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 HR-MSI, 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]
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 image features. However, the training of these methods generally requires plenty of paired LR-HSI and HR-MSI, 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 defined to decompose a high-order tensor as a sequence of 2-D matrices 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 exploit 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 $G^{(1)} \in \mathbb{R}^{2 \times 100 \times 2}$, $G^{(2)} \in \mathbb{R}^{2 \times 100 \times 2}$, and $G^{(3)} \in \mathbb{R}^{2 \times 100 \times 2}$, respectively. (d) Distributions of row, column, and spectral signatures randomly extracted from the tensor $X = \Phi(G^{(1)}, G^{(2)}, G^{(3)})$.

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 stronger 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.
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 \( i \)-th mode. Thus, vectors and matrices are the first- and second-order tensors, respectively. \( \mathcal{X}(i_1, i_2, \ldots, i_n) \) or \( x_{i_1i_2\ldots i_n} \) 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 \( \mathcal{G}^{(k)} \in \mathbb{R}^{I_k \times R_k \times \cdots \times R_{k-1}} \) [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

\[
\mathcal{X}(i_1, i_2, \ldots, i_n) = \text{Tr}(\mathcal{G}^{(1)}(i_1)\mathcal{G}^{(2)}(i_2), \ldots, \mathcal{G}^{(n)}(i_n)) = \text{Tr} \left( \prod_{k=1}^{n} \mathcal{G}^{(k)}(i_k) \right) \tag{1}
\]

where \( \mathcal{G}^{(k)}(i_k) \in \mathbb{R}^{I_k \times R_k \times \cdots \times R_{k-1}} \) denotes the \( i_k \)-th lateral slice matrix of factor tensor \( \mathcal{G}^{(k)} \) and \( \text{Tr}() \) is the matrix trace operation.

The multilinear product of two adjacent factor tensors \( \mathcal{G}^{(k)} \in \mathbb{R}^{I_k \times I_{k+1} \times \cdots \times I_{k+1}} \) and \( \mathcal{G}^{(k+1)} \in \mathbb{R}^{I_{k+1} \times I_{k+2} \times \cdots \times I_n} \) is denoted as \( \mathcal{G}^{(k,k+1)} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_n} \) and each lateral slice of \( \mathcal{G}^{(k,k+1)} \) is defined as

\[
\mathcal{G}^{(k,k+1)}((j_k-1)I_k + i_k) = \mathcal{G}^{(k)}(i_k)\mathcal{G}^{(k+1)}(j_k) \tag{2}
\]

for \( i_k = 1, \ldots, I_k, j_k = 1, \ldots, I_{k+1} \), where \( \mathcal{G}^{(k,k+1)} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_n} \). Based on the multilinear product operation, \( \mathcal{G}^{(1,\ldots,k)} \in \mathbb{R}^{I_1 \times \cdots \times I_k \times I_{k+1} \times \cdots \times I_n} \) is the multilinear product of the first \( k \) factor tensors, \( \mathcal{G}^{(k,1,\ldots,n)} \in \mathbb{R}^{I_1 \times \cdots \times I_k \times \cdots \times I_n} \) is the multilinear product of the last \( n-k \) factor tensors, and \( \mathcal{G}^{(k,k+1)} \) is the multilinear product of all factor tensors except \( k \)th factor tensor.

Let \( \overrightarrow{\mathcal{X}}_k \in \mathbb{R}^{I_k \times I_{k+1} \times \cdots \times I_n} \) be a \( n \)-order tensor; then, \( \overrightarrow{\mathcal{X}}_k \) 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 \( \overrightarrow{\mathcal{X}}_k \) can be expressed as

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

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

\[
\mathbf{X}_{<k>} = \overrightarrow{\mathbf{X}}_{(k,2)} = \mathbf{G}_{(k)}^{(k)}(\mathcal{G}_{<2>}^{(k)})^T \tag{4}
\]
TABLE I  
**Tensor Operations**

