

Learning a Low Tensor-Train Rank Representation for Hyperspectral Image Super-Resolution

Renwei Dian¹, Student Member, IEEE, Shutao Li², Fellow, IEEE, and Leyuan Fang³, Senior Member, IEEE

Abstract—Hyperspectral images (HSIs) with high spectral resolution only have the low spatial resolution. On the contrary, multispectral images (MSIs) with much lower spectral resolution can be obtained with higher spatial resolution. Therefore, fusing the high-spatial-resolution MSI (HR-MSI) with low-spatial-resolution HSI of the same scene has become the very popular HSI super-resolution scheme. In this paper, a novel low tensor-train (TT) rank (LTTR)-based HSI super-resolution method is proposed, where an LTTR prior is designed to learn the correlations among the spatial, spectral, and nonlocal modes of the nonlocal similar high-spatial-resolution HSI (HR-HSI) cubes. First, we cluster the HR-MSI cubes as many groups based on their similarities, and the HR-HSI cubes are also clustered according to the learned cluster structure in the HR-MSI cubes. The HR-HSI cubes in each group are much similar to each other and can constitute a 4-D tensor, whose four modes are highly correlated. Therefore, we impose the LTTR constraint on these 4-D tensors, which can effectively learn the correlations among the spatial, spectral, and nonlocal modes because of the well-balanced matricization scheme of TT rank. We formulate the super-resolution problem as TT rank regularized optimization problem, which is solved via the scheme of alternating direction method of multipliers. Experiments on HSI data sets indicate the effectiveness of the LTTR-based method.

Index Terms—Hyperspectral imaging, image fusion, low tensor-train (TT) rank (LTTR) learning, superresolution.

I. INTRODUCTION

HYPERSPECTRAL imaging technique can acquire images across a number of different wavelengths. Hyperspectral images (HSIs) have the high spectral resolution, which can provide advantages for a precise identification of the materials. Therefore, HSIs have gained more and more usage in a lot of earth remote sensing tasks [1], [2], including object classification [3]–[7] and change detection [8]. However, owing to the finite sun irradiance, there is the inevitable tradeoff between the spatial resolution and spectral

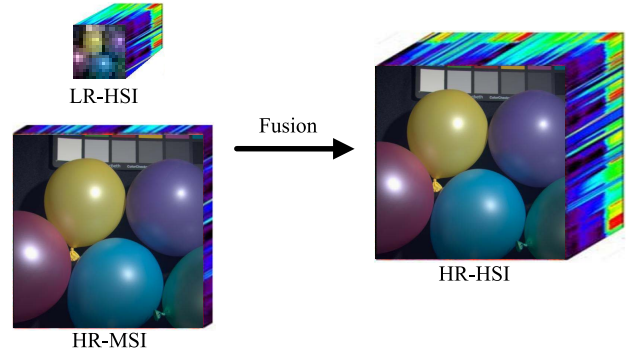


Fig. 1. HSI super-resolution. Reconstruct the HR-HSI from the LR-HSI and HR-MSI of the same scenario.

resolution for existing cameras [9]. Hence, to keep the high spectral resolution, HSIs only have the low spatial resolution, which limits their applications. In contrast, compared with HSIs, multispectral images (MSIs) usually have higher spatial resolution and much lower spectral resolution. As illustrated in Fig. 1, fusing high-spatial-resolution MSI (HR-MSI) with low-spatial-resolution HSI (LR-HSI) of the same scene has become a very popular HSI super-resolution scheme.

HSIs are 3-D data, which have two spatial modes and one spectral mode. Tensors provide a natural way to represent the high-dimensional data, and hence the HSIs can be processed from a point of the tensor. In the past years, tensor representation has been successfully used for multi-dimensional data classification [2], [10], [11], completion [12]–[16], and denoising [17].

As one of the most popular tensor representation methods, tensor-train (TT) method [18] defines a novel tensor rank, termed as TT rank. The TT rank consists of ranks of matrices formed by a well-balanced matricization scheme, i.e., matricizes the tensor along permutations of modes. The low TT rank (LTTR) prior has been demonstrated to be very effective in the color image and videos completion [19] and outperforms the Tucker rank based methods.

