Graph-Based Logarithmic Low-Rank Tensor Decomposition for the Fusion of Remotely Sensed Images

Fei Ma, Shuai Huo, and Feixia Yang

***Abstract*—Hyperspectral images with high spatial resolution play an important role in material classification, change detection, and others. However, owing to the limitation of imaging sensors, it is difficult to directly acquire images with both high spatial resolution and high spectral resolution. Therefore, the fusion of remotely sensed images is an effective way to obtain high-resolution desired data, which is usually an ill-posed inverse problem and susceptible to noise corruption. To address these issues, a low-rank model based on tensor decomposition is proposed to fuse hyperspectral and multispectral images by incorporating graph regularization, in which the logarithmic low-rank function is utilized to suppress the small components for denoising. Furthermore, this article takes advantage of the local spatial similarity of remotely sensed images to enhance the reconstruction performance by constructing spatial graphs, and also promotes signature smoothing between adjacent endmember spectra using the neighborhood-based spectral graph regularization. Finally, a set of efficient solvers is carefully designed via alternating optimization for closed-from solutions and computational reduction, in which vector-matrix operators are adapted to solve the 3-D core tensor. Experimental tests on several real datasets illustrate that the proposed fusion method yields better reconstruction performance than the current state-of-the-art methods, and can significantly suppress noise at the same time.**

***Index Terms*—Graph regularization, hyperspectral image (HSI) super-resolution, image fusion, low rank, tensor decomposition.**

I. Introduction

With the rapid development of imaging sensors in remote sensing, hyperspectral images (HSIs) are widely used in many fields, such as objective classification and target detection [1], [2]. However, they are susceptible to noise corruption and spectral variability [3], which seriously affect the image quality. Generally speaking, the common remotely sensed data include panchromatic (PAN), multispectral, and hyperspectral images. PAN images have high spatial resolution but only one spectral band. Furthermore, multispectral images (MSIs) usually take several to a dozen bands and relatively high spatial resolution, while HSIs occupy dozens to several hundred bands but low spatial resolution. Remote sensing images with both high spatial resolution and high spectral resolution can perform material identification and classification more accurately. However, owing to the limitation of mutual constraint between spatial and spectral resolution, the images by remote sensing satellites can hardly be obtained in both high spatial resolution and high spectral resolution at the same time [1], [4]. Therefore, image fusion is an effective way to enhance the spatial resolution of hyperspectral images, which has become a research hotspot in remote sensing.

So far, there have been various research results on the fusion of HSIs and MSIs, which can be roughly divided into three categories from the perspective of signal models, including matrix factorization, deep learning, and tensor decomposition. The representative approaches based on matrix factorization are panchromatic sharpening [5], Bayesian inference [6], and coupled nonnegative matrix factorization (CNMF) [7]. Such methods usually need to reshape the original cubic data into matrix form, and then extract spectral and spatial signatures separately from the observed data for image reconstruction. Obviously, data fusion is an ill-posed inverse problem, which can be solved by the regularization method. Therefore, the regularized terms of low-rankness [8], sparsity [9], smoothness [10], and minimum-volume [11] were developed to improve the quality of reconstruction images. For example, Veganzones *et al.* [12] proposed a low-rank super-resolution method to partition the image into subgraphs and solve the data fusion problem of each subgraph independently in the low-dimensional subspace by exploiting the local low-rank property of real datasets. Zare *et al.* [13] brought forward a fusion model based on smoothed graph signals, and took the clustering method to redefine the spatial similarity of HSIs. In general, matrix factorization-based methods inevitably have to unfold the original 3-D data into matrices, so that the structure information of the observed HSI is lost, resulting in a certain degree of spatial and/or spectral distortion of the reconstructed image. With the development of artificial intelligence, some algorithms based on deep neural networks have emerged. Deep learning-based fusion models were designed to train the observation data and hyperparameters, and then reconstruct the super-resolution images using

Manuscript received August 8, 2021; revised September 25, 2021; accepted October 12, 2021. Date of publication October 27, 2021; date of current version November 15, 2021. This work was supported in part by the Scientific Research Project of Colleges from Liaoning Department of Education (P.R.C) under Grant LJKZ0357, Grant LJ2019QL006, and Grant LJ2019JL022, in part by the NSFC General Project under Grant 61971210, and in part by the Discipline Innovation Team of Liaoning Technical University under Grant LNTU20TD-20 and Grant LNTU20TD-25. *(Corresponding author: Shuai Huo.)*

Fei Ma and Shuai Huo are with the School of Electronic and Information Engineering, Liaoning Technical University, Huludao 125105, China (e-mail: femircom@gmail.com; jingmengzhiyue@gmail.com).

Feixia Yang is with the School of Electrical and Control Engineering, Liaoning Technical University, Huludao 125105, China (e-mail: yangfx091011@gmail.com).

Digital Object Identifier 10.1109/JSTARS.2021.3123466

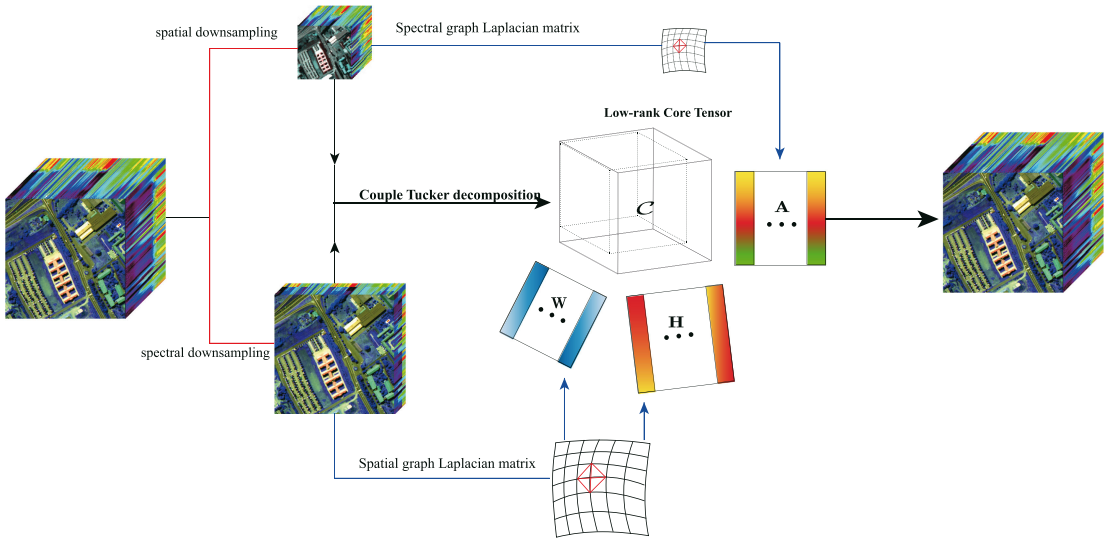

Fig. 1. Flowchart of the proposed HSI-MSI fusion based on low-rank Tucker decomposition with graph regularization.

the related constraint conditions [14], [15]. For instance, Yao *et al.* [16] developed an unsupervised cross-attentive neural network that performed feature extraction via two convolutional networks with shared parameters for resolution enhancement. Hong *et al.* [17], [18] proposed a minibatch graph convolutional network (miniGCN) to drop the computational cost by training large-scale GCNs in a minibatch fashion, and also developed a general multimodal deep learning framework. Compared with matrix factorization, deep learning-based approaches can address nonconvex models in remote sensing [19]. However, these methods generally require the establishment of training samples (i.e., image elements or features), and involve the adjustment of a large number of hyperparameters. Meanwhile, the performance of network model is often limited by the size of datasets.

Another popular methods can achieve the HSI-MSI fusion via tensor decomposition without destroying the original data structure. The common forms of tensor decomposition include Canonical polyadic decomposition (CPD) [20] and Tucker decomposition [21]. For example, Kanatsoulis *et al.* [22] proposed a CPD-based coupled tensor decomposition model and performed the fusion performance in the case where the degrading operator was unknown. CPD with rank-1 property is a special case of Tucker decomposition, so the Tucker-based tensor model has received more attention due to its universality. However, the rank of a tensor has multiple different definitions and is not exactly equal in value, which is different from that of a matrix [23]. Imposing prior information or regularization may address the nonuniqueness of tensor rank. For instance, Bu *et al.* [24] brought forward a Tucker-based fusion model by exploiting the smoothness of factor matrices and the sparsity of core tensor. Zhang *et al.* [25] designed a super-resolution method based on Tucker tensor decomposition and joint spatial-spectral-graph regularization. Ma *et al.* [21] put forward a Tucker-based model by imposing low-rank and sparsity regularization on factor matrices and core tensor, respectively. Since Tucker decomposition can keep the 3-D data cube unchanged during processing, there is no loss of structural information.

To utilize the advantages of tensor models, this article incorporates low-rank and graph regularizers into Tucker decomposition to redefine the fusion of HSIs and MSIs as an optimization problem. Our contributions to this article mainly include the following aspects. First, owing to local similarity in spatial and spectral dimensions, the logarithmic sum function is leveraged to promote the low-rankness. Second, the graph matrices in the spatial and spectral subspaces are constructed respectively to smooth the spatial-spectral signatures for the removal of redundant information and spectral variability. Finally, an efficient algorithm is carefully developed to obtain the closed-form solution to each variable using alternative optimization (AO) [26] and alternating direction multiplier method (ADMM) [27], especially vector-matrix operators for the core tensor. The flowchart of simulations is given in Fig. 1, and the experimental results on several widely used real datasets show that our proposed fusion method performs better performance than the state-of-the-art approaches, which also verifies the effectiveness of incorporated regularization for super-resolution image reconstruction.

The rest is organized as follows. Section II presents the observation model and regularization terms. In Section III, we redefine the fusion method as a regularized optimization problem and design a set of efficient solvers to get the closed-form solutions via ADMM. Then, the experimental tests are conducted on different datasets in Section IV. Finally, Section V concludes the articles.

In addition, two proofs are given in the appendix. Some key notations are defined as follows. Scalar, vector, and matrix are represented as lower case letters, bold lower case letters, and bold upper case letters, respectively. The Nth order tensor is given as $\mathcal{A} \in \mathbb{R}^{I_1 \times \ldots \times I_k \times \ldots \times I_N}$, the element of which is $a_{i_1 \ldots i_k \ldots i_N} (1 \leq i_k \leq I_K)$. vec$(\mathcal{A})$ indicates a vectorization of tensor variable $\mathcal{A}$. The matrices $\mathbf{A}_{(n)} \in \mathbb{R}^{I_n \times I_1 I_2, \ldots, I_{n-1} I_{n+1}, \ldots, I_N}$ and $\mathbf{A}_{<n>} \in$

$\mathbb{R}^{I_1 I_2,\ldots,I_n \times I_{n+1},\ldots,I_N}$ denote the mode-n unfolding and the mode-n canonical unfolding of the tensor $\mathcal{A}$, respectively.

II. SIGNAL MODELS AND REGULARIZATION TERMS

This section presents the observed models and Tucker decomposition, and then reformulates the fusion of remotely sensed images as a regularized optimization problem. At the same time, the logarithmic low-rank and graph regularizers are given in the following sections.

A. Observation Models and Tucker Decomposition

In this article, a desired HSI (also called target or reference image) with high spatial resolution is denoted as $\mathcal{Z} \in \mathbb{R}^{N_w \times N_h \times N_a}$, where N_w, N_h, and N_a represent the width, height, and spectral bands of $\mathcal{Z}$, respectively. The observed HSI $\mathcal{Y} \in \mathbb{R}^{n_w \times n_h \times N_a}$ has low spatial resolution, where n_w and n_h are the width and height, respectively, and the number of spectral bands N_a is the same in $\mathcal{Z}$. The observed MSI $\mathcal{X} \in \mathbb{R}^{N_w \times N_h \times n_a}$ has high spatial resolution, where n_a is the number of spectral bands, and the spatial resolution $N_w \times N_h$ is the same in $\mathcal{Z}$. Image fusion aims to reconstruct the hyperspectral super-resolution image by integrating the observed hyperspectral and multispectral data. Conversely, these two observation images can be viewed as the degraded versions of the reference image separately in spatial and spectral subspaces, as shown in Fig. 1. In this way, the tensor-based observation model can be expressed as

$$\mathcal{Y} = \mathcal{Z} \times_1 \mathbf{P}_1 \times_2 \mathbf{P}_2 + \mathcal{E}_y \tag{1}$$

$$\mathcal{X} = \mathcal{Z} \times_3 \mathbf{P}_3 + \mathcal{E}_x \tag{2}$$

where $r = \frac{N_w}{n_w} = \frac{N_h}{n_h}$ represents the degradation ratio in the spatial dimensions; $\mathbf{P}_1 \in \mathbb{R}^{n_w \times N_n}$ and $\mathbf{P}_2 \in \mathbb{R}^{n_h \times N_h}$ denote the degradation matrices in the width and height, which can be constructed by downsampling the corresponding Toeplitz matrix from the point spread function (PSF), respectively; $\mathbf{P}_3 \in \mathbb{R}^{n_a \times N_a}$ denote the degradation matrices in the spectral dimensions, which can be constructed by the spectral response functions (SFRs) of hyperspectral and multispectral sensors; $\mathcal{E}_y$ and $\mathcal{E}_x$ are the error terms.

