Hyperspectral and Multispectral Image Fusion Using Factor Smoothed Tensor Ring Decomposition

Yong Chen, Jinshan Zeng, Wei He, *Member, IEEE*, Xi-Le Zhao, and Ting-Zhu Huang

***Abstract*— Fusing a pair of low-spatial-resolution hyperspectral image (LR-HSI) and high-spatial-resolution multispectral image (HR-MSI) has been regarded as an effective and economical strategy to achieve HR-HSI, which is essential to many applications. Among existing fusion models, the tensor ring (TR) decomposition-based model has attracted rising attention due to its superiority in approximating high-dimensional data compared to other traditional matrix/tensor decomposition models. Unlike directly estimating HR-HSI in traditional models, the TR fusion model translates the fusion procedure into an estimate of the TR factor of HR-HSI, which can efficiently capture the spatial–spectral correlation of HR-HSI. Although the spatial–spectral correlation has been preserved well by TR decomposition, the spatial–spectral continuity of HR-HSI is ignored in existing TR decomposition models, sometimes resulting in poor quality of reconstructed images. In this article, we introduce a factor smoothed regularization for TR decomposition to capture the spatial–spectral continuity of HR-HSI. As a result, our proposed model is called *factor smoothed TR decomposition* model, dubbed *FSTRD*. In order to solve the suggested model, we develop an efficient proximal alternating minimization algorithm. A series of experiments on four synthetic datasets and one real-world dataset show that the quality of reconstructed images can be significantly improved by the introduced factor smoothed regularization, and thus, the suggested method yields the best performance by comparing it to state-of-the-art methods.**

***Index Terms*— Hyperspectral image (HSI) and multispectral image (MSI) fusion, proximal alternating minimization (PAM), smoothed regularization, tensor ring (TR) decomposition.**

I. Introduction

HYPERSPECTRAL imaging has attracted an amount of attention since it can simultaneously collect images of the same scene across visible and infrared wavelengths. The collected hyperspectral image (HSI) with dense bands has found a wide range of applications on remote sensing and computer vision, such as classification [1], unmixing [2], [3], and recognition [4]. However, due to the critical tradeoffs between the spatial and spectral resolutions of imaging techniques, HSI with a high spectral resolution usually has a low spatial resolution (LR). Directly improving the spatial resolution in the hyperspectral imaging system is costly and difficult due to various hardware limitations. Fortunately, multispectral sensors can acquire an image with a higher spatial resolution but with a lower spectral resolution, such as the RGB image, multispectral image (MSI), and panchromatic image. Since both LR-HSI and high-spatial-resolution MSI (HR-MSI) contain valuable information, LR-HSI and HR-MSI fusion has become an effective technique to enhance the spatial resolution of LR-HSI. The fused image is called HR-HSI, which is desired to have both high-spectral and high-spatial resolutions.

To date, extensive methods have been proposed for fusing LR-HSI and HR-MSI, and they can be generally divided into three categories [5]–[7]: pansharpening-based approaches, deep learning (DL)-based approaches, and factorization-based approaches. The pansharpening techniques aim to obtain an HR-MSI by fusing an LR-MSI with a high-resolution panchromatic image. The fusion problem of LR-HSI and HR-MSI can be treated as many pansharpening subproblems, where the role of each band of HR-MSI is the same as that of panchromatic image. Pansharpening-based approaches are mainly based on two categories: component substitution (CS) [8], [9] and multiresolution analysis (MRA) [10], [11]. The advantages of these techniques are low computational cost and fast implementation. However, the spectral information is a lack in a single panchromatic image compared with the desired HR-HSI; thus, these methods usually cause spectral distortion.

Due to the high efficiency and potential performance of deep convolution neural network (CNN) in computer vision tasks, DL-based fusion methods have attracted rising attention. DL-based pansharpening methods [12]–[16] yielded the HR-MSI by learning a nonlinear mapping function under the input of the original MSI and high-resolution panchromatic image, and these methods can be extended for LR-HSI and LR-MSI fusion via modifying the first and last convolution layers. Recently, the DL methods that directly combine the characteristics of LR-HSI and HR-HSI as the input and then map to the HR-HSI were widely extended to the fusion of LR-HSI and HR-MSI [17]–[24]. For example, Wang *et al.* [20]

Manuscript received June 2, 2021; revised July 17, 2021 and August 26, 2021; accepted September 13, 2021. This work was supported in part by NSFC under Grant 12171072, Grant 62101222, Grant 61977038, Grant 61772003, and Grant 61876203; in part by the Key Projects of Applied Basic Research in Sichuan Province under Grant 2020YJ0216 and Grant 2021YJ0107; in part by the National Key Research and Development Program of China under Grant 2020YFA0714001; in part by the Japan Society for the Promotion of Science (KAKENHI) under Grant 19K20308; and in part by the Thousand Talents Plan of Jiangxi Province under Grant jxsq2019201124.
(Corresponding author: Jinshan Zeng.)

Yong Chen and Jinshan Zeng are with the School of Computer and Information Engineering, Jiangxi Normal University, Nanchang 330022, China (e-mail: chenyong1872008@163.com; jinshanzeng@jxnu.edu.cn).

Wei He is with the Geoinformatics Unit, RIKEN Center for Advanced Intelligence Project, RIKEN, Tokyo 103-0027, Japan (e-mail: wei.he@riken.jp).

Xi-Le Zhao and Ting-Zhu Huang are with the School of Mathematical Sciences/Research Center for Image and Vision Computing, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China (e-mail: xlzhao122003@163.com; tingzhuhuang@126.com).
Digital Object Identifier 10.1109/TGRS.2021.3114197

proposed a deep iterative network for blind LR-HSI and HR-MSI fusion. One major advantage of DL-based methods is that they can obtain satisfactory results since CNN possesses a strong ability to explore the image features. However, the training of these methods generally requires plenty of paired LR-HSI and HR-HSI, which is difficult to collect. To tackle such an issue, unsupervised DL methods are proposed to fuse LR-HSI and HR-MSI [25]–[27]. For example, Wang *et al.* [27] presented a variational probabilistic autoencoder framework for unsupervised LR-HSI and HR-MSI fusion.

Factorization-based approaches have been widely presented for LR-HSI and HR-MSI fusion since these methods are unsupervised and effective. This kind of method regards the reconstruction of HR-HSI as an ill-posed problem and recovers the HR-HSI by minimizing an energy function. By modeling the HR-HSI as 2-D or high-dimensional data, there are two ways to factorize the HR-HSI, i.e., matrix factorization-based and tensor factorization-based approaches. Matrix factorization-based approaches assume that each spectral signature of HR-HSI can be mathematically represented as a linear combination of several endmembers, and they factorize the desired HR-HSI as the spectral basis and coefficient [28]–[37]. Based on the matrix factorization strategy, the reconstruction of HR-HSI is transformed to the estimation of spectral basis and coefficient. However, the strategy of vectorizing each band will destroy the intrinsic structure in the process of reconstructing the HR-HSI. Since LR-HSI, HR-MSI, and HR-HSI are intrinsically 3-D, they can be expressed by tensor data. Motivated by this observation, the idea of tensor factorization becomes very popular in the fusion of LR-HSI and HR-MSI. Benefit from the strong representation ability of the tensor factorization, this kind of tensor factorization-based method achieves state-of-the-art results. So far, there are some well-known tensor factorization schemes, including Tucker decomposition [38]–[42], Canonical polyadic (CP) decomposition [43], [44], tensor singular value decomposition (t-SVD) [45]–[48], tensor-train decomposition [49], and block-term decomposition [50], applied to LR-HSI and HR-MSI fusion. Besides these, a new tensor factorization called tensor ring (TR) decomposition was recently suggested in [51]. Specifically, TR decomposition is defined to decompose a high-order tensor as a sequence of cyclically contracted third-order tensors. TR decomposition inherits the representation ability of traditional matrix/CP decomposition and extends the low-rank exploration of Tucker decomposition [52]–[54]. Inspired by the effectiveness of TR decomposition, the tensor methods for the fusion of LR-HSI and HR-MSI are based on the TR framework [55], [56].

Matrix/tensor decomposition can effectively explore the global and local redundant information of HR-HSI and translate the high-dimensional data in a relatively low-dimensional space by decomposing the HR-HSI as a set of factors. However, when these models are introduced for the fusion of LR-HSI and HR-MSI, some other factor prior knowledge, such as the smoothed structure, was ignored, resulting in the loss of spatial–spectral piecewise smoothness of reconstructed HR-HSI. There are many methods to improve the fusion performance by considering the factor property under the framework of matrix factorization. For example, Simões *et al.* [29] first introduced factor vector total variation (TV) to preserve the spatial smoothness of HR-HSI; Wei *et al.* [31] incorporated a factor sparse regularization into matrix factorization framework for LR-HSI and HR-MSI fusion. From the perspective of tensor factorization framework, Li *et al.* [57] presented the coupled Tucker decomposition with factor core tensor sparse prior for fusing LR-HSI and HR-MSI; Ding *et al.* [50] proposed a fusion method by using latent factor TV and low-rank regularized block-term decomposition. Hence, incorporating the additional factor regularization into the factorization-based framework is an effective way to improve the accuracy of reconstructed results, especially when the data is seriously polluted by noise [58]. However, although TR decomposition is an impressive representation for high-dimensional data and successfully applied to fuse LR-HSI and HR-MSI [55], the existing TR decomposition-based methods ignored the TR factor regularization, resulting in the lack of consideration of spatial–spectral continuity in HR-HSI. A toy example is presented in Fig. 1 to illustrate the necessity of TR factor regularization if the prior of spatial–spectral continuity exists in HR-HSI. It can be observed from the first row of Fig. 1 that, when the row fibers of the TR factor are nonsmooth, then the signatures of the resulted tensor in different dimensions are nonsmooth. In contrast, the smoothed TR factor will generate smoothed tensor shown in the second row of Fig. 1. Thus, the TR decomposition model still has the potential to improve.

A. Contributions

In this article, we propose a factor-smoothed TR decomposition (FSTRD)-based method for the fusion of LR-HSI and HR-MSI. To fully explore the high spatial–spectral correlation of HR-HSI, we introduce the representative and flexible TR decomposition to represent it. Under the TR representation of HR-HSI, the degradation process of LR-HSI and HR-MSI from HR-HSI is formulated as the TR decomposition format. Moreover, we explore the TR factor to inherit the potential characteristics of the original HR-HSI. As the HR-HSI has the piecewise smooth characteristics in the spatial–spectral dimension, it is necessary to incorporate the smoothed regularization into the TR factor to preserve the spatial–spectral information of HR-HSI. Based on these analyses, the reconstruction of HR-HSI is transformed to estimate the TR factor from the input LR-HSI and HR-MSI. An efficient proximal alternating minimization (PAM) algorithm is designed to solve the proposed model. A series of experimental results on simulated and real datasets illustrate that the proposed FSTRD method outperforms the state-of-the-art factorization-based approaches.

B. Related Work

In the literature, three representative tensor models have been proposed for LR-HSI and HR-MSI fusion, including CP decomposition [43], [44], Tucker decomposition [38], [40], [41], [57], and t-SVD [45]. The tensor decomposition is extended from the matrix factorization [28]–[34], [59] [see Fig. 2(a)] to explore the spatial–spectral correlation of

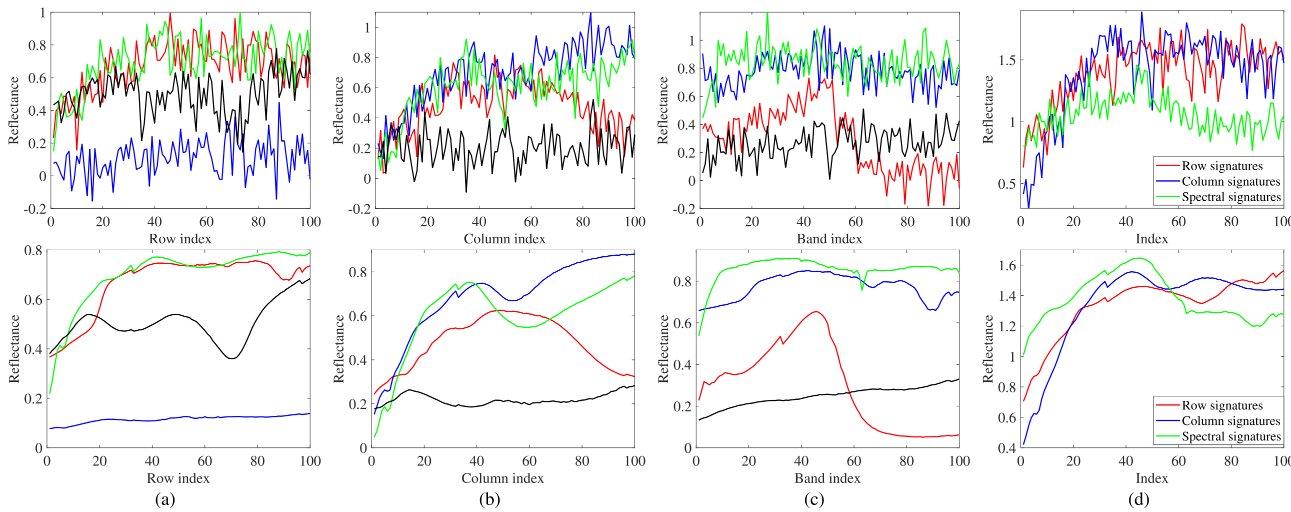

Fig. 1. Signatures curve of two tensors generated by different TR factors. First row: the row fibers of TR factors without smooth structure. Second row: the row fibers of TR factors with smooth structure. (a)–(c) Distribution of each row fiber of the TR factors $\mathcal{G}^{(1)} \in \mathbb{R}^{2\times100\times2}$, $\mathcal{G}^{(2)} \in \mathbb{R}^{2\times100\times2}$, and $\mathcal{G}^{(3)} \in \mathbb{R}^{2\times100\times2}$, respectively. (d) Distributions of row, column, and spectral signatures randomly extracted from the tensor $\mathcal{X} = \Phi(\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \mathcal{G}^{(3)})$.