This paper presents an LTTR-based HSI super-resolution method by fusing the HR-MSI and LR-HSI, which belongs to the scope of low-rank learning. Since low-rank prior is effective to learn the data redundancies and self-similarities, low-rank learning has been popularly used in matrix completion [20], image fusion [21], clustering [22], and pattern recognition [23]. In this paper, the LTTR estimates the

Manuscript received April 7, 2018; revised July 16, 2018 and October 30, 2018; accepted December 2, 2018. Date of publication January 7, 2019; date of current version August 21, 2019. This work was supported in part by the National Natural Science Fund of China for International Cooperation and Exchanges under Grant 61520106001, in part by the Fund of Hunan Province for Science and Technology Plan Project under Grant 2017RS3024, in part by the Hunan Provincial Innovation Foundation for Postgraduate, and in part by the China Scholarship Council. (Corresponding author: Shutao Li.)

The authors are with the College of Electrical and Information Engineering, Hunan University, Changsha 410082, China, and also with the Key Laboratory of Visual Perception and Artificial Intelligence of Hunan Province, Hunan University, Changsha 410082, China (e-mail: drw@hnu.edu.cn; shutao_li@hnu.edu.cn; fangleyuan@gmail.com).

Color versions of one or more of the figures in this article are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TNNLS.2018.2885616

high-spatial-resolution HSI (HR-HSI) from the LR-HSI and HR-MSI with LTTR priors, which can effectively learn the correlations among the spatial, spectral, and nonlocal modes of the similar HR-HSI cubes. A typical natural scene is self-similar, and therefore, it often contains many nonlocal similar patches [24]. To exploit this prior, we partition the HR-MSI as many cubes of the same size, and then, we group similar cubes together. Based on the learned cluster in the HR-MSI, the HR-HSI cubes are also grouped. Similar HR-HSI cubes in each group can constitute a 4-D tensor, which has two spatial modes: one spectral mode and one non-local mode. (We call the last mode as nonlocal mode.) Since the 4-D tensor is constituted by similar HR-HSI cubes, the four modes of the tensor have strong correlations. Therefore, the LTTR prior is imposed on these 4-D tensors, which can effectively learn the correlations of the four modes due to the well-balanced matricization scheme of TT rank. We formulate the super-resolution problem as LTTR regularized optimization problem and solve it via the alternating direction method of multipliers (ADMMs) [25] algorithm.

We organize the remaining part of this paper as follows. Section II presents the review of the recent HSI super-resolution approaches. In Section III, the preliminaries on tensors are given. We introduce the proposed LTTR method for HSI super-resolution in Section IV. Section V gives the experimental results and corresponding discussions. In Section VI, we give conclusions and potential research directions.

II. RELATED WORK

Recently, fusing the HR-MSI with LR-HSI to acquire the HR-HSI has become a very popular HSI super-resolution scheme. We can classify recent HSI super-resolution approaches as three groups: matrix factorization-based approaches [26]–[34], tensor factorization-based approaches [35]–[37], and deep convolution neural network (CNN)-based approaches [8], [38].

Based on the assumption that each spectral can be represented by linearly represented by several spectral atoms, matrix factorization-based approaches approximate the desired HR-HSI by the spectral basis and corresponding coefficients. Specifically, methods [27]–[29] formulate the fusion problem as the coupled matrix factorization problem, and they alternately update spectral basis and coefficients from the HR-MSI and LR-HSI with some specific priors. Besides, works in [30]–[33] also try to make use of the spatial structures of the HR-HSI to solve the super-resolution problem. For example, Akhtar *et al.* [30] first learn a nonnegative dictionary, and then simultaneously estimate coefficients for all pixels in each local patch by using the simultaneous greedy pursuit algorithm, which exploits the prior that nearby pixels are may well represent the same materials. Simoes *et al.* [31] use the total variation to promote the spatial smoothness. Wei *et al.* [32] assume that the HSIs lie in a low-dimensional subspace and use the spatial dictionaries learned from the HR-MSI to promote the spatial similarities. Dong *et al.* [33] first learn the nonnegative spectral basis via nonnegative dictionary approach, and then estimate the coefficients via the structured

sparse coding approach, where the non-local similarities of the HR-HSI are exploited.

