two groups of parameters need to be considered: trace-norm parameters  $\alpha_k$  and graph regularization parameters  $\gamma_c$ ,  $\gamma_r$ , and  $\gamma_b$ . These two groups of the parameters can be set empirically, which makes the restoration model flexible and easily extended to different cases. The  $\alpha_k$  parameters in (6) control the relative intensity of the low-rank constraints on each mode of the restoration result  $\mathcal{X}$ . Hence, the values of  $\alpha_k$  influence the restoration results significantly. We set  $\alpha_1 = \alpha_2 = 1$  as they correspond to the two spatial directions. Typically, for HSIs, the spatial-mode ranks are much larger than that of the spectral mode [as illustrated in Fig. 3(a)]. We experimentally tune  $\alpha_3$  based on HaLRTC in [34] and, finally, set  $\alpha_3 = 1e3$ . To validate the generalization of the proposed MGLRTA, experiments on multispectral imagery (i.e., Landsat 7) are performed, where the spatial-mode ranks are almost equal to that of the spectral mode [as illustrated in Fig. 3(b)], so we set  $\alpha_1 = \alpha_2 = \alpha_3 = 1$ . This effect is further observed in Fig. 4. Here, nonnormalized  $\alpha_k$ 's offer better performance than the normalized constraints with  $\sum_{k=1}^3 \alpha_k = 1$ .

After fixing parameters  $\alpha_k$ , graph regularization parameters  $\gamma_c$ ,  $\gamma_r$ , and  $\gamma_b$  are tuned accordingly. We set  $\gamma_c = \gamma_r$ , empirically. As illustrated in Fig. 5, the restoration performance is sensitive to  $\gamma_b$ , i.e., spectral graph construction; while the change of  $\gamma_c$  and  $\gamma_r$  does not influence the restoration performance too much. The penalty parameter  $\beta$  controls the weighting of the fidelity term in (9), which typically relates

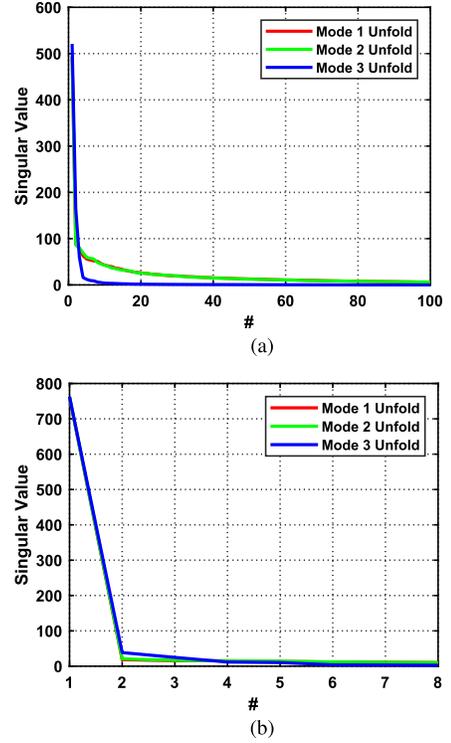


Fig. 3. Comparative analysis of low rankness on each mode using HOSVD. (a) Pavia University dataset. (b) Landsat 7 dataset.

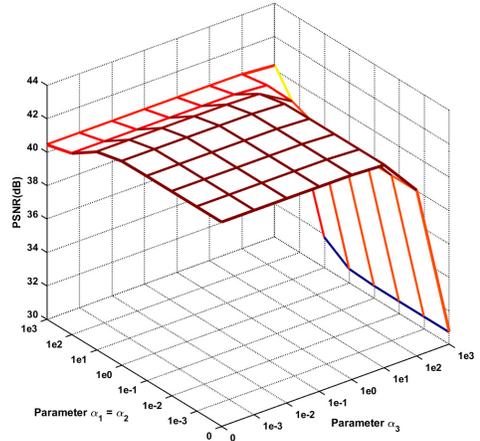


Fig. 4. Parameter tuning of  $\alpha$  for the proposed MGLRTA on the Landsat 7 dataset.

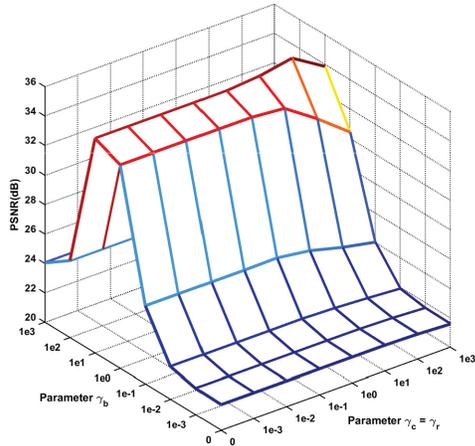
more to the speed of convergence than to the final optimization result, so we seek appropriate values that yield convergence without excessive iteration. In this work, we experimentally set  $\beta = 1e-1$  and error parameter  $\tau = 1e-3$  for all experiments. Here,  $\beta = 1e-1$  produces a tradeoff of better results and convergence than  $1e-6$  of the penalty parameter for HaLRTC in [34]. Then,  $\beta$  is updated with  $\beta = \mu\beta$  with  $\mu$  being 1.05.

#### B. Destriping Performance

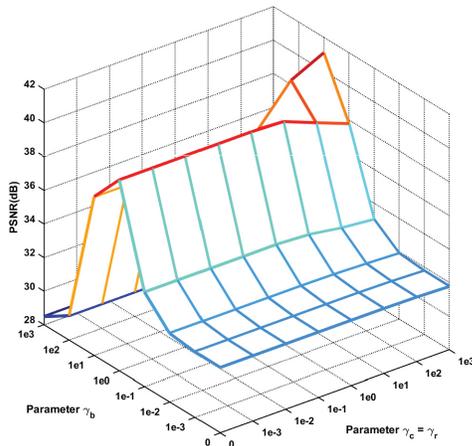
Two HSI datasets are adopted to add synthetic stripes, and a multispectral image corrupted with oblique stripes arising in