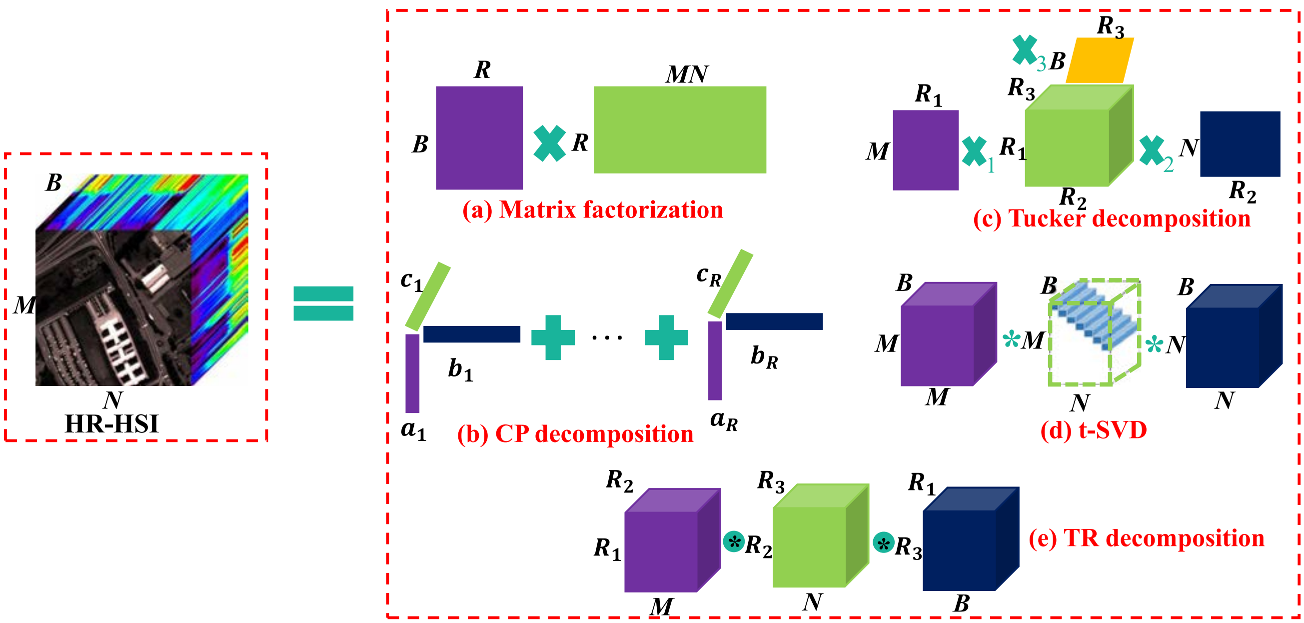

Fig. 2. Matrix/tensor decomposition of a 3-D tensor with the size of $\mathbb{R}^{M\times N\times B}$. For (a), the tensor is reshaped to the matrix. (a) Matrix factorization. (b) CP decomposition. (c) Tucker decomposition. (d) t-SVD. (e) TR decomposition.

HR-HSI. The CP decomposition-based approaches decompose the HR-HSI into the sum of rank-1 tensors, as presented in Fig. 2(b), which assumes that the correlations of spatial–spectral dimensions are the same. However, the correlation of spectral dimension is strong than the spatial dimension in true HR-HSI. The Tucker decomposition [see Fig. 2(c)] decomposes the HR-HSI by using one core tensor and a set of factor matrices. However, the core tensor is independent of the LR-HSI and HR-MSI degradation processing from the HR-HSI. Moreover, since the existence of the core tensor, the number of decomposed parameters increases exponentially following its dimensions. t-SVD [see Fig. 2(d)] is based on a new definition of tensor–tensor product, which maintains some properties that are similar to the matrix case. However, the degradation process of LR-HSI and HR-MSI cannot be represented under the t-SVD framework, and it mainly explores the correlation of one mode of high-dimensional data. In this work, we propose to employ the TR decomposition for approximating the HR-HSI. The number of decomposed variables is much smaller than that of Tucker decomposition. Moreover, all TR factors participate in the degradation process of LR-HSI and HR-MSI, as shown in Fig. 3. Furthermore, according to (3), TR factors can be circularly shifted and treated equivalently; thus, it can effectively balance the correlations of all dimensions than CP decomposition and t-SVD.

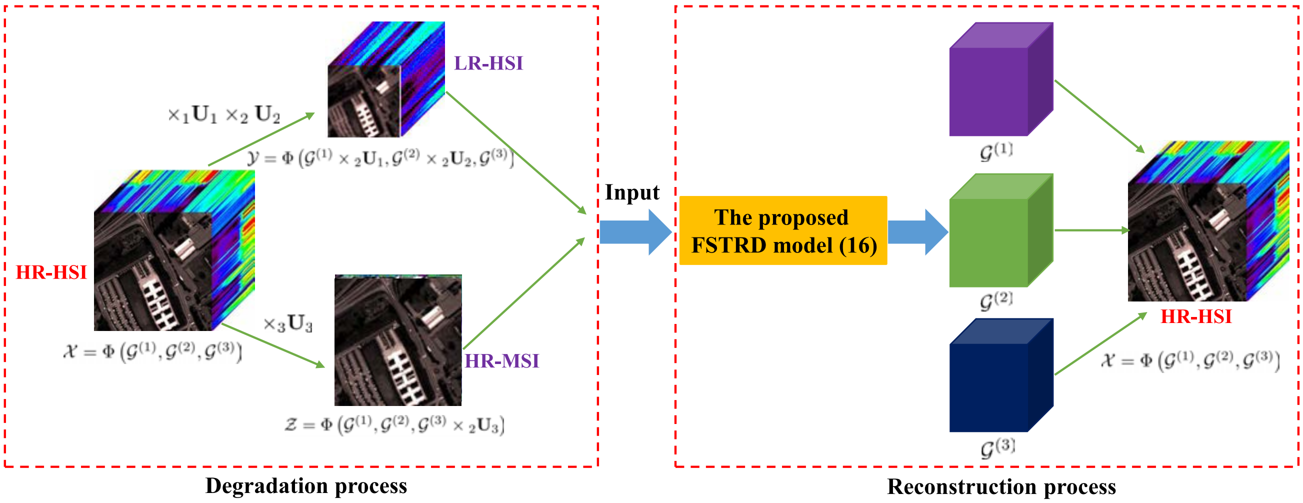

Fig. 3. Illustration of the proposed method.

Although the TR decomposition has been widely used in other fields recently, such as high-dimensional image completion [52], [53] and HSI restoration [60], these applications that use TR decomposition try to restore the image from the missing or noisy data. The proposed FSTRD, however, tries to reconstruct the HR-HSI from the paired LR-HSI and HR-MSI and illustrates the degradation model of fusion problem from the perspective of TR decomposition (see Fig. 3). Moreover, the proposed method first explores that each TR factor inherits the potential piecewise smoothness of original HR-HSI in each dimension and then incorporates the factor smoothed regularization to the TR framework, which can improve the application ability of TR decomposition. The framework of the proposed method is shown in Fig. 3.

The remainder of this article is organized as follows. Section II describes the notations and related fusion framework. The proposed fusion model and its optimization are presented in Section III. Section IV shows the extensive experimental results and discussions. The conclusion is given in Section V.

II. Notations and Problem Formulation

A. Notations

In this article, we use the same tensor notations mainly from the literature [51], [61]. Lowercase and uppercase are employed to denote scalars, i.e., $m, M \in \mathbb{R}$. Boldface lowercase is employed to denote vectors, i.e., $\mathbf{x} \in \mathbb{R}^M$. Matrices are denoted by boldface capital letter, i.e., $\mathbf{X} \in \mathbb{R}^{M\times N}$. Tensors with n-order ($n \geq 3$) are denoted by calligraphic letter, i.e., $\mathcal{X} \in \mathbb{R}^{I_1\times I_2\times\cdots\times I_n}$, where I_i is the dimension of the ith mode. Thus, vectors and matrices are the first- and second-order tensors, respectively. $\mathcal{X}(i_1, i_2, \ldots, i_n)$ or $x_{i_1 i_2,\ldots,i_n}$ is the element value of $\mathcal{X}$ in location $(i_1, i_2, \ldots, i_n)$. Moreover, we summarize the operations of tensor in Table I. In the next, we give the definition of TR decomposition used in this article.

TR decomposition decomposes a tensor into a series of third-order factor tensors $\mathcal{G} = \{\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \ldots, \mathcal{G}^{(n)}\}$, where $\mathcal{G}^{(k)} \in \mathbb{R}^{r_k\times I_k\times r_{k+1}}$ [51]. For a n-order tensor $\mathcal{X} \in \mathbb{R}^{I_1\times I_2\times\cdots\times I_n}$, its TR decomposition is defined as

$$\begin{aligned}\mathcal{X}(i_1, i_2, \ldots, i_n) &= \mathrm{Tr}(\mathbf{G}^{(1)}(i_1)\mathbf{G}^{(2)}(i_2), \ldots, \mathbf{G}^{(n)}(i_n))\\ &= \mathrm{Tr}\left(\prod_{k=1}^{n}\mathbf{G}^{(k)}(i_k)\right)\end{aligned} \tag{1}$$

where $\mathbf{G}^{(k)}(i_k) \in \mathbb{R}^{r_k\times r_{k+1}}$ denotes the i_kth lateral slice matrix of factor tensor $\mathcal{G}^{(k)}$ and $\mathrm{Tr}(\cdot)$ is the matrix trace operation. Since trace operation is defined for square matrix, TR decomposition sets $r_1 = r_{n+1}$, and $r = [r_1, r_2, \ldots, r_n]$ is the TR-rank of tensor $\mathcal{X}$. In this article, TR decomposition is briefly denoted as notation $\mathcal{X} = \Phi(\mathcal{G})$.

The multilinear product of two adjacent factor tensors $\mathcal{G}^{(k)} \in \mathbb{R}^{r_k\times I_k\times r_{k+1}}$ and $\mathcal{G}^{(k+1)} \in \mathbb{R}^{r_{k+1}\times I_{k+1}\times r_{k+2}}$ is denoted as $\mathcal{G}^{(k,k+1)} \in \mathbb{R}^{r_k\times I_k I_{k+1}\times r_{k+2}}$, and each lateral slice of $\mathcal{G}^{(k,k+1)}$ is defined as

$$\mathbf{G}^{(k,k+1)}((j_k-1)I_k + i_k) = \mathbf{G}^{(k)}(i_k)\mathbf{G}^{(k+1)}(j_k) \tag{2}$$

for $i_k = 1, \ldots, I_k$, $j_k = 1, \ldots, I_{k+1}$, where $\mathbf{G}^{(k,k+1)} \in \mathbb{R}^{r_k\times r_{k+2}}$. Based on the multilinear product operation, $\mathcal{G}^{(1,\ldots,k)} \in \mathbb{R}^{r_1\times I_1,\ldots,I_k\times r_{k+1}}$ is the multilinear product of the first k factor tensors, $\mathcal{G}^{(k+1,\ldots,n)} \in \mathbb{R}^{r_{k+1}\times I_{k+1},\ldots,I_n\times r_1}$ is the multilinear product of the last $n-k$ factor tensors, and $\mathcal{G}^{\neq k} \in \mathbb{R}^{r_{k+1}\times I_{k+1},\ldots,I_n I_1,\ldots,I_{k-1}\times r_k}$ is the multilinear product of all factor tensors except kth factor tensor.

Let $\overleftarrow{\mathcal{X}}_k \in \mathbb{R}^{I_k\times I_{k+1}\cdots\times I_n\times I_1\times\cdots\times I_{k-1}}$ be a n-order tensor; then, $\overleftarrow{\mathcal{X}}_k$ can be regarded as the circularly shifts of the dimensions of $\mathcal{X}$ by k. If the TR decomposition of $\mathcal{X}$ is $\mathcal{X} = \Phi(\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \ldots, \mathcal{G}^{(n)})$, then the TR decomposition of $\overleftarrow{\mathcal{X}}_k$ can be expressed as

$$\overleftarrow{\mathcal{X}}_k = \Phi(\mathcal{G}^{(k)}, \mathcal{G}^{(k+1)}, \ldots, \mathcal{G}^{(n)}, \mathcal{G}^{(1)}, \ldots, \mathcal{G}^{(k-1)}). \tag{3}$$

Under the relation of (3), each $\mathcal{G}^{(k)}$ can be shifted to the first position. Based on this property, the matrix representation of TR decomposition can be formulated as

$$\mathbf{X}_{<k>} = \overleftarrow{\mathbf{X}}_{(k,2)} = \mathbf{G}^{(k)}_{(2)}\left(\mathbf{G}^{(\neq k)}_{<2>}\right)^T \tag{4}$$

TABLE I
TENSOR OPERATIONS

Notations	operations
$\mathbf{X}_{(k)} \in \mathbb{R}^{I_k \times I_1 I_2 \cdots I_n}$	the first mode-k matricization of tensor $\mathcal{X}$.
$\mathbf{X}_{<k>} \in \mathbb{R}^{I_k \times I_{k+1} \cdots I_n I_1 \cdots I_{k-1}}$	the second mode-k matricization of tensor $\mathcal{X}$.
$\text{Fold}_k(\cdot)$	the inverse operator of mode-k matricization.
$< \mathcal{X}, \mathcal{Y} >$	inner product of two tensors: $\sum_{i_1,i_2,\cdots,i_n} x_{i_1,i_2,\cdots,i_n} y_{i_1,i_2,\cdots,i_n}$.
$\|\mathcal{X}\|_F$	Frobenius-norm of tensor $\mathcal{X}$: $\sqrt{< \mathcal{X}, \mathcal{X} >}$.
$\|\mathcal{X}\|_1$	ℓ_1-norm of tensor $\mathcal{X}$: $\sum_{i_1,i_2,\cdots,i_n} \|x_{i_1,i_2,\cdots,i_n}\|$.
$\mathcal{Y} = \mathcal{X} \times_k \mathbf{A}$	mode-k multiplication of $\mathcal{X}$: $\mathcal{Y}_{i_1,\cdots,i_{k-1},j,i_{k+1},\cdots,i_n} = \sum_{i_k} x_{i_1,i_2,\cdots,i_n} a_{j,i_k}$.
$\mathbf{Y}_{(k)} = \mathbf{A}\mathbf{X}_{(k)}$	matrix representation of $\mathcal{Y}$: $\mathcal{Y} = \mathcal{X} \times_k \mathbf{A}$.

which is an important conversion for updating each factor tensor. The other efficient property is that, if the TR decomposition of $\mathcal{X}$ is $\mathcal{X} = \Phi(\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \ldots, \mathcal{G}^{(n)})$, then the mode-$k$ multiplication of $\mathcal{X}$ can be formulated as

$$\mathcal{X} \times_k \mathbf{A} = \Phi(\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \ldots, \mathcal{G}^{(k)} \times_2 \mathbf{A}, \ldots, \mathcal{G}^{(n)}). \tag{5}$$

B. Problem Formulation

In the real scene, HSI and MSI are third-order tensor data, where the first two modes are spatial sizes of the image, and the third mode is the spectral dimension. Thus, the given LR-HSI and HR-MSI can be denoted as $\mathcal{Y} \in \mathbb{R}^{m \times n \times B}$ and $\mathcal{Z} \in \mathbb{R}^{M \times N \times b}$, where $m \times n$ and $M \times N$ are spatial sizes of LR-HSI and HR-MSI, respectively, and B and b are the number of the spectral band of LR-HSI and HR-MSI, respectively. The goal of HSI and MSI fusion is to estimate the HR-HSI $\mathcal{X} \in \mathbb{R}^{M \times N \times B}$ from input LR-HSI and HR-MSI, where $M > m$, $N > n$, and $B > b$. As the imaging object and the spectral number between LR-HSI and HR-HSI are the same, then the acquisition process of LR-HSI can be regarded as the spatially downsampling from the HR-HSI, which can be formulated as

$$\mathcal{Y} = \mathcal{X} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 + \mathcal{N}_y \tag{6}$$

where $\mathbf{U}_1 \in \mathbb{R}^{m \times M}$ and $\mathbf{U}_2 \in \mathbb{R}^{n \times N}$ are the degradation operator of blurring and downsampling in spatial height and width modes, respectively, and $\mathcal{N}_y \in \mathbb{R}^{m \times n \times B}$ is the Gaussian noise contained in LR-HSI. Analogously, the imaging object and spatial sizes of HR-MSI and HR-HSI are the same; then, HR-MSI can be formulated as the spectrally downsampling from the HR-HSI, and thus, the following relationship is satisfied:

$$\mathcal{Z} = \mathcal{X} \times_3 \mathbf{U}_3 + \mathcal{N}_z \tag{7}$$

where $\mathbf{U}_3 \in \mathbb{R}^{b \times B}$ is the spectral downsampling matrix of the multispectral imaging sensor and $\mathcal{N}_z \in \mathbb{R}^{M \times N \times b}$ is the Gaussian noise contained in HR-MSI. Base on two degradation models of (6) and (7), directly estimating the HR-HSI $\mathcal{X}$ can be solved the following minimization problem:

$$\min_{\mathcal{X}} \frac{1}{2}\|\mathcal{Y} - \mathcal{X} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2\|_F^2 + \frac{\lambda}{2}\|\mathcal{Z} - \mathcal{X} \times_3 \mathbf{U}_3\|_F^2 \tag{8}$$

where λ is the parameter used to balance two terms.