Since HSI can be naturally represented by a 3-D tensor, another group of approaches solves the super-resolution problem from the point of the tensor. Our previous work [35] first proposes a nonlocal sparse Tucker factorization for HSI super-resolution. Based on Tucker decomposition, each cube of HR-HSI is approximated by core tensor multiplication by dictionaries of three modes. We learn the dictionaries of width and height modes from the HR-MSI and learn the spectral dictionary from the LR-HSI. To better model the non-local similarities, the similar cubes in the HR-HSI are assumed to admit sparse representation on the same dictionaries. In general, there are two fundamental differences between reference [35] and the proposed method. On the one hand, they are based on different tensor representations. More specifically, reference [35] is based on Tucker decomposition [39], which decomposes each HR-HSI cube as a core tensor multiplication by dictionaries of three modes. However, the proposed method is based on TT decomposition [18], and the LTTR prior is imposed on the grouped 4-D tensors, that is, the canonical unfolding matrices of grouped 4-D tensors are assumed to have low-rank property. On the other hand, the ways to exploit nonlocal similarities are different. Dian *et al.* [35] assume that the similar cubes admit sparse representation on the same dictionaries to use nonlocal self-similarities. This paper proposes to group the similar cubes as a 4-D tensor and imposes the LTTR prior on the grouped 4-D tensor to explicitly exploit the nonlocal similarities of the HR-HSI. Li *et al.* [36] formulate the super-resolution problem as the coupled sparse Tucker decomposition, which alternatively updates the dictionaries of three modes and sparse core tensor. The proposed LTTR-based method is also intrinsically different from this method. In addition, Zhang *et al.* [37] proposed a graph-regularized low-rank Tucker representation-based HSI super-resolution method, which infers the spatial consistency and spectral smoothness from the HR-MSI and LR-HSI, respectively.

Recently, deep CNN has become an emerging tool for HSI super-resolution due to its good learning performance and high speed. Dian *et al.* [8] proposed a deep residual learning-based HSI super-resolution method, which uses the deep CNN to learn the map between initialized HR-HSI and ground truth. This method is supervised and needs data to pretraining the CNN. Besides, Qu *et al.* [38] propose an unsupervised deep CNN for HSI super-resolution, which does not need pretraining and contains two neural networks with a shared decoder.

III. PRELIMINARIES ON TENSORS

An N -dimensional tensor is denoted by $\mathcal{M} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$. We define the n -mode unfolding vectors of \mathcal{M} as I_n -dimensional vectors by changing index i_n and keeping the other indices fixed. The Frobenius norm of tensor \mathcal{M} is defined as $\|\mathcal{M}\|_F = (\sum_{i_1, \dots, i_N} |m_{i_1 \dots i_N}|^2)^{1/2}$.

Mode- n matricization of tensor \mathcal{M} is the matrix $\mathbf{M}_{(n)} \in \mathbb{R}^{I_n \times I_1 I_2 \dots I_{n-1} I_{n+1} \dots I_N}$, whose columns are all n -mode vectors.

Mode- n canonical matricization [40] of tensor \mathcal{M} is defined as

$$\mathbf{M}_{\langle n \rangle} \in \mathbb{R}^{(I_1 I_2, \dots, I_n) \times (I_{n+1} I_{n+2}, \dots, I_N)} \quad (1)$$

and elements in $\mathbf{M}_{\langle n \rangle}$ satisfy

$$\begin{aligned} \mathbf{M}_{\langle n \rangle} \left(i_1 + (i_2 - 1)I_1 + \dots + (i_n - 1) \prod_{j=1}^{n-1} I_j, \right. \\ \left. i_{n+1} + (i_{n+2} - 1)I_{n+1} + \dots + (i_N - 1) \prod_{j=n+1}^{N-1} I_j \right) \\ = \mathcal{M}(i_1, i_2, \dots, i_N). \end{aligned} \quad (2)$$

Conversely, the mode- n canonical matrix $\mathbf{M}_{\langle n \rangle}$ can be transformed back to the tensor \mathcal{M} by $\mathcal{M} = \text{fold}_n(\mathbf{M}_{\langle n \rangle})$.