TABLE II  
RECONSTRUCTION PERFORMANCE FOR THE STRIPED PAVIA UNIVERSITY DATASET OF VARYING DENSITY  $d$

	$d = 0.1$				$d = 0.5$				$d = 0.9$			
	PSNR(dB)	SSIM	ERGAS	SAM	PSNR(dB)	SSIM	ERGAS	SAM	PSNR(dB)	SSIM	ERGAS	SAM
Striped HSI	23.49	0.6932	262.2060	0.3336	16.93	0.2487	554.6417	0.7970	14.41	0.0537	740.4027	1.2605
LRTV [58]	<b>46.83</b>	0.9866	31.6676	0.0385	35.22	0.9016	86.2972	0.1168	26.29	0.6379	195.4620	0.2632
LRTDGS [15]	28.67	0.8083	144.6118	0.1218	17.36	0.2888	531.0589	0.6248	13.97	0.0070	778.2541	1.4028
HaLRTC [34]	32.18	0.9501	95.2884	0.0260	32.06	0.9458	96.6492	<u>0.0384</u>	29.96	0.8944	127.8235	0.0669
TVTR [13]	33.31	0.9528	87.4341	<u>0.0253</u>	27.18	0.7864	201.2838	0.1740	13.32	0.2956	930.8700	0.4046
WLRTR [23]	41.44	0.9874	33.1668	0.0327	31.89	0.9172	5527.9000	0.1016	18.59	0.6531	25850.8348	0.2908
DIP [30]	37.15	0.9506	55.7212	0.0578	<u>40.47</u>	0.9694	40.7500	0.0473	<b>37.35</b>	<u>0.9538</u>	<u>55.0573</u>	<u>0.0569</u>
T3SC [60]	42.15	<u>0.9939</u>	<u>23.0688</u>	0.0290	38.58	<b>0.9897</b>	<b>34.8635</b>	0.0431	<u>37.10</u>	<b>0.9870</b>	<b>41.4354</b>	<b>0.0493</b>
MGLRTA	<u>44.75</u>	<b>0.9942</b>	<b>22.7926</b>	<b>0.0090</b>	<b>41.08</b>	<u>0.9838</u>	<u>36.8796</u>	<b>0.0306</b>	31.15	0.9065	111.9950	0.0669



(a)



(b)

Fig. 5. Parameter tuning of  $\gamma_r$ ,  $\gamma_c$ , and  $\gamma_b$  for the proposed MGLRTA on two synthetic experimental datasets. (a) Synthetic Indian Pines dataset. (b) Pavia University dataset

a real sensor is employed for testing the generalization of the proposed MGLRTA.

1) *Synthetic Datasets*: The Pavia University dataset<sup>1</sup> acquired by a Reflective Optics System Imaging Spectrometer (ROSIS) sensor over the University of Pavia in Northern Italy is used. The top-left  $256 \times 256 \times 103$  subimage of the

<sup>1</sup>[http://www.ehu.es/ccwintco/index.php?title=Hyperspectral\\_Remote\\_Sensing\\_Scenes](http://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes)

dataset is adopted as the reference imagery. The pixel values are normalized to the range  $[0, 1]$ . The synthetic Indian Pines dataset with the size of  $145 \times 145 \times 224$  [16], which simulates the noise-free Indian Pines, is also adopted. Specifically, stripes are added according to the process outlined in [4]. Given an HSI  $\mathcal{X}$  of size  $M \times N \times D$ , synthetic stripes are generated by randomly selecting  $dN$  columns in a given band to be striped. Here,  $d$  is the parameter that controls the spatial density of the striping,  $0 < d < 1$ . This process is repeated independently in each of the  $D$  bands of the HSI. Thus, the synthetic stripes are aperiodic spatially and uncorrelated across the spectral bands. In addition, the [50, 51, 100, 200]th and [50, 51, 100, 140]th columns across all bands are selected as deadlines to cover a more complex case for Pavia University and synthetic Indian Pines datasets, respectively.

2) *Real Dataset*: The real dataset<sup>2</sup> is multispectral imagery acquired by Landsat 7 ETM+. Due to the scan line corrector (SLC) failure since May 31, 2003, the acquired imagery has oblique data gaps. However, the SLC-off data are still provided as they maintain the same radiometric and geometric corrections as data collected prior to the SLC failure. The band-specific gap mask files are also provided with each Landsat 7 SLC-off level-1 data product. To quantitatively and visually compare the reconstruction performance of the proposed MGLRTA with other methods, an SLC-on dataset before May 31, 2003, was downloaded as the reference image, and the SLC-off mask file was used to generate the SLC-off image. The degradation mask is from the real world; hence, the generated image represents the real-world dataset.

3) *Results Analysis*: Tables II and III list the quantitative performance measures of the various algorithms on the synthetically striped Pavia University dataset and Indian Pines dataset for stripes of varying spatial density  $d$ . It is observed that the proposed MGLRTA achieves the best performance compared with traditional machine-learning-based methods. Although the restoration performance of the proposed MGLRTA is worse than that of the deep-learning-based DIP<sup>3</sup> and T3SC<sup>4</sup> sometimes, MGLRTA has relatively stable performance. That is, for machine-learning-based restoration

<sup>2</sup><https://earthexplorer.usgs.gov/>

<sup>3</sup>The implementation details may be found in [30].

<sup>4</sup>For the Pavia University dataset, the training set is cropped from the bottom right of the Pavia Center dataset with the size being  $596 \times 307 \times 102$  and the bottom of the Pavia University dataset with the size of  $353 \times 340 \times 102$ . For Indian Pines dataset, the Cuprite dataset of the size  $512 \times 614 \times 224$  is adopted as the training set. Other parameters remain the same as in [60].