According to the definition of Tucker decomposition [28], the target image can be decomposed as

$$\mathcal{Z} = \mathcal{C} \times_1 \mathbf{W} \times_2 \mathbf{H} \times_3 \mathbf{A} + \mathcal{E}_z \tag{3}$$

where $\mathbf{W} \in \mathbb{R}^{N_w \times r_w}$, $\mathbf{H} \in \mathbb{R}^{N_h \times r_h}$, and $\mathbf{A} \in \mathbb{R}^{N_a \times r_a}$ denote the factor matrices of width, height, and spectrum, respectively; $\mathcal{C} \in \mathbb{R}^{r_w \times r_h \times r_a}$ is the decomposed core tensor; r_w, r_h, and r_a represent three dimensions of the core tensor; $\mathcal{E}_z$ means the error term.

B. Fusion Model

Combining the observation models (1) and (2) [29], HSI-MSI fusion can be reformulated as

$$\begin{aligned} \min_{\mathcal{Z}} \quad & \frac{1}{2}||\mathcal{Y} - \mathcal{Z} \times_1 \mathbf{P}_1 \times_2 \mathbf{P}_2||_F^2 + \frac{1}{2}||\mathcal{X} - \mathcal{Z} \times_3 \mathbf{P}_3||_F^2 \\ \text{s.t.} \quad & \mathcal{Z} = \mathcal{C} \times_1 \mathbf{W} \times_2 \mathbf{H} \times_3 \mathbf{A}. \end{aligned} \tag{4}$$

Since image reconstruction is an ill-posed inverse problem, this article incorporates prior information to regularize the model (4). In hyperspectral data, each group of adjacent bands has close correlation, and the neighboring pixels also have strong similarities. So imposing graph regularization on spatial and spectral matrices separately aims to reduce the impacts of distortions for performance enhancement. Instead of the accurate value of tensor rank, we exert the low-rank regularization on the core tensor for denoising and addressing the nonuniqueness of tensor rank. Thus, (4) can be rewritten as

$$\begin{aligned} \min_{\mathcal{Z}} \quad & \frac{1}{2}||\mathcal{Y} - \mathcal{Z} \times_1 \mathbf{P}_1 \times_2 \mathbf{P}_2||_F^2 + \frac{1}{2}||\mathcal{X} - \mathcal{Z} \times_3 \mathbf{P}_3||_F^2 \\ & + \alpha\phi_a(\mathcal{C}) + \beta\phi_b(\mathbf{W}, \mathbf{H}) + \gamma\phi_c(\mathbf{A}) \\ \text{s.t.} \quad & \mathcal{Z} = \mathcal{C} \times_1 \mathbf{W} \times_2 \mathbf{H} \times_3 \mathbf{A} \end{aligned} \tag{5}$$

where $\phi_a(\mathcal{C})$ denotes the low-rank regularizer of the core tensor; $\phi_b(\mathbf{W}, \mathbf{H})$ and $\phi_c(\mathbf{A})$ indicate the spatial and spectral graph regularization terms, respectively; α, β, and γ denote the corresponding weights. As shown in Fig. 1, the HSR hyperspectral data can be reconstructed by the factor matrices and core tensor under the low-rank regularization on basis of Tucker decomposition.

C. Regularization Terms

1) Logarithmic Low-Rank Function: Generally, the rank of a matrix $\mathbf{M}$ is equal to the number of positive singular values, but it is difficult to solve directly due to nonconvexity. In order to address this issue, the nuclear norm is often a convex surrogate of low rank, and it is equal to the sum of singular values of $\mathbf{M}$.

A recent study [30] has shown that convex relaxations for low-rank problems can be adapted to the more general case, i.e., the logarithmic sum function can be more efficient in strengthening the low-rankness of a matrix. It is found that the larger singular values of the matrix represent dominant features or structural information, while the smaller singular values mainly contain noise or interference. The logarithmic sum function [31], [32] is defined as

$$\text{logsum}(\mathbf{M}) = \sum_i \log\left(\sigma_i(\mathbf{M}) + \varepsilon\right) \tag{6}$$

where ε is a small positive number to avoid the case of zero values. We can observe that the logarithmic function can further shrink the matrix rank to preserve the structural information and suppress the noise in the smaller singular values. Let $0 < \alpha, 0 < \varepsilon < \min(\sqrt{\alpha}, (\alpha/\sigma_1))$ and define the low-rank recovery problem as

$$\min_{\mathbf{M}} \quad \frac{1}{2}\|\mathbf{\Gamma} - \mathbf{M}\|_F^2 + \alpha \sum_i \log\left(\sigma_i(\mathbf{M}) + \varepsilon\right) \tag{7}$$

$\mathbf{\Gamma}$ denotes a known observation matrix, we can obtain the optimal solution as

$$\mathbf{M} = \mathbf{U}\text{diag}\left(d_1, d_2, \ldots, d_n\right)\mathbf{V} \tag{8}$$

where $d_i = \mathrm{D}_{\alpha,\varepsilon}(\sigma_i)$, and the operator $\mathrm{D}_{\alpha,\varepsilon}(\sigma_i)$ is defined as

$$\mathrm{D}_{\alpha,\varepsilon}(x) = \begin{cases} 0 & c_2 \le 0 \\ \frac{c_1 + \sqrt{c_2}}{2} & c_2 > 0 \end{cases} \tag{9}$$

where $c_1 = |x| - \varepsilon$ and $c_2 = c_1^2 - 4(\alpha - \varepsilon|x|)$. Among the various definitions of tensor rank, the tensor-train rank is adopted in this article.

2) Tensor-Train Rank: Given an N-dimensional tensor $\mathcal{G} \in \mathbb{R}^{I_1 \times I_2, \ldots, \times I_N}$, and decompose it into a tensor train to obtain $\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \ldots, \mathcal{G}^{(N)}$, where $\mathcal{G}^{(n)} \in \mathbb{R}^{R^{n-1} \times I_n \times R^n}$ and $R^0 = R^N = 1$ $(n = 1, 2, \ldots, N)$. The elements in $\mathcal{G}$ are expressed as

$$\mathcal{G}_{i_1, i_2, \ldots, i_N} = \sum_{r_1, r_2, \ldots, r_{N-1}=1}^{R_1, R_2, \ldots, R_{N-1}} \mathcal{G}^{(1)}_{1, i_1, r_1} \mathcal{G}^{(2)}_{r_1, i_2, r_2}, \ldots, \mathcal{G}^{(N)}_{r_{N-1}, i_N, 1} \tag{10}$$

where $\{R_1, R_2, \ldots, R_{N-1}\}$ is called the tensor-train rank of $\mathcal{G}$ [33]. However, the train decomposition of a tensor is also not unique due to the different parameters [34]. In practice, the lower bound of tensor-train rank is more interesting, especially for 3-D hyperspectral data.

Proposition 1: If the mode-n canonical unfolding of an N-dimensional tensor $\mathcal{G}^n$ is $\mathbf{G}_{<n>}$, then the tensor-train rank R_n of $\mathcal{G}^n$ satisfies $R_n \geq T_n$, where $T_n = \text{rank}(\mathbf{G}_{<n>})$.

More specifically, the rank of $\mathbf{G}_{<n>}$ is the lower bound of tensor-train rank R_n. The detailed proof is given in Appendix V. We employ the lower bound to denote the tensor-train rank in low-rank problems. Thus, the tensor-train rank is defined as

$$\sum_{n=1}^{N-1} \alpha_n \text{rank}\left(\mathbf{G}_{\langle n \rangle}\right) \tag{11}$$

where α_n is the weight and satisfies $\sum_{n=1}^{N-1} \alpha_n = 1$. The matrix $\mathbf{G}_{<n>}$ can be obtained by unfolding the tensor along the first n modes, so that its rank can model the correlation between the first n modes and next $N - n$ modes. Owing to the structure of hyperspectral datacube, this article employs the mode-2 canonical unfolding to partition the spatial and spectral information into different subspaces [35], with the corresponding weight $\alpha_2 = 1$. Incorporating the logarithmic sum function with the tensor-train rank, the low-rank regularization on core tensor can be expressed as

$$\phi_a(\mathcal{C}) = \sum_i \log\left(\sigma_i(\mathbf{C}_{<2>}) + \varepsilon\right). \tag{12}$$

3) Spatial Graph Regularizer: It is well known that the observed hyperspectral images in a low-dimensional subspace contain high redundancy [36]–[38]. To further drop the redundancy, this article takes advantage of the manifold information of pixels in the MSI [39] and constructs a spatial graph to evaluate the local correlation in the pixel subspace, which can transfer the spatial information from the observed MSI to the fused image. Thus, this can maintain the spatial structure better in the reconstructed image. The spatial graph is represented as $\mathbf{G}_W = (\mathbf{V}_W, \mathbf{E}_W, \mathbf{S}_W)$, where $\mathbf{V}_W$ denotes the vertexes (i.e., the pixels in the MSI), $\mathbf{E}_W$ represents the corresponding set of edges, and $\mathbf{S}_W$ involves the connection weights between vertexes. Thus, the adjacency matrix $\mathbf{S}_W$ can be calculated by

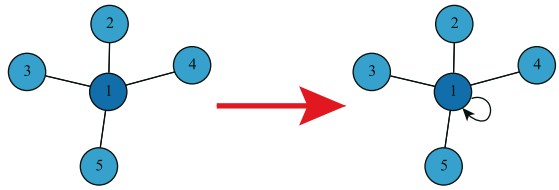

Fig. 2. Augmented graph Laplacian matrix.

the following:

$$\mathbf{S}_W(i,j) = \exp\left(-\frac{\|\mathbf{m}(x_i, y_i, :) - \mathbf{m}(x_j, y_j, :)\|^2}{\sigma^2}\right) \tag{13}$$

where $\mathbf{m}(x_i, y_i, :)$ denotes the fiber of the MSI, and σ is the smoothing factor. From the spatial graph $\mathbf{G}_W$, the normalized spatial Laplacian matrix can be defined by

$$\widehat{\mathbf{L}}_W = \mathbf{D}_W^{-\frac{1}{2}}(\mathbf{D}_W - \mathbf{S}_W)\mathbf{D}_W^{-\frac{1}{2}} = \mathbf{D}_W^{-\frac{1}{2}}\mathbf{L}_W\mathbf{D}_W^{-\frac{1}{2}} \tag{14}$$

where $\mathbf{L}_W = \mathbf{D}_W - \mathbf{S}_W$ denotes the graph Laplacian matrix; $\mathbf{D}_W = \text{diag}(d_1, d_2, \ldots, d_n)$ is the degree matrix, and $d_i = \sum_j \mathbf{S}_{W_{i,j}}$; $\mathbf{I}_W$ is the unit matrix with proper sizes. According to the properties of normalized Laplacian matrix [40], it is a semi-positive definite matrix and its eigenvalues are between 0 and 2, which is prone to over-smoothing in multiple iterations. Therefore, in order to suppress the over-smoothing phenomenon, this article constructs an augmented Laplacian matrix by enhancing the self-connection, i.e., $\widetilde{\mathbf{L}}_W = \widehat{\mathbf{L}}_W + \mathbf{I}$. The physical meaning of the augmented Laplacian matrix is shown in Fig. 2, which is equivalent to adding a self-connected edge to each node on the original graph structure. Thus, the spatial graph regularizer [41] can be defined as

$$\phi_b(\mathbf{W}, \mathbf{H}) = \text{tr}\left((\mathbf{H} \otimes \mathbf{W})^T \widetilde{\mathbf{L}}_W (\mathbf{H} \otimes \mathbf{W})\right). \tag{15}$$

4) Spectral Graph Regularizer: As we all know, the endmember signature describes the spectral reflection of a material that varies gradually with the spectra, so it is usually a smooth curve. Similar to the endmember signature in the linear unmixing model [9], the third factor matrix also represents the spectral signature in the tensor decomposition. Thus, this article develops a neighborhood graph to regularize the spectral factor matrix for smoothing [42], [43]. Based on the nearest neighbor spectrum, the spectral graph can be defined as $\mathbf{G}_A = (\mathbf{V}_A, \mathbf{E}_A, \mathbf{S}_A)$, where $\mathbf{V}_A$ denotes the spectral bands, $\mathbf{E}_A$ is the set of edges between the neighborhood-based spectral bands, and $\mathbf{S}_A$ is the adjacency matrix.

The smoothing regularizer is often expressed as

$$\phi_c(\mathbf{A}) \triangleq \sum_{j=1}^{N} \sum_{\{m,n\} \in \epsilon} |\mathbf{a}_j(m) - \mathbf{a}_j(n)| \tag{16}$$

where ϵ denotes the column neighborhood set of endmember matrix, and $\mathbf{a}_j(m)$ denotes the mth element of the jth endmember $\mathbf{a}_j$. Thus, (16) can be reshaped into a neighborhood-based

adjacency matrix as

$$\mathbf{S}_A(m) = \exp\left(-\frac{\|\mathbf{A}(:,m) - \mathbf{A}(:,m+1)\|^2}{\sigma^2}\right) \tag{17}$$

where $\mathbf{A}(:,m)$ denotes the mth spectral signature of endmember matrix, and σ is the smoothing factor.