<table>
<thead>
<tr>
<th>Notations</th>
<th>Operations</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_{(k)} \in \mathbb{R}^{l_1 \times l_2 \cdots l_n}$</td>
<td>the first mode-$k$ matricization of tensor $X$.</td>
</tr>
<tr>
<td>$X_{&lt;k&gt;} \in \mathbb{R}^{l_1 \times l_2 \cdots l_n \times \cdots \times l_{k-1}}$</td>
<td>the second mode-$k$ matricization of tensor $X$.</td>
</tr>
<tr>
<td>Fold$_k(\cdot)$</td>
<td>the inverse operator of mode-$k$ matricization.</td>
</tr>
<tr>
<td>$&lt;X,Y&gt;$</td>
<td>inner product of two tensors: $\sum_{i_1,i_2,\ldots,i_n} x_{i_1,i_2,\ldots,i_n} y_{i_1,i_2,\ldots,i_n}$.</td>
</tr>
<tr>
<td>$</td>
<td></td>
</tr>
<tr>
<td>$\ell_1$-norm of tensor $X$: $\sum_{i_1,i_2,\ldots,i_n}</td>
<td>x_{i_1,i_2,\ldots,i_n}</td>
</tr>
</tbody>
</table>

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 $Y \in \mathbb{R}^{m \times n \times b}$ and $Z \in \mathbb{R}^{M \times N \times b}$, where $m \times n$ and $M \times N$ are spatial sizes of LR-HSI and HR-MSI, respectively, and $b$ is 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 $X \in \mathbb{R}^{M \times N \times b}$ 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-MSI 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

$$Y = X \times_1 U_1 \times_2 U_2 + \mathcal{N}_Y$$  \hspace{1cm} (6)

where $U_1 \in \mathbb{R}^{m \times M}$ and $U_2 \in \mathbb{R}^{n \times N}$ 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:

$$Z = X \times_3 U_3 + \mathcal{N}_Z$$  \hspace{1cm} (7)

where $U_3 \in \mathbb{R}^{b \times B}$ 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 $X$ can be solved the following minimization problem:

$$\min_X \frac{1}{2} ||Y - X \times_1 U_1 \times_2 U_2||_F^2 + \frac{\lambda}{2} ||Z - X \times_3 U_3||_F^2$$  \hspace{1cm} (8)

where $\lambda$ is the parameter used to balance two terms.

C. Factorization-Based Methods

Directly estimating the HR-HSI $X$ from problem (8) is an ill-posed inverse problem; the reason is that the number of total measurements from $Y$ and $Z$ 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 $X$.

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 regularized on $X^{(3)}$ (the mode-3 matricization of tensor $X$) 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

$$\min_{E,A} \frac{1}{2} \|Y^{(3)} - EA(U_1 \otimes U_2)\|_F^2 + \frac{\lambda}{2} \|Z^{(3)} - U_3 EA\|_F^2$$  \hspace{1cm} (9)

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

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:

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

where $G^{(1)} \in \mathbb{R}^{r_1 \times M \times r_2}$, $G^{(2)} \in \mathbb{R}^{r_2 \times N \times r_1}$, and $G^{(3)} \in \mathbb{R}^{r_1 \times B \times r_1}$ 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

\[
\mathcal{Y} = \Phi(G^{(1)} \times_2 U_1, G^{(2)} \times_2 U_2, G^{(3)}) + \mathcal{N},
\]

(11)

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

\[
Z = \Phi(G^{(1)}, G^{(2)}, G^{(3)} \times_2 U_3) + \mathcal{N}.
\]

(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

\[
\min_{G^{(1)}, G^{(2)}, G^{(3)}} \frac{1}{2} \| \mathcal{Y} - \Phi(G^{(1)} \times_2 U_1, G^{(2)} \times_2 U_2, G^{(3)}) \|^2_F + \frac{\lambda}{2} \| Z - \Phi(G^{(1)}, G^{(2)}, G^{(3)} \times_2 U_3) \|^2_F.
\]

(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

\[
\min_{G^{(1)}, G^{(2)}, G^{(3)}} \frac{1}{2} \| \mathcal{Y} - \Phi(G^{(1)} \times_2 U_1, G^{(2)} \times_2 U_2, G^{(3)}) \|^2_F + \frac{\lambda}{2} \| Z - \Phi(G^{(1)}, G^{(2)}, G^{(3)} \times_2 U_3) \|^2_F + \tau \sum_{k=1}^{3} F_k(G^{(k)})
\]

(14)

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

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

\[
X_{<k>} = G_{(2)}(G_{<k>}^{(k)})^T, \quad (k = 1, 2, 3).
\]

(15)