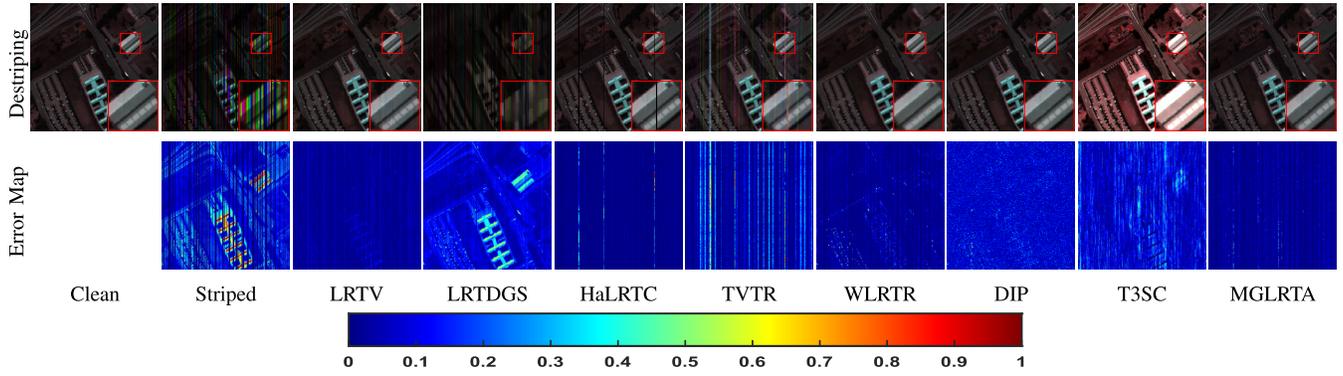


Fig. 6. Pseudocolor visualization of inpainting results for the Pavia University dataset with stripes density being 0.5; pseudocolor composed of bands 45, 15, and 20. The color bar gives the scale for the error maps. Detailed subimages are enlarged in red squares.

TABLE III  
RECONSTRUCTION PERFORMANCE FOR THE STRIPED SYNTHETIC INDIAN PINES DATASET OF VARYING DENSITY  $d$

	$d = 0.1$				$d = 0.5$				$d = 0.9$			
	PSNR(dB)	SSIM	ERGAS	SAM	PSNR(dB)	SSIM	ERGAS	SAM	PSNR(dB)	SSIM	ERGAS	SAM
Striped HSI	14.91	0.5048	420.7976	0.3530	8.63	0.1057	861.2600	0.8062	6.15	0.0173	1143.7898	1.2601
LRTV [58]	<u>39.86</u>	<u>0.9906</u>	35.7349	0.0095	28.12	0.7839	109.0579	0.0760	17.71	0.3500	314.0800	0.2601
LRTDGS [15]	21.33	0.8579	199.7278	0.0544	9.86	0.1726	756.0714	0.6102	5.76	0.0031	1196.4328	1.4945
HaLRTC [34]	30.82	0.9367	69.4574	0.0154	21.75	0.8801	189.6996	0.0126	21.28	0.8224	199.2031	0.0834
TVTR [13]	32.11	0.9485	60.0080	0.0110	31.70	0.9387	62.8745	0.0172	18.19	0.4597	323.1344	0.2438
WLRTR [23]	<b>41.17</b>	<b>0.9953</b>	<b>21.1750</b>	<u>0.0072</u>	33.28	0.9704	52.1509	0.0245	18.11	0.7245	19135.6852	0.5160
DIP [30]	37.33	0.982	32.0546	0.0144	<b>40.90</b>	<b>0.9964</b>	<b>21.6948</b>	<b>0.0044</b>	<b>32.39</b>	<b>0.9735</b>	<b>60.6570</b>	<b>0.0227</b>
T3SC [60]	28.67	0.8982	164.58	0.1248	27.18	0.8805	176.6608	0.1344	26.12	0.8644	188.0795	0.1426
MGLRTA	38.29	0.9891	<u>29.3919</u>	<b>0.0032</b>	<u>34.56</u>	<u>0.9771</u>	<u>47.3774</u>	<u>0.0228</u>	<u>24.33</u>	<u>0.8448</u>	<u>140.5459</u>	<u>0.0394</u>

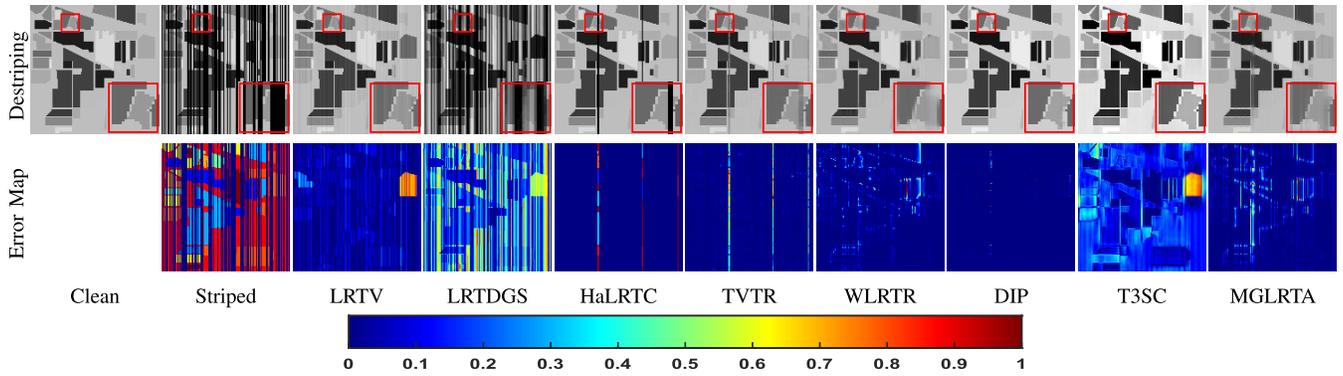


Fig. 7. Visualization of destriping results for the synthetic Indian Pines dataset with stripes density being 0.5; the band number is 45. The color bar gives the scale for the error maps. Detailed subimages are enlarged in red squares.

methods, with the increasing of the stripes' density, the restoration performance decays correspondingly. However, the performance of DIP is hard to interpret physically and intuitively. As can be observed, the performance is unusually high when the density is  $d = 0.5$ . This effect is further shown in Figs. 6 and 7. Detailed subimages are presented in red squares. Note that T3SC adopts bandwise min-max normalization, and its pseudocolor image is brighter than the others. In particular, WLRTR results in oversmoothing in that some spatial details are lost. HaLRTC and LRTV yield generally better results but still contain some weak stripe noise compared with the original clean imagery.

However, it is observed from Table IV that MGLRTA performs best among all methods for the oblique gaps inpainting. It is obvious that HaLRTC, WLRTR, and MGLRTA have more considerable improvement compared with other methods. Meanwhile, it can also be observed from Fig. 8 that MGLRTA has a clear reconstruction result with less error. Although LRTV has a good reconstruction result, as shown in Fig. 6, it results in a poor reconstruction in Fig. 8. This observation validates that the proposed MGLRTA has good generalization performance to solve many real-world degradation problems while achieving better preservation of edge and texture details.