Similar to the spatial graph, the spectral Laplacian matrix can be obtained by

$$\widehat{\mathbf{L}}_A = \mathbf{D}_A^{-\frac{1}{2}} \mathbf{L}_A \mathbf{D}_A^{-\frac{1}{2}} \tag{18}$$

where $\mathbf{L}_A = \mathbf{D}_A - \mathbf{S}_A$ denotes the graph Laplacian matrix, $\mathbf{D}_A = \text{diag}(g_1, g_2, \ldots, g_n)$ is the degree matrix and $g_i = \sum_j \mathbf{S}_{A_{i,j}}$. Similarly, we can obtain the augmented Laplacian matrix $\widetilde{\mathbf{L}}_A = \widehat{\mathbf{L}}_A + \mathbf{I}$. Thus, the regularization expression of spectral graph can be rewritten as

$$\phi_c(\mathbf{A}) \triangleq \text{tr}(\mathbf{A}^T \widetilde{\mathbf{L}}_A \mathbf{A}). \tag{19}$$

III. Optimization Algorithm

After redefining the models and regularization terms, the fusion model in (5) can be reformulated as

$$\begin{aligned}\min_{\mathcal{Z}} \quad & \frac{1}{2}||\mathcal{Y} - \mathcal{C} \times_1 (\mathbf{P}_1\mathbf{W}) \times_2 (\mathbf{P}_2\mathbf{H}) \times_3 \mathbf{A}||_F^2 \\ & + \frac{1}{2}||\mathcal{X} - \mathcal{C} \times_1 \mathbf{W} \times_2 \mathbf{H} \times_3 (\mathbf{P}_3\mathbf{A})\,||_F^2 \\ & + \beta \text{tr}\left((\mathbf{H}\otimes\mathbf{W})^T \widetilde{\mathbf{L}}_W (\mathbf{H}\otimes\mathbf{W})\right) + \gamma \text{tr}(\mathbf{A}^T \widetilde{\mathbf{L}}_A \mathbf{A}) \\ & + \alpha \text{logsum}(\mathcal{C}).\end{aligned} \tag{20}$$

Then, (20) is solved using AO, which can be decoupled into several univariate alternating iterations as

$$\mathbf{W}^{k+1} \in \arg\min f\left(\mathbf{W}, \mathbf{H}^k, \mathbf{A}^k, \mathcal{C}^k\right) \tag{21a}$$

$$\mathbf{H}^{k+1} \in \arg\min f\left(\mathbf{W}^{k+1}, \mathbf{H}, \mathbf{A}^k, \mathcal{C}^k\right) \tag{21b}$$

$$\mathbf{A}^{k+1} \in \arg\min f\left(\mathbf{A}^{k+1}, \mathbf{H}^{k+1}, \mathbf{A}, \mathcal{C}^k\right) \tag{21c}$$

$$\mathcal{C}^{k+1} \in \arg\min f\left(\mathbf{W}^{k+1}, \mathbf{H}^{k+1}, \mathbf{A}^{k+1}, \mathcal{C}\right) \tag{21d}$$

where k denotes the number of iterations, $f(\mathbf{W}, \mathbf{H}, \mathbf{A}, \mathcal{C})$ denotes the objective function.

A. Estimation of Factor Matrix $\mathbf{W}$

Since (21a) contains the graph regularization and data fitting terms, it cannot be solved directly. Therefore, ADMM is adopted to conduct variable splitting by adding equation constraints. Substituting the spatial graph regularizer into (21a), we have

$$\begin{aligned}\min_{\mathbf{W}} \quad & \frac{1}{2}||\mathcal{Y} - \mathcal{C} \times_1 (\mathbf{P}_1\mathbf{W}) \times_2 (\mathbf{P}_2\mathbf{H}) \times_3 \mathbf{A}||_F^2 \\ & + \frac{1}{2}||\mathcal{X} - \mathcal{C} \times_1 \mathbf{W} \times_2 \mathbf{H} \times_3 (\mathbf{P}_3\mathbf{A})\,||_F^2 \\ & + \beta \text{tr}\left((\mathbf{H}\otimes\mathbf{W})^T \widetilde{\mathbf{L}}_W (\mathbf{H}\otimes\mathbf{W})\right).\end{aligned} \tag{22}$$

In general, the calculation of factor variable requires the unfolding of a tensor into 2-D form. Thus, (22) can be unfolded along mode-1 into matrix form as

$$\begin{aligned}\min \quad & \frac{1}{2}\left\|\mathbf{Y}_{(1)} - \mathbf{P}_1\mathbf{W}\mathbf{U}_W\right\|_F^2 + \frac{1}{2}\left\|\mathbf{X}_{(1)} - \mathbf{W}\mathbf{V}_W\right\|_F^2 \\ & + \beta \text{tr}\left((\mathbf{H}\otimes\mathbf{W}_1)^T \widetilde{\mathbf{L}}_W (\mathbf{H}\otimes\mathbf{W}_1)\right) \\ \text{s.t.} \quad & \mathbf{W}_1 = \mathbf{W}\end{aligned} \tag{23}$$

where $\mathbf{U}_W = (\mathcal{C} \times_2 (\mathbf{P_2}\mathbf{H}) \times_3 \mathbf{A})_{(1)}$ and $\mathbf{V}_W = (\mathcal{C} \times_2 \mathbf{H} \times_3 (\mathbf{P_3}\mathbf{A}))_{(1)}$.

Then, the augmented Lagrangian function of (23) can be written as

$$\begin{aligned}\mathcal{L}(\mathbf{W}, \mathbf{W}_1, \mathbf{M}_1) = & \frac{1}{2}\|\mathcal{Y} - \mathcal{C} \times_1 (\mathbf{P}_1\mathbf{W}) \times_2 (\mathbf{P}_2\mathbf{H}) \times_3 \mathbf{A}\|_F^2 \\ & + \frac{1}{2}\|\mathcal{X} - \mathcal{C} \times_1 \mathbf{W} \times_2 \mathbf{H} \times_3 (\mathbf{P}_3\mathbf{A})\|_F^2 \\ & + \beta \text{tr}\left((\mathbf{H}\otimes\mathbf{W}_1)^T \widetilde{\mathbf{L}}_W (\mathbf{H}\otimes\mathbf{W}_1)\right) \\ & + \frac{\eta}{2}||\mathbf{W_1} - \mathbf{W} + \frac{\mathbf{M}_{w1}}{\eta}||_F^2\end{aligned} \tag{24}$$

where $\mathbf{M}_{w1}$ is the dual variable and η is the weight. Using ADMM to update the original variables $\mathbf{W}$, $\mathbf{W}_1$ and the dual variable $\mathbf{M}_{w1}$ iteratively until convergence, the closed-form solution of each variable is obtained as follows.

1) Update $\mathbf{W}$*:* Obviously, (24) is a convex quadratic programming problem. However, since it often involves heavy computational burden, the Sylvester equation can be obtained as Formula (25) via least-squares criterion. Therefore, the conjugate gradient algorithm is utilized to update $\mathbf{W}$ for the optimal solution

$$\begin{aligned}& \mathbf{P}_1^T\mathbf{P}_1\mathbf{W}\mathbf{U}_w\mathbf{U}_w^T + \mathbf{W}\left(\mathbf{V}_w\mathbf{V}_w^T + \eta\mathbf{I}\right) \\ & = \mathbf{P}_1^T\mathbf{Y}_{(1)}\mathbf{U}_W^T + \mathbf{X}_{(1)}\mathbf{V}_W^T + \eta\mathbf{W}_1 + \mathbf{M}_{w1}.\end{aligned} \tag{25}$$

2) Update $\mathbf{W}_1$*:* The update of $\mathbf{W}_1$ is hard to optimize because of the Kronecker product contained in the spatial graph regularization term. The problem can be simplified by the graph Laplacian matrix, and the detailed procedure is shown in Appendix VI. Thus, the subproblem of $\mathbf{W}_1$ can be rewritten as

$$\min \frac{\beta}{2}\left\|\mathbf{W}_1^T\mathbf{N}_W\right\|_F^2 + \frac{\eta}{2}\left\|\mathbf{W}_1 - \mathbf{W} + \frac{\mathbf{M}_{w1}}{\eta}\right\|_F^2 \tag{26}$$

where $\mathbf{N}_w$ is the unfolding form of $\mathcal{U} \times_2 \mathbf{H}_1^T$ along mode-1, $\mathcal{U}$ is the tensor form of the upper triangular matrix derived from the decomposition of graph Laplacian matrix. The closed-form solution of $\mathbf{W}_1$ can be obtained as

$$\mathbf{W}_1^{j+1} = (\beta\mathbf{N}_w\mathbf{N}_w^T + \eta\mathbf{I})^{-1}(\eta\mathbf{W}^{j+1} - \mathbf{M}_{w1}). \tag{27}$$

3) Update $\mathbf{M}_{w1}$*:*

$$\mathbf{M}_{w1}^{j+1} = \mathbf{M}_{w1}^j + \eta(\mathbf{W}_1^{j+1} - \mathbf{W}^{j+1}). \tag{28}$$

B. Estimate of Factor Matrix $\mathbf{H}$

The solving process of factor matrix $\mathbf{H}$ is similar to that of $\mathbf{W}$, so (21b) is rewritten in the tensor form as

$$\begin{aligned}\min_{\mathbf{H}} \frac{1}{2}||\mathcal{Y}-\mathcal{C}\times_1(\mathbf{P}_1\mathbf{W})\times_2(\mathbf{P}_2\mathbf{H})\times_3\mathbf{A}||_F^2 \\ +\frac{1}{2}||\mathcal{X}-\mathcal{C}\times_1\mathbf{W}\times_2\mathbf{H}\times_3(\mathbf{P}_3\mathbf{A})||_F^2 \\ +\beta\text{tr}\left((\mathbf{H}\otimes\mathbf{W})^T\widetilde{\mathbf{L}}_W(\mathbf{H}\otimes\mathbf{W})\right). \end{aligned} \tag{29}$$

Owing to the symmetry of spatial factor matrices, the solution of $\mathbf{H}$ is similar to $\mathbf{W}$ as follows.

1) Update $\mathbf{H}$*:* The Sylvester equation (30) with respect to $\mathbf{H}$ can be obtained via least squares criterion, and the conjugate gradient algorithm is adopted to solve the variable $\mathbf{H}$

$$\begin{aligned}\mathbf{P}_2^T\mathbf{P}_2\mathbf{H}\mathbf{U}_h\mathbf{U}_h^T+\mathbf{H}\left(\mathbf{V}_h\mathbf{V}_h^T+\eta\mathbf{I}\right) \\ =\mathbf{P}_2^T\mathbf{Y}_{(2)}\mathbf{U}_h^T+\mathbf{X}_{(2)}\mathbf{V}_h^T+\eta\mathbf{H}_1+\mathbf{M}_{h1}\end{aligned} \tag{30}$$

where the structure matrices $\mathbf{U}_h=(\mathcal{C}\times_1(\mathbf{P}_1\mathbf{W})\times_3\mathbf{A})_{(2)}$, and $\mathbf{V}_h=(\mathcal{C}\times_1\mathbf{W}\times_3(\mathbf{P}_3\mathbf{A}))_{(2)}$.

2) Update Other Variables:

$$\mathbf{H}_1^{j+1}=(\gamma\mathbf{N}_h\mathbf{N}_h^T+\eta I)^{-1}(\eta\mathbf{H}^{j+1}-\mathbf{M}_{h1}^j) \tag{31}$$

$$\mathbf{M}_{h1}^{j+1}=\mathbf{M}_{h1}^j+\eta\left(\mathbf{H}_1^{j+1}-\mathbf{H}^{j+1}\right) \tag{32}$$

where $\mathbf{N}_h$ is the mode-2 unfolding form of $\mathcal{U}\times_1\mathbf{W}_1^T$. In sum, owing to spatial symmetry, $\mathbf{H}$ and $\mathbf{W}$ can be updated in the same solvers.

C. Estimation of Factor Matrix $\mathbf{A}$

For spectral factor matrix, by incorporating spectral graph regularization into two data fitting terms, (21c) can be redefined as

$$\begin{aligned}\min\frac{1}{2}\left\|\mathcal{Y}-\mathcal{C}\times_1(\mathbf{P}_1\mathbf{W})\times_2(\mathbf{P}_2\mathbf{H})\times_3\mathbf{A}\right\|_F^2 \\ +\frac{1}{2}\left\|\mathcal{X}-\mathcal{C}\times_1\mathbf{W}\times_2\mathbf{H}\times_3(\mathbf{P}_3\mathbf{A})\right\|_F^2 \\ +\gamma\text{tr}(\mathbf{A}^T\widetilde{\mathbf{L}}_A\mathbf{A}).\end{aligned} \tag{33}$$

Similar to the spatial graph, the Laplace matrix $\widetilde{\mathbf{L}}_A$ can be factorized into the product of two matrices via Cholesky decomposition, i.e., $\widetilde{\mathbf{L}}_A=\mathbf{V}^\mathbf{T}\mathbf{V}$. Thus, $\text{tr}(\mathbf{A}^T\widetilde{\mathbf{L}}_A\mathbf{A})$ can be rewritten as $\frac{\gamma}{2}||\mathbf{V}\mathbf{A}||_F^2$. It is worth mentioning that the spectral graph regularization is equivalent to the vertical total variation of mode-3 factor matrix [11]. Using ADMM, (33) is unfolded into matrix form along mode-3. We have

$$\begin{aligned}\min_{\mathbf{A},\mathbf{A}_1}\frac{1}{2}\left\|\mathbf{X}_{(3)}-\mathbf{P}_3\mathbf{A}\mathbf{V}_a\right\|_F^2+\frac{1}{2}\left\|\mathbf{Y}_{(3)}-\mathbf{A}\mathbf{U}_a\right\|_F^2+\frac{\gamma}{2}||\mathbf{V}\mathbf{A}_1||_F^2 \\ \text{s.t. }\mathbf{A}_1=\mathbf{A}\end{aligned} \tag{34}$$

where $\mathbf{U}_a=(\mathcal{C}\times_1(\mathbf{P}_1\mathbf{W})\times_2(\mathbf{P}_2\mathbf{H}))_{(3)}$, $\mathbf{V}_a=(\mathcal{C}\times_1\mathbf{W}\times_2\mathbf{H})_{(3)}$. The solutions of these variables are computed separately using the conjugate gradient and least square methods, which are updated iteratively in turn.