C. Factorization-Based Methods

Directly estimating the HR-HSI $\mathcal{X}$ from problem (8) is an ill-posed inverse problem; the reason is that the number of total measurements from $\mathcal{Y}$ and $\mathcal{Z}$ is much smaller than that of the unknown variable. To solve the ill-posed fusion problem, regularization is an effective tool by exploring the prior knowledge about the desired HR-HSI $\mathcal{X}$.

As HR-HSI has a strong correlation in spectral dimension, i.e., each spectral signature of HR-HSI can be denoted as the linear relationships of a small number of endmembers, low-rank matrix factorization regularization on $\mathbf{X}_{(3)}$ (the mode-3 matricization of tensor $\mathcal{X}$) has been widely employed to regularize the ill-posed fusion problem (8). Therefore, classical matrix factorization [28], [30] was first proposed to estimate the HR-HSI from a pair of LR-HSI and HR-MSI, which can be formulated as

$$\min_{\mathbf{E},\mathbf{A}} \frac{1}{2}\|\mathbf{Y}_{(3)} - \mathbf{E}\mathbf{A}(\mathbf{U}_1 \otimes \mathbf{U}_2)^T\|_F^2 + \frac{\lambda}{2}\|\mathbf{Z}_{(3)} - \mathbf{U}_3\mathbf{E}\mathbf{A}\|_F^2 \tag{9}$$

where $\mathbf{E} \in \mathbb{R}^{B \times r}$ and $\mathbf{A} \in \mathbb{R}^{r \times MN}$ ($r \ll B$) are the basis matrix and corresponding coefficient matrix, respectively. In fusion problem (9), HR-HSI is factorized into two factors $\mathbf{E}$ and $\mathbf{A}$, i.e., $\mathbf{X}_{(3)} = \mathbf{E}\mathbf{A}$.

Notice that existing matrix factorization-based fusion methods only capture the correlation of HR-HSI in the spectral dimension but ignore the spatial correlation. Moreover, due to the third-order tensor nature of HR-HSI and some strong correlations in different dimensions, tensor regularization-based approaches have been suggested in the literature [34], [43], [57], [62] and achieved some superior performance via comparing to the matrix factorization-based methods. Since classic Tucker and CP ranks can depict the correlation in different dimensions for high-order data, Tucker decomposition [34], [57] and CP decomposition [43] methods are popular for fusing the LR-HSI and HR-MSI. Recently, TR decomposition is a novel tensor rank characterization, which decomposes a tensor into a series of third-order factor tensors. The decomposition form and relationship between tensor element and factor can be found in (1). Based on the notation of TR decomposition, the HR-HSI can be represented as TR decomposition as follows:

$$\mathcal{X} = \Phi(\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \mathcal{G}^{(3)}) \tag{10}$$

where $\mathcal{G}^{(1)} \in \mathbb{R}^{r_1 \times M \times r_2}$, $\mathcal{G}^{(2)} \in \mathbb{R}^{r_2 \times N \times r_3}$, and $\mathcal{G}^{(3)} \in \mathbb{R}^{r_3 \times B \times r_1}$ are three TR factors.

Based on the efficient representation of TR decomposition, previous works widely employed TR decomposition for characterizing the tensor rank under the context of HSI processing [53], [60]. The superiority of TR decomposition to approximate a high-order tensor is analyzed in [60]. Based on the relationship of mode-k multiplication of $\mathcal{X}$ presented in (5), the degradation process (6) can be rewritten as

$$\mathcal{Y} = \Phi(\mathcal{G}^{(1)} \times_2 \mathbf{U}_1, \mathcal{G}^{(2)} \times_2 \mathbf{U}_2, \mathcal{G}^{(3)}) + \mathcal{N}_y. \tag{11}$$

Moreover, under the framework of TR decomposition, the HR-MSI can be expressed as follows:

$$\mathcal{Z} = \Phi(\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \mathcal{G}^{(3)} \times_2 \mathbf{U}_3) + \mathcal{N}_z. \tag{12}$$

By exploring the TR representation of $\mathcal{X}$, the TR decomposition was employed to fuse the LR-HSI and HR-MSI [55]. By combing (11) and (12), the model can be formulated as

$$\min_{\mathcal{G}^{(1)},\mathcal{G}^{(2)},\mathcal{G}^{(3)}} \frac{1}{2}\|\mathcal{Y} - \Phi(\mathcal{G}^{(1)} \times_2 \mathbf{U}_1, \mathcal{G}^{(2)} \times_2 \mathbf{U}_2, \mathcal{G}^{(3)})\|_F^2 + \frac{\lambda}{2}\|\mathcal{Z} - \Phi(\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \mathcal{G}^{(3)} \times_2 \mathbf{U}_3)\|_F^2. \tag{13}$$

Matrix/tensor decomposition-based approaches solve the fusion problem by estimating the decomposition factors of HR-HSI from the LR-HSI and HR-MSI. These methods can take full advantage of the low-rank characteristic of HR-HSI and transform the high-dimensional data into a low-dimensional subspace [63], which can significantly reduce redundancy and computational cost.

III. Proposed Factor Smoothed Tensor Ring Decomposition Method

Since TR representation gives a compact and efficient approximation for high-order tensor data HR-HSI, TR decomposition model (13) can replace other matrix and tensor representations for LR-HSI and HR-MSI fusion. However, the original TR decomposition model (13) only considers the high-correlation of HR-HSI in the spatial–spectral dimension. Moreover, directly estimating the TR factors from model (13) is an unstable problem. To obtain a stable solution and better reconstruct the HR-HSI, additional prior knowledge about the unknown variables should be taken into consideration.

A. Proposed Model

Besides the high-correlation property of HR-HSI, the spatial–spectral piecewise smooth structure is also an important prior for reconstructing the HR-HSI [64]. TR decomposition uses a set of latent tensor factors to represent the HR-HSI; thus, additional spatial–spectral piecewise smooth prior knowledge cannot be directly designed to HR-HSI itself. To keep the original prior of HR-HSI in the TR decomposition, the factor regularization should be investigated in the TR decomposition framework for LR-HSI and HR-MSI fusion

$$\min_{\mathcal{G}^{(1)},\mathcal{G}^{(2)},\mathcal{G}^{(3)}} \frac{1}{2}\|\mathcal{Y} - \Phi(\mathcal{G}^{(1)} \times_2 \mathbf{U}_1, \mathcal{G}^{(2)} \times_2 \mathbf{U}_2, \mathcal{G}^{(3)})\|_F^2 + \frac{\lambda}{2}\|\mathcal{Z} - \Phi(\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \mathcal{G}^{(3)} \times_2 \mathbf{U}_3)\|_F^2 + \tau\sum_{k=1}^{3} F_k(\mathcal{G}^{(k)}) \tag{14}$$

where $F_k(\mathcal{G}^{(k)})$ is the regularization term for $\mathcal{G}^{(k)}$, which is used to capture the spatial–spectral piecewise smooth structure of HR-HSI. τ is the regularization parameter employed to balance the fidelity term and the regularization term.

To explore the property of tensor factor $\mathcal{G}^{(k)}$ and design effective regularization, we should establish the relationship between the original HR-HSI and factor $\mathcal{G}^{(k)}$. Based on the matrix representation of TR decomposition, the TR decomposition can be transformed as

$$\mathbf{X}_{<k>} = \mathbf{G}_{(2)}^{(k)}\left(\mathbf{G}_{<2>}^{(\neq k)}\right)^T, \quad (k = 1, 2, 3). \tag{15}$$

From the perspective of matrix factorization, it is easy to understand that each column vector of $\mathbf{X}_{<k>}$ can be regarded as the representation of a linear combination of all column of the factor $\mathbf{G}_{(2)}^{(k)}$, indicating that all columns of $\mathbf{G}_{(2)}^{(k)}$ are a set of basis of the low-dimensional space of $\mathbf{X}_{<k>}$. Since the HR-HSI has the piecewise smooth structure in two spatial dimensions and one spectral dimension, each column of $\mathbf{X}_{<k>}$ is a continuous data. Based on the fact that continuous bases can represent continuous data, hence, the constraint of continuity of all columns of three factors $\mathbf{G}_{(2)}^{(k)}$ $(k = 1, 2, 3)$ can preserve the piecewise smooth structure of $\mathcal{X}$ in two spatial dimensions and one spectral dimension. To maintain the factor continuity and regularize the TR decomposition model, we introduce TV regularization [29], [65] to constrain the TR factors. Moreover, the weighted TV and iteratively updating the weights strategies are employed to promote the continuity of each column of $\mathbf{G}_{(2)}^{(k)}$. In summary, the proposed FSTRD model for LR-HSI and HR-MSI fusion can be formulated as

$$\min_{\mathcal{G}^{(1)},\mathcal{G}^{(2)},\mathcal{G}^{(3)}} \frac{1}{2}\|\mathcal{Y} - \Phi(\mathcal{G}^{(1)} \times_2 \mathbf{U}_1, \mathcal{G}^{(2)} \times_2 \mathbf{U}_2, \mathcal{G}^{(3)})\|_F^2 + \frac{\lambda}{2}\|\mathcal{Z} - \Phi(\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \mathcal{G}^{(3)} \times_2 \mathbf{U}_3)\|_F^2 + \tau\sum_{k=1}^{3}\|\mathcal{W}^{(k)} \odot (\mathcal{G}^{(k)} \times_2 \mathbf{D})\|_1 \tag{16}$$

where $\mathbf{D}$ is the first-order difference square matrix, whose dimension is associated with the second dimension of $\mathcal{G}^{(k)}$, and $\mathcal{W}^{(k)} \in \mathbb{R}^{r_k \times I_k \times r_{k+1}}$ is the nonnegative weighted tensor.

The proposed model can thoroughly capture prior knowledge of HR-HSI. The first two terms are the data-fitting terms, imposing that the HR-HSI $\mathcal{X}$ should be able to represent the input LR-HSI data $\mathcal{Y}$ and HR-MSI data $\mathcal{Z}$ according to the spatial and spectral degraded model formulated in (11) and (12). Moreover, the data-fitting terms also hide that the TR representation is introduced to explore the strong correlation of HR-HSI in all dimensions. The third term of factor smoothed regularization, in particular, can help to preserve the piecewise smooth structure and suppress the discontinuity caused by the noise. It is worth noting that although He *et al.* [53] employed TR decomposition with TV regularization for remote sensing inpainting, this is very different from our work. He *et al.* [53] introduced TR decomposition to capture the low-rank prior of remote sensing image and directly applied TV regularization to the image itself. In our work, we explore the degraded relationship between the HR-HSI with LR-HSI and HR-MSI under the TR framework and further excavate the smoothed prior for the TR factor rather than the image itself.

B. Optimization

In this section, we design an efficient optimization algorithm for estimating three TR factors from the proposed FSTRD model (16). The optimization of the FSTRD model is not jointly convex for three factors $\mathcal{G}^{(k)}$ $(k = 1, 2, 3)$, but it is convex for each separable variable. Therefore, we employ the PAM framework [66], [67] to solve it, which can be guaranteed that the solution converges to a critical point of the objective function.

Let $f(\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \mathcal{G}^{(3)})$ be the objective function (16); then, we introduce proximal term for updating each factor under the solved framework of PAM. Therefore, the optimization of FSTRD model can be alternately solved by the following three subproblems:

$$\begin{cases} \mathcal{G}^{(1),i+1} = \arg\min\limits_{\mathcal{G}^{(1)}} f(\mathcal{G}^{(1)}, \mathcal{G}^{(2),i}, \mathcal{G}^{(3),i}) \\ \qquad\qquad + \frac{\rho}{2}\|\mathcal{G}^{(1)} - \mathcal{G}^{(1),i}\|_F^2, \\ \mathcal{G}^{(2),i+1} = \arg\min\limits_{\mathcal{G}^{(2)}} f(\mathcal{G}^{(1),i+1}, \mathcal{G}^{(2)}, \mathcal{G}^{(3),i}) \\ \qquad\qquad + \frac{\rho}{2}\|\mathcal{G}^{(2)} - \mathcal{G}^{(2),i}\|_F^2, \\ \mathcal{G}^{(3),i+1} = \arg\min\limits_{\mathcal{G}^{(3)}} f(\mathcal{G}^{(1),i+1}, \mathcal{G}^{(2),i+1}, \mathcal{G}^{(3)}) \\ \qquad\qquad + \frac{\rho}{2}\|\mathcal{G}^{(3)} - \mathcal{G}^{(3),i}\|_F^2 \end{cases} \tag{17}$$

where i, $(\rho/2)\|\cdot\|_F^2$, and ρ are iteration number, proximal term, and positive proximal parameter, respectively. In the next, the detailed solution of each subproblem is presented.