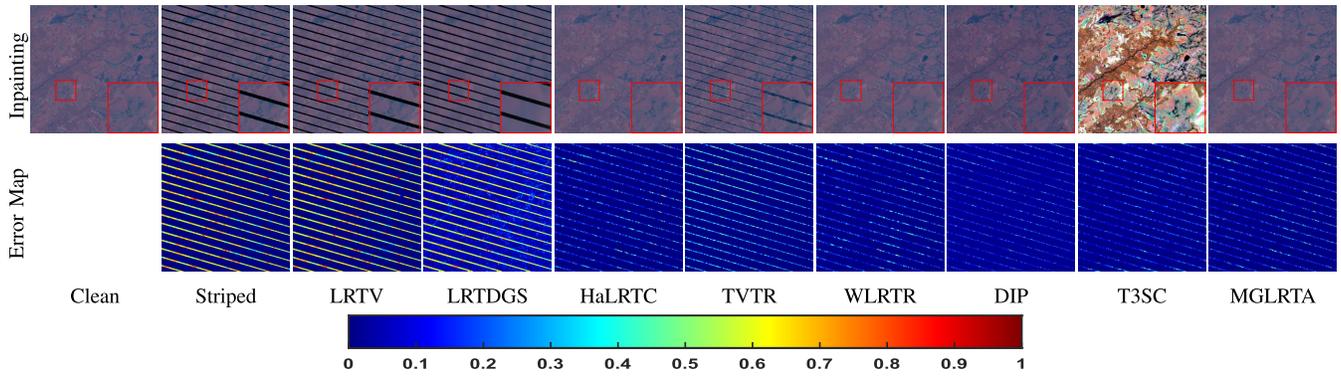


Fig. 8. Pseudocolor visualization of inpainting results for the Landsat 7 dataset; pseudocolor composed of bands 3, 2, and 1. The color bar gives the scale for the error maps. Detailed subimages are enlarged in red squares.

TABLE IV

RECONSTRUCTION PERFORMANCE FOR THE SLC-OFF LANDSAT DATASET

	PSNR(dB)	SSIM	ERGAS	SAM
SLC-off Dataset	15.39	0.6400	437.7463	0.2480
LRTV [58]	15.40	0.6400	437.7463	0.2377
LRTDGS [15]	16.11	0.4885	395.1153	0.0835
HaLRTC [34]	40.27	0.9612	<u>32.2204</u>	<u>0.0066</u>
TVTR [13]	20.58	0.7330	269.7683	0.0812
WLRTR [23]	<u>40.50</u>	<u>0.9626</u>	32.8028	0.0081
DIP [30]	39.97	0.9461	35.3285	0.0161
T3SC [60]	24.33	0.9014	133.5152	0.0441
MGLRTA	<b>41.85</b>	<b>0.9688</b>	<b>27.8182</b>	<b>0.0058</b>

### C. Cloud Removal Performance

Apart from stripes, another serious issue that optical remotely sensed imageries have is the cloud shadow cover. Cloud shadow cover leads to the loss of consecutive-region pixels. In most existing research [13], [63], a series of multitemporal images are expected to provide for the removal of the cloud and recover the cloud covering the area in the image. However, it tends to be difficult to acquire multitemporal images in real applications. In this section, we analyze cloud removal performance and clarify that the proposed MGLRTA has competitive reconstruction performance only with two time points.

1) *Dataset*: Here, the Bay Area scene<sup>5</sup> was taken in the years 2013 and 2015 with the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor surrounding the city of Patterson (California). The spatial dimensions are  $600 \times 500$  pixels with 224 spectral bands. The real cloud mask from Gaofen-5 data is used to generate that cloudy Bay Area dataset.

2) *Results Analysis*: Because the SOTA reconstruction methods mentioned above do not report their abilities in cloud removal for short-term HSIs, we will not compare them in this part. We adequately analyze the advantages of the introduction of geometric side information for cloud removal. Visual restoration results and corresponding quantitative performance of classic HaLRTC, deep-learning-based

<sup>5</sup><https://citius.usc.es/investigacion/datasets/hyperspectral-change-detection-dataset>

TABLE V

RECONSTRUCTION PERFORMANCE OF CLOUD REMOVAL RESULTS FOR THE BAY AREA DATASET

	PSNR(dB)	SSIM	ERGAS	SAM
Cloudy Dataset	47.36	0.9574	99.2665	0.0707
HaLRTC [34]	60.74	0.9964	16.1055	0.0045
DIP [30]	54.43	0.9922	21.3424	0.0208
MGLRTA, case 1	<b>Inf</b>	<b>0.9982</b>	<b>11.2741</b>	<b>0.0037</b>
MGLRTA, case 2	<b>Inf</b>	<b>0.9982</b>	<u>11.3685</u>	<b>0.0037</b>

DIP, and the proposed MGLRTA (considering two different cases) are illustrated in Fig. 9 and Table V. It is observed that both HaLRTC and MGLRTA are capable of recovering the majority of the missing values in the area covered by clouds only using the HSI with two time points. However, MGLRTA's quantitative performance is better than that of HaLRTC. By tuning the row and column graph regularization parameters  $\gamma_r$  and  $\gamma_c$  (i.e., case 1 and case 2), the proposed MGLRTA recovers the cloudy image with errors in different regions.

### D. Spectral Recovery of Atmospheric Absorption Bands

The effect of the intervening atmosphere is also one of the most severe and common factors that influence the image [64]. An entire spectral band may have to be removed due to unwanted atmospheric contamination, such as water vapor and mixed gases. Although atmospheric correction and compensation can be conducted to the level-2 data products, the produced data may not fulfill the requirements for real applications. For example, the classic Indian Pines and Pavia University datasets suffer from the same problems, and the researchers have to remove the bands covering the region of water absorption. Thus, the proposed MGLRTA is validated on the ability to recover the successive atmosphere absorption bands.