IV. PROPOSED LTTR APPROACH

In this section, we present the proposed LTTR method for HSI super-resolution.

A. Problem Formulation

In this paper, we denote the LR-HSI, HR-MSI, and HR-HSI as the 3-D tensors, where the first, second, and third modes denote the width, height, and spectral modes of images, respectively. The desired HR-HSI is denoted by $\mathcal{Z} \in \mathbb{R}^{N_W \times N_H \times N_S}$, where N_W and N_H stand for the dimension of two spatial modes and N_S is the dimension of spectral mode.

$\mathcal{X} \in \mathbb{R}^{N_w \times N_h \times N_S}$ represents the acquired LR-HSI. \mathcal{X} can be modeled as the spatially downsampled version of \mathcal{Z} , and this relationship can be written as

$$\mathbf{X}_{(3)} = \mathbf{Z}_{(3)} \mathbf{B} \mathbf{S} \quad (3)$$

where $\mathbf{X}_{(3)} \in \mathbb{R}^{N_S \times N_w N_h}$ and $\mathbf{Z}_{(3)} \in \mathbb{R}^{N_S \times N_W N_H}$ are the 3-mode matrices of \mathcal{X} and \mathcal{Z} , respectively. $\mathbf{B} \in \mathbb{R}^{N_W N_H \times N_w N_h}$ represents the point spread function of the hyperspectral imaging sensor, and is a block circulant matrix, which can be diagonalized by fast Fourier transform (FFT), i.e.,

$$\mathbf{B} = \mathbf{F} \mathbf{D} \mathbf{F}^H. \quad (4)$$

Here, the matrix \mathbf{F} denotes the FFT operation and \mathbf{F}^H represents the inverse FFT operation. The diagonal matrix \mathbf{D} holds eigenvalues of \mathbf{B} . The matrix $\mathbf{S} \in \mathbb{R}^{N_W N_H \times N_w N_h}$ is the spatial downsampling matrix, which selects the corresponding spectral pixels.

$\mathcal{Y} \in \mathbb{R}^{N_W \times N_H \times N_S}$ denote the acquired HR-MSI in the same scene by the multispectral imaging sensor. The HR-MSI can be modeled as the spectrally downsampled version of \mathcal{Z} , and the relationship can be written as

$$\mathbf{Y}_{(3)} = \mathbf{R} \mathbf{Z}_{(3)} \quad (5)$$

where $\mathbf{Y}_{(3)} \in \mathbb{R}^{N_S \times N_W N_H}$ is the three-mode matrices of \mathcal{Y} , and $\mathbf{R} \in \mathbb{R}^{N_S \times N_S}$ is the spectral response matrix of the multispectral imaging sensor, which downsamples the HR-HSI along the spectral mode.

B. Relaxed TT Rank

Given an N -dimensional tensor $\mathcal{M} \in \mathbb{R}^{I_1 \times I_2, \dots, \times I_N}$, its TT decomposition [18] factors \mathcal{M} as core tensors $\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \dots, \mathcal{G}^{(N)}$, where $\mathcal{G}^{(n)} \in \mathbb{R}^{R^{n-1} \times I_n \times R^n}$ for $n = 1, 2, \dots, N$ with $R^0 = R^N = 1$. Each element in \mathcal{M} is calculated as

$$\mathcal{M}_{i_1, i_2, \dots, i_N} = \sum_{r_1, r_2, \dots, r_{N-1}=1}^{R_1, R_2, \dots, R_{N-1}} \mathcal{G}_{1, i_1, r_1}^{(1)} \mathcal{G}_{r_1, i_2, r_2}^{(2)} \dots \mathcal{G}_{r_{N-1}, i_N, 1}^{(N)} \quad (6)$$

where $\{R_1, R_2, \dots, R_{N-1}\}$ is called as the TT rank of \mathcal{M} . However, the TT decomposition for a tensor is not unique, and different TT decompositions often have different TT ranks. In practice, we often focus on the lower bound of TT ranks.