1) Update $\mathbf{A}$*:* Because of high computational complexity, the Sylvester equation (35) can be solved via the conjugate gradient method

$$\begin{aligned}\mathbf{P}_3^T\mathbf{P}_3\mathbf{A}\mathbf{V}_a\mathbf{V}_a^T+\mathbf{A}\left(\mathbf{U}_a\mathbf{U}_a^T+\eta\mathbf{I}\right) \\ =\mathbf{P}_3^T\mathbf{X}_{(3)}\mathbf{V}_a^T+\mathbf{Y}_{(3)}\mathbf{U}_a^T+\eta\mathbf{A}_1+\mathbf{M}_{a1}.\end{aligned} \tag{35}$$

2) Update Other Variables:

$$\mathbf{A}_1^{j+1}=(\gamma\mathbf{V}^T\mathbf{V}+\eta\mathbf{I})^{-1}(\eta\mathbf{A}^{j+1}-\mathbf{M}_{a1}^j) \tag{36}$$

$$\mathbf{M}_{a1}^{j+1}=\mathbf{M}_{a1}^j+\eta(\mathbf{A}^{j+1}-\mathbf{A}_1^{j+1}). \tag{37}$$

D. Estimate of Core Tensor $\mathcal{C}$

For the core tensor, (21d) with low-rank regularization can be rewritten as

$$\begin{aligned}\min_{\mathcal{C}}\frac{1}{2}\left\|\mathcal{Y}-\mathcal{C}\times_1(\mathbf{P}_1\mathbf{W})\times_2(\mathbf{P}_2\mathbf{H})\times_3\mathbf{A}\right\|_F^2 \\ +\frac{1}{2}\left\|\mathcal{X}-\mathcal{C}\times_1\mathbf{W}\times_2\mathbf{H}\times_3(\mathbf{P}_3\mathbf{A})\right\|_F^2 \\ +\alpha\text{logsum}(\mathcal{C}).\end{aligned} \tag{38}$$

To simplify two data fitting terms in the objective function, the core tensor should be reshaped into a vector as $\mathbf{c}=\text{vec}(\mathcal{C})$. After splitting variables with equation constraints, (38) can be expressed as

$$\begin{aligned}\min\frac{1}{2}\left\|\mathbf{y}-\mathbf{B}_1\mathbf{c}\right\|_F^2+\frac{1}{2}\left\|\mathbf{x}-\mathbf{B}_2\mathbf{c}_1\right\|_F^2+\alpha\text{logsum}(\mathbf{F}_{<2>}) \\ \text{s.t. }\mathbf{c}_1=\mathbf{c},\mathbf{F}_{<2>}=\mathbf{C}_{<2>}\end{aligned} \tag{39}$$

where $\mathbf{B}_1=\mathbf{A}\otimes(\mathbf{P_2}\mathbf{H})\otimes(\mathbf{P_1}\mathbf{W})$; $\mathbf{B}_2=(\mathbf{P_3}\mathbf{A})\otimes\mathbf{H}\otimes\mathbf{W}$; $\mathbf{x}$ and $\mathbf{y}$ denote the vectorized versions of observed images $\mathcal{X}$ and $\mathcal{Y}$, respectively, i.e., $\mathbf{x}=\text{vec}(\mathcal{X})$ and $\mathbf{y}=\text{vec}(\mathcal{Y})$.

Obviously, the update of $\mathbf{F}_{<2>}$ is a low-rank regularization problem. According to (8), the closed-form solution of $\mathbf{F}_{<2>}$ can be obtained as

$$\mathbf{F}_{<2>}^{j+1}=\widetilde{\mathbf{U}}_F\text{diag}(q_1,q_2,\ldots,q_n)\widetilde{\mathbf{V}}_F^T \tag{40}$$

where $q_i=Q_{\alpha\epsilon}(\sigma_i)$, and $\widetilde{\mathbf{U}}_F\text{diag}(q_1,q_2,\ldots,q_n)\widetilde{\mathbf{V}}_F^T$ is the singular value decomposition of $\mathbf{C}_{<2>}-\frac{\mathbf{M}_{c2}}{\eta}$. The closed-form solutions of the other variables can be obtained directly via ADMM, which are then updated iteratively in the following order:

$$\begin{aligned}\mathbf{c}^{j+1}=(\mathbf{B}_1^T\mathbf{B}_1+\eta\mathbf{I})^{-1}(\mathbf{B_1}^T\mathbf{y}+\eta\mathbf{c}_1^j+\mathbf{m}_{c1} \\ +\eta\text{vec}(\mathbf{F}_{<2>}^{j+1})+\eta\mathbf{M}_{c2}^j)\end{aligned} \tag{41a}$$

$$\mathbf{c}_1^{j+1}=(\mathbf{B}_2^T\mathbf{B}_2+\eta\mathbf{I})^{-1}(\eta\mathbf{c}^{j+1}-\mathbf{m}_{c1}^j+\mathbf{B}_2^T\mathbf{x}) \tag{41b}$$

$$\mathbf{m}_{c1}^{j+1}=\mathbf{m}_{c1}^j+\eta(\mathbf{c}^{j+1}-\mathbf{c}_1^{j+1}) \tag{41c}$$

$$\mathbf{M}_{c2}^{j+1}=\mathbf{M}_{c2}^j+\eta(\mathbf{C}_{<2>}^{j+1}-\mathbf{F}_{<2>}^{j+1}). \tag{41d}$$

To sum up, the solving process of the proposed fusion method in this article is summarized in Algorithm 1. First of all, $\mathbf{W}$, $\mathbf{H}$, and $\mathcal{C}$ are initialized separately by $\mathbf{0}$ with proper sizes, and the initial value of $\mathbf{A}$ can be evaluated by VCA [44]. Then, all the factor matrices and the core tensor are iterated alternatively

Algorithm 1: A graph-regularized low-rank tensor fusion algorithm (G-LRTF).

Input: $\mathcal{Y}$,$\mathcal{X}$,$\mathbf{P}_1$,$\mathbf{P}_2$,$\mathbf{P}_3$
Output: $\mathcal{Z}^{k+1}$
Initialization:
$\mathbf{W}^0 = 0$,$\mathbf{H}^0 = 0$,$\mathbf{A}^0$,$\mathcal{C}^0 = 0$,$k = 0$,$maxIter = 30$;
while *the stopping criterion is not met* **do**
  Update $\mathbf{W}$ by (21a);
  while *not convergence* **do**
    Update $\mathbf{W}$ by (25);
    Update $\mathbf{W}_1$ by (27);
    Update $\mathbf{M}_{w1}$ by (28);
  Update $\mathbf{H}$ by (21b);
  while *not convergence* **do**
    Update $\mathbf{H}$ by (30);
    Update $\mathbf{H}_1$ by (31);
    Update $\mathbf{M}_{h1}$ by (32);
  Update $\mathbf{A}$ by (21c);
  while *not convergence* **do**
    Update $\mathbf{A}$ by (35);
    Update $\mathbf{A}_1$ by (36);
    Update $\mathbf{M}_{a1}$ by (37);
  Update $\mathcal{C}$ by (21d);
  while *not convergence* **do**
    Update $\mathcal{C}$ by (41a);
    Update $\mathcal{C}_1$ by (41b);
    Update $\mathbf{F}_{<2>}$ by (40);
    Update $\mathbf{M}_{c1}$ by (41c);
    Update $\mathbf{M}_{c2}$ by (41d);
  $k = k + 1$;
return $\mathcal{Z}^{k+1} = \mathcal{C} \times_1 \mathbf{W} \times_2 \mathbf{H} \times_3 \mathbf{A}$;

until convergence. Finally, the super-resolution image is reconstructed after each subproblem is solved.

E. Convergence Rules

In this article, the fusion method decomposes the optimization problem with low-rank and graph regularization into alternating iterations of three factor matrices and a core tensor, i.e., (21a)–(21d). The stopping rule satisfies the relative difference threshold between the successive updates of the objective function $f(\mathbf{W}, \mathbf{H}, \mathbf{A}, \mathcal{C})$ is less than 0.001. The experiments show that changing the number of iterations in the ADMM-based algorithms has little effect on the convergence of the whole algorithm [20], [21]. That is, the variables of $\mathbf{W}$, $\mathbf{H}$, $\mathbf{A}$, and $\mathcal{C}$ do not run exhaustively for convergence. Therefore, the iterations in outer and inner loops are set to 50 and 10, respectively.

IV. EXPERIMENTS

We evaluate the quality of hyperspectral super-resolution images from two perspectives of reconstruction performance and classification accuracy. Three widely used real datasets are adopted to simulate and test the proposed low-rank tensor fusion algorithm using Wald's protocol [45]–[47]. This flowchart can not only achieve the coregistration of two observed images, but also provide the reference image for performance comparison. According to the hyperspectral observation model, the hyperspectral and multispectral data in the same scene, i.e., $\mathcal{Y}$ and $\mathcal{X}$, can be regarded as the degraded versions of the reference image $\mathcal{Z}$ in the spatial and spectral subspaces, respectively. The fusion method aims at integrating $\mathcal{Y}$ and $\mathcal{X}$ to yield the high-resolution image $\widehat{\mathcal{Z}}$, and measure the similarity between $\widehat{\mathcal{Z}}$ and $\mathcal{Z}$ in quality metrics [48]–[50], including peak signal to noise ratio (PSNR), root-mean-square error (rmse), relative dimensionless global error in synthesis (ERGAS), spectral angle mapper (SAM), degree of distortion (DD), and structural similarity (SSIM). Moreover, we evaluate the classification performance of the super-resolution images reconstructed by the fusion methods in several metrics, including average accuracy (AA), overall accuracy (OA), Kappa coefficient ($\mathcal{K}$), and per-class accuracy.

A. Experimental Datasets

This section presents three real datasets, including Pavia University,[1] Chikusei,[2] and Columbia Computer Vision Laboratory (CAVE).[3] Then, we evaluate the performance of fusion methods from the perspectives of reconstruction quality and classification accuracy.

1) Pavia University: The first dataset was captured at Pavia University [51] in northern Italy using the hyperspectral sensor ROSIS, with 115 bands and a spatial resolution of 1.3 m. After removing the pollution by water vapor absorption and strong noise, 103 spectral bands covering the wavelength of 430–860 nm were retained. Simultaneously, a 600×300 region in Pavia University dataset is selected as the original reference image $\mathcal{Z}$. Thus, the multispectral data with four spectral bands in the same area were collected by the sensor IKONOS, covering 445–516, 516–595, 632–698, and 757–853 nm. The spectral downsampling operator was computed as $\mathbf{P}_3 \in \mathbb{R}^{4\times 103}$.

2) Chikusei: The second dataset was taken by a Headwall Hyperspec-VNIR-C image sensor on July 29, 2014, in the rural and urban areas of Chikusei[52], Japan. In the spectral dimension, this dataset retains 128 bands ranging from 363 to 1018 nm. In the spatial dimension, a subscene of 300×300 pixels was selected as the region of interest (RoI) with a ground sampling resolution of 2.5 m. The MSI was obtained by the spectral degradation operator $\mathbf{P}_3 \in \mathbb{R}^{4\times 128}$.

3) CAVE: The third dataset was CAVE dataset [53], which was collected widely for hyperspectral image super-resolution including 32 indoor images, which were captured by a camera. This original image had 31 bands covering the spectrum range from 400 to 700 nm, with the target region of 400×400 pixels, which was graded to get the observed MSI by the spectral downsampling matrix $\mathbf{P}_3 \in \mathbb{R}^{3\times 31}$.

To verify the local similarity, each target image is divided into several 100×100 nonoverlapping subblocks to test the performance. With degradation ratio of 4, the spatial downsampling matrix $\mathbf{P}_1 \in \mathbb{R}^{25\times 100}$ and $\mathbf{P}_2 \in \mathbb{R}^{25\times 100}$ can be constructed to

[1][Online]. Available: http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes
[2][Online]. Available: https://www.sal.t.u-tokyo.ac.jp/hyperdata/
[3][Online]. Available: https://www.cs.columbia.edu/CAVE/databases/

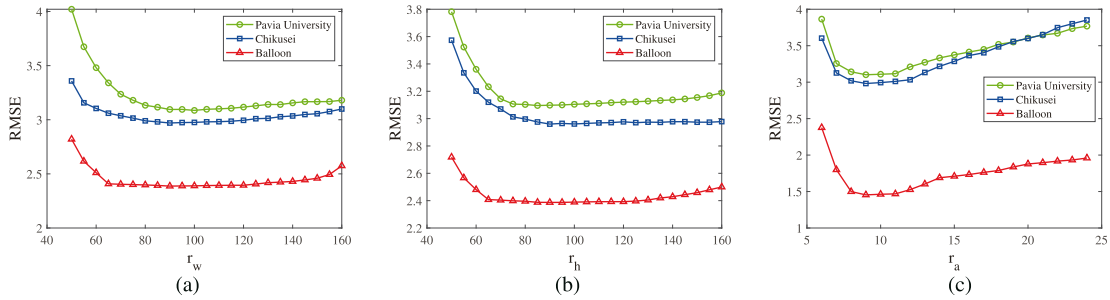

Fig. 3. Estimation of core tensor dimensions for two datasets (a) r_w, (b) r_h, and (c) r_a, with two datasets distributed as Pavia University (green line); Chikusei (blue line) and Balloon (red line).