1) *Dataset*: The Gaofen-5 dataset covering shortwave infrared (SWIR) spectral wavelength with 180 spectral bands and a 30-m spatial resolution was taken in the year 2019 with the Advanced Hyperspectral Imager (AHSI) aboard China's GaoFen-5 satellite [65]. The spatial dimensions are  $512 \times 512$  pixels with 180 spectral bands covering short and infrared

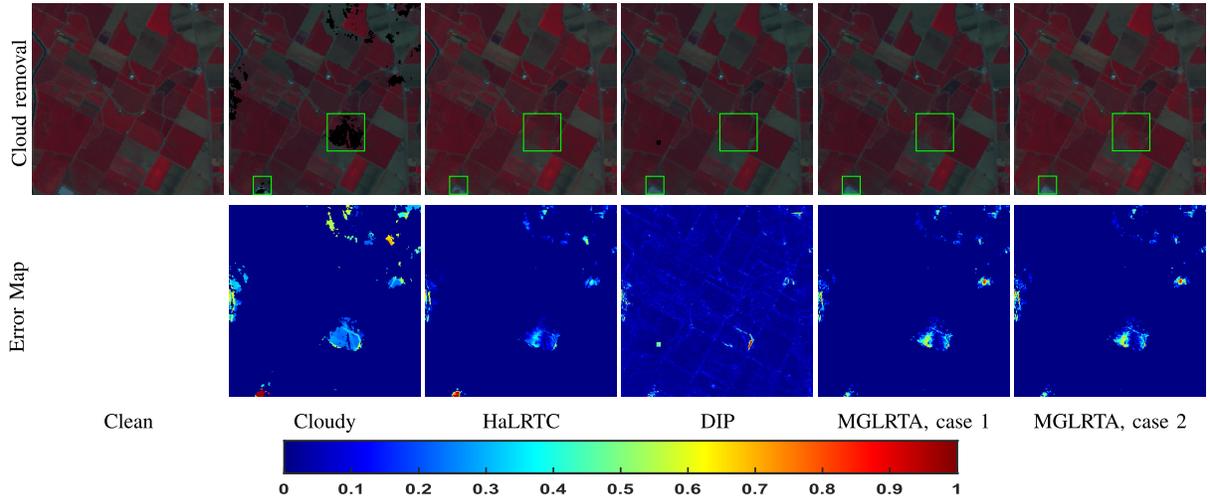


Fig. 9. Pseudocolor visualization of cloud removal results for the Bay Area dataset; pseudocolor composed of bands 55, 30, and 15. The color bar gives the scale for the error maps. Green squares detail those areas that we need to observe.  $\alpha_k = 1$  for both HaLRTC and MGLRTA; MGLRTA under two cases: case 1:  $\gamma_{c,r,b,t} = 1e-3$  and case 2:  $\gamma_{c,r} = 1e-1$  and  $\gamma_{b,t} = 1e-3$ .

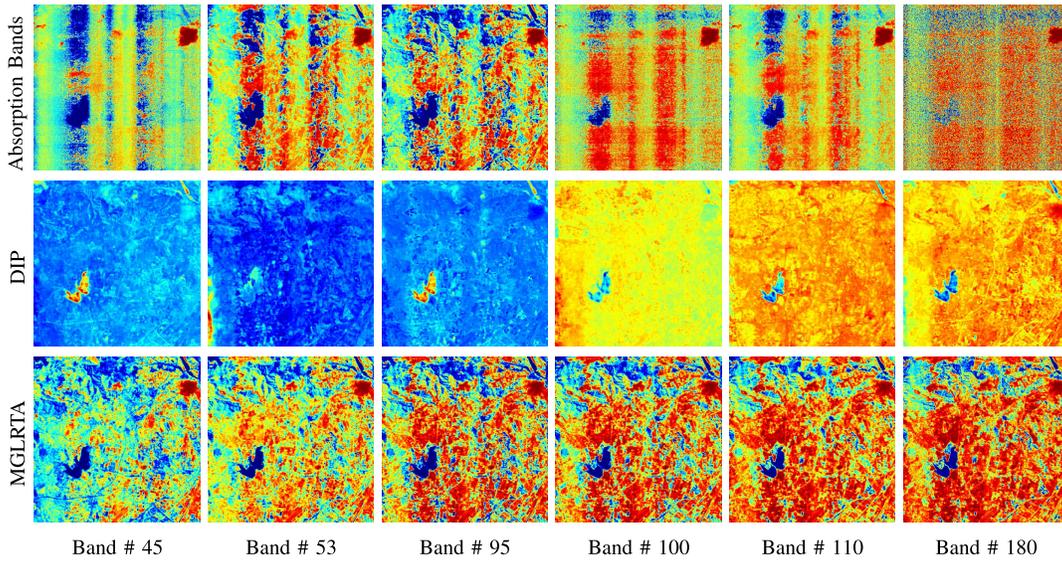


Fig. 10. Visual restoration results of the absorption bands for the Gaofen-5 dataset.

spectral wavelengths. The {42–58, 95–114, 176–180}th bands are water absorption bands needed to be recovered.

2) *Results Analysis*: As shown in Fig. 10, several atmospheric absorption bands are exhibited before and after restoration using the proposed MGLRTA and DIP. To restore the absorption bands,  $\Omega$  is given, wherein the elements in the entire {42–58, 95–114, 176–180}th bands equal to 0, which makes the observations of the known elements nonuniform and successively missing, while, in such a severe case, it is further observed that the proposed MGLRTA recovers the degraded image in a satisfactory visual assessment.

### E. Execution Time

Table VI summarizes the computational complexity of all comparison methods under consideration. All traditional experiments are carried out using MATLAB on an AMD

Ryzen 7 5800X 8-Core Processor with 3.80 GHz and 64 GB of RAM. DIP<sup>6</sup> is carried out using Pytorch<sup>7</sup> on 8-GB GPU of Nvidia Geforce RTX 3070 Ti. T3SC<sup>8</sup> is carried out using Pytorch on 12-GB GPU of Nvidia Geforce RTX 3080 Ti. Unlike DIP and other model-based methods, T3SC requires a training procedure with several hours for each kind of noise distribution, which is time-consuming. Thus, we do not compare it with the others here. The computational cost of TVTR is the highest, and the DIP ranks second. Computational costs of LRTV, LRTDGS, and HaLRT are much lower than other restoration techniques. In contrast, although costs the most time, TVTR may not achieve the best restoration performance because, without any training process, DIP's computational