From the perspective of matrix factorization, it is easy to understand that each column vector of \( X_{<k>} \) can be regarded as the representation of a linear combination of all columns of the factor \( G_{(2)}^{(k)} \), indicating that all columns of \( G_{(2)}^{(k)} \) are a set of basis of the low-dimensional space of \( X_{<k>} \). Since the HR-HSI has the piecewise smooth structure in two spatial dimensions and one spectral dimension, each column of \( X_{<k>} \) 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 \( G_{(2)}^{(k)} (k = 1, 2, 3) \) 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 \( G_{(2)} \). In summary, the proposed FSTRD model for LR-HSI and HR-MSI fusion can be formulated as

\[
\min_{G^{(1)}, G^{(2)}, G^{(3)}} \frac{1}{2} \| \mathcal{Y} - \Phi(G^{(1)} \times_2 U_1, G^{(2)} \times_2 U_2, G^{(3)}) \|^2_F + \frac{\lambda}{2} \| Z - \Phi(G^{(1)}, G^{(2)}, G^{(3)} \times_2 U_3) \|^2_F + \tau \sum_{k=1}^{3} \| \mathcal{W}^{(k)} \odot (G^{(k)} \times_2 D) \|_1
\]

(16)

where \( D \) is the first-order difference square matrix, whose dimension is associated with the second dimension of \( G^{(k)} \), and \( \mathcal{W}^{(k)} \in \mathbb{R}^{n \times k \times n+1} \) 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 $G^{(k)}$ ($k = 1, 2, 3$), 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 $f(G^{(1)}, G^{(2)}, G^{(3)})$ 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:

$$
\begin{align*}
\lambda^{(1),i+1} &= \arg \min_{G^{(1)}} \frac{1}{2} \left\| \nabla \phi(G^{(1)}) \times \times_2 U_1, G^{(2),i} \times \times_2 U_2, G^{(3),i} \right\|^2_F \\
\lambda^{(2),i+1} &= \arg \min_{G^{(2)}} \frac{1}{2} \left\| \nabla \phi(G^{(2)}) \times \times_2 U_1, G^{(2),i} \times \times_2 U_2, G^{(3),i} \right\|^2_F \\
\lambda^{(3),i+1} &= \arg \min_{G^{(3)}} \frac{1}{2} \left\| \nabla \phi(G^{(3)}) \times \times_2 U_1, G^{(2),i} \times \times_2 U_2, G^{(3),i} \right\|^2_F 
\end{align*}
$$

where $i$, $(\rho/2) \cdot \cdot \cdot_\|_F$, and $\rho$ 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 $G^{(1)}$:** The subproblem of $G^{(1)}$ can be formulated as

$$
\begin{align*}
\min_{G^{(1)}} & \frac{1}{2} \left\| \nabla \phi(G^{(1)}) \times \times_2 U_1, G^{(2),i} \times \times_2 U_2, G^{(3),i} \right\|^2_F \\
& + \frac{\lambda}{2} \left\| \nabla \phi(G^{(1)}) \times \times_2 U_1, G^{(2),i} \times \times_2 U_2, G^{(3),i} \right\|^2_F \\
& + \tau \|W^{(1)} \odot (G^{(1)} \times \times_2 D)\|_1 + \frac{\rho}{2} \left\| \nabla \phi(G^{(1)}) \times \times_2 D \right\|^2_F 
\end{align*}
$$

Because of the nonsmooth term of $\ell_1$-norm, the closeform solution of $G^{(1)}$ 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 $R_1$; the unconstrained problem is equivalent to the following minimization problem:

$$
\begin{align*}
\min_{G^{(1)}, R_1} & \frac{1}{2} \left\| \nabla \phi(G^{(1)}) \times \times_2 U_1, G^{(2),i} \times \times_2 U_2, G^{(3),i} \right\|^2_F \\
& + \frac{\lambda}{2} \left\| \nabla \phi(G^{(1)}) \times \times_2 U_1, G^{(2),i} \times \times_2 U_2, G^{(3),i} \right\|^2_F \\
& + \tau \|W^{(1)} \odot R_1\|_1 + \frac{\rho}{2} \left\| \nabla \phi(G^{(1)}) \times \times_2 D \right\|^2_F , \text{ s.t. } R_1 = G^{(1)} \times \times_2 D 
\end{align*}
$$