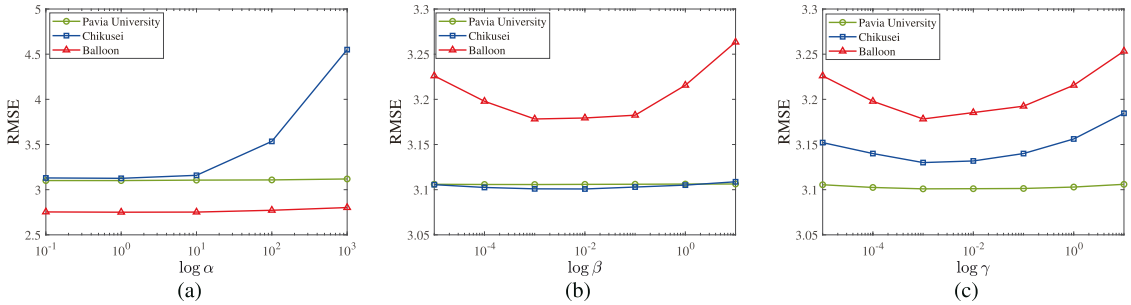

Fig. 4. RMSE curves w.r.t regularization parameters including (a) α, (b) β, and (c) γ.

degrade the reference images for the observed hyperspectral data with 25×25 pixels in each subblock.

B. Experimental Settings

1) Dimension Setting of Core Tensor: The hyperspectral image data is decomposed into two spatial factor matrices $\mathbf{W}$, $\mathbf{H}$, a spectral factor matrix $\mathbf{A}$, and a core tensor $\mathcal{C}$ by Tucker decomposition. The factor matrices can be regarded as principal components in each dimension, and the core tensor contains the corresponding coefficients of each factor matrix, which indicates the degree of intercorrelation between different modes. To drop the computational complexity, the proposed G-LRTF segments the original dataset into several pieces, so the dimension setting of the core tensor is discussed in each subblock [36]. In this article, we set appropriate dimensions for the subscenes and leverage the low-rank prior to suppress noise corruption and redundant information. The dimensions of core tensor are r_w, r_h, and r_a, which denote two spatial dimensions and the spectral dimension, respectively. The performance rmse is taken to test the effect of tensor dimensions on reconstruction error. The curves of rmse with respect to each dimension separately by fixing the other two dimensions, as shown in Fig. 3. Considering the stability of performance, the spatial dimensions r_w and r_h are set to 100, and the spectral dimension is set to 9.

2) Setting of SNRs: To evaluate the antinoise performance of the proposed fusion method, we set three groups of SNRs for the observed hyperspectral and multispectral images to simulate different noise environments, including (20, 25 dB), (25, 25 dB), and (25, 30 dB), and then discuss the performance of the proposed G-LRTF under the different noise conditions.

3) Setting of Regularization Coefficients: Based on the Tucker tensor decomposition, this article incorporates the logarithmic sum function and spatial-spectral graphs to regularize the core tensor and factor matrices with the corresponding weights α, β, and γ. In general, the regularization coefficient is set to balance the relationship between the data fitting terms and the regularization term in the objective function, and the weights have strong correlations with the noise power of the observed images [20], [34]. The variation curves of rmse with respect to the regularization weights are illustrated in Fig. 4. Thus, the regularization weights are set to $\alpha = 1$, $\beta = 0.001$, and $\gamma = 0.001$.

4) Comparison of Low-Rank Regularizers: Generally, the nuclear norm is often a convex surrogate of the rank function in low-rank minimization problems. As a convex relaxation, logarithmic sum function is imposed on core tensor to promote the property of low rank in this article. We conducted the experiments to compare the nuclear norm with the logarithmic sum function on three datasets, where the other regularization terms and settings remained the same. From the experimental results in Table I, one can see that logarithmic sum function achieves slightly better performance than nuclear norm, except that the performance of SAM and SSIM is almost the same.

C. Experimental Results

To make full use of the local spatial information and drop the computational complexity, each dataset was segmented by a fixed-size window of 100×100 pixels into several nonoverlapping blocks, which were handled separately and then stitched together after fusion. Using the downsampling matrix, the

TABLE I
PERFORMANCE COMPARISON BETWEEN NUCLEAR NORM (NN) AND LOGARITHM SUM FUNCTION (LOGSUM) (THE BEST VALUE IS IN BOLD)

Dataset	SNRs	Low rank	DD	PSNR	RMSE	SAM	ERGAS	SSIM
PaviaU	(20,25) dB	Ideal	0	∞	0	0	0	1
		NN	2.832	36.824	3.904	**0.074**	2.507	**0.934**
		logsum	**2.762**	**37.108**	**3.725**	0.075	**2.497**	**0.934**
	(25,25) dB	NN	2.792	36.168	3.749	**0.073**	2.461	**0.937**
		logsum	**2.680**	**37.377**	**3.614**	**0.073**	**2.426**	0.936
	(25,30) dB	NN	2.126	38.867	2.987	**0.059**	1.974	**0.962**
		logsum	**2.002**	**39.733**	**2.865**	**0.059**	**1.939**	**0.962**
Chikusei	(20,25) dB	NN	2.117	39.766	2.930	**0.041**	3.503	**0.952**
		logsum	**2.106**	**39.782**	**2.913**	**0.041**	**3.481**	**0.952**
	(25,25) dB	NN	1.973	40.354	2.759	**0.037**	3.260	**0.956**
		logsum	**1.967**	**40.369**	**2.745**	**0.037**	**3.258**	**0.956**
	(25,30) dB	NN	1.562	42.286	2.283	**0.033**	2.995	**0.972**
		logsum	**1.549**	**42.343**	**2.270**	**0.033**	**2.983**	**0.972**
Balloon	(20,25) dB	NN	2.502	37.731	3.583	0.066	1.996	0.935
		logsum	**1.638**	**40.090**	**2.576**	**0.065**	**1.676**	**0.940**
	(25,25) dB	NN	2.220	39.109	3.136	**0.049**	1.669	0.957
		logsum	**1.242**	**42.232**	**1.999**	**0.049**	**1.300**	**0.958**
	(25,30) dB	NN	2.471	38.758	3.161	0.046	1.524	0.978
		logsum	**0.916**	**44.908**	**1.478**	**0.045**	**0.961**	**0.980**

TABLE II
QUANTITATIVE RESULTS ON PAVIA UNIVERSITY (THE BEST VALUE IS IN BOLD)

Method		DD	PSNR	RMSE	SAM	ERGAS	SSIM
Case I (20,25) dB	Ideal	0	∞	0	0	0	1
	STEREO	3.187	35.845	4.249	0.088	2.822	0.919
	BSR	3.627	33.950	5.318	**0.073**	3.203	0.927
	CSTF	3.089	36.129	4.103	0.081	2.680	0.921
	NLSTF	5.306	31.588	6.808	0.153	4.576	0.824
	CNMF	5.232	31.482	6.949	0.148	4.513	0.843
	Proposed	**2.762**	**37.108**	**3.725**	0.075	**2.497**	**0.934**
Case II (25,25) dB	STEREO	2.945	36.560	3.949	0.080	2.628	0.929
	BSR	3.593	34.022	5.273	**0.073**	3.178	0.928
	CSTF	2.969	36.429	3.938	0.077	2.592	0.925
	NLSTF	3.891	34.318	5.086	0.113	3.480	0.881
	CNMF	3.475	35.112	4.622	0.096	3.091	0.910
	Proposed	**2.680**	**37.377**	**3.614**	**0.073**	**2.426**	**0.936**
Case III (25,30) dB	STEREO	2.449	38.016	3.408	0.071	2.303	0.950
	BSR	3.276	34.592	4.945	0.066	2.996	0.942
	CSTF	2.542	37.635	3.471	0.067	2.276	0.944
	NLSTF	3.082	36.289	4.032	0.093	2.746	0.925
	CNMF	2.729	37.054	3.751	0.080	2.536	0.946
	Proposed	**2.002**	**39.733**	**2.865**	**0.059**	**1.939**	**0.962**

reference image was degraded in the spectral dimension to obtain the observed MSI $\mathcal{X}$. In the same way, the spatial degradation operators $\mathbf{P}_1$ and $\mathbf{P}_2$ were designed to generate the observed hyperspectral data $\mathcal{Y}$. For performance comparison, five state-of-the-art fusion methods were selected as the benchmarks, including the original CNMF (coupled nonnegative matrix factorization) [46], BSR (Bayesian sparse representation) [6], CSTF (coupled sparse tensor factorization) [29], NLSTF (nonlocal sparse tensor factorization) [54], and STEREO (super-resolution tensor-reconstruction) [22].

1) Reconstruction Quality: This section conducts experimental simulations on three datasets and analyzes the performance of the proposed G-LRTF with the benchmark algorithms. First of all, the overall performance of the algorithms are tested under three different noise conditions, as shown in Tables II– IV for Pavia University, Chikusei, and Balloon, respectively. On the whole, the performance of each algorithm drops as the SNRs decrease. For three datasets, the performance of the proposed G-LRTF performs significantly better than the benchmark algorithms. For example, the proposed G-LRTF achieves great performance improvement on the CAVE dataset in low-SNR environments, where the rmse and DD are reduced by more than 27% and 40%, respectively. Meanwhile, the performance on Chikusei and Pavia University datasets is improved substantially, i.e., the PSNR is raised by 2 and 1 dB, respectively.

The reconstruction and difference images for three datasets are provided in this experiment, as shown in Figs. 5–7. We can observe that the fused data by the proposed G-LRTF is closer to the corresponding reference image, although these reconstructed images are visually indistinguishable. Especially, STEREO, CSTF and G-LRTF are all based on tensor decomposition methods and they have comparable performance. The STEREO algorithm achieves good performance because CPD is more suitable for linear features such as roads and building edges. Both CSTF and G-LRTF are based on Tucker decomposition, so their difference images are very similar and close to the noise image. However, local feature-based processing methods often lead to the nonsmoothness of difference images, especially

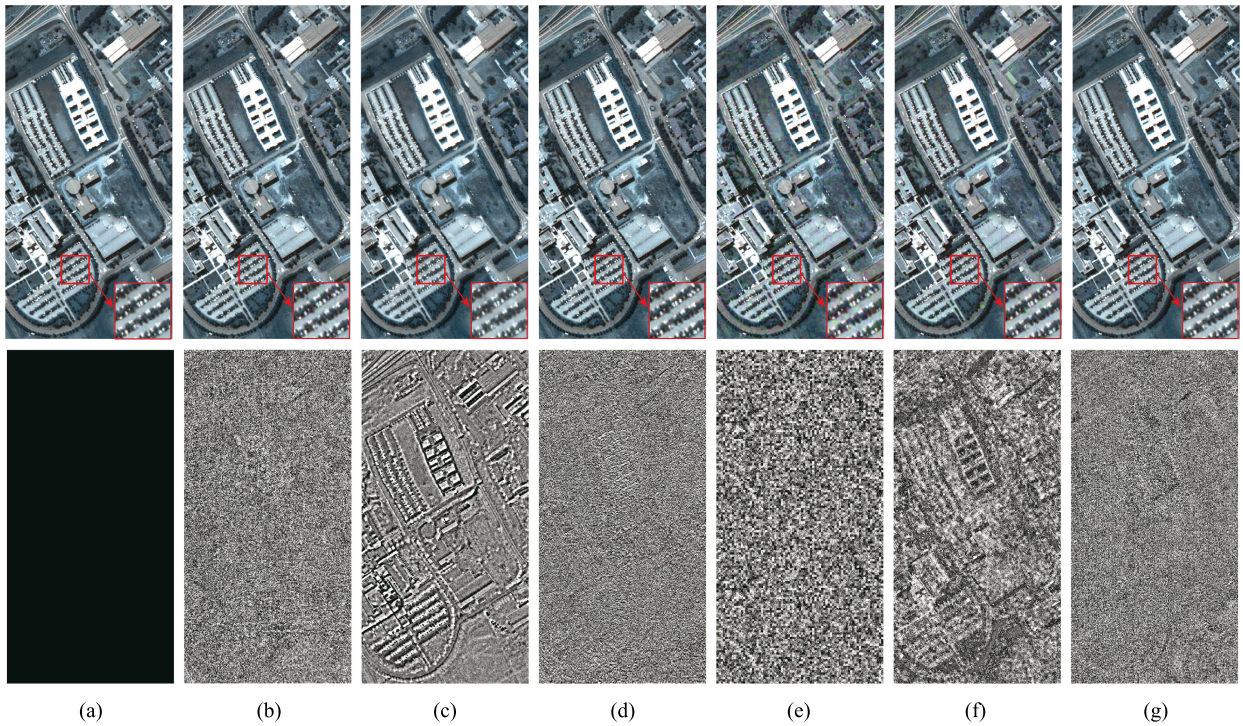

Fig. 5. Reference/reconstruction images (top row) and the corresponding difference images (bottom row) on Pavia University dataset by the different algorithms, including (a) reference image, (b) STERRO, (c) BSR, (d) CSTF, (e) NLSTF, (f) CNMF, and (g) proposed G-LRTF.