Proposition 1: If the mode- n canonical matrix $\mathbf{M}_{\langle n \rangle}$ has $\text{rank}(\mathbf{M}_{\langle n \rangle}) = T_n$, then TT rank R_n satisfies that $R_n \geq T_n$.

Proof: For the mode- n TT rank is equal to R_n ; therefore, \mathcal{M} can be represented as

$$\mathcal{M}_{i_1, i_2, \dots, i_N} = \sum_{r_n=1}^{R_n} \mathcal{F}_{i_1, i_2, \dots, i_n, r_n} \mathcal{N}_{r_n, i_{n+1}, i_{n+2}, \dots, i_N} \quad (7)$$

where $\mathcal{F} \in \mathbb{R}^{I_1 \times I_2, \dots, \times I_n \times R_n}$ and $\mathcal{N} \in \mathbb{R}^{R_n \times I_{n+1}, I_{n+2}, \dots, \times I_N}$ are two tensors, and their elements are calculated via

$$\begin{aligned} \mathcal{F}_{i_1, i_2, \dots, i_n, r_n} &= \sum_{r_1, \dots, r_{n-1}=1}^{R_1, \dots, R_{n-1}} \mathcal{G}_{1, i_1, r_1}^{(1)} \mathcal{G}_{r_1, i_2, r_2}^{(2)} \dots \mathcal{G}_{r_{n-1}, i_n, r_n}^{(n)} \\ \mathcal{N}_{r_n, i_{n+1}, \dots, i_N} &= \sum_{r_{n+1}, \dots, r_{N-1}=1}^{R_{n+1}, \dots, R_{N-1}} \mathcal{G}_{r_n, i_{n+1}, r_{n+1}}^{(n+1)} \dots \mathcal{G}_{r_{N-1}, i_N, 1}^{(N)}. \end{aligned} \quad (8)$$

Equation (7) is equivalent as

$$\mathbf{M}_{\langle n \rangle} = \mathbf{F}_{\langle n \rangle} \mathbf{N}_{\langle 1 \rangle} \quad (9)$$

where $\mathbf{F}_{\langle n \rangle} \in \mathbb{R}^{I_1 \times I_2, \dots, I_n \times R_n}$ and $\mathbf{N}_{\langle 1 \rangle} \in \mathbb{R}^{R_n \times I_{n+1}, \dots, I_N}$ are the mode- n canonical matrix of \mathcal{F} and mode-1 canonical matrix of \mathcal{N} , respectively. Therefore, we have $\text{rank}(\mathbf{M}_{\langle n \rangle}) \leq R_n$. ■

More specifically, the $\text{rank}(\mathbf{M}_{\langle n \rangle}) = T_n$ is the lower bound of TT rank R_n . The lower bound $\text{rank}(\mathbf{M}_{\langle n \rangle})$ is often used to represent the TT rank in the TT rank minimization problems. Hence, we define the TT rank of tensor \mathcal{M} as

$$\sum_{n=1}^{N-1} \alpha_n \text{rank}(\mathbf{M}_{\langle n \rangle}) \quad (10)$$

where α_n is the weight and satisfies $\sum_{n=1}^{N-1} \alpha_n = 1$. The matrix $\mathbf{M}_{\langle n \rangle}$ is obtained by matricizing along the first n modes, and its rank can model the correlations between the first n modes and the last $N-n$ modes. Since $\text{rank}(\mathbf{M}_{\langle n \rangle})$ ($n = 1, 2, \dots, N-1$) contains correlations between permutations of all modes, it provides an effective way to capture the global correlations of a tensor.