The augmented Lagrangian function of the problem is formulated as

$$
L_\beta(G^{(1)}, R_1, M_1) = \frac{1}{2} \left\| \nabla \phi(G^{(1)}) \times \times_2 U_1, G^{(2),i} \times \times_2 U_2, G^{(3),i} \right\|^2_F \\
+ \frac{\lambda}{2} \left\| \nabla \phi(G^{(1)}, G^{(2),i}, G^{(3),i} \times \times_2 U_3) \right\|^2_F + \tau \|W^{(1)} \odot R_1\|_1 + \frac{\rho}{2} \left\| \nabla \phi(G^{(1)} - G^{(1),i}) \right\|^2_F \\
+ \frac{\beta}{2} \left\| R_1 - G^{(1)} \times \times_2 D + \frac{M_1}{\beta} \right\|^2_F 
$$

where $M_1$ is the Lagrangian multiplier and $\beta$ is a positive penalty parameter. Then, the solution of $G^{(1)}$ can be achieved by iteratively optimized the following subproblems:

1) **$R_1$-subproblem:** The minimization of $R_1$ is

$$
\min_{R_1} \tau \|W^{(1)} \odot R_1\|_1 + \frac{\beta}{2} \left\| R_1 - G^{(1)} \times \times_2 D + \frac{M_1}{\beta} \right\|^2_F 
$$

which is a weighted $\ell_1$-norm minimization, and the closedform solution is obtained by soft-threshold shrinkage operator

$$
R_1 = \operatorname{sign}(J_1) \max \left( |J_1| - \frac{\tau}{\beta} |\lambda^{(1)}|, 0 \right) 
$$

where $\epsilon$ is a positive small constant avoiding singularity.

2) **$G^{(1)}$-subproblem:** The minimization of $G^{(1)}$ is

$$
\min_{G^{(1)}} \frac{1}{2} \left\| \nabla \phi(G^{(1)}) \times \times_2 U_1, G^{(2),i} \times \times_2 U_2, G^{(3),i} \right\|^2_F \\
+ \frac{\lambda}{2} \left\| \nabla \phi(G^{(1)}, G^{(2),i}, G^{(3),i} \times \times_2 U_3) \right\|^2_F \\
+ \frac{\rho}{2} \left\| \nabla \phi(G^{(1)} - G^{(1),i}) \right\|^2_F + \frac{\beta}{2} \left\| R_1 - G^{(1)} \times \times_2 D + \frac{M_1}{\beta} \right\|^2_F 
$$

Based on the matrix representation of TR decomposition (4), let $P_1 = (G^{(1),i} \times \times_2 U_2)G^{(2),i}T_{<2,3}$ and $P_2 = (G^{(1),i} \times \times_2 U_3)T_{<2,3}$, and the problem of $G^{(1)}$ can be converted to the following minimization:

$$
\min_{G^{(1)}, R_1} \frac{1}{2} \left\| \nabla \phi(G^{(1)}) \times \times_2 U_1, G^{(2),i} \times \times_2 U_2, G^{(3),i} \right\|^2_F \\
+ \frac{\lambda}{2} \left\| \nabla \phi(G^{(1)}, G^{(2),i}, G^{(3),i} \times \times_2 U_3) \right\|^2_F \\
+ \frac{\rho}{2} \left\| \nabla \phi(G^{(1)} - G^{(1),i}) \right\|^2_F + \frac{\beta}{2} \left\| R_1 - G^{(1)} \times \times_2 D + \frac{M_1}{\beta} \right\|^2_F 
$$

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

$$
\begin{align*}
U_1T_1U_1G^{(1),i}P_1P_1^T + \lambda Z_{<1>}P_2P_2^T + \rho G^{(1)} + \beta D^TED^{(1)} \\
= U_1T_1Y_{<>1}P_1^T + \lambda Z_{<1>}P_2^T + \rho G^{(1)} + \beta D^TED^{(1)} \\
+ \rho G^{(1),i} + \beta D^T \left( R_{<1>} + \frac{M_1}{\beta} \right). 
\end{align*}
$$

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

---

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.
The optimization of Lagrangian multiplier $\mathcal{M}_1$ is formulated as follows:

$$\mathcal{M}_1 \leftarrow \mathcal{M}_1 + \beta(R_1 - G_{1}(1) \times 2 \mathbf{D}).$$