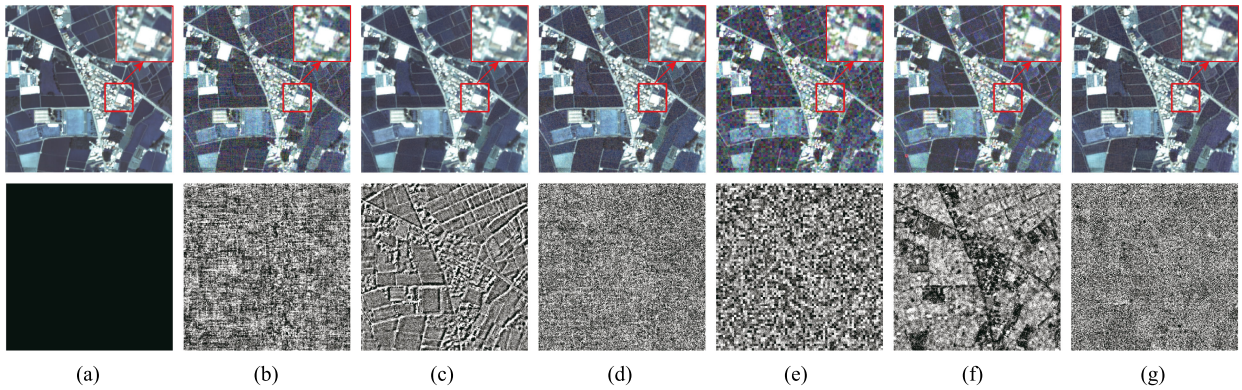

Fig. 6. Reference/reconstruction images (top row) and the corresponding difference images (bottom row) on Chikusei dataset by the different algorithms, including (a) reference image, (b) STERRO, (c) BSR, (d) CSTF, (e) NLSTF, (f) CNMF, and (g) proposed G-LRTF. Second row, difference images corresponding to different algorithms.

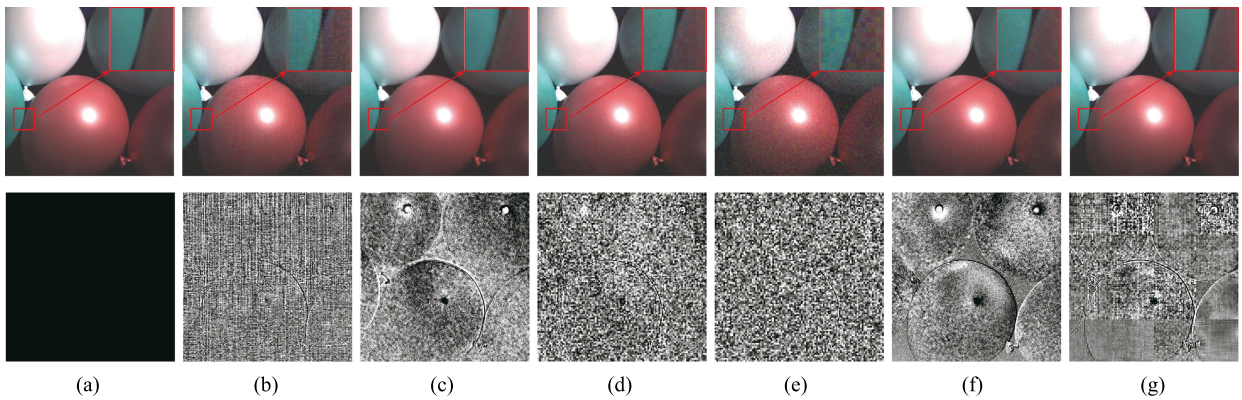

Fig. 7. Reference/reconstruction images (top row) and the corresponding difference images (bottom row) on Balloon dataset by the different algorithms, including (a) reference image, (b) STERRO, (c) BSR, (d) CSTF, (e) NLSTF, (f) CNMF, and (g) proposed G-LRTF.

TABLE III
QUANTITATIVE RESULTS ON CHIKUSEI DATASET (THE BEST VALUE IS IN BOLD)

Method		DD	PSNR	RMSE	SAM	ERGAS	SSIM
Case I (20,25) dB	Ideal	0	∞	0	0	0	1
	STEREO	4.318	33.547	5.597	0.093	7.289	0.818
	BSR	2.716	37.455	4.238	0.042	3.841	0.939
	CSTF	2.630	37.997	3.476	0.053	4.513	0.919
	NLSTF	5.347	31.624	6.732	0.117	9.187	0.777
	CNMF	2.797	36.950	3.781	0.057	4.740	0.922
	Proposed	**2.106**	**39.782**	**2.913**	**0.041**	**3.481**	**0.952**
Case II (25,25) dB	STEREO	3.365	35.856	4.469	0.071	5.608	0.876
	BSR	2.660	37.601	4.148	0.041	3.787	0.940
	CSTF	2.341	38.914	3.091	0.045	3.866	0.934
	NLSTF	3.607	35.033	4.569	0.076	6.165	0.869
	CNMF	3.115	35.959	4.183	0.065	5.208	0.901
	Proposed	**1.967**	**40.369**	**2.745**	**0.037**	**3.258**	**0.956**
Case III (25,30) dB	STEREO	2.529	38.342	3.451	0.056	4.489	0.926
	BSR	2.289	38.371	3.697	0.038	3.660	0.958
	CSTF	1.631	42.254	2.319	0.036	3.027	0.965
	NLSTF	3.149	36.192	3.989	0.070	5.667	0.899
	CNMF	2.415	38.044	3.311	0.053	4.502	0.942
	Proposed	**1.549**	**42.343**	**2.270**	**0.033**	**2.983**	**0.972**

TABLE IV
QUANTITATIVE RESULTS ON BALLOON DATASET (THE BEST VALUE IS IN BOLD)

Method		DD	PSNR	RMSE	SAM	ERGAS	SSIM
Case I (20,25) dB	Ideal	0	∞	0	0	0	1
	STEREO	4.273	33.501	5.474	0.212	3.577	0.719
	BSR	2.290	38.662	3.049	0.127	1.969	0.907
	CSTF	2.760	37.380	3.559	0.148	2.312	0.872
	NLSTF	4.892	32.339	6.198	0.246	4.080	0.698
	CNMF	3.027	36.187	4.010	0.139	2.593	0.844
	Proposed	**1.638**	**40.090**	**2.576**	**0.065**	**1.676**	**0.940**
Case II (25,25) dB	STEREO	3.564	35.140	4.600	0.182	2.998	0.776
	BSR	2.158	39.104	2.891	0.119	1.868	0.914
	CSTF	2.038	39.873	2.625	0.106	1.708	0.918
	NLSTF	3.507	35.282	4.486	0.183	2.946	0.795
	CNMF	2.398	38.165	3.182	0.107	2.072	0.886
	Proposed	**1.242**	**42.232**	**1.999**	**0.049**	**1.300**	**0.958**
Case III (25,30) dB	STEREO	2.562	37.926	3.341	0.139	2.177	0.872
	BSR	1.542	41.669	2.178	0.094	1.416	0.959
	CSTF	1.548	42.187	2.042	0.090	1.325	0.955
	NLSTF	2.816	37.109	3.578	0.158	2.359	0.867
	CNMF	1.565	41.727	2.114	0.082	1.375	0.953
	Proposed	**0.916**	**44.908**	**1.478**	**0.045**	**0.961**	**0.980**

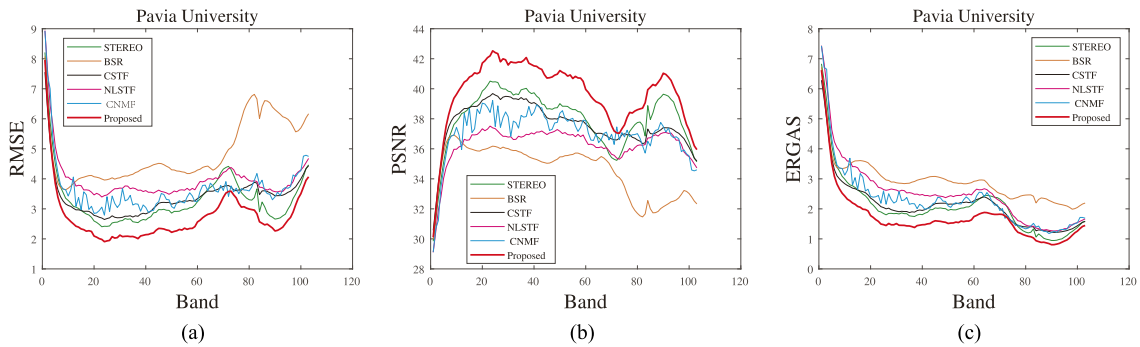

Fig. 8. Performance comparison curve for Pavia University. (a) RMSE, (b) PSNR, and (c) ERGAS.

as shown in Balloon dataset, but it does not affect the overall performance.

To evaluate the fusion performance of the algorithms with respect to spectral bands, the performance curves of rmse, PSNR, and ERGAS are given in this experiment, as shown in Figs. 8– 10. From these figures, the proposed G-LRTF outperforms the benchmarks on three datasets in most cases. In addition, in terms of the curve fluctuation, the proposed G-LRTF yields better smoothness than the baseline algorithms.

Since both the proposed G-LRTF and the CSTF are handled by tensor decomposition, the third factor matrices represent the spectral signatures, which are similar in matrix factorization.

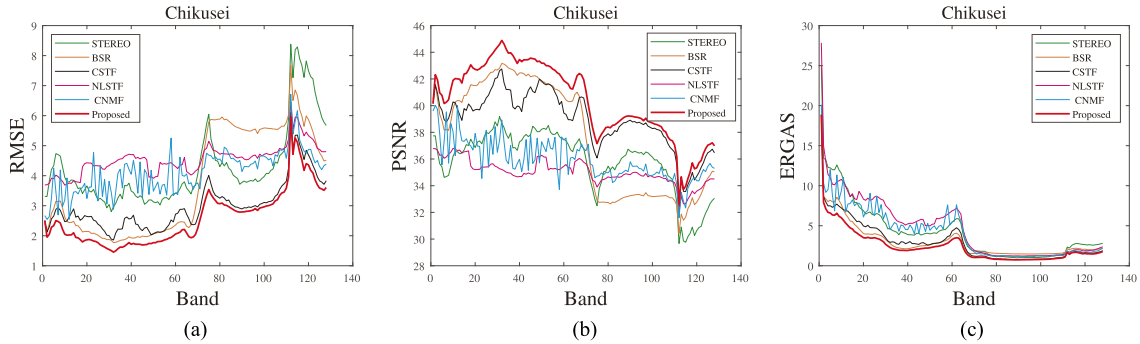

Fig. 9. Performance comparison curve for Chikusei. (a) RMSE, (b) PSNR, and (c) ERGAS.

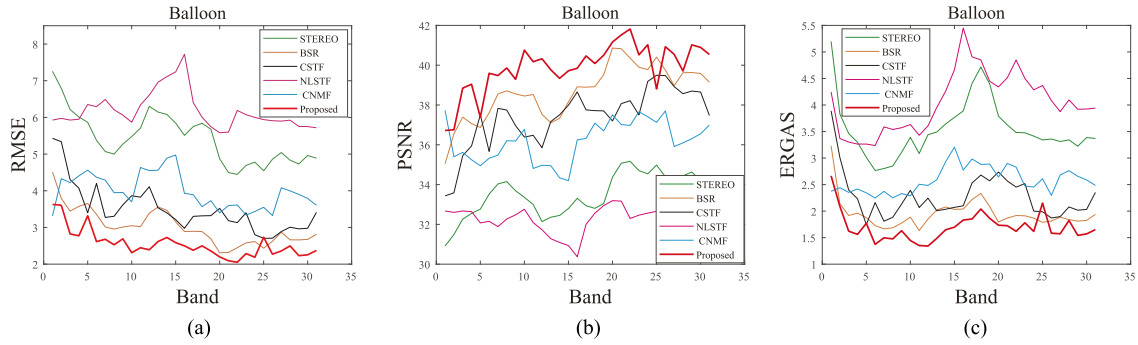

Fig. 10. Performance comparison curve for Balloon. (a) RMSE, (b) PSNR, and (c) ERGAS.