1) Optimization With Respect to $\mathcal{G}^{(1)}$: The subproblem of $\mathcal{G}^{(1)}$ can be formulated as

$$\begin{aligned} \min_{\mathcal{G}^{(1)}} \ & \frac{1}{2}\|\mathcal{Y} - \Phi(\mathcal{G}^{(1)} \times_2 \mathbf{U}_1, \mathcal{G}^{(2),i} \times_2 \mathbf{U}_2, \mathcal{G}^{(3),i})\|_F^2 \\ & + \frac{\lambda}{2}\|\mathcal{Z} - \Phi(\mathcal{G}^{(1)}, \mathcal{G}^{(2),i}, \mathcal{G}^{(3),i} \times_2 \mathbf{U}_3)\|_F^2 \\ & + \tau\|\mathcal{W}^{(1)} \odot (\mathcal{G}^{(1)} \times_2 \mathbf{D})\|_1 + \frac{\rho}{2}\|\mathcal{G}^{(1)} - \mathcal{G}^{(1),i}\|_F^2. \end{aligned} \tag{18}$$

Because of the nonsmooth term of ℓ_1-norm, the close-form solution of $\mathcal{G}^{(1)}$ cannot be directly obtained. The alternating direction method of multipliers (ADMM) algorithm [68]–[71] can efficiently solve the nonsmooth problem. First, we introduce one auxiliary variable $\mathcal{R}_1$; the unconstrained problem is equivalent to the following minimization problem:

$$\begin{aligned} \min_{\mathcal{G}^{(1)}, \mathcal{R}_1} \ & \frac{1}{2}\|\mathcal{Y} - \Phi(\mathcal{G}^{(1)} \times_2 \mathbf{U}_1, \mathcal{G}^{(2),i} \times_2 \mathbf{U}_2, \mathcal{G}^{(3),i})\|_F^2 \\ & + \frac{\lambda}{2}\|\mathcal{Z} - \Phi(\mathcal{G}^{(1)}, \mathcal{G}^{(2),i}, \mathcal{G}^{(3),i} \times_2 \mathbf{U}_3)\|_F^2 + \tau\|\mathcal{W}^{(1)} \odot \mathcal{R}_1\|_1 \\ & + \frac{\rho}{2}\|\mathcal{G}^{(1)} - \mathcal{G}^{(1),i}\|_F^2, \quad \text{s.t.}, \ \mathcal{R}_1 = \mathcal{G}^{(1)} \times_2 \mathbf{D}. \end{aligned} \tag{19}$$

The augmented Lagrangian function of the problem is formulated as

$$\begin{aligned} & L_\beta(\mathcal{G}^{(1)}, \mathcal{R}_1, \mathcal{M}_1) \\ & = \frac{1}{2}\|\mathcal{Y} - \Phi(\mathcal{G}^{(1)} \times_2 \mathbf{U}_1, \mathcal{G}^{(2),i} \times_2 \mathbf{U}_2, \mathcal{G}^{(3),i})\|_F^2 \\ & \quad + \frac{\lambda}{2}\|\mathcal{Z} - \Phi(\mathcal{G}^{(1)}, \mathcal{G}^{(2),i}, \mathcal{G}^{(3),i} \times_2 \mathbf{U}_3)\|_F^2 + \tau\|\mathcal{W}^{(1)} \odot \mathcal{R}_1\|_1 \\ & \quad + \frac{\rho}{2}\|\mathcal{G}^{(1)} - \mathcal{G}^{(1),i}\|_F^2 + \frac{\beta}{2}\left\|\mathcal{R}_1 - \mathcal{G}^{(1)} \times_2 \mathbf{D} + \frac{\mathcal{M}_1}{\beta}\right\|_F^2 \end{aligned} \tag{20}$$

where $\mathcal{M}_1$ is the Lagrangian multiplier and β is a positive penalty parameter. Then, the solution of $\mathcal{G}^{(1)}$ can be achieved by iteratively optimized the following subproblems:

a) $\mathcal{R}_1$-subproblem: The minimization of $\mathcal{R}_1$ is

$$\min_{\mathcal{R}_1} \ \tau\|\mathcal{W}^{(1)} \odot \mathcal{R}_1\|_1 + \frac{\beta}{2}\left\|\mathcal{R}_1 - \mathcal{G}^{(1)} \times_2 \mathbf{D} + \frac{\mathcal{M}_1}{\beta}\right\|_F^2 \tag{21}$$

which is a weighted ℓ_1-norm minimization, and the closed-form solution is obtained by soft-threshold shrinkage operator

$$\mathcal{R}_1 = \text{sign}(\mathcal{J}_1)\max\left(|\mathcal{J}_1| - \frac{\tau}{\beta}\mathcal{W}^{(1)}, 0\right) \tag{22}$$

where $\mathcal{J}_1 = \mathcal{G}^{(1)} \times_2 \mathbf{D} - (\mathcal{M}_1/\beta)$, and $\text{sign}(\cdot)$ is symbolic function. According to the reweighted strategy [72], the weight coefficient is set as

$$\mathcal{W}^{(1)} = 1/(|\mathcal{J}_1| + \text{eps}) \tag{23}$$

where eps is a positive small constant avoiding singularity.

b) $\mathcal{G}^{(1)}$-subproblem: The minimization of $\mathcal{G}^{(1)}$ is

$$\begin{aligned} \min_{\mathcal{G}^{(1)}} \ & \frac{1}{2}\|\mathcal{Y} - \Phi(\mathcal{G}^{(1)} \times_2 \mathbf{U}_1, \mathcal{G}^{(2),i} \times_2 \mathbf{U}_2, \mathcal{G}^{(3),i})\|_F^2 \\ & + \frac{\lambda}{2}\|\mathcal{Z} - \Phi(\mathcal{G}^{(1)}, \mathcal{G}^{(2),i}, \mathcal{G}^{(3),i} \times_2 \mathbf{U}_3)\|_F^2 \\ & + \frac{\rho}{2}\|\mathcal{G}^{(1)} - \mathcal{G}^{(1),i}\|_F^2 + \frac{\beta}{2}\left\|\mathcal{R}_1 - \mathcal{G}^{(1)} \times_2 \mathbf{D} + \frac{\mathcal{M}_1}{\beta}\right\|_F^2. \end{aligned} \tag{24}$$

Based on the matrix representation of TR decomposition (4), let $\mathbf{P}_1 = ((\mathcal{G}^{(2),i} \times_2 \mathbf{U}_2)\mathcal{G}^{(3),i})_{<2>}^T$ and $\mathbf{P}_2 = (\mathcal{G}^{(2),i}(\mathcal{G}^{(3),i} \times_2 \mathbf{U}_3))_{<2>}^T$, and the problem of $\mathcal{G}^{(1)}$ can be converted to the following minimization:

$$\begin{aligned} \min_{\mathbf{G}_{(2)}^{(1)}} \ & \frac{1}{2}\|\mathbf{Y}_{<1>} - \mathbf{U}_1\mathbf{G}_{(2)}^{(1)}\mathbf{P}_1\|_F^2 + \frac{\lambda}{2}\|\mathbf{Z}_{<1>} - \mathbf{G}_{(2)}^{(1)}\mathbf{P}_2\|_F^2 \\ & + \frac{\rho}{2}\|\mathbf{G}_{(2)}^{(1)} - \mathbf{G}_{(2)}^{(1),i}\|_F^2 + \frac{\beta}{2}\left\|\mathbf{R}_{\mathbf{1}(2)} - \mathbf{D}\mathbf{G}_{(2)}^{(1)} + \frac{\mathbf{M}_{\mathbf{1}(2)}}{\beta}\right\|_F^2 \end{aligned} \tag{25}$$

which is a quadratic problem, and the closed-form solution is obtained by solving the following linear system:

$$\begin{aligned} & \mathbf{U}_1^T\mathbf{U}_1\mathbf{G}_{(2)}^{(1)}\mathbf{P}_1\mathbf{P}_1^T + \lambda\mathbf{G}_{(2)}^{(1)}\mathbf{P}_2\mathbf{P}_2^T + \rho\mathbf{G}_{(2)}^{(1)} + \beta\mathbf{D}^T\mathbf{D}\mathbf{G}_{(2)}^{(1)} \\ & \quad = \mathbf{U}_1^T\mathbf{Y}_{<1>}\mathbf{P}_1^T + \lambda\mathbf{Z}_{<1>}\mathbf{P}_2^T \\ & \qquad + \rho\mathbf{G}_{(2)}^{(1),i} + \beta\mathbf{D}^T\left(\mathbf{R}_{\mathbf{1}(2)} + \frac{\mathbf{M}_{\mathbf{1}(2)}}{\beta}\right). \end{aligned} \tag{26}$$

This liner system is a general Sylvester equation, which can be efficiently solved by the conjugate gradient (CG) method. Then, the factor $\mathcal{G}^{(1)}$ is achieved by folding the solution $\mathbf{G}_{(2)}^{(1)}$, i.e., $\mathcal{G}^{(1)} = \text{Fold}_2(\mathbf{G}_{(2)}^{(1)})$.

The update of Lagrangian multiplier $\mathcal{M}_1$ is formulated as follows:

$$\mathcal{M}_1 \leftarrow \mathcal{M}_1 + \beta(\mathcal{R}_1 - \mathcal{G}^{(1)} \times_2 \mathbf{D}). \tag{27}$$

By iteratively updating $\mathcal{R}_1$, $\mathcal{G}^{(1)}$, and $\mathcal{M}_1$, we can obtain the first factor $\mathcal{G}^{(1),i+1}$.

2) Optimization With Respect to $\mathcal{G}^{(2)}$: The subproblem of $\mathcal{G}^{(2)}$ can be formulated as

$$\begin{aligned}\min_{\mathcal{G}^{(2)}}\ & \frac{1}{2}\|\mathcal{Y} - \Phi(\mathcal{G}^{(1),i+1} \times_2 \mathbf{U}_1, \mathcal{G}^{(2)} \times_2 \mathbf{U}_2, \mathcal{G}^{(3),i})\|_F^2 \\ & + \frac{\lambda}{2}\|\mathcal{Z} - \Phi(\mathcal{G}^{(1),i+1}, \mathcal{G}^{(2)}, \mathcal{G}^{(3),i} \times_2 \mathbf{U}_3)\|_F^2 \\ & + \tau\|\mathcal{W}^{(2)} \odot (\mathcal{G}^{(2)} \times_2 \mathbf{D})\|_1 + \frac{\rho}{2}\|\mathcal{G}^{(2)} - \mathcal{G}^{(2),i}\|_F^2. \end{aligned} \tag{28}$$

The optimization of $\mathcal{G}^{(2)}$ is similar to $\mathcal{G}^{(1)}$. By using ADMM algorithm, the update of each subproblem is formulated as follows:

$$\mathcal{R}_2 = \mathrm{sign}(\mathcal{J}_2)\max\left(|\mathcal{J}_2| - \frac{\tau}{\beta}\mathcal{W}^{(2)}, 0\right) \tag{29}$$

$$\begin{aligned}&\mathbf{U}_2^T\mathbf{U}_2\mathbf{G}^{(2)}_{(2)}\mathbf{Q}_1\mathbf{Q}_1^T + \lambda\mathbf{G}^{(2)}_{(2)}\mathbf{Q}_2\mathbf{Q}_2^T + \rho\mathbf{G}^{(2)}_{(2)} + \beta\mathbf{D}^T\mathbf{D}\mathbf{G}^{(2)}_{(2)} \\ &\quad = \mathbf{U}_2^T\mathbf{Y}_{<2>}\mathbf{Q}_1^T + \lambda\mathbf{Z}_{<2>}\mathbf{Q}_2^T + \rho\mathbf{G}^{(2),i}_{(2)} \\ &\qquad + \beta\mathbf{D}^T\left(\mathbf{R_2}_{(2)} + \frac{\mathbf{M_2}_{(2)}}{\beta}\right)\end{aligned} \tag{30}$$

$$\mathcal{M}_2 \leftarrow \mathcal{M}_2 + \beta(\mathcal{R}_2 - \mathcal{G}^{(2)} \times_2 \mathbf{D}) \tag{31}$$

where $\mathcal{R}_2$ and $\mathcal{M}_2$ are an auxiliary variable and the Lagrangian multiplier, respectively. Moreover, the variables of above equations are defined as

$$\begin{cases}\mathcal{J}_2 = \mathcal{G}^{(2)} \times_2 \mathbf{D} - \dfrac{\mathcal{M}_2}{\beta} \\ \mathcal{W}^{(2)} = 1/(|\mathcal{J}_2| + \mathrm{eps}) \\ \mathbf{Q}_1 = (\mathcal{G}^{(3),i}(\mathcal{G}^{(1),i+1} \times_2 \mathbf{U}_1))^T_{<2>} \\ \mathbf{Q}_2 = ((\mathcal{G}^{(3),i} \times_2 \mathbf{U}_3)\mathcal{G}^{(1),i+1})^T_{<2>}.\end{cases} \tag{32}$$

3) Optimization With Respect to $\mathcal{G}^{(3)}$: The subproblem of $\mathcal{G}^{(3)}$ can be formulated as

$$\begin{aligned}\min_{\mathcal{G}^{(3)}}\ & \frac{1}{2}\|\mathcal{Y} - \Phi(\mathcal{G}^{(1),i+1} \times_2 \mathbf{U}_1, \mathcal{G}^{(2),i+1} \times_2 \mathbf{U}_2, \mathcal{G}^{(3)})\|_F^2 \\ & + \frac{\lambda}{2}\|\mathcal{Z} - \Phi(\mathcal{G}^{(1),i+1}, \mathcal{G}^{(2),i+1}, \mathcal{G}^{(3)} \times_2 \mathbf{U}_3)\|_F^2 \\ & + \tau\|\mathcal{W}^{(3)} \odot (\mathcal{G}^{(3)} \times_2 \mathbf{D})\|_1 + \frac{\rho}{2}\|\mathcal{G}^{(3)} - \mathcal{G}^{(3),i}\|_F^2. \end{aligned} \tag{33}$$

We can also use ADMM to optimize factor $\mathcal{G}^{(3)}$; the solution of each subproblem can be iteratively updated as follows:

$$\mathcal{R}_3 = \mathrm{sign}(\mathcal{J}_3)\max\left(|\mathcal{J}_3| - \frac{\tau}{\beta}\mathcal{W}^{(3)}, 0\right) \tag{34}$$

$$\begin{aligned}&\lambda\mathbf{U}_3^T\mathbf{U}_3\mathbf{G}^{(3)}_{(2)}\mathbf{T}_2\mathbf{T}_2^T + \mathbf{G}^{(2)}_{(2)}\mathbf{T}_1\mathbf{T}_1^T + \rho\mathbf{G}^{(3)}_{(2)} + \beta\mathbf{D}^T\mathbf{D}\mathbf{G}^{(3)}_{(2)} \\ &\quad = \mathbf{Y}_{<3>}\mathbf{T}_1^T + \lambda\mathbf{U}_3^T\mathbf{Z}_{<3>}\mathbf{T}_2^T + \rho\mathbf{G}^{(3),i}_{(2)} \\ &\qquad + \beta\mathbf{D}^T\left(\mathbf{R_3}_{(2)} + \frac{\mathbf{M_3}_{(2)}}{\beta}\right)\end{aligned} \tag{35}$$

$$\mathcal{M}_3 \leftarrow \mathcal{M}_3 + \beta(\mathcal{R}_3 - \mathcal{G}^{(3)} \times_2 \mathbf{D}) \tag{36}$$