<sup>6</sup><https://github.com/acecreamu/deep-hs-prior>

<sup>7</sup><https://pytorch.org/>

<sup>8</sup><https://github.com/inria-thoth/T3SC>

TABLE VI  
EXECUTION TIME (IN SECONDS) FOR THE TWO SYNTHETIC DATASETS

	Indian Pines			Pavia University		
	0.1	0.5	0.9	0.1	0.5	0.9
LRTV [58]	80.0	61.3	69.3	68.6	174.7	65.2
LRTDGS [15]	36.8	46.1	99.3	68.9	60.5	82.4
HaLRTC [34]	53.2	65.4	77.8	88.2	96.5	107.8
TVTR [13]	558.7	621.1	531.2	729.8	686.4	737.6
WLRTR [23]	150.3	145.5	154.2	180.2	179.9	181.5
DIP [30]	350.6	520.3	294.3	484.9	443.2	409.3
MGLRTA	105.3	137.8	129.5	146.0	174.7	160.3

time is greatly reduced. However, compared with the proposed MGLRTA, the DIP costs more than three times as much time as MGLRTA does. Overall, the proposed MGLRTA achieves satisfactory restoration results with relatively less computation than deep-learning-based DIP.

## V. CONCLUSIONS

In this article, an efficient MGLRTA algorithm for HSI restoration has been proposed, leveraging a graph regularized LRTA with the geometric side information learned from the graph signal defined on horizontal, lateral, and frontal slices of the reconstructed image tensor. The potential proximity induced by the low-rank constraint is explored and integrated with the LRTA to empower the reconstruction capability of the low-rank induced HSI restoration algorithm. The formulated convex nonsmooth optimization problem is solved with a well-posed iterative ADMM scheme. Synthetical and real data experiments are conducted to validate the proposed MGLRTA, which demonstrates that the integration of graph structure information and low rankness is effective in HSI restoration and can outperform the SOTA methods. In particular, the proposed MGLRTA is robust to noisy observations and harsh degradation cases.

## REFERENCES

- [1] P. Ghamisi *et al.*, "Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 37–78, Dec. 2017.
- [2] M. Zhang, W. Li, Q. Du, L. Gao, and B. Zhang, "Feature extraction for classification of hyperspectral and LiDAR data using patch-to-patch CNN," *IEEE Trans. Cybern.*, vol. 50, no. 1, pp. 100–111, Jan. 2020.
- [3] B. Rasti, Y. Chang, E. Dalsasso, L. Denis, and P. Ghamisi, "Image restoration for remote sensing: Overview and toolbox," 2021, *arXiv:2107.00557*.
- [4] N. Liu, W. Li, R. Tao, and J. E. Fowler, "Wavelet-domain low-rank/group-sparse destriping for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 10310–10321, Dec. 2019.
- [5] R. Pande-Chhetri and A. Abd-Elrahman, "De-stripping hyperspectral imagery using wavelet transform and adaptive frequency domain filtering," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 5, pp. 620–636, 2011.
- [6] L. Zhuang, X. Fu, M. K. Ng, and J. M. Bioucas-Dias, "Hyperspectral image denoising based on global and nonlocal low-rank factorizations," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 12, pp. 10438–10454, Dec. 2021.
- [7] 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.
- [8] 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. 176–188, Jan. 2016.
- [9] Y. Chang, L. Yan, B. Chen, S. Zhong, and Y. Tian, "Hyperspectral image restoration: Where does the low-rank property exist," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 8, pp. 6869–6884, Aug. 2021.
- [10] T. Hu, W. Li, N. Liu, R. Tao, F. Zhang, and P. Scheunders, "Hyperspectral image restoration using adaptive anisotropy total variation and nuclear norms," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1516–1533, Feb. 2021.
- [11] T. Xie, S. Li, and B. Sun, "Hyperspectral images denoising via non-convex regularized low-rank and sparse matrix decomposition," *IEEE Trans. Image Process.*, vol. 29, pp. 44–56, 2020.
- [12] 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.
- [13] W. He, N. Yokoya, L. Yuan, and Q. Zhao, "Remote sensing image reconstruction using tensor ring completion and total variation," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 8998–9009, Nov. 2019.
- [14] 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.
- [15] Y. Chen, W. He, N. Yokoya, and T.-Z. Huang, "Hyperspectral image restoration using weighted group sparsity-regularized low-rank tensor decomposition," *IEEE Trans. Cyber.*, vol. 50, no. 8, pp. 3556–3570, Aug. 2019.
- [16] Y. Wang, J. Peng, Q. Zhao, Y. Leung, X.-L. Zhao, and D. Meng, "Hyperspectral image restoration via total variation regularized low-rank tensor decomposition," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1227–1243, Apr. 2018.
- [17] Y. Chen, W. He, N. Yokoya, T.-Z. Huang, and X.-L. Zhao, "Nonlocal tensor-ring decomposition for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 2, pp. 1348–1362, Oct. 2019.
- [18] Y. Zheng, T. Huang, X. 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.
- [19] W. He *et al.*, "Non-local meets global: An iterative paradigm for hyperspectral image restoration," 2020, *arXiv:2010.12921*.
- [20] J. Xu, L. Zhang, W. Zuo, D. Zhang, and X. Feng, "Patch group based nonlocal self-similarity prior learning for image denoising," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 244–252.
- [21] J.-L. Lisani and J.-M. Morel, "Exploring patch similarity in an image," *Image Process. Line*, vol. 11, pp. 284–316, Sep. 2021.
- [22] T. Xie, S. Li, L. Fang, and L. Liu, "Tensor completion via non-local low-rank regularization," *IEEE Trans. Cybern.*, vol. 49, no. 6, pp. 2344–2354, Jun. 2019.
- [23] Y. Chang, L. Yan, X. Zhao, H. Fang, Z. Zhang, and S. Zhong, "Weighted low-rank tensor recovery for hyperspectral image restoration," *IEEE Trans. Cybern.*, vol. 50, no. 11, pp. 4558–4572, Nov. 2020.
- [24] J. Xue, Y. Zhao, W. Liao, and J. C.-W. Chan, "Nonlocal low-rank regularized tensor decomposition for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 5174–5189, Jul. 2019.
- [25] 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.
- [26] Y. Chang, L. Yan, and S. Zhong, "Hyper-Laplacian regularized unidirectional low-rank tensor recovery for multispectral image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4260–4268.
- [27] N. Yair and T. Michaeli, "Multi-scale weighted nuclear norm image restoration," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3165–3174.
- [28] W. Cai, J. Jiang, and S. Ouyang, "Hyperspectral image denoising using adaptive weight graph total variation regularization and low-rank matrix recovery," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
- [29] 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.
- [30] O. Sidorov and J. Y. Hardeberg, "Deep hyperspectral prior: Single-image denoising, inpainting, super-resolution," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019, pp. 3844–3851.