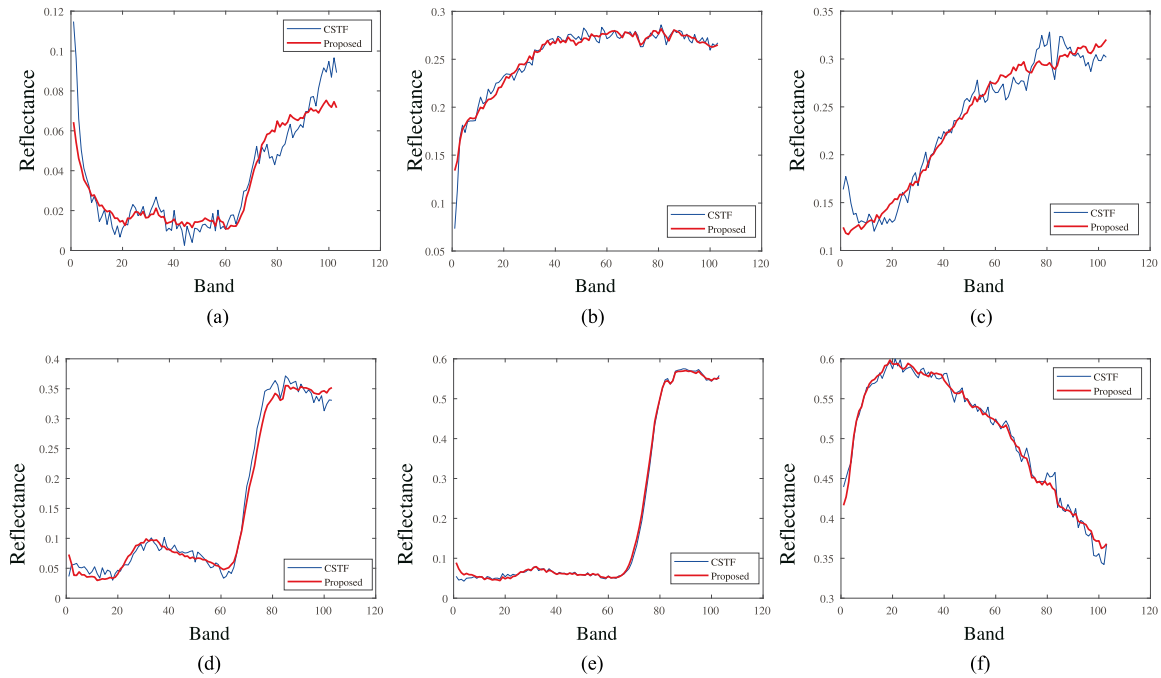

Fig. 11. Signature curves of the fifth block via CSTF (thin blue line) and the proposed G-LRTF (bold red line).

The proposed G-LRTF takes advantage of the gradual variability of spectral signature, and adopts graph regularization to remove the effect of spectral variability for smoothing. The curves of spectral signatures by CSTF and G-LRTF are shown in Fig. 11 for Pavia University dataset. On the whole, one can see that the signature curves of the proposed G-LRTF are much smoother than that of CSTF, as depicted in subfigures (a) and (c). The proposed G-LRTF achieves obvious advantages in the lower and higher reflectance parts in subfigures (b) and (f), while CSTF and G-LRTF algorithms have similar performance at the

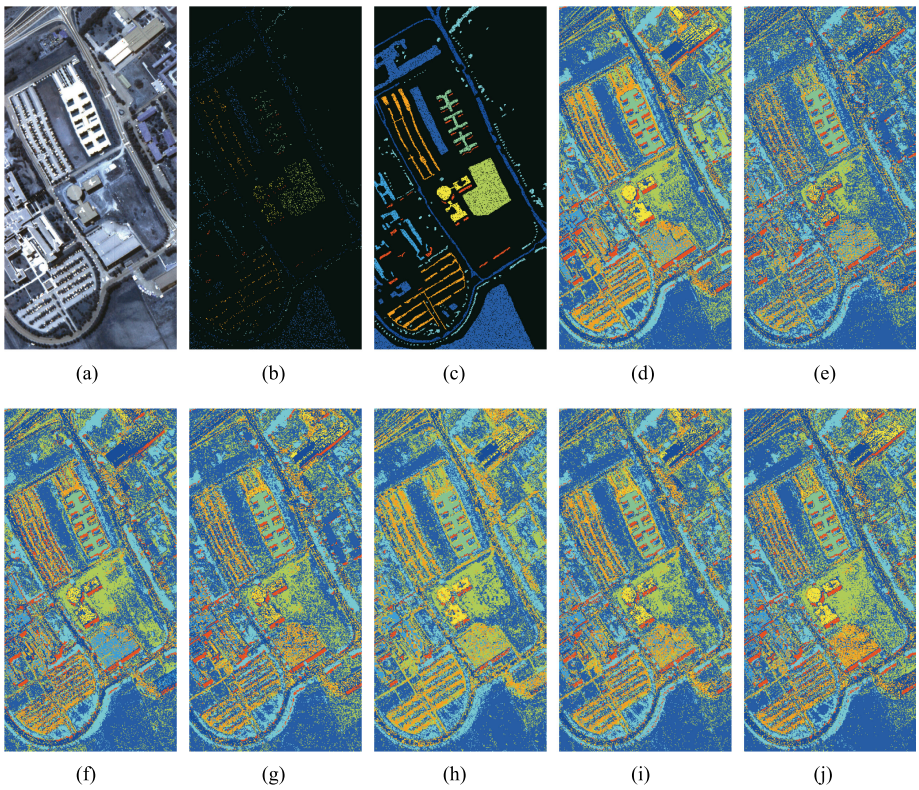

Fig. 12. Visualization of (a) false-color images, (b) training and (c) test sample distribution, and classification maps by using different algorithms on Pavia University dataset, including (d) reference image, (e) STEREO, (f) BSR, (g) CSTF, (h) NLSTF, (i) CNMF, and (j) proposed G-LRTF.

TABLE V
CLASSIFICATION PERFORMANCE OF SUPER-RESOLUTION IMAGES BY DIFFERENT METHODS ON PAVIA UNIVERSITY (THE BEST VALUE IS IN BOLD)

	STEREO	BSR	CSTF	NLSTF	CNMF	Proposed	Real HS
OA	0.7137	0.776	0.752	0.800	0.804	**0.832**	0.844
AA	0.6559	0.747	0.713	0.751	0.766	**0.812**	0.835
$\mathcal{K}$	0.6254	0.711	0.678	0.739	0.744	**0.783**	0.797
C1	0.7002	0.737	0.672	0.582	0.665	**0.731**	0.779
C2	0.8802	0.880	0.891	**0.948**	0.934	0.938	0.943
C3	0.4729	0.523	0.492	0.665	0.631	**0.667**	0.639
C4	0.7368	**0.885**	0.812	0.786	0.828	0.837	0.892
C5	0.9992	**1.000**	**1.000**	**1.000**	0.998	**1.000**	1.000
C6	0.4570	0.608	0.569	0.696	0.634	**0.759**	0.595
C7	0.2911	0.573	0.417	0.669	0.615	**0.767**	0.842
C8	0.4970	0.645	0.612	0.741	**0.761**	0.658	0.855
C9	0.8693	0.875	0.949	0.671	0.832	**0.952**	0.973

reflectance between 0.2 and 0.4. It is evident that the spectral graph regularization significantly improves the smoothness of endmember signature.

2) Classification Accuracy: As mentioned above, the quality of hyperspectral images has a critical impact on classification accuracy, which often serves as the performance metric. Furthermore, the widely-used K-neighborhood classifier [55] is adopted to evaluate the quality of reconstructed images. The classification maps of super-resolution images fused by the benchmarks and the proposed G-LRTF are drawn in Fig. 12 for Pavia University dataset. One can observe that the classification performance of the proposed G-LRTF is much better than that of the baseline algorithms in some continuous regions, and it is much closer to the reference image. By incorporating graph regularizers, the proposed G-LRTF can preserve the spectral signatures of local regions. More specifically, the classification results of each class of substances are shown in Table V. The proposed G-LRTF achieves higher accuracy than the benchmarks in most independent classes, and has significant advantages especially in three metrics including AA, OA and $\mathcal{K}$ [18]. The experimental results demonstrate that the fusion image reconstructed by the proposed G-LRTF has very high quality, and can greatly improve the classification accuracy.

D. Complexity Analysis

The computation burden of the proposed G-LRTF includes the construction of spatial graph, factor matrices and core tensor. The complexity of graph construction is $\mathcal{O}(N_w N_h)$, and the iteration part mainly includes the estimation of four variables,

TABLE VI
RUNNING TIME COMPARISON OF FUSION ALGORITHMS (THE BEST VALUE IS IN BOLD)

Datasets	Running time(s)					
	STEREO	BSR	CSTF	NLSTF	CNMF	Proposed
Pavia University	**21.20**	149.00	53.02	33.74	147.20	58.50
Chikusei	**12.32**	87.29	43.86	16.65	68.35	62.18
Balloon	**18.84**	127.44	42.75	25.99	22.86	50.05

i.e., $\mathbf{W}$, $\mathbf{H}$, $\mathbf{A}$, and $\mathcal{C}$. First, the Sylvester equations of the factor matrices $\mathbf{W}$, $\mathbf{H}$, and $\mathbf{A}$ are computed by the conjugate gradient, and their computational complexity consumes $\mathcal{O}(n_w^2 N_w^2)$, $\mathcal{O}(n_h^2 N_h^2)$, and $\mathcal{O}(n_a^2 N_a^2)$, respectively. After complexity reduction via vector-matrix operators, the computational cost can be reduced to $\mathcal{O}(n_w^2 N_w + n_w N_w^2)$, $\mathcal{O}(n_h^2 N_h + n_h N_h^2)$, and $\mathcal{O}(n_a^2 N_a + n_a N_a^2)$. For the computation of Laplacian matrix in factor matrices, the Cholesky decomposition requires $\mathcal{O}(N_w^3)$. Furthermore, in the iteration of core tensor $\mathcal{C}$, the computational burden occupies $\mathcal{O}(n_w^2 n_h n_a + n_w n_h^2 n_a + n_w n_h n_a^2)$. According to the above analysis, the bottleneck of the proposed G-LRTF in this article lies in the graph regularization terms, which involve high complexity. The running time is also a metric to evaluate the complexity of an algorithm, and Table VI lists the running times of different algorithms on the three datasets. From these results, we can observe that STEREO consumes the least time due to the rank-1 property of CPD, which is a special case of Tucker decomposition. Compared with CSTF, the G-LRTF proposed takes longer running time than CSTF, but has obvious advantages over BSR. Although the local segmentation can decrease the computational burden to some extent, it increases the number of iterations. Complexity reduction will be the focus of further research in the future.

V. CONCLUSION

This article proposes a low-rank tensor decomposition approach, termed as G-LRTF, by incorporating spatial-spectral graph regularization for hyperspectral image super-resolution. To capture the correlation of different dimensions in hyperspectral data, we focus on the local manifold information in hyperspectral and multispectral data to build the spatial-spectral graph Laplacian matrices, and reconstruct the images based on coupled Tucker decomposition. In the simulation, three widely used real datasets are taken to conduct the experimental tests. For performance evaluation, we discuss the setting of parameters and noise conditions, and then test the fusion performance in visual and quantitative forms including the reconstructed and difference images, spectral curves, performance metrics, and classification accuracy. The proposed G-LRTF achieves much better than the state-of-the-art methods in performance metrics, and especially yields the higher classification accuracy in AA, OA, and $\mathcal{K}$. The experimental results illustrate that the proposed fusion method can significantly enhance the quality of reconstructed super-resolution images, which also verifies the validation of regularization terms. In future work, we will focus on the fusion method in several aspects including the following: 1) further studying the complexity reduction for high efficiency; 2) further exploring the stopping rules for convergence; 3) further investigating nonlocal similarity to reduce spatial distortion.

APPENDIX A
PROOF OF PROPOSITION 1

Proof: Assume that the tensor-train rank of a tensor $\mathcal{M}$ is R_n, the tensor $\mathcal{M}$ can be expressed as

$$\mathcal{M}_{i_1,i_2,\ldots,i_N} = \sum_{r_n=1}^{R_n} \mathcal{F}_{i_1,i_2,\ldots,i_n,r_n} \mathcal{N}_{r_n,i_{n+1},i_{n+2},\ldots,i_N} \tag{42}$$

where the elements of tensor factors $\mathcal{F} \in \mathbb{R}^{I_1 \times I_2,\ldots,\times I_n \times R_n}$ and $\mathcal{N} \in \mathbb{R}^{R_n \times I_{n+1}, I_{n+2}\ldots,\times I_N}$ are represented as

$$\mathcal{F}_{i_1,i_2,\ldots,i_n,r_n} = \sum_{r_1,\ldots,r_{n-1}=1}^{R_1,\ldots,R_{n-1}} \mathcal{G}^{(1)}_{1,i_1,r_1} \mathcal{G}^{(2)}_{r_1,i_2,r_2}, \ldots, \mathcal{G}^{(n)}_{r_{n-1},i_n,r_n}$$

$$\mathcal{N}_{r_n,i_{n+1},\ldots,i_N} = \sum_{r_{n+1},\ldots,r_{N-1}=1}^{R_{n+1},\ldots,R_{N-1}} \mathcal{G}^{(n+1)}_{r_n,i_{n+1},r_{n+1}}, \ldots, \mathcal{G}^{(N)}_{r_{N-1},i_N,1}. \tag{43}$$

Thus, (42) can be reformulated as

$$\mathbf{M}_{\langle n \rangle} = \mathbf{F}_{\langle n \rangle} \mathbf{N}_{\langle 1 \rangle} \tag{44}$$

where $\mathbf{F}_{\langle n \rangle} \in \mathbb{R}^{I_1 I_2,\ldots,I_n \times R_n}$ and $\mathbf{N}_{\langle 1 \rangle} \in \mathbb{R}^{R_n \times I_{n+1},\ldots,I_N}$ are the mode-n canonical matrix unfolding [56] of the tensor $\mathcal{F}$ and the mode-1 canonical matrix unfolding of the tensor $\mathcal{N}$, respectively. Finally, we can obtain $\text{rank}(\mathbf{M}_{<n>}) \leq R_n$. ■