Algorithm 1 PAM Algorithm for Solving Model (16)

Input: LR-HSI $\mathcal{Y}$, HR-MSI $\mathcal{Z}$, degraded operators $\mathbf{U}_1$, $\mathbf{U}_2$, and $\mathbf{U}_3$, parameters λ, τ, ρ, β, TR rank $r = [r_1, r_2, r_3]$, and $i_{\max}$.

```
Initialize: G^(1),0, G^(2),0, G^(3),0, and i = 0.
while not coverged do
  Initialize: G^(1) = G^(1),i and M_1 = 0.
  while not coverged do
    Update R_1 via (22).
    Update G^(1) via (26).
    Update M_1 via (27).
  end while
  Output: G^(1),i+1.
  Initialize: G^(2) = G^(2),i and M_2 = 0.
  while not coverged do
    Update R_2 via (29).
    Update G^(2) via (30).
    Update M_2 via (31).
  end while
  Output: G^(2),i+1.
  Initialize: G^(3) = G^(3),i and M_3 = 0.
  while not coverged do
    Update R_3 via (34).
    Update G^(3) via (35).
    Update M_3 via (36).
  end while
  Output: G^(3),i+1.
  Check ||Φ(G)^(i+1) − Φ(G)^i||_F / ||Φ(G)^i||_F ≤ ε and i < i_max.
  i = i + 1.
end while
```

Output: HR-HSI $\mathcal{X} = \Phi(\mathcal{G}^{(1),i+1}, \mathcal{G}^{(2),i+1}, \mathcal{G}^{(3),i+1})$.

where the variables are computed by

$$\begin{cases}\mathcal{J}_3 = \mathcal{G}^{(3)} \times_2 \mathbf{D} - \dfrac{\mathcal{M}_3}{\beta} \\ \mathcal{W}^{(3)} = 1/(|\mathcal{J}_3| + \mathrm{eps}) \\ \mathbf{T}_1 = ((\mathcal{G}^{(1),i+1} \times_2 \mathbf{U}_1)(\mathcal{G}^{(2),i+1} \times_2 \mathbf{U}_2))^T_{<2>} \\ \mathbf{T}_2 = (\mathcal{G}^{(1),i+1}\mathcal{G}^{(2),i+1})^T_{<2>}.\end{cases} \tag{37}$$

Summarizing the optimization procedure of each factor $\mathcal{G}^{(1)}$, $\mathcal{G}^{(2)}$, and $\mathcal{G}^{(3)}$, we present the whole PAM algorithm for solving model (16) in Algorithm 1. By introducing the proximal operator in the alternating minimization framework, the sequence $\{(\mathcal{G}^{(1),i}, \mathcal{G}^{(2),i}, \mathcal{G}^{(3),i})\}$ generated by Algorithm 1 converges to a critical point if it is bounded [67].

C. Computational Complexity

We analyze the computational complexity of the proposed algorithm as follows. For achieving the HR-HSI $\mathcal{X} \in \mathbb{R}^{M \times N \times B}$, we assume that the TR rank is equal and set as $r_1 = r_2 = r_3 = R$. In the PAM framework, the TR factors $\mathcal{G}^{(1)}$, $\mathcal{G}^{(2)}$, and $\mathcal{G}^{(3)}$ are optimized by ADMM. For the $\mathcal{G}^{(1)}$ subproblem, the updates of auxiliary variable $\mathcal{R}_1$ and Lagrangian multiplier $\mathcal{M}_1$ are simple algebraic operation, and the total computation complexity is $O(2MR^2)$. The main computational cost of solving the $\mathcal{G}^{(1)}$ subproblem is the multiplication of the system matrix times a vector on CG algorithm, and it can be implemented efficiently by the matrix representation with complexity $O(M^2R^2 + MR^4)$. Thus, the computational complexity of each $\mathcal{G}^{(1)}$ iteration is $O(K_{\mathrm{CG}}(M^2R^2 + MR^4) + 2MR^2)$. Analogously, the computational costs of the $\mathcal{G}^{(2)}$ and $\mathcal{G}^{(3)}$ iterations are $O(K_{\mathrm{CG}}(N^2R^2 + NR^4) + 2NR^2)$ and $O(K_{\mathrm{CG}}(B^2R^2 + BR^4) + 2BR^2)$, respectively. Therefore, the total computation complexity of each iteration in Algorithm 1 is $O(K_{\mathrm{ADMM}}K_{\mathrm{CG}}((M^2+N^2+B^2)R^2+(M+N+B)R^4)+2K_{\mathrm{ADMM}}(M+N+B)R^2)$, where K_{ADMM} and K_{CG} are the iteration numbers of ADMM and CG algorithms, respectively.

IV. Experimental Results and Discussion

In this section, extensive datasets are used to illustrate the performance of the proposed FSTRD method. Both quantitative and visual results are employed to compare the superiority of our method. We compare the result with other DL-based state-of-the-art LR-HSI and HR-MSI fusion methods, including factor smoothed matrix factorization (Hysure) [29], fast fusion method (FUSE) [30], nonlocal sparse tensor factorization (NLSTF) [38], coupled CP factorization (STEREO) [43], coupled sparse tensor factorization (CSTF) [57], low tensor train rank (LTTR)-based method [49], region-based low-rank matrix decomposition fusion (RLRMDF) method [35], convolutional neural network (CNN)-based denoiser fusion (CNN-Fus) method [21], deep spatiospectral attention CNN-based fusion method (HSRnet) [23], and coupled TR factorization (CTRF) [55]. The model parameters of all compared methods are selected according to the author's suggestions in their paper and released code to achieve the best results. The parameter selection of the proposed method will give in the discussion. Before the simulated process, the pixel value of HR-HSI is scaled in $[0, 1]$.

A. Simulated Data Experiment

1) Dataset: To thoroughly demonstrate the effectiveness of the proposed method, we select four different datasets to test. Two datasets are from computer version society, while another two datasets are real remote sensing HSI. These four datasets are often used as benchmark datasets. The detailed description of experimental datasets is presented as follows.

1) The first is CAVE dataset,[1] which contains 32 indoor HSIs imaged by generalized assorted pixel camera in real-world scenes. The size of each HSI is $512 \times 512 \times 31$, where 512×512 is the number of spatial pixels, and 31 is the number of spectral bands. The wavelength of 31 spectral bands is from 400 to 700 nm with an interval of 10 nm. Five different HSI scenes (Balloons, Toy, Peppers, Flowers, and Painting) are selected as the ground-truth datasets to test.
2) The second dataset is Harvard,[2] which has 50 HSIs of indoor and outdoor scenes under daylight illumination. The spatial and spectral sizes of each HSI in Harvard datasets are 1040×1392 and 31, respectively, where

[1]http://www1.cs.columbia.edu/CAVE/databases/multispectral
[2]http://vision.seas.harvard.edu/hyperspec/download.html

TABLE II
Quantitative Indices Comparison of Different Methods Under Different Noise Cases on the CAVE Dataset

Method	PSNR	SSIM	ERGAS	SAM	CC
SNR = 10 dB					
HySure	25.48	0.499	3.374	26.811	0.925
FUSE	21.42	0.195	4.934	47.418	0.866
NLSTF	28.93	0.569	2.301	35.564	0.970
STEREO	30.63	0.726	1.856	27.534	0.977
CSTF	31.34	0.720	1.750	29.393	0.982
LTTR	23.97	0.292	4.197	45.634	0.913
RLRMDF	22.43	0.230	4.893	48.926	0.885
CNN-Fus	27.04	0.440	2.855	32.676	0.949
HSRnet	**33.82**	**0.866**	**1.264**	**15.425**	**0.990**
CTRF	30.97	0.720	1.808	28.714	0.980
FSTRD	32.80	0.768	1.487	27.015	0.987
SNR = 15 dB					
HySure	27.49	0.693	3.153	25.350	0.941
FUSE	26.36	0.388	2.904	38.184	0.950
NLSTF	32.80	0.767	1.482	27.625	0.987
STEREO	32.83	0.816	1.473	22.885	0.985
CSTF	33.73	0.833	1.415	24.899	0.989
LTTR	29.78	0.577	2.097	35.378	0.975
RLRMDF	27.31	0.439	2.767	40.076	0.957
CNN-Fus	31.76	0.689	1.631	27.256	0.983
HSRnet	**35.57**	**0.912**	**1.045**	**11.967**	**0.993**
CTRF	33.17	0.799	1.449	24.916	0.986
FSTRD	35.33	0.862	1.130	21.464	0.991
SNR = 20 dB					
HySure	32.68	0.779	1.562	24.604	0.989
FUSE	30.87	0.611	1.782	29.768	0.981
NLSTF	34.10	0.805	1.306	25.996	0.989
STEREO	34.71	0.867	1.250	20.137	0.989
CSTF	35.69	0.885	1.160	21.295	0.991
LTTR	34.98	0.825	1.165	25.197	0.992
RLRMDF	32.09	0.681	1.629	31.894	0.985
CNN-Fus	34.59	0.832	1.235	23.869	0.990
HSRnet	36.14	0.904	0.966	16.688	0.993
CTRF	35.07	0.841	1.244	22.924	0.988
FSTRD	**37.80**	**0.921**	**0.870**	**16.000**	**0.994**
SNR = 25 dB					
HySure	35.41	0.881	1.167	17.741	0.994
FUSE	34.30	0.779	1.235	22.959	0.991
NLSTF	37.53	0.911	0.928	19.186	0.993
STEREO	36.19	0.899	1.119	17.931	0.990
CSTF	37.53	0.917	0.968	17.824	0.993
LTTR	38.36	0.922	0.805	17.782	0.994
RLRMDF	36.39	0.857	1.091	24.709	0.993
CNN-Fus	37.14	0.882	0.923	16.561	0.993
HSRnet	37.91	0.940	0.801	13.368	0.994
CTRF	37.06	0.891	1.028	19.436	0.991
FSTRD	**39.00**	**0.948**	**0.782**	**12.972**	**0.995**
SNR = 30 dB					
HySure	36.97	0.920	1.031	13.979	0.995
FUSE	37.02	0.881	0.944	17.006	0.995
NLSTF	39.14	0.933	0.813	17.582	0.995
STEREO	37.33	0.916	1.004	15.630	0.991
CSTF	38.74	0.937	0.914	15.879	0.994
LTTR	40.17	0.961	0.670	12.353	0.995
RLRMDF	39.55	0.934	0.830	18.923	0.995
CNN-Fus	40.23	0.942	0.704	12.273	**0.996**
HSRnet	40.08	0.958	0.687	11.404	**0.996**
CTRF	38.71	0.928	0.927	16.173	0.993
FSTRD	**40.78**	**0.963**	**0.668**	**10.969**	**0.996**

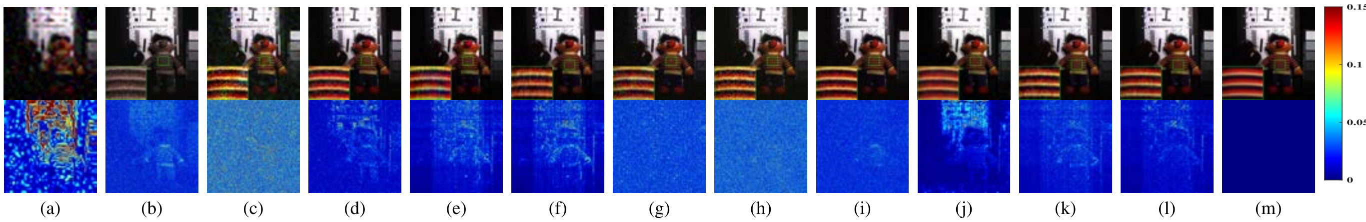

Fig. 4. First row: reconstructed results of Toy under SNR = 15 dB on the CAVE dataset. The false color image is composed by bands (R: 27, G: 17, and B: 11). Second row: corresponding error maps between the original and reconstructed images in averaging three bands. (a) LR-HSI. (b) HySure. (c) FUSE. (d) NLSTF. (e) STEREO. (f) CSTF. (g) LTTR. (h) RLRMDF. (i) CNN-Fus. (j) HSRnet. (k) CTRF. (l) FSTRD. (m) Original.

31 spectral bands are ranged from 420 to 720 nm with an incremental of 10 nm. We choose five HSIs (numbers are img1, imgb8, imgc4, imgd3, and imgh0) with different textures and details as the HR-HSI, and the spatial size of 1024×1024 is cropped.

3) The third dataset is the University of Pavia,[3] which was collected by using Reflective Optics System Imaging Spectrometer (ROSIS). The original Pavia dataset has 115 spectral bands with a spatial size of 610×340. After removing the low SNR bands and cropping the subregion, we choose the up-left 256×256 spatial pixels with 93 spectral bands as the HR-HSI.

4) The fourth dataset is Indian Pines,[4] which was captured by the NASA AVIRIS instrument over the Indian Pines test site. The original image consists of 145×145 pixels with the number of 220 spectral bands, and some bands are seriously degraded by noise. In our experiment, we extract these data to $128 \times 128 \times 184$ in the whole dimension as the ground truth.

2) Generation of LR-HSI and HR-MSI: To generate LR-HSI, we first filter the HR-HSI by averaged blurring kernel with a size of 9×9. Then, the LR-HSI is obtained by downsampling the blurred image, and each pixel of LR-HSI is selected from $s \times s$ pixels, where s is the downsampling factor. In our experiments, the downsampling factors s of CAVE, Harvard, Pavia, and Indian Pines datasets are set as 16, 32, 8, and 4, respectively. The generation of HR-MSI is simulated by downsampling the HR-HSI along the spectral mode using the spectral response matrix. We employ the spectral response matrix from Nikon D700 camera[5] to generate HR-MSI (RGB image) for CAVE and Harvard datasets. The Pavia dataset is degraded by the IKONOS-like reflectance spectral response filter [30] to simulate HR-MSI with four bands. The HR-MSI with six bands of the Indian Pines dataset is generated by choosing the spectral bands from the original HR-HSI [57]. When we obtain the LR-HSI and HR-MSI, the Gaussian noise with the same signal-to-noise ratios (SNRs) is simultaneously added to the LR-HSI and HR-MSI, where SNR is varied from 10 to 30 dB with an interval of 5 dB.

3) Quantitative Indices: To thoroughly evaluate the performance of reconstructed HR-HSI from LR-HSI and HR-MSI, we employ peak signal-to-noise ratio (PSNR), structure similarity (SSIM) [73], erreur relative global adimensionnelle de synthese (ERGAS) [64], spectral angle mapper (SAM) [74], and cross correlation (CC) [44] as quantitative indices. These indices contain the evaluation of spatial and spectral information preservation. In general, the larger PSNR, SSIM, and CC, and smaller SAM and ERGAS values indicate the better result.