For a matrix \mathbf{A} , $\text{rank}(\mathbf{A})$ is equal to the number of positive singular values of \mathbf{A} . The rank constraint is nonconvex and hard to optimize. To solve this problem, the nuclear norm is usually introduced to replace the rank constraint since it is the

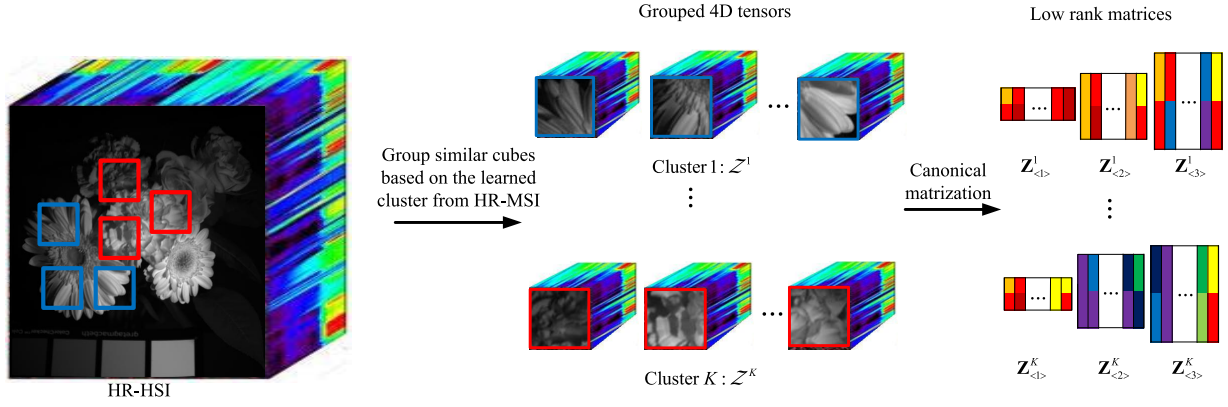


Fig. 2. Illustration of the proposed nonlocal LTTR prior for the HR-MSI.

tightest convex relaxation to the nonconvex rank minimization problem. The nuclear norm $\|\mathbf{A}\|_*$ for the matrix \mathbf{A} is defined as

$$\|\mathbf{A}\|_* = \sum_i \sigma_i(\mathbf{A}) \quad (11)$$

where $\sigma_i(\mathbf{A})$ denotes the i^{th} singular value of the matrix \mathbf{A} . Given the matrix $\mathbf{C} \in \mathbb{R}^{m \times n} (m > n)$, let $\mathbf{C} = \mathbf{U} \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n) \mathbf{V}$ be the singular value decomposition (SVD) of \mathbf{C} . For the following low-rank recovery problem:

$$\min_{\mathbf{A}} \frac{1}{2} \|\mathbf{C} - \mathbf{A}\|_F^2 + \alpha \|\mathbf{A}\|_*. \quad (12)$$

We have the closed-form solution

$$\mathbf{A} = \mathbf{U} \text{diag}(a_1, a_2, \dots, a_n) \mathbf{V} \quad (13)$$

where $a_i = \max(\sigma_i - \alpha, 0)$. However, larger singular values of a data matrix provide the major structural information and we should shrink them less. However, the smaller singular values mainly keep the noise, and we should shrink them more. Motivated by this, the log sum of singular values has been demonstrated to be more effective to solve the rank minimization problems [41], which is defined as

$$\text{LS}(\mathbf{A}) = \sum_i \log(\sigma_i(\mathbf{A}) + \varepsilon) \quad (14)$$

where ε is a very small positive number to avoid zero. The logarithm function can shrink the larger singular values less and maintains more meaningful information. On the contrary, the logarithm function can shrink the smaller singular values more to remove the noise. Let $0 < \alpha, 0 < \varepsilon < \min(\sqrt{\alpha}, (\alpha/\sigma_1))$. For the following low rank recovery problem:

$$\min_{\mathbf{A}} \frac{1}{2} \|\mathbf{C} - \mathbf{A}\|_F^2 + \alpha \sum_i \log(\sigma_i(\mathbf{A}) + \varepsilon). \quad (15)$$