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

2) **Optimization With Respect to $G_{2}(2)$**: The subproblem of $G_{2}(2)$ can be formulated as

$$\min_{G_{2}(2)} \frac{1}{2} \|Y - \Phi(G_{2}(i), i+1) \times 2 \mathbf{U}_1, G_{2}(2) \times 2 \mathbf{U}_2, G_{3}(3)\|_F^2 + \frac{\lambda}{2} \|Z - \Phi(G_{2}(i+1), G_{2}(3), i+1 \times 2 \mathbf{U}_3)\|_F^2 + \tau \|W_{2} \odot (G_{2}(2) \times 2 \mathbf{D})\|_1 + \frac{\rho}{2} \|G_{2}(2) - G_{2}(i), i+1\|_F^2.$$

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

$$\mathcal{R}_2 = \text{sign}(\mathcal{J}_2) \max \left( \left| \mathcal{J}_2 \right| - \frac{\tau}{\lambda} \|W_{2}\|, 0 \right)$$

$$\mathbf{U}_1^T \mathbf{U}_2 G_{2}(2) \mathbf{Q}_1 T + \lambda G_{2}(2) \mathbf{Q}_2 T + \rho G_{2}(2) + \beta \mathbf{D}^T \mathbf{D} G_{2}(2)$$

$$= \mathbf{U}_1^T \mathbf{Y}_{<2}, \mathbf{Q}_1 T + \lambda Z_{<2}, \mathbf{Q}_2 T + \rho G_{2}(2) + \beta \mathbf{D}^T \left( R_{2} + \frac{M_{2}(2)}{\beta} \right)$$

$$\mathbf{M}_2 \leftarrow \mathbf{M}_2 + \beta(\mathcal{R}_2 - G_{2}(2) \times 2 \mathbf{D})$$

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

$$\left\{ \begin{array}{l}
\mathcal{J}_2 = G_{2}(2) \times 2 \mathbf{D} - \frac{\mathcal{M}_2}{\beta} \\
W_{2} = 1 / ((\mathcal{J}_2) + \epsilon) \\
\mathbf{Q}_1 = (G_{3}(3), i+1 \times 2 \mathbf{U}_1)^{T}_{<2}, \mathbf{Q}_2 = ((G_{3}(3), i+1 \times 2 \mathbf{U}_3)^{T}_{<2}.
\end{array} \right.$$ (32)

3) **Optimization With Respect to $G_{3}(3)$**: The subproblem of $G_{3}(3)$ can be formulated as

$$\min_{G_{3}(3)} \frac{1}{2} \|Y - \Phi(G_{1}(i), i+1 \times 2 \mathbf{U}_1, G_{2}(i), i+1 \times 2 \mathbf{U}_2, G_{3}(3))\|_F^2 + \frac{\lambda}{2} \|Z - \Phi(G_{2}(i), G_{2}(3), i+1 \times 2 \mathbf{U}_3)\|_F^2 + \tau \|W_{3} \odot (G_{3}(3) \times 2 \mathbf{D})\|_1 + \frac{\rho}{2} \|G_{3}(3) - G_{3}(i), i+1\|_F^2.$$

The we can also use ADMM to optimize factor $G_{3}(3)$; the solution of each subproblem can be iteratively updated as follows:

$$\mathcal{R}_3 = \text{sign}(\mathcal{J}_3) \max \left( \left| \mathcal{J}_3 \right| - \frac{\tau}{\lambda} \|W_{3}\|, 0 \right)$$

$$\lambda \mathbf{U}_1^T \mathbf{U}_2 G_{3}(3) T_1 T + \rho G_{3}(3) + \beta \mathbf{D}^T \mathbf{D} G_{3}(3)$$

$$= \mathbf{Y}_{<3}, T_1 + \lambda \mathbf{U}_1^T \mathbf{Z}_{<3}, T_3 + \rho G_{3}(3) + \beta \mathbf{D}^T \left( R_{3} + \frac{M_{3}(2)}{\beta} \right)$$

$$\mathcal{M}_3 \leftarrow \mathcal{M}_3 + \beta(\mathcal{R}_3 - G_{3}(3) \times 2 \mathbf{D})$$

where the variables are computed by

$$\left\{ \begin{array}{l}
\mathcal{J}_3 = G_{3}(3) \times 2 \mathbf{D} - \frac{\mathcal{M}_3}{\beta} \\
W_{3} = 1 / ((\mathcal{J}_3) + \epsilon) \\
T_1 = ((G_{1}(i), i+1 \times 2 \mathbf{U}_1, G_{2}(i), i+1 \times 2 \mathbf{U}_2)^{T}_{<3}, \\
T_2 = (G_{1}(i), i+1 \times 2 \mathbf{D} G_{3}(i), i+1 \times 2 \mathbf{D})^{T}_{<3}.
\end{array} \right.$$ (37)

Summarizing the optimization procedure of each factor $G_{1}(1)$, $G_{2}(2)$, and $G_{3}(3)$, 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 $\{G_{1}(i), i+1, G_{2}(i), i+1, G_{3}(i), i+1\}$ generated by Algorithm 1 converges to a critical point if it is bounded [67].