B. Experimental Results on Simulated Data

1) Results on CAVE Dataset: There are a total of 25 experiments in the CAVE dataset since we choose five different scenes with five different noise cases. A representative case is used to compare the visual reconstructed result of various methods. Fig. 4 shows the reconstructed result of the Toy scene under SNR = 15 dB, and the false-color image is composed of bands 27, 17, and 11. The first row presents the reconstructed HR-HSI by different methods, and the second row illustrates corresponding error maps (the difference between original and reconstructed) obtained by averaging three bands. To better present the visual comparison, we enlarge a detailed region of reconstructed results. As we can see from the result, although HySure can reconstruct the HR-HSI, the contrast is destroyed compared with the original image. The result obtained by FUSE exists noise and is distorted. Tensor decomposition-based and DL-based fusion methods achieve better results than matrix factorization-based approaches, but the details cannot be better preserved, as shown in the enlarged region. By incorporating the factor smoothed prior to the TR decomposition, the proposed FSTRD achieves the best-reconstructed result, preserving most of the details and removing high-intensity noises. From error map results, the reconstructed HSR-HSI produced by our FSTRD has fewer errors than that of other compared methods.

Table II presents the quantitative indices comparison of different methods under different noise cases on the CAVE dataset. The values of all indices are obtained by averaging the result of five different HSI scenes, and the best results are highlighted in bold. From the table, matrix factorization-based methods, HySure and FUSE, obtain poor results compared with tensor-and DL-based approaches. NLSTF, LTTR, RLRMDF, and CNN-Fus cannot obtain satisfactory results in low SNR cases, which illustrates the sensitivity to noise. DL-based method HSRnet obtains better results than other compared methods in low SNR cases because it employs extensive training data. It is clear to see that the proposed

[3]http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes

[4]https://engineering.purdue.edu/biehl/MultiSpec/hyperspectral.html

[5]https://www.maxmax.com/spectral_response.htm

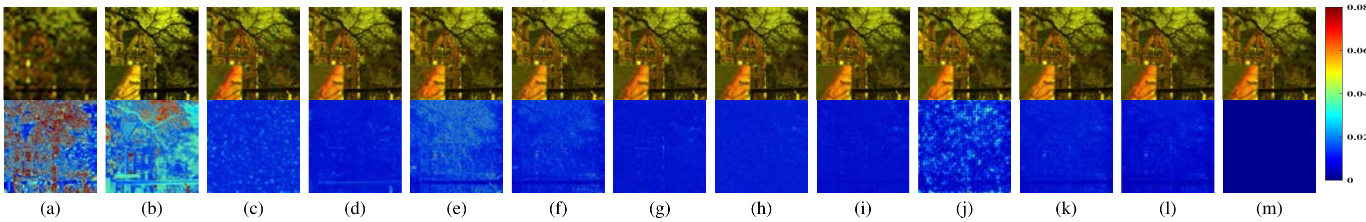

Fig. 5. First row: reconstructed results of imgb8 under SNR = 20 dB on the Harvard dataset. The false color image is composed by bands (R: 28, G: 16, and B: 1). Second row: corresponding error maps between the original and reconstructed images in averaging three bands. (a) LR-HSI. (b) HySure. (c) FUSE. (d) NLSTF. (e) STEREO. (f) CSTF. (g) LTTR. (h) RLRMDF. (i) CNN-Fus. (j) HSRnet. (k) CTRF. (l) FSTRD. (m) Original.

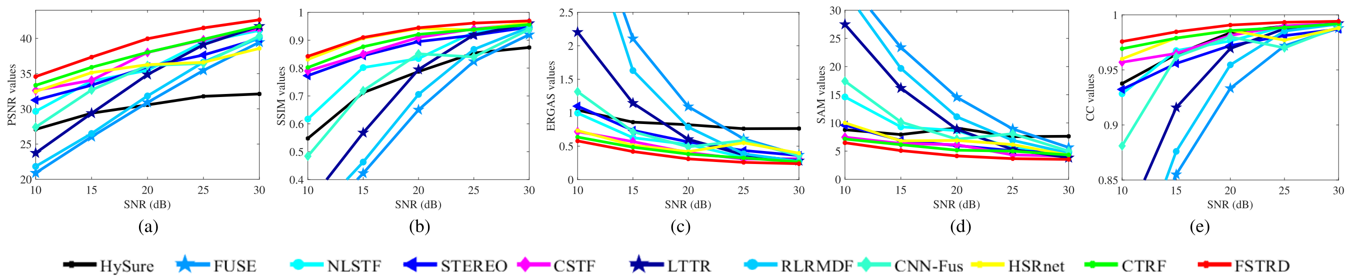

Fig. 6. Quantitative indices comparison of different methods under different noise cases on the Harvard dataset. (a) PSNR. (b) SSIM. (c) ERGAS. (d) SAM. (e) CC.

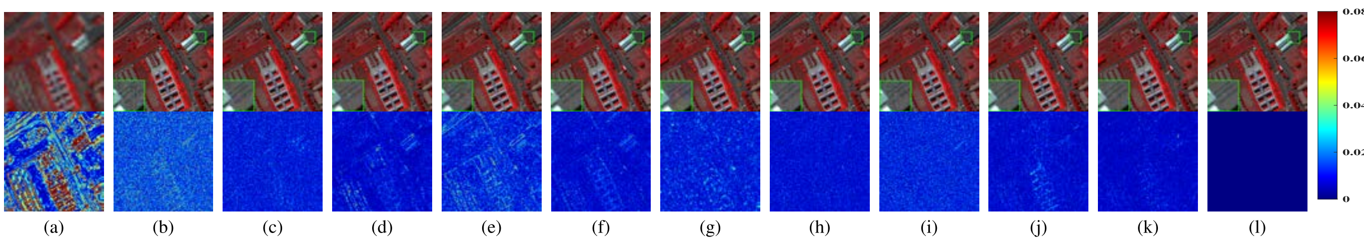

Fig. 7. First row: reconstructed results of SNR = 25 dB on the Pavia dataset. The false color image is composed by bands (R: 82, G: 38, and B: 34). Second row: corresponding error maps between the original and reconstructed images in averaging three bands. (a) LR-HSI. (b) HySure. (c) FUSE. (d) NLSTF. (e) STEREO. (f) CSTF. (g) LTTR. (h) RLRMDF. (i) CNN-Fus. (j) CTRF. (k) FSTRD. (l) Original.

FSTRD achieves competitive results in low SNR cases and outperforms all comparison methods in high SNR cases; the reason is that our method explores the factor smoothness in the framework of tensor decomposition.

2) Results on Harvard Dataset: The total number of experimental cases in the Harvard dataset is the same as the CAVE dataset; we also present a representative case to compare the reconstructed result. Fig. 5 shows the reconstructed imgb8 image with noise case SNR = 20 dB by the test approaches. It can be seen that HySure destroys the image contrast, and FUSE, NLSTF, STEREO, RLRMDF, and CNN-Fus cannot eliminate the noise in the images. From the results of the enlarged region, we can find that HSRnet blurs the image details, while the proposed FSTRD achieves the best results in recovering the image details and restraining noise. The second row of Fig. 5 shows that FSTRD obtains a smaller error compared with other methods, which again demonstrates the superiority of our method.

Fig. 6 illustrates the five average quantitative indices of different methods in the conditions of different noise intensities. Overall, it can be observed that the proposed FSTRD outperforms other fusion approaches, with regard to the five indices in all SNR cases. Meanwhile, compared with other algorithms, HSRnet, CTRF, and the proposed FSTRD achieve obvious improvement in low SNRs. This is mainly because HSRnet uses many training data, while CTRF and our FSTRD methods are based on efficient TR decomposition. Moreover, FSTRD is much better than HSRnet and CTRF in all noise cases due to the sparse prior of latent TR factor.

3) Results on Pavia and Indian Pines Datasets: Due to the lack of a large number of training data in remote sensing datasets, we abandon the comparison of the HSRnet method on the next remote sensing datasets. However, it has been demonstrated that the proposed method is competitive with the HSRnet method in CAVE and Harvard datasets. Figs. 7 and 8 show the reconstructed results and residual images obtained by different methods for Pavia and Indian Pines datasets with SNR = 25 dB, respectively. HySure and STEREO are hard to remove the noise in the Pavia dataset. Although other methods can obtain satisfactory reconstructed results, the error maps are shown that our FSTRD achieves a smaller error, indicating the superiorities of preserving most of the details and removing noise of the proposed method. For the results of the Indian Pines dataset, FUSE cannot completely reconstruct the image

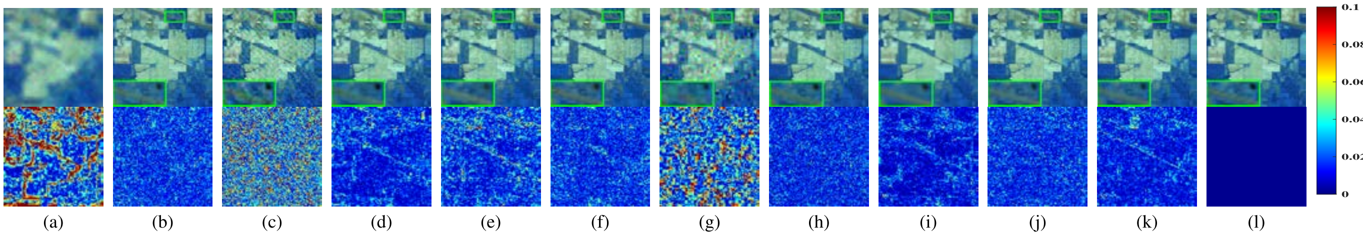

Fig. 8. First row: reconstructed results of SNR = 25 dB on the Indian Pines dataset. The false color image is composed of bands (R: 167, G: 117, and B: 93). Second row: corresponding error maps between the original and reconstructed images in averaging three bands. (a) LR-HSI. (b) HySure. (c) FUSE. (d) NLSTF. (e) STEREO. (f) CSTF. (g) LTTR. (h) RLRMDF. (i) CNN-Fus. (j) CTRF. (k) FSTRD. (l) Original.

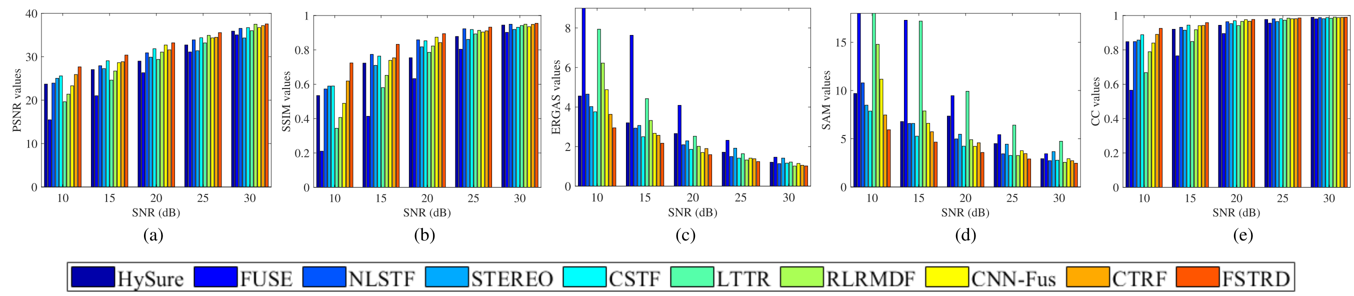

Fig. 9. Quantitative indices comparison of different methods under different noise cases on Pavia and Indian Pines datasets. (a) PSNR. (b) SSIM. (c) ERGAS. (d) SAM. (e) CC.

detail and obtains artifact. The results of the first row present that most of the methods can recover the HR-HSI, but the enlarged region and error map illustrate the improvement of our method.

Fig. 9 presents the quantitative results of different methods by averaging the Pavia and Indian Pines datasets. Similar to CAVE and Harvard datasets, FUSE, LTTR, and RLRMDF obtain poor results in low SNRs, but satisfactory results are achieved in high SNRs, illustrating that these methods are sensitive to noise again. In both low and high SNRs cases, FSTRD always achieves the best results in terms of five quantitative indices. In summary, the above results thoroughly illustrate that the proposed method can achieve the best result for fusing the LR-HSI and HR-MSI, and our method is robust to different datasets and noise intensities. Moreover, the comparisons of five different quantitative indices demonstrate that the proposed method is better to reconstruct the image detail and preserve spectral reflectance.

C. Real Data Experiment

To further demonstrate the effectiveness of the proposed method, we implement a real LR-HSI and HR-MSI fusion experiment. The LR-HSI is collected by the Hyperion sensor, which is of the size $120 \times 120 \times 89$. The HR-MSI with 13 bands is taken by the Sentinel-2A satellite. Four bands with the 10-m spatial resolution are employed for the test, and the spatial downsampling factor is 3, i.e., the size of HR-MSI is $360 \times 360 \times 4$. These four bands are extracted from bands 2, 3, 4, and 8 with the central wavelengths being 490, 560, 665, and 842 nm, respectively. It is worth noting that both the ground-truth spatial and spectral degradation operators are unobtainable for real data. Thus, we estimate the spatial and spectral degradation operators by using the method suggested in [29]. For the parameters selection in real data, we first employ the classical noise estimation algorithm proposed in [75] to roughly estimate the noise intensity and then set the parameters according to the empirical determination presented in the discussion. In our experiment, by using the algorithm [75], the estimation of the noise intensity of the LR-HSI is SNR = 32.75 dB. Since the estimated noise intensity is close to SNR = 30 dB, we set parameters as $\lambda = 0.5$, $\tau = 0.0001$, and $r = [6, 300, 6]$ for the real data.

Fig. 10 shows the false color image of reconstructed results for the real dataset. One can see from the results that all comparison methods can obviously obtain a higher spatial resolution compared with the original LR-HSI, as shown in Fig. 10(a). However, FUSE, NLSTF, STEREO, CSTF, LTTR, and CNN-Fus algorithms cannot completely recover the image details and reconstruct the artifact as shown in the enlarged region. Compared with the image structures of the reference HR-MSI, the proposed method achieves the best result, which indicates the better preservation of the image structure.

To present the quantitative evaluation of different fusion methods on the real dataset, we employ one blind image quality assessment (BIQA) [76] for comparison. Table III lists the nonreference index values of the real dataset, and the results are obtained by averaging all bands. It can be seen that HySure, FUSE, RLRMDF, and CNN-Fus methods obtain better results. However, from the visual results shown in Fig. 10, there are serious spectral distortion and artifacts in their results. Combining the visual comparison and nonreference image assessment index, we can observe that the proposed method obtains a better result.

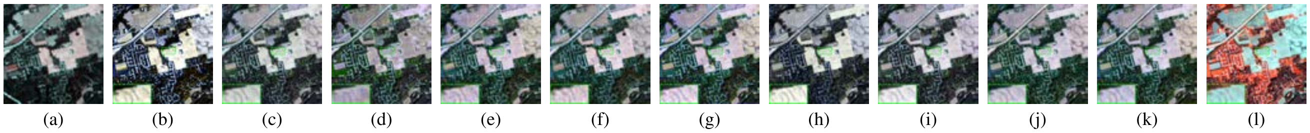

Fig. 10. False color image of the reconstructed results for the real dataset. The false color image is composed by bands (R: 21, G: 15, and B: 1). (a) LR-HSI. (b) HySure. (c) FUSE. (d) NLSTF. (e) STEREO. (f) CSTF. (g) LTTR. (h) RLRMDF. (i) CNN-Fus. (j) CTRF. (k) FSTRD. (l) HR-MSI.