APPENDIX B
PROOF OF $\mathbf{W}_1$ UPDATE

Proof: The optimization problem (45) of $\mathbf{W}_1$ is obtained from the augmented Lagrangian function in (24) as follows:

$$\mathbf{W}_1^{k+1} \in \arg \min_{\mathbf{W}_1 \in \mathbb{R}_w^N \times r_w} \beta \text{tr}\left((\mathbf{H} \otimes \mathbf{W}_1)^T \widetilde{\mathbf{L}}_W (\mathbf{H} \otimes \mathbf{W}_1) \right) + \frac{\eta}{2} ||\mathbf{W}_1^k - \mathbf{W}^{k+1} + \frac{\mathbf{M}_{w1}}{\eta}||_F^2. \tag{45}$$

Considering the symmetry and semipositive characterization of the Laplacian matrix $\widetilde{\mathbf{L}}_W$, we perform the Cholesky decomposition of $\widetilde{\mathbf{L}}_W$ to obtain $\widetilde{\mathbf{L}}_W = \mathbf{U}^T \mathbf{U}$, where $\mathbf{U}$ is the upper triangular matrix. Therefore, $\text{tr}((\mathbf{H}_1 \otimes \mathbf{W}_1)^T \widetilde{\mathbf{L}}_W (\mathbf{H}_1 \otimes \mathbf{W}_1))$ is reformulated as $\|\mathbf{U}(\mathbf{H}_1 \otimes \mathbf{W}_1)\|_F^2$, and then it can be further simplified as $||\mathbf{W}_1^T \mathbf{N}_w||_F^2$, where $\mathbf{N}_w$ is the unfolding form of $\mathcal{U} \times_2 \mathbf{H}_1^T$ along mode-1. Thus, (45) can be rewritten as

$$\min \frac{\beta}{2} \left\| \mathbf{W}_1^T \mathbf{N}_W \right\|_F^2 + \frac{\eta}{2} \left\| \mathbf{W}_1 - \mathbf{W} + \frac{\mathbf{M}_{w1}}{\eta} \right\|_F^2. \tag{46}$$

The closed-form solution of $\mathbf{W}_1$ can be obtained as

$$\mathbf{W}_1 = (\beta \mathbf{N}\mathbf{N}^T + \eta \mathbf{I})^{-1} (\eta \mathbf{W} - \mathbf{M}_{w1}). \tag{47}$$

■

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] H. Albanwan and R. Qin, "Spatiotemporal fusion in remote sensing," in *Proc. Recent Adv. Image Restoration Appl. Real World Problems*, 2020, pp. 95–114.

[3] D. Hong, N. Yokoya, J. Chanussot, and X. X. Zhu, "An augmented linear mixing model to address spectral variability for hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1923–1938, Apr. 2019.

[4] M. J. Khan, H. S. Khan, A. Yousaf, K. Khurshid, and A. Abbas, "Modern trends in hyperspectral image analysis: A review," *IEEE Access*, vol. 6, pp. 14 118–14 129, 2018.

[5] J. Liu, "Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3461–3472, 2000.

[6] Q. Wei, J. Bioucas-Dias, N. Dobigeon, and J. Tourneret, "Hyperspectral and multispectral image fusion based on a sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3658–3668, Jul. 2015.

[7] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled non-negative matrix factorization (CNMF) for hyperspectral and multispectral data fusion: Application to pasture classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2011, pp. 1779–1782.

[8] F. Ma, F. Yang, Z. Ping, and W. Wang, "Joint spatial-spectral smoothing in a minimum-volume simplex for hyperspectral image super-resolution," *Appl. Sci.*, vol. 10, no. 1, 2020, Art. no. 237.

[9] F. Yang, Z. Ping, F. Ma, and Y. Wang, "Fusion of hyperspectral and multispectral images with sparse and proximal regularization," *IEEE Access*, vol. 7, pp. 186 352–186363, 2019.

[10] F. Yang, F. Ma, Z. Ping, and G. Xu, "Total variation and signature-based regularizations on coupled nonnegative matrix factorization for data fusion," *IEEE Access*, vol. 7, pp. 2695–2706, 2018.

[11] C.-H. Lin, F. Ma, C.-Y. Chi, and C.-H. Hsieh, "A convex optimization-based coupled nonnegative matrix factorization algorithm for hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1652–1667, Mar. 2018.

[12] M. A. Veganzones, M. Simões, G. Licciardi, J. Bioucas, and J. Chanussot, "Hyperspectral super-resolution of locally low rank images from complementary multisource data," in *Proc. IEEE Int. Conf. Image Process.*, 2014, pp. 703–707.

[13] M. Zare, M. S. Helfroush, and K. Kazemi, "Fusing hyperspectral and multispectral images using smooth graph signal modelling," *Int. J. Remote Sens.*, vol. 41, no. 22, pp. 8610–8630, 2020.

[14] Y. Qu, H. Qi, and C. Kwan, "Unsupervised sparse dirichlet-net for hyperspectral image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 2511–2520.

[15] J.-F. Hu, T.-Z. Huang, L.-J. Deng, T.-X. Jiang, G. Vivone, and J. Chanussot, "Hyperspectral image super-resolution via deep spatiospectral attention convolutional neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, 2021.

[16] J. Yao, D. Hong, J. Chanussot, D. Meng, X. Zhu, and Z. Xu, "Cross-attention in coupled unmixing nets for unsupervised hyperspectral super-resolution," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 208–224.

[17] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, and J. Chanussot, "Graph convolutional networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 5966–5978, Jul. 2021.

[18] D. Hong *et al.*, "More diverse means better: Multimodal deep learning meets remote-sensing imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 4340–4354, May 2021.

[19] D. Hong *et al.*, "Interpretable hyperspectral artificial intelligence: When nonconvex modeling meets hyperspectral remote sensing," *IEEE Geosci. Remote Sens. Mag.*, vol. 9, no. 2, pp. 52–87, Jun. 2021.

[20] Q. Wei, J. Bioucas-Dias, N. Dobigeon, J. Tourneret, M. Chen, and S. Godsill, "Multiband image fusion based on spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7236–7249, Dec. 2016.

[21] F. Ma, F. Yang, and Y. Wang, "Low-rank tensor decomposition with smooth and sparse regularization for hyperspectral and multispectral data fusion," *IEEE Access*, vol. 8, pp. 129 842–129856, 2020.

[22] C. I. Kanatsoulis, X. Fu, N. D. Sidiropoulos, and W. Ma, "Hyperspectral super-resolution: A coupled tensor factorization approach," *IEEE Trans. Signal Process.*, vol. 66, no. 24, pp. 6503–6517, Dec. 2018.

[23] H. Zeng, X. Xie, and J. Ning, "Hyperspectral image denoising via global spatial-spectral total variation regularized nonconvex local low-rank tensor approximation," *Signal Process.*, vol. 178, 2021, Art. no. 107805.

[24] Y. Bu *et al.*, "Hyperspectral and multispectral image fusion via graph Laplacian-guided coupled tensor decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 1, pp. 648–662, Jan. 2021.

[25] K. Zhang, M. Wang, S. Yang, and L. Jiao, "Spatial-spectral-graph-regularized low-rank tensor decomposition for multispectral and hyperspectral image fusion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1030–1040, Apr. 2018.

[26] S. Boyd, N. Parikh, and E. Chu, *Distributed Optimization and Statistical Learning Via the Alternating Direction Method of Multipliers*. Boston, MA, USA: Now, 2011.

[27] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[28] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3551–3582, Jul. 2017.

[29] S. Li, R. Dian, L. Fang, and J. M. Bioucas-Dias, "Fusing hyperspectral and multispectral images via coupled sparse tensor factorization," *IEEE Trans. Image Process.*, vol. 27, no. 8, pp. 4118–4130, Aug. 2018.

[30] Z. Hu, F. Nie, R. Wang, and X. Li, "Low rank regularization: A review," *Neural Netw.*, vol. 136, pp. 218–232, 2020.

[31] R. Dian and S. Li, "Hyperspectral image super-resolution via subspace-based low tensor multi-rank regularization," *IEEE Trans. Image Process.*, vol. 28, no. 10, pp. 5135–5146, Oct. 2019.

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

[33] I. V. Oseledets, "Tensor-train decomposition," *SIAM J. Sci. Comput.*, vol. 33, no. 5, pp. 2295–2317, 2011.

[34] 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.

[35] W. He, Y. Chen, N. Yokoya, C. Li, and Q. Zhao, "Hyperspectral super-resolution via coupled tensor ring factorization," 2020, *arXiv:2001.01547*.

[36] W. Sun, L. Zhang, B. Du, W. Li, and Y. Mark Lai, "Band selection using improved sparse subspace clustering for hyperspectral imagery classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2784–2797, Jun. 2015.

[37] C. Wu, B. Du, and L. Zhang, "A subspace-based change detection method for hyperspectral images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 815–830, Apr. 2013.

[38] W. Sun, J. Ma, G. Yang, B. Du, and L. Zhang, "A poisson nonnegative matrix factorization method with parameter subspace clustering constraint for endmember extraction in hyperspectral imagery," *ISPRS J. Photogrammetry Remote Sens.*, vol. 128, pp. 27–39, 2017.

[39] Q. Wang, J. Lin, and Y. Yuan, "Salient band selection for hyperspectral image classification via manifold ranking," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 6, pp. 1279–1289, Jun. 2016.

[40] W.-L. Chiang, X. Liu, S. Si, Y. Li, S. Bengio, and C.-J. Hsieh, "Cluster-GCN: An efficient algorithm for training deep and large graph convolutional networks," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2019, pp. 257–266.

[41] F. Shang, L. Jiao, and F. Wang, "Graph dual regularization non-negative matrix factorization for co-clustering," *Pattern Recognit.*, vol. 45, no. 6, pp. 2237–2250, 2012.

[42] M. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4484–4502, Nov. 2012.

[43] J. Sigurdsson, M. O. Ulfarsson, and J. R. Sveinsson, "Blind hyperspectral unmixing using total variation and ℓ_q sparse regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 11, pp. 6371–6384, Nov. 2016.

[44] J. M. Nascimento and J. M. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, 2005.

[45] L. Loncan *et al.*, "Hyperspectral pansharpening: A review," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 3, pp. 27–46, Sep. 2015.

[46] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 528–537, Feb. 2012.

[47] L. Wald, T. Ranchin, and M. Mangolini, "Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images," *Photogrammetric Eng. Remote Sensing*, vol. 63, no. 6, pp. 691–699, 1997.

[48] F. Zhang, G. Yang, and J.-H. Xue, "Hyperspectral image denoising based on low-rank coefficients and orthonormal dictionary," *Signal Process.*, vol. 177, 2020, Art. no. 107738.

[49] R. H. Yuhas, A. F. Goetz, and J. W. Boardman, "Discrimination among semi-arid landscape endmembers using the spectral angle mapper (sam) algorithm," in *Proc. Summaries 3rd Annu. JPL Airborne Geosci. Workshop*, 1992, vol. 1, pp. 147–149.
[50] L. Wald, "Quality of high resolution synthesised images: Is there a simple criterion?," in *Proc. 3rd Conf. Fusion Earth Data: Merging Point Meas.*, Raster Maps Remotely Sensed Images, 2000, pp. 99–103.
[51] F. Dell' Acqua, P. Gamba, and A. Ferrari, "Exploiting spectral and spatial information for classifying hyperspectral data in urban areas," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2003, vol. 1, pp. 464–466.
[52] N. Yokoya and A. Iwasaki, "Airborne hyperspectral data over Chikusei," Japan, Tech. Rep. SAL-2016-05-27, May 2016. [Online]. Available: http://park.itc.u-tokyo.ac.jp/sal/hyperdata/TechRepSAL20160527.pdf
[53] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, "Generalized assorted pixel camera: Postcapture control of resolution, dynamic range, and spectrum," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2241–2253, Sep. 2010.
[54] R. Dian, L. Fang, and S. Li, "Hyperspectral image super-resolution via non-local sparse tensor factorization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 5344–5353.
[55] L. Gao, D. Hong, J. Yao, B. Zhang, P. Gamba, and J. Chanussot, "Spectral superresolution of multispectral imagery with joint sparse and low-rank learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 3, pp. 2269–2280, Mar. 2021.
[56] A. Cichocki, N. Lee, I. Oseledets, A.-H. Phan, Q. Zhao, and D. P. Mandic, "Tensor networks for dimensionality reduction and large-scale optimization: Part 1 low-rank tensor decompositions," *Found. Trends Mach. Learn.*, vol. 9, no. 4/5, pp. 249–429, 2016.

Fei Ma received the Ph.D. degree in circuit and system from the Beijing University of Posts and Telecommunications, Beijing, China, in 2010.

He held a Postdoctoral position with the University of Electronic Science and Technology of China, Chengdu, China, from 2014 to 2020. In the meanwhile, he conducted research as a Visiting Scholar with National Tsing Hua University, Hsinchu, Taiwan, from 2015 to 2016. He is currently an Associate Professor with the School of Electronic and Information Engineering, Liaoning Technical University, Huludao, China. His research interests include hyperspectral image processing, radar signal processing, pattern recognition, and convex optimization.

Shuai Huo received the B.S. degree in communication engineering in 2020 from the School of Electronics and Information Engineering, Liaoning Technical University, Huludao, China, where he is currently working toward the master's degree in communication and information system.

His research interests include hyperspectral image processing, convex optimization, and machine learning.

Feixia Yang received the Ph.D. degree in electronics science and technology from the Beijing University of Posts and Telecommunications, Beijing, China, in 2020.

Since 2006, she has been a Faculty Member with the School of Electrical and Control Engineering, Liaoning Technical University, Huludao, China. Her research interests include digital image processing, remote sensing imaging, and pattern recognition.