**Algorithm 1** PAM Algorithm for Solving Model (16)

**Input:** LR-HSI $Y$, HR-MSI $Z$, degraded operators $U_1, U_2, U_3$, parameters $\lambda, \tau, \rho, \beta$, TR rank $r = [r_1, r_2, r_3]$, and $i_{\text{max}}$.

1. **Initialize:** $G_{1}(0), G_{2}(0), G_{3}(0)$, and $i = 0$.

2. **while** not converged **do**

3. **Initialize:** $G_{1}(i) = G_{1}(i), i$ and $\mathcal{M}_1 = 0$.

4. **while** not converged **do**

5. Update $\mathcal{R}_1$ via (22).

6. Update $G_{1}(i)$ via (26).

7. Update $\mathcal{M}_1$ via (27).

8. **end while**

9. Output: $G_{1}(i), i+1$.

10. **Initialize:** $G_{2}(i) = G_{2}(i), i$ and $\mathcal{M}_2 = 0$.

11. **while** not converged **do**

12. Update $\mathcal{R}_2$ via (29).

13. Update $G_{2}(i)$ via (30).

14. Update $\mathcal{M}_2$ via (31).

15. **end while**

16. Output: $G_{2}(i), i+1$.

17. **Initialize:** $G_{3}(i) = G_{3}(i), i$ and $\mathcal{M}_3 = 0$.

18. **while** not converged **do**

19. Update $\mathcal{R}_3$ via (34).

20. Update $G_{3}(i)$ via (35).

21. Update $\mathcal{M}_3$ via (36).

22. **end while**

23. Output: $G_{3}(i), i+1$.

24. Check $\|G_{1}(i), i+1 - G_{0}\|_F \leq \epsilon$ and $i < i_{\text{max}}$.

25. $i = i + 1$.

26. **end while**

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

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 $G_{1}(1)$, $G_{2}(2)$, and $G_{3}(3)$ are optimized by ADMM. For the $G_{1}(1)$ 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 $O(2M R^2)$. The main computational cost of solving the $G_{1}(1)$ 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 $O(M^2R^2 + M^2R^2)$. Analogously, the computational costs of the $G^{(2)}$ and $G^{(3)}$ iterations are $O(K_{CG}(N^2R^2 + N^2R^4) + 2MR^2)$ and $O(K_{CG}(B^2R^2 + B^2R^4) + 2MR^2)$, respectively. Therefore, the total computation complexity of each iteration in Algorithm 1 is $O(K_{ADMM}K_{CG}(M^2 + N^2 + B^2)R^2 + (M + N + B)R^2 + 2K_{ADMM}(M + N + B)R^2)$, 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], non-local sparse tensor factorization (NLSTF) [38], coupled CP factorization (STEREO) [43], coupled sparse tensor factorization (CSTF) [57], low tensor train rank (LTT)R)-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 spatiotemporal 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, 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 $512 \times 512$ 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, which has 50 HSIs of indoor and outdoor scenes under daylight illumination.

The spatial and spectral sizes of each HSI in Harvard datasets are $1040 \times 1392$ and 31, respectively, where

---

1http://www1.cs.columbia.edu/CAVE/databases/multispectral
2http://vision.seas.harvard.edu/hyperspec/download.html
31 spectral bands are ranged from 420 to 720 nm with an increment 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, 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, 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 128×128×184 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×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 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, RLRMD, 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

![Image](https://example.com/image.png)

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) RLRMD. (i) CNN-Fus. (j) HSRnet. (k) CTRF. (l) FSTRD. (m) Original.