TABLE III
BIQA COMPARISON OF THE REAL DATA

HySure	FUSE	NLSTF	STEREO	CSTF	LTTR	RLRMDF	CNN-Fus	CTRF	FSTRD
0.019	0.015	0.007	0.010	0.008	0.011	0.013	0.027	0.011	0.012

TABLE IV
QUANTITATIVE INDICES COMPARISON OF DIFFERENT METHODS UNDER GAUSSIAN KERNEL ON THE PAVIA DATASET

Method	PSNR	SSIM	ERGAS	SAM	CC
			SNR = 10 dB		
HySure	21.46	0.405	6.326	20.225	0.817
FUSE	20.14	0.330	8.169	27.409	0.697
NLSTF	26.00	0.627	3.703	12.323	0.876
STEREO	26.56	0.623	3.404	10.191	0.878
CSTF	27.30	0.635	3.153	9.598	0.902
LTTR	21.16	0.398	6.741	24.355	0.731
RLRMDF	20.30	0.336	8.719	27.619	0.698
CNN-Fus	24.09	0.488	4.676	15.133	0.832
CTRF	27.34	0.646	3.088	9.342	0.901
FSTRD	**28.33**	**0.713**	**2.845**	**8.013**	**0.920**
			SNR = 30 dB		
HySure	39.23	0.965	0.841	3.179	0.993
FUSE	39.25	0.964	0.828	3.214	0.993
NLSTF	39.42	0.966	0.806	3.040	0.993
STEREO	36.43	0.942	1.139	4.191	0.987
CSTF	39.38	0.965	0.802	2.976	0.993
LTTR	36.56	0.947	1.160	4.405	0.986
RLRMDF	39.58	0.966	0.796	3.050	0.993
CNN-Fus	38.95	0.957	0.859	3.412	0.992
CTRF	39.41	0.967	0.817	3.122	0.993
FSTRD	**40.14**	**0.974**	**0.752**	**2.734**	**0.994**

TABLE V
QUANTITATIVE INDICES COMPARISON OF DIFFERENT METHODS UNDER TWO SUPER-RESOLUTION FACTORS ON THE PAVIA DATASET

Method	PSNR	SSIM	ERGAS	SAM	CC
			s = 8		
Hysure	29.07	0.716	2.863	10.972	0.933
FUSE	30.77	0.790	2.220	8.593	0.955
NLSTF	32.65	0.867	1.747	6.190	0.972
STEREO	31.08	0.833	2.046	6.913	0.957
CSTF	33.90	0.885	1.499	5.081	0.978
LTTR	29.64	0.787	2.452	9.477	0.944
RLRMDF	32.50	0.843	1.794	6.540	0.970
CNN-Fus	33.14	0.860	1.662	5.754	0.974
CTRF	33.38	0.874	1.581	5.705	0.974
FSTRD	**34.88**	**0.913**	**1.345**	**4.370**	**0.982**
			s = 16		
Hysure	24.32	0.624	2.328	15.398	0.941
FUSE	30.51	0.782	1.160	8.892	0.952
NLSTF	32.05	0.859	0.939	6.566	0.968
STEREO	27.72	0.713	1.492	9.100	0.904
CSTF	33.82	0.885	0.757	5.140	0.978
LTTR	29.92	0.805	1.197	8.938	0.948
RLRMDF	32.20	0.832	0.929	6.862	0.968
CNN-Fus	33.06	0.859	0.835	5.810	0.974
CTRF	33.30	0.873	0.800	5.755	0.974
FSTRD	**34.72**	**0.913**	**0.684**	**4.471**	**0.982**

D. Discussion

There are five regularization parameters λ, τ, r_1, r_2, and r_3 and two algorithm parameters ρ and β in our FSTRD method. For the algorithmic parameter, we empirically fix $\rho = 1$ and $\beta = 0.1$ in all experiments. Next, we will analyze the influence of regularization parameters on fusion performance. Moreover, we further test the experiments by blurring the original image with the Gaussian kernel to illustrate the robustness of the proposed method for blurring kernel settings. The above analyses are conducted on the Pavia dataset under the noise levels SNR $=$ 10 and 30 dB.

1) Analysis of λ and τ: Fig. 11 plots the PSNR value of the reconstructed result as a function of parameters λ and τ under two noise SNRs. As we can see from the results, when the noise intensity is high, i.e., SNR $=$ 10 dB, we should choose the combination of larger τ and smaller λ, illustrating the effectiveness of factor smoothed regularization in high noise cases. On the contrary, smaller τ and larger λ should be selected in low noise intensity cases. Hence, according to the noise intensity, we empirically select λ and τ in the set of $\{0.01, 0.5\}$ and $\{0.0001, 0.01, 1\}$, respectively.

TABLE VI
QUANTITATIVE INDICES COMPARISON OF ABLATION EXPERIMENTS ON THE PAVIA DATASET

Method	PSNR	SSIM	ERGAS	SAM	CC
			SNR = 10 dB		
FSTRD w/o spa-spe	27.17	0.630	3.194	9.734	0.899
FSTRD w/o spa	27.40	0.648	3.069	9.241	0.902
FSTRD w/o spe	28.23	0.708	2.853	8.179	0.919
FSTRD	**28.61**	**0.732**	**2.729**	**7.436**	**0.924**
			SNR = 30 dB		
FSTRD w/o spa-spe	39.11	0.963	0.851	3.319	0.992
FSTRD w/o spa	39.89	0.971	0.772	2.862	0.993
FSTRD w/o spe	39.59	0.972	0.799	3.012	0.993
FSTRD	**40.01**	**0.973**	**0.765**	**2.798**	**0.994**

TABLE VII
RUNNING TIME IN SECONDS OF DIFFERENT METHODS

Dataset	HySure	FUSE	NLSTF	STEREO	CSTF	LTTR	RLRMDF	CNN-Fus	CTRF	FSTRD
CAVE	28.61	1.34	20.39	2.00	0.67	159.01	18.32	45.48	6.73	12.53
Harvard	265.39	5.53	104.14	6.00	1.83	673.55	212.88	169.39	21.89	40.10
Pavia	9.55	0.41	10.69	2.32	0.54	109.45	3.82	35.86	4.49	8.24
Indian	1.25	0.18	5.39	1.12	0.74	114.15	2.96	18.13	2.42	7.83
Real	47.80	0.48	61.09	10.71	2.78	200.54	23.09	76.30	17.00	22.92
Average	70.52	1.59	40.34	4.43	1.31	251.34	52.21	69.03	10.51	18.32

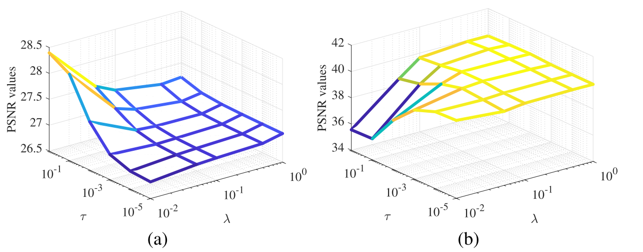

Fig. 11. Sensitivity analysis of parameters λ and τ. (a) SNR = 10 dB. (b) SNR = 30 dB.

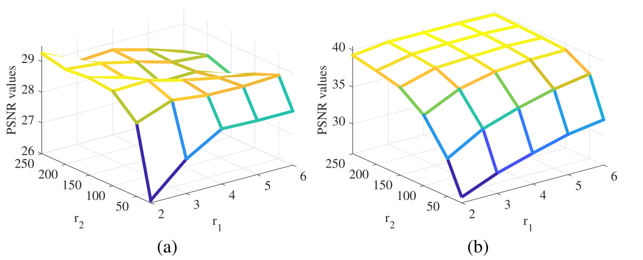

Fig. 12. Sensitivity analysis of parameters TR rank. (a) SNR = 10 dB. (b) SNR = 30 dB.

2) Analysis of TR Rank: The TR rank contains three parameters $r = [r_1, r_2, r_3]$, and the estimation of TR rank is still an open problem. To reduce the parameter selection of TR rank, we first set $r_1 = r_3$; thus, the TR rank $r = [r_1, r_2, r_1]$. Fig. 12 presents the PSNR curves of reconstructed Pavia dataset as a function of r_1 and r_2. As we can see from the results, the smaller r_1 can obtain better results in the low SNR case, and in the high SNR, we should select larger r_1. For the parameter r_2, we should choose larger r_2 in all cases. Therefore, we empirically select r_1 and r_2 in the range of $[2, 6]$ and $[100, 300]$, respectively.

3) Analysis of Blurring Kernel: Table IV lists the reconstructed result on the Pavia dataset under two SNRs by using a Gaussian blurring kernel of size 7×7 with standard deviation 2. We can see that the proposed method can still achieve the best performance compared with other methods, indicating the robustness of our method for different blurring kernels.

4) Analysis of Super-Resolution Factors s: Table V reports the quantitative indices comparison of two super-resolution factors on the Pavia dataset with SNR = 20 dB. We can observe that the proposed method still outperforms other compared approaches as the super-resolution factor increases, and the super-resolution factors cannot largely influence the result.

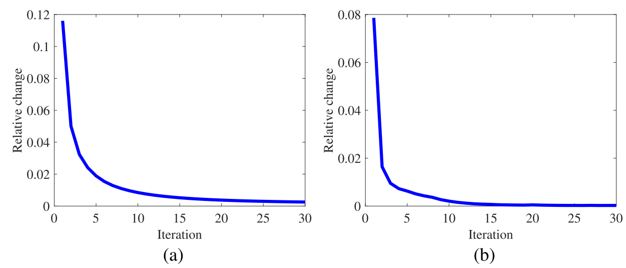

Fig. 13. Relative change values ($\|\mathcal{X}^{k+1} - \mathcal{X}^k\|_F / \|\mathcal{X}^k\|_F$) versus the iteration number of the FSTRD solver in the simulated Pavia dataset. (a) SNR = 10 dB. (b) SNR = 30 dB.

5) Effect of Three TR Factor Smoothed Regularization Terms: To further illustrate the effectiveness of the proposed factor smoothed regularization, we add ablation experiments by disabling in turn spatial and spectral TR factor regularizations in the proposed method. We disable the two spatial and one spectral smoothed regularizations (i.e., $\|\mathcal{W}^{(1)} \odot (\mathcal{G}^{(1)} \times_2 \mathbf{D})\|_1$, $\|\mathcal{W}^{(2)} \odot (\mathcal{G}^{(2)} \times_2 \mathbf{D})\|_1$, and $\|\mathcal{W}^{(3)} \odot (\mathcal{G}^{(3)} \times_2 \mathbf{D})\|_1$), two spatial smoothed regularizations (i.e., $\|\mathcal{W}^{(1)} \odot (\mathcal{G}^{(1)} \times_2 \mathbf{D})\|_1$ and $\|\mathcal{W}^{(2)} \odot (\mathcal{G}^{(2)} \times_2 \mathbf{D})\|_1$), and the spectral smoothed regularization (i.e., $\|\mathcal{W}^{(3)} \odot (\mathcal{G}^{(3)} \times_2 \mathbf{D})\|_1$), which are referred as FSTRD w/o spa-spe, FSTRD w/o spa, and FSTRD w/o spe, respectively. Table VI lists the quantitative indices comparison of ablation experiments on the Pavia dataset by using the proposed FSTRD and its variants. It is clear to see that FSTRD outperforms its variants, and the three TR factor smoothed components contribute significantly to the performance of the proposed method.

6) Numerical Convergence: According to the existing theory of the PAM algorithm, the proposed method can converge to a critical point. Fig. 13 shows the numerical convergence of the proposed method on the simulated Pavia dataset. From the curves, we can see that the relative change value converges to zero as the iteration number increases, indicating the numerical convergence guarantee of the proposed method.

7) Computational Efficiency: Table VII reports the computational time of different methods on the CAVE, Harvard, Pavia, and Indian Pines datasets with SNR = 10 dB and the real dataset. It is worth noting that the time results of CAVE and Harvard datasets are obtained by averaging the five different scenes. The running time results are recorded in MATLAB R2020b using a desktop of 32-GB RAM, with an Intel Core i9-10850K CPU@3.60 GHz. As can be seen from the results, FUSE and CSTF take the shortest running time. The proposed method needs relatively more computation time than FUSE, STEREO, CSTF, and CTRF. The reason is that the PAM algorithm is relatively slow, but the stability and convergence of the algorithm can be guaranteed. In general, the proposed method can achieve better fusion results, and the time cost is acceptable and competitive compared with other methods.

V. Conclusion

In this article, we propose an effective FSTRD method to reconstruct an HR-HSI from a pair of LR-HSI and HR-MSI of the same scene. For better representing the spatial–spectral correlation, the TR decomposition is designed to approximate the HR-HSI. Based on the TR representation, the degradation of LR-HSI and HR-MSI can be obtained by downsampling the TR factor. Thus, the reconstruction of HR-HSI is transformed to estimate the TR factor from the LR-HSI and HR-HSI. Moreover, to further capture the spatial–spectral continuity of HR-HSI, the smoothed regularization is introduced to constrain the TR factor. The optimization of each TR factor is based on an efficient PAM algorithm. Numerical results demonstrate that the proposed method outperforms the state-of-the-art methods, in particular, yielding reconstructed images with higher quality.

Although the proposed FSTRD method can perform better fusion results, there is still room for improvement. For example, automatic regularization and TR rank parameters selection should be overcome in the future, which is expected to improve the application and practicality of the proposed method in related fields. In addition, the organic combination of a TR network with data-driven DL ideas is employed to improve the fusion result.

Acknowledgment

The authors would like to thank the editors and the anonymous reviewers for their constructive comments that significantly improve the quality of this article. They would also like to thank Mr. Jin-Fan Hu for providing all the results of the HSRnet method [23]. They would also like to thank Ms. Jian-Li Wang's help.

References

[1] D. Hong, X. Wu, P. Ghamisi, J. Chanussot, N. Yokoya, and X. X. Zhu, "Invariant attribute profiles: A spatial-frequency joint feature extractor for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 6, pp. 3791–3808, Jun. 2020.

[2] 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.

[3] J. Liu, S. Yuan, X. Zhu, Y. Huang, and Q. Zhao, "Nonnegative matrix factorization with entropy regularization for hyperspectral unmixing," *Int. J. Remote Sens.*, vol. 42, no. 16, pp. 6362–6393, 2021.

[4] M. Uzair, A. Mahmood, and A. Mian, "Hyperspectral face recognition with spatiospectral information fusion and PLS regression," *IEEE Trans. Image Process.*, vol. 24, no. 3, pp. 1127–1137, Mar. 2015.