- [31] Q. Zhang, Q. Yuan, J. Li, X. Liu, H. Shen, and L. Zhang, "Hybrid noise removal in hyperspectral imagery with a spatial-spectral gradient network," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 7317–7329, Dec. 2019.
- [32] H. V. Nguyen, M. O. Ulfarsson, and J. R. Sveinsson, "Hyperspectral image denoising using SURE-based unsupervised convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 4, pp. 3369–3382, Apr. 2021.
- [33] P. Azimpour, T. Bahraïni, and H. S. Yazdi, "Hyperspectral image denoising via clustering-based latent variable in variational Bayesian framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 4, pp. 3266–3276, Apr. 2021.
- [34] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208–220, Jan. 2013.
- [35] Z. Gong, W. Hu, X. Du, P. Zhong, and P. Hu, "Deep manifold embedding for hyperspectral image classification," *IEEE Trans. Cybern.*, early access, Apr. 19, 2021, doi: [10.1109/TCYB.2021.3069790](https://doi.org/10.1109/TCYB.2021.3069790).
- [36] H. Huang, G. Shi, H. He, Y. Duan, and F. Luo, "Dimensionality reduction of hyperspectral imagery based on spatial-spectral manifold learning," *IEEE Trans. Cybern.*, vol. 50, no. 6, pp. 2604–2616, Jun. 2020.
- [37] N. Liu, W. Li, and Q. Du, "Unsupervised feature extraction for hyperspectral imagery using collaboration-competition graph," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 6, pp. 1491–1503, Dec. 2018.
- [38] X. Lu, Y. Wang, and Y. Yuan, "Graph-regularized low-rank representation for destriping of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 4009–4018, Jul. 2013.
- [39] E. J. Candès and T. Tao, "The power of convex relaxation: Near-optimal matrix completion," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2053–2080, May 2010.
- [40] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, May 2013.
- [41] A. Sandryhaila and J. M. F. Moura, "Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 80–90, Sep. 2014.
- [42] V. Kalofolias, X. Bresson, M. Bronstein, and P. Vandergheynst, "Matrix completion on graphs," 2014, [arXiv:1408.1717](https://arxiv.org/abs/1408.1717).
- [43] E. J. Candès and B. Recht, "Exact low-rank matrix completion via convex optimization," in *Proc. 46th Annu. Allerton Conf. Commun., Control, Comput.*, Sep. 2008, pp. 806–812.
- [44] A. Sharma and M. Ovsjanikov, "Matrix decomposition on graphs: A functional view," 2021, [arXiv:2102.03233](https://arxiv.org/abs/2102.03233).
- [45] F. Monti, M. M. Bronstein, and X. Bresson, "Geometric matrix completion with recurrent multi-graph neural networks," 2017, [arXiv:1704.06803](https://arxiv.org/abs/1704.06803).
- [46] W. Dai, E. Kerman, and O. Milenkovic, "A geometric approach to low-rank matrix completion," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 237–247, Jan. 2012.
- [47] R. van den Berg, T. N. Kipf, and M. Welling, "Graph convolutional matrix completion," 2017, [arXiv:1706.02263](https://arxiv.org/abs/1706.02263).
- [48] R. Dian, S. Li, and L. Fang, "Learning a low tensor-train rank representation for hyperspectral image super-resolution," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 9, pp. 2672–2683, Sep. 2019.
- [49] Y.-Y. Liu, X.-L. Zhao, Y.-B. Zheng, T.-H. Ma, and H. Zhang, "Hyperspectral image restoration by tensor fibered rank constrained optimization and plug-and-play regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–17, 2022.
- [50] G. Bergqvist and E. G. Larsson, "The higher-order singular value decomposition: Theory and an application [lecture notes]," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 151–154, May 2010.
- [51] H. Cai, V. W. Zheng, and K. C. Chang, "A comprehensive survey of graph embedding: Problems, techniques, and applications," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 9, pp. 1616–1637, Feb. 2018.
- [52] S. Yan, D. Xu, B. Zhang, H.-J. Zhang, Q. Yang, and S. Lin, "Graph embedding and extensions: A general framework for dimensionality reduction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 1, pp. 40–51, Jan. 2007.
- [53] C. M. Bachmann, T. L. Ainsworth, and R. A. Fusina, "Exploiting manifold geometry in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 441–454, Mar. 2005.
- [54] S. Dong, P.-A. Absil, and K. A. Gallivan, "Graph learning for regularized low-rank matrix completion," in *Proc. 23rd Int. Symp. Math. Theory Netw. Syst. (MTNS)*, 2018, pp. 1–8.
- [55] Z. Kang, H. Pan, S. C. H. Hoi, and Z. Xu, "Robust graph learning from noisy data," *IEEE Trans. Cybern.*, vol. 50, no. 5, pp. 1833–1843, May 2020.
- [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, Feb. 2010.
- [57] G. Golub and C. Van Loan, *Matrix Computations*, vol. 3. Baltimore, MD, USA: Johns Hopkins Univ. Press, 2012.
- [58] 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.
- [59] V. Lempitsky, A. Vedaldi, and D. Ulyanov, "Deep image prior," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 9446–9454.
- [60] T. Bodrito, A. Zouaoui, J. Chanussot, and J. Mairal, "A trainable spectral-spatial sparse coding model for hyperspectral image restoration," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 34, 2021, pp. 5430–5442.
- [61] L. Wald, *Data Fusion: Definitions and Architectures*. Paris, France: Presses de l'École des Mines de Paris, 2002.
- [62] 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.
- [63] Y. Chen, W. He, N. Yokoya, and T.-Z. Huang, "Blind cloud and cloud shadow removal of multitemporal images based on total variation regularized low-rank sparsity decomposition," *ISPRS J. Photogramm. Remote Sens.*, vol. 157, pp. 93–107, Nov. 2019.
- [64] M. K. Griffin and H. K. Burke, "Compensation of hyperspectral data for atmospheric effects," *Lincoln Lab. J.*, vol. 14, no. 1, pp. 29–54, 2003.
- [65] Y.-N. Liu *et al.*, "The advanced hyperspectral imager: Aboard China's GaoFen-5 satellite," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 4, pp. 23–32, Dec. 2019.



include remote sensing feature extraction.