---

2. [https://engineering.purdue.edu/biehl/MultiSpec/hyperspectral.html](https://engineering.purdue.edu/biehl/MultiSpec/hyperspectral.html)
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 × 120 × 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 × 360 × 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 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.
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>
<thead>
<tr>
<th>Method</th>
<th>PSNR (SNR = 10 dB)</th>
<th>PSNR (SNR = 30 dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HySure</td>
<td>21.46</td>
<td>39.23</td>
</tr>
<tr>
<td>FUSE</td>
<td>20.14</td>
<td>39.25</td>
</tr>
<tr>
<td>NLSTF</td>
<td>20.00</td>
<td>39.42</td>
</tr>
<tr>
<td>STEREO</td>
<td>26.56</td>
<td>39.42</td>
</tr>
<tr>
<td>CSTF</td>
<td>27.30</td>
<td>39.42</td>
</tr>
<tr>
<td>LTTR</td>
<td>21.16</td>
<td>39.42</td>
</tr>
<tr>
<td>RLRMDF</td>
<td>20.30</td>
<td>39.58</td>
</tr>
<tr>
<td>CNN-Pus</td>
<td>24.09</td>
<td>38.95</td>
</tr>
<tr>
<td>CTRF</td>
<td>27.34</td>
<td>39.41</td>
</tr>
<tr>
<td>FSTRD</td>
<td>28.33</td>
<td>40.14</td>
</tr>
</tbody>
</table>

TABLE III
BIQA COMPARISON OF THE REAL DATA

<table>
<thead>
<tr>
<th>Method</th>
<th>PSNR</th>
<th>FUSE</th>
<th>NLSTF</th>
<th>STEREO</th>
<th>CSTF</th>
<th>LTTR</th>
<th>RLRMDF</th>
<th>CNN-Pus</th>
<th>CTRF</th>
<th>FSTRD</th>
</tr>
</thead>
<tbody>
<tr>
<td>HySure</td>
<td>0.019</td>
<td>0.015</td>
<td>0.007</td>
<td>0.010</td>
<td>0.008</td>
<td>0.011</td>
<td>0.013</td>
<td>0.027</td>
<td>0.011</td>
<td>0.012</td>
</tr>
</tbody>
</table>

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

D. Discussion

There are five regularization parameters $\lambda$, $\tau$, $r_1$, $r_2$, and $r_3$ and two algorithm parameters $\rho$ and $\beta$ 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 $\lambda$ and $\tau$: Fig. 11 plots the PSNR value of the reconstructed result as a function of parameters $\lambda$ and $\tau$ 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 $\tau$ and smaller $\lambda$, illustrating the effectiveness of factor smoothed regularization in high noise cases. On the contrary, smaller $\tau$ and larger $\lambda$ should be selected in low noise intensity cases. Hence, according to the noise intensity, we empirically select $\lambda$ and $\tau$ in the set of $\{0.01, 0.5\}$ and $\{0.0001, 0.01, 1\}$, respectively.