[5] 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.

[6] N. Yokoya, C. Grohnfeldt, and J. Chanussot, "Hyperspectral and multispectral data fusion: A comparative review of the recent literature," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 2, pp. 29–56, Jun. 2017.

[7] R. Dian, S. Li, B. Sun, and A. Guo, "Recent advances and new guidelines on hyperspectral and multispectral image fusion," *Inf. Fusion*, vol. 69, pp. 40–51, May 2021.

[8] B. Aiazzi, S. Baronti, and M. Selva, "Improving component substitution pansharpening through multivariate regression of MS + pan data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3230–3239, Oct. 2007.

[9] J. Liu, C. Zhou, R. Fei, C. Zhang, and J. Zhang, "Pansharpening via neighbor embedding of spatial details," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 4028–4042, 2021.

[10] J. G. 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.

[11] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "MTF-tailored multiscale fusion of high-resolution MS and PAN imagery," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 5, pp. 591–596, 2006.

[12] J. Yang, X. Fu, Y. Hu, Y. Huang, X. Ding, and J. Paisley, "PanNet: A deep network architecture for pan-sharpening," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 5449–5457.

[13] W. Xie, J. Lei, Y. Cui, Y. Li, and Q. Du, "Hyperspectral pansharpening with deep priors," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 5, pp. 1529–1543, May 2020.

[14] W. Dong, T. Zhang, J. Qu, S. Xiao, J. Liang, and Y. Li, "Laplacian pyramid dense network for hyperspectral pansharpening," *IEEE Trans. Geosci. Remote Sens.*, early access, May 14, 2021, doi: 10.1109/TGRS.2021.3076768.

[15] L. Zhang, J. Zhang, J. Ma, and X. Jia, "SC-PNN: Saliency cascade convolutional neural network for pansharpening," *IEEE Trans. Geosci. Remote Sens.*, early access, Feb. 9, 2021, doi: 10.1109/TGRS.2021.3054641.

[16] W. Dong, S. Hou, S. Xiao, J. Qu, Q. Du, and Y. Li, "Generative dual-adversarial network with spectral fidelity and spatial enhancement for hyperspectral pansharpening," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jun. 10, 2021, doi: 10.1109/TNNLS.2021.3084745.

[17] R. Dian, S. Li, A. Guo, and L. Fang, "Deep hyperspectral image sharpening," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 11, pp. 5345–5355, Nov. 2018.

[18] W. Xie, X. Jia, Y. Li, and J. Lei, "Hyperspectral image super-resolution using deep feature matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 6055–6067, Aug. 2019.

[19] Q. Xie, M. Zhou, Q. Zhao, Z. Xu, and D. Meng, "MHF-Net: An interpretable deep network for multispectral and hyperspectral image fusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Aug. 11, 2020, doi: 10.1109/TPAMI.2020.3015691.

[20] W. Wang, W. Zeng, Y. Huang, X. Ding, and J. Paisley, "Deep blind hyperspectral image fusion," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2019, pp. 4150–4159.

[21] R. Dian, S. Li, and X. Kang, "Regularizing hyperspectral and multispectral image fusion by CNN denoiser," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 3, pp. 1124–1135, Mar. 2021.

[22] Z. Zhu, J. Hou, J. Chen, H. Zeng, and J. Zhou, "Hyperspectral image super-resolution via deep progressive zero-centric residual learning," *IEEE Trans. Image Process.*, vol. 30, pp. 1423–1438, 2021.

[23] 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.*, early access, Jun. 9, 2021, doi: 10.1109/TNNLS.2021.3084682.

[24] X. Zhang, W. Huang, Q. Wang, and X. Li, "SSR-NET: Spatial–spectral reconstruction network for hyperspectral and multispectral image fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 5953–5965, Jul. 2021.

[25] Y. Fu, T. Zhang, Y. Zheng, D. Zhang, and H. Huang, "Hyperspectral image super-resolution with optimized RGB guidance," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 11653–11662.

[26] T. Uezato, D. Hong, N. Yokoya, and W. He, "Guided deep decoder: Unsupervised image pair fusion," in *Proc. ECCV*, 2020, pp. 87–102.

[27] Z. Wang, B. Chen, R. Lu, H. Zhang, H. Liu, and P. K. Varshney, "FusionNet: An unsupervised convolutional variational network for hyperspectral and multispectral image fusion," *IEEE Trans. Image Process.*, vol. 29, pp. 7565–7577, 2020.

[28] 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.

[29] M. Simões, J. Bioucas-Dias, L. B. Almeida, and J. Chanussot, "A convex formulation for hyperspectral image superresolution via subspace-based regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 6, pp. 3373–3388, Jun. 2015.

[30] Q. Wei, N. Dobigeon, and J. Tourneret, "Fast fusion of multi-band images based on solving a Sylvester equation," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4109–4121, Nov. 2015.

[31] Q. Wei, J. Bioucas-Dias, N. Dobigeon, and J. Y. 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.

[32] W. Dong *et al.*, "Hyperspectral image super-resolution via non-negative structured sparse representation," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 2337–2352, May 2016.

[33] 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.

[34] K. Wang, Y. Wang, X.-L. Zhao, J. C.-W. Chan, Z. Xu, and D. Meng, "Hyperspectral and multispectral image fusion via nonlocal low-rank tensor decomposition and spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 11, pp. 7654–7671, Nov. 2020.

[35] J. Liu, Z. Wu, L. Xiao, J. Sun, and H. Yan, "A truncated matrix decomposition for hyperspectral image super-resolution," *IEEE Trans. Image Process.*, vol. 29, pp. 8028–8042, 2020.

[36] W. He, N. Yokoya, and X. Yuan, "Fast hyperspectral image recovery of dual-camera compressive hyperspectral imaging via non-iterative subspace-based fusion," *IEEE Trans. Image Process.*, vol. 30, pp. 7170–7183, 2021.

[37] Y. Chen, T.-Z. Huang, W. He, X.-L. Zhao, H. Zhang, and J. Zeng, "Hyperspectral image denoising using factor group sparsity-regularized nonconvex low-rank approximation," *IEEE Trans. Geosci. Remote Sens.*, early access, Sep. 16, 2021, doi: 10.1109/TGRS.2021.3110769.

[38] 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.*, Jul. 2017, pp. 5344–5353.

[39] 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 2019.

[40] T. Xu, T.-Z. Huang, L.-J. Deng, X.-L. Zhao, and J. Huang, "Hyperspectral image superresolution using unidirectional total variation with tucker decomposition," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 4381–4398, 2020.

[41] 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.

[42] N. Liu, L. Li, W. Li, R. Tao, J. E. Fowler, and J. Chanussot, "Hyperspectral restoration and fusion with multispectral imagery via low-rank tensor-approximation," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 9, pp. 7817–7830, Sep. 2021.

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

[44] Y. Xu, Z. Wu, J. Chanussot, P. Comon, and Z. Wei, "Nonlocal coupled tensor CP decomposition for hyperspectral and multispectral image fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 348–362, Jan. 2020.

[45] Y. Xu, Z. Wu, J. Chanussot, and Z. Wei, "Nonlocal patch tensor sparse representation for hyperspectral image super-resolution," *IEEE Trans. Image Process.*, vol. 28, no. 6, pp. 3034–3047, Jun. 2019.

[46] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor robust principal component analysis with a new tensor nuclear norm," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 4, pp. 925–938, Jan. 2020.

[47] J. Hou, F. Zhang, H. Qiu, J. Wang, Y. Wang, and D. Meng, "Robust low-tubal-rank tensor recovery from binary measurements," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Mar. 3, 2021, doi: 10.1109/TPAMI.2021.3063527.

[48] T.-X. Jiang, X.-L. Zhao, H. Zhang, and M. K. Ng, "Dictionary learning with low-rank coding coefficients for tensor completion," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Aug. 31, 2021, doi: 10.1109/TNNLS.2021.3104837.

[49] 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.

[50] M. Ding, X. Fu, T. Z. Huang, J. Wang, and X. L. Zhao, "Hyperspectral super-resolution via interpretable block-term tensor modeling," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 3, pp. 641–656, Apr. 2021.

[51] Q. Zhao, G. Zhou, S. Xie, L. Zhang, and A. Cichocki, "Tensor ring decomposition," 2016, *arXiv:1606.05535*. [Online]. Available: http://arxiv.org/abs/1606.05535

[52] L. Yuan, J. Cao, X. Zhao, Q. Wu, and Q. Zhao, "Higher-dimension tensor completion via low-rank tensor ring decomposition," in *Proc. APSIPA ASC*, Nov. 2018, pp. 1071–1076.

[53] 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.

[54] Y. Chen, T. Huang, W. He, N. Yokoya, and X. Zhao, "Hyperspectral image compressive sensing reconstruction using subspace-based nonlocal tensor ring decomposition," *IEEE Trans. Image Process.*, vol. 29, pp. 6813–6828, 2020.

[55] W. He, Y. Chen, N. Yokoya, C. Li, and Q. Zhao, "Hyperspectral super-resolution via coupled tensor ring factorization," *Pattern Recognit.*, vol. 122, Feb. 2021, Art. no. 108280.

[56] Y. Xu, Z. Wu, J. Chanussot, and Z. Wei, "Hyperspectral images super-resolution via learning high-order coupled tensor ring representation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 11, pp. 4747–4760, Nov. 2020.

[57] 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.

[58] J. Li, X. Liu, Q. Yuan, H. Shen, and L. Zhang, "Antinoise hyperspectral image fusion by mining tensor low-multilinear-rank and variational properties," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 7832–7848, Oct. 2019.

[59] N. Akhtar, F. Shafait, and A. Mian, "Bayesian sparse representation for hyperspectral image super resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 3631–3640.

[60] 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, Feb. 2020.

[61] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.

[62] Y.-C. Miao, X.-L. Zhao, X. Fu, J.-L. Wang, and Y.-B. Zheng, "Hyperspectral denoising using unsupervised disentangled spatiospectral deep priors," *IEEE Trans. Geosci. Remote Sens.*, early access, Aug. 27, 2021, doi: 10.1109/TGRS.2021.3106380.

[63] W. He *et al.*, "Non-local meets global: An integrated paradigm for hyperspectral image restoration," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Sep. 29, 2020, doi: 10.1109/TPAMI.2020.3027563.

[64] Y. Wang, J. Peng, Q. Zhao, D. Meng, Y. Leung, and X.-L. Zhao, "Hyperspectral image restoration via total variation regularized low-rank tensor decomposition," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1227–1243, Apr. 2018.

[65] 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.

[66] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran, "Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka–Łojasiewicz inequality," *Math. Oper. Res.*, vol. 35, no. 2, pp. 438–457, May 2010.

[67] H. Attouch, J. Bolte, and B. F. Svaiter, "Convergence of descent methods for semi-algebraic and tame problems: Proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods," *Math. Program.*, vol. 137, nos. 1–2, pp. 91–129, Apr. 2013.

[68] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jul. 2011.

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

[70] 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.*, early access, Jan. 5, 2021, doi: 10.1109/TGRS.2020.3045169.

[71] Y. Wang, W. Yin, and J. Zeng, "Global convergence of ADMM in nonconvex nonsmooth optimization," *J. Sci. Comput.*, vol. 78, no. 1, pp. 29–63, 2019.

[72] E. Candes, M. Wakin, and S. Boyd, "Enhancing sparsity by reweighted l_1 minimization," *J. Fourier Anal. Appl.*, vol. 14, no. 5, pp. 877–905, 2008.

[73] 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.
[74] Y. Chen, W. He, N. Yokoya, and T.-Z. Huang, “Hyperspectral image restoration using weighted group sparsity-regularized low-rank tensor decomposition,” *IEEE Trans. Cybern.*, vol. 50, no. 8, pp. 3556–3570, Aug. 2020.
[75] J. M. Bioucas-Dias and J. M. P. Nascimento, “Hyperspectral subspace identification,” *IEEE Trans. Geosci. Remote Sensing*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.
[76] S. Gabarda and G. Cristóbal, “Blind image quality assessment through anisotropy,” *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 24, no. 12, pp. 42–51, 2007.

Yong Chen received the B.S. degree from the School of Science, East China University of Technology, Nanchang, China, in 2015, and the Ph.D. degree from the School of Mathematical Sciences, University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2020.

From 2018 to 2019, he was a Research Intern with the Geoinformatics Unit, RIKEN Center for Advanced Intelligence Project, RIKEN, Tokyo, Japan. He is currently with the School of Computer and Information Engineering, Jiangxi Normal University, Nanchang. His research interests include remote sensing image processing and low-rank matrix/tensor representation.

Jinshan Zeng received the Ph.D. degree in mathematics from Xi’an Jiaotong University, Xi’an, China, in 2015.

He is currently a Distinguished Professor with the School of Computer and Information Engineering, Jiangxi Normal University, Nanchang, China, where he also serves as the Director of the Department of Data Science and Big Data. He has published over 40 articles in high-impact journals and conferences, such as IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE (TPAMI), *Journal of Machine Learning Research* (JMLR), IEEE TRANSACTIONS ON SIGNAL PROCESSING (TSP), International Conference on Machine Learning (ICML), and the Association for the Advancement of Artificial Intelligence (AAAI). His research interests include nonconvex optimization, machine learning (in particular deep learning), and remote sensing.

Dr. Zeng has had two papers coauthored with collaborators that received the International Consortium of Chinese Mathematicians (ICCM) Best Paper Awards in 2018 and 2020.

Wei He (Member, IEEE) received the B.S. degree from the School of Mathematics and Statistics and the Ph.D. degree in surveying, mapping, and remote sensing (LIESMARS) from Wuhan University, Wuhan, China, in 2012 and 2017, respectively.

He is currently a Research Scientist with the Geoinformatics Unit, RIKEN Center for Advanced Intelligence Project, RIKEN, Tokyo, Japan, where he was a Researcher from 2018 to 2020. His research interests include image quality improvement, remote sensing image processing, and low-rank representation and deep learning.

Xi-Le Zhao received the M.S. and Ph.D. degrees from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2009 and 2012, respectively.

He is currently a Professor with the School of Mathematical Sciences, UESTC. His research interests are model-driven and data-driven methods for image processing problems. His homepage is https://zhaoxile.github.io/

Ting-Zhu Huang received the B.S., M.S., and Ph.D. degrees in computational mathematics from the Department of Mathematics, Xi’an Jiaotong University, Xi’an, China, in 1986, 1992, and 2001, respectively.

He is currently a Professor with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, China. His research interests include scientific computation and applications, numerical algorithms for image processing, numerical linear algebra, preconditioning technologies, and matrix analysis with applications.

Prof. Huang is also an Editor of *The Scientific World Journal*, *Advances in Numerical Analysis*, the *Journal of Applied Mathematics*, the *Journal of Pure and Applied Mathematics: Advances in Applied Mathematics*, and the *Journal of Electronic Science and Technology*, China.