**Na Liu** (Member, IEEE) received the B.E. degree in electronic information engineering and the Ph.D. degree in control science and engineering from the Beijing University of Chemical Technology, Beijing, China, in 2016 and 2021, respectively.

She was also a Visiting Ph.D. Student with the University of Grenoble Alpes, Grenoble, France, funded by the Chinese Scholarship Council. She is currently a Post-Doctoral Researcher with the School of Information and Electronics, Beijing Institute of Technology, Beijing. Her research interests include remote sensing data analysis, hyperspectral data enhancement, and feature extraction.



**Wei Li** (Senior Member, IEEE) received the B.E. degree in telecommunications engineering from Xidian University, Xi'an, China, in 2007, the M.S. degree in information science and technology from Sun Yat-sen University, Guangzhou, China, in 2009, and the Ph.D. degree in electrical and computer engineering from Mississippi State University, Starkville, MS, USA, in 2012.

Subsequently, he spent one year as a Post-Doctoral Researcher at the University of California at Davis, Davis, CA, USA. He is currently a Professor with the School of Information and Electronics, Beijing Institute of Technology, Beijing, China. His research interests include hyperspectral image analysis, pattern recognition, and target detection.

Dr. Li received the JSTARS Best Reviewer in 2016 and the TGRS Best Reviewer Award in 2020 from the IEEE Geoscience and Remote Sensing Society (GRSS) and the Outstanding Paper Award at the IEEE International Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (Whispers) in 2019. He has served as an Associate Editor for the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (JSTARS) and IEEE SIGNAL PROCESSING LETTERS (SPL). He is also serving as an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS).



**Ran Tao** (Senior Member, IEEE) received the B.S. degree from the Electronic Engineering Institute of PLA, Hefei, China, in 1985, and the M.S. and Ph.D. degrees from the Harbin Institute of Technology, Harbin, China, in 1990 and 1993, respectively.

He was a Senior Visiting Scholar with the University of Michigan, Ann Arbor, MI, USA, and the University of Delaware, Newark, DE, USA, in 2001 and 2016, respectively. He was a Chief Professor of the Program for Changjiang Scholars and Innovative Research Team in University from 2010 to 2012.

He has been a Chief Professor of the Creative Research Groups with the National Natural Science Foundation of China, Beijing, China, since 2014. He is currently a Professor with the School of Information and Electronics, Beijing Institute of Technology, Beijing. His research interests include fractional Fourier transform and its applications, theory, and technology for radar and communication systems.

Dr. Tao is also a fellow of the Institute of Engineering and Technology and the Chinese Institute of Electronics. He was a recipient of the National Science Foundation of China for Distinguished Young Scholars in 2006 and a Distinguished Professor of the Changjiang Scholars Program in 2009. He is also the Vice-Chair of the IEEE China Council.



**Qian Du** (Fellow, IEEE) received the Ph.D. degree in electrical engineering from the University of Maryland, Baltimore, MD, USA, in 2000.

She is currently the Bobby Shackouls Professor with the Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS, USA. Her research interests include hyperspectral remote sensing image analysis and applications, pattern classification, data compression, and neural networks.

Dr. Du is also a fellow of the SPIE-International Society for Optics and Photonics. She received the 2010 Best Reviewer Award from the IEEE Geoscience and Remote Sensing Society. She was the Co-Chair of the Data Fusion Technical Committee of the IEEE Geoscience and Remote Sensing Society from 2009 to 2013 and the Chair of the Remote Sensing and Mapping Technical Committee of the International Association for Pattern Recognition from 2010 to 2014. She is also the General Chair of the 4th IEEE GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, Shanghai, China, in 2012. She has served as an Associate

Editor for the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, the *Journal of Applied Remote Sensing*, and the IEEE SIGNAL PROCESSING LETTERS. Since 2016, she has been the Editor-in-Chief of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING.



**Jocelyn Chanussot** (Fellow, IEEE) received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (Grenoble INP), Grenoble, France, in 1995, and the Ph.D. degree from the Université de Savoie, Annecy, France, in 1998.

Since 1999, he has been with Grenoble INP, where he is currently a Professor of signal and image processing. He has been a Visiting Scholar with Stanford University, Stanford, CA, USA, the KTH Royal Institute of Technology, Stockholm, Sweden, and the National University of Singapore (NUS), Singapore. Since 2013, he has been an Adjunct Professor with the University of Iceland, Reykjavik, Iceland. From 2015 to 2017, he was a Visiting Professor with the University of California at Los Angeles (UCLA), Los Angeles, CA. He holds the AXA Chair in remote sensing and is an Adjunct Professor with the Chinese Academy of Sciences, Aerospace Information Research Institute, Beijing, China. His research interests include image analysis, hyperspectral remote sensing, data fusion, machine learning, and artificial intelligence.

Dr. Chanussot was a member of the Institut Universitaire de France from 2012 to 2017 and a Highly Cited Researcher (Clarivate Analytics/Thomson Reuters). He was a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society from 2006 to 2008 and the Program Chair of the IEEE International Workshop on Machine Learning for Signal Processing in 2009. He is also the Founding President of the IEEE Geoscience and Remote Sensing French Chapter, which received the 2010 IEEE GRS-S Chapter Excellence Award. He received multiple outstanding paper awards. He was the Co-Chair and the Chair of the GRS Data Fusion Technical Committee from 2005 to 2008 and 2009 to 2011, respectively. He is also the General Chair of the first IEEE GRSS Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote sensing (WHISPERS). He was the Vice-President of the IEEE Geoscience and Remote Sensing Society (GRSS), in charge of meetings and symposia from 2017 to 2019. He is also an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE TRANSACTIONS ON IMAGE PROCESSING, and the PROCEEDINGS OF THE IEEE.