<table>
<thead>
<tr>
<th>Method</th>
<th>PSNR</th>
<th>SSIM</th>
<th>ERGAS</th>
<th>SAM</th>
<th>CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>HySure</td>
<td>29.07</td>
<td>0.716</td>
<td>2.663</td>
<td>10.972</td>
<td>0.933</td>
</tr>
<tr>
<td>FUSE</td>
<td>30.77</td>
<td>0.790</td>
<td>2.220</td>
<td>8.593</td>
<td>0.955</td>
</tr>
<tr>
<td>NLSTF</td>
<td>32.65</td>
<td>0.867</td>
<td>1.747</td>
<td>6.190</td>
<td>0.972</td>
</tr>
<tr>
<td>STEREO</td>
<td>31.08</td>
<td>0.833</td>
<td>2.046</td>
<td>6.913</td>
<td>0.957</td>
</tr>
<tr>
<td>CSTF</td>
<td>33.90</td>
<td>0.885</td>
<td>1.499</td>
<td>5.081</td>
<td>0.978</td>
</tr>
<tr>
<td>LTTR</td>
<td>29.64</td>
<td>0.787</td>
<td>2.452</td>
<td>9.477</td>
<td>0.944</td>
</tr>
<tr>
<td>RLRMDF</td>
<td>32.50</td>
<td>0.843</td>
<td>1.794</td>
<td>6.540</td>
<td>0.970</td>
</tr>
<tr>
<td>CNN-Pus</td>
<td>33.14</td>
<td>0.860</td>
<td>1.662</td>
<td>5.754</td>
<td>0.974</td>
</tr>
<tr>
<td>CTRF</td>
<td>33.38</td>
<td>0.874</td>
<td>1.581</td>
<td>5.705</td>
<td>0.974</td>
</tr>
<tr>
<td>FSTRD</td>
<td>34.88</td>
<td>0.913</td>
<td>1.345</td>
<td>4.370</td>
<td>0.982</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Method</th>
<th>PSNR</th>
<th>SSIM</th>
<th>ERGAS</th>
<th>SAM</th>
<th>CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>HySure</td>
<td>24.32</td>
<td>0.624</td>
<td>2.328</td>
<td>15.398</td>
<td>0.941</td>
</tr>
<tr>
<td>FUSE</td>
<td>30.51</td>
<td>0.782</td>
<td>1.160</td>
<td>8.892</td>
<td>0.952</td>
</tr>
<tr>
<td>NLSTF</td>
<td>32.05</td>
<td>0.859</td>
<td>0.939</td>
<td>6.566</td>
<td>0.968</td>
</tr>
<tr>
<td>STEREO</td>
<td>27.72</td>
<td>0.713</td>
<td>1.492</td>
<td>9.100</td>
<td>0.904</td>
</tr>
<tr>
<td>CSTF</td>
<td>33.82</td>
<td>0.885</td>
<td>0.757</td>
<td>5.140</td>
<td>0.978</td>
</tr>
<tr>
<td>LTTR</td>
<td>29.92</td>
<td>0.805</td>
<td>1.197</td>
<td>8.938</td>
<td>0.948</td>
</tr>
<tr>
<td>RLRMDF</td>
<td>32.20</td>
<td>0.832</td>
<td>0.929</td>
<td>6.862</td>
<td>0.968</td>
</tr>
<tr>
<td>CNN-Pus</td>
<td>33.06</td>
<td>0.859</td>
<td>0.835</td>
<td>5.810</td>
<td>0.974</td>
</tr>
<tr>
<td>CTRF</td>
<td>33.30</td>
<td>0.873</td>
<td>0.800</td>
<td>5.755</td>
<td>0.974</td>
</tr>
<tr>
<td>FSTRD</td>
<td>34.72</td>
<td>0.913</td>
<td>0.684</td>
<td>4.471</td>
<td>0.982</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Method</th>
<th>PSNR</th>
<th>SSIM</th>
<th>ERGAS</th>
<th>SAM</th>
<th>CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSTRD w/o spa-spe</td>
<td>27.17</td>
<td>0.630</td>
<td>3.194</td>
<td>9.734</td>
<td>0.899</td>
</tr>
<tr>
<td>FSTRD w/o spa-spe</td>
<td>27.40</td>
<td>0.648</td>
<td>3.069</td>
<td>9.241</td>
<td>0.902</td>
</tr>
<tr>
<td>FSTRD w/o spe-spe</td>
<td>28.23</td>
<td>0.708</td>
<td>2.853</td>
<td>8.179</td>
<td>0.919</td>
</tr>
<tr>
<td>FSTRD</td>
<td>28.61</td>
<td>0.732</td>
<td>2.729</td>
<td>7.436</td>
<td>0.924</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>PSNR</th>
<th>SSIM</th>
<th>ERGAS</th>
<th>SAM</th>
<th>CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSTRD w/o spa-spe</td>
<td>39.11</td>
<td>0.963</td>
<td>0.851</td>
<td>3.319</td>
<td>0.992</td>
</tr>
<tr>
<td>FSTRD w/o spa-spe</td>
<td>39.89</td>
<td>0.971</td>
<td>0.772</td>
<td>2.862</td>
<td>0.993</td>
</tr>
<tr>
<td>FSTRD w/o spe-spe</td>
<td>39.59</td>
<td>0.972</td>
<td>0.799</td>
<td>3.012</td>
<td>0.993</td>
</tr>
<tr>
<td>FSTRD</td>
<td>40.01</td>
<td>0.973</td>
<td>0.765</td>
<td>2.798</td>
<td>0.994</td>
</tr>
</tbody>
</table>
Authorized licensed use limited to: University of Electronic Science and Tech of China. Downloaded on October 09,2021 at 06:05:20 UTC from IEEE Xplore. Restrictions apply.
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 HSNet method [23]. They would also like to thank Ms. Jian-Li Wang’s help.

REFERENCES


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


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.

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.

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.

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/

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.

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.

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/

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.

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.

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.