We have the following local minimum:

$$\mathbf{A} = \mathbf{U} \text{diag}(d_1, d_2, \dots, d_n) \mathbf{V} \quad (16)$$

where $d_i = D_{\alpha, \varepsilon}(\sigma_i)$. Here, $D_{\alpha, \varepsilon}(\cdot)$ is the operator defined as

$$D_{\alpha, \varepsilon}(x) = \begin{cases} 0 & c_2 \leq 0 \\ \frac{c_1 + \sqrt{c_2}}{2} & c_2 > 0 \end{cases} \quad (17)$$

where $c_1 = |x| - \varepsilon$ and $c_2 = c_1^2 - 4(\alpha - \varepsilon|x|)$.

By replacing $\text{rank}(\mathbf{M}_{(n)})$ in (10) with $\text{LS}(\mathbf{M}_{(n)})$, the relaxed TT rank is defined as

$$\|\mathcal{M}\|_{\text{TT}} = \sum_{n=1}^{N-1} \alpha_n \text{LS}(\mathbf{M}_{(n)}). \quad (18)$$

C. HSI-MSI Fusion via LTTR

In the above-mentioned formulation, the goal of HSI super-resolution is estimating the HR-MSI \mathcal{Z} from \mathcal{X} and \mathcal{Y} . According to (1) and (2), we write the fusion problem as

$$\min_{\mathbf{X}_{(3)}} \|\mathbf{X}_{(3)} - \mathbf{Z}_{(3)} \mathbf{B} \mathbf{S}\|_F^2 + \|\mathbf{Y}_{(3)} - \mathbf{R} \mathbf{Z}_{(3)}\|_F^2 \quad (19)$$

where $\|\cdot\|_F$ stands for Frobenius norm. Since both of the HR-MSI and LR-MSI are the downsampled version of desired HR-MSI, the optimization problem (19) is severely ill-posed, and its solution may not be unique. Hence, some prior information of the HR-MSI is needed to regularize the super-resolution problem. Recently, various priors have been used for the super-resolution problem, including sparse priors [26], [32], [35], priors of spectral unmixing [27]–[29], low-rank priors [42], nonlocal spatial similarities [33], and spatial smoothness [31]. Here, we design a nonlocal LTTR prior for HR-MSI to regularize the fusion problem. As illustrated in Fig. 2, similar cubes in the HR-MSI are grouped together, and then, cubes in the same group constitute a 4-D tensor, which has LTTR property, that is, its canonical unfolding matrices are low rank. By imposing low-rank property on the canonical unfolding matrices of grouped 4-D tensors, we can effectively learn the correlations among the spatial, spectral, and nonlocal modes. However, the HR-MSI is unknown; we cannot directly group its cubes. To solve this problem, the group operation is conducted on the HR-MSI, since the spatial information mainly exists in the HR-MSI. We first spatially partition the HR-MSI into several cubes with the full spectral size of HR-MSI N_s and spatial size of $d_W \times d_H$, and we group similar HR-MSI cubes into K clusters $\mathcal{Y}^{(k)} = \{\mathcal{Y}^{(k,j)}\}_{j=1}^{N_k}$, $k = 1, 2, \dots, K$, where N_k is the number of cubes in the k^{th} cluster. The LR-MSI is not used for the nonlocal clustering for two reasons. On the one hand, the LR-MSI is spatially downsampled, and the spatial structures have been severely destroyed. On the

other hand, it is hard to match the cubes in the LR-HSI and HR-HSI, since they have the different spatial resolution. In the clustering process, we employ the efficient K -means++ method [43], which chooses random beginning centers with very specific probabilities and dramatically improve the speed. Based on the learned cluster structure, we build clusters of HR-HSIs cubes with the same spatial structure, that is, $\mathcal{Z}^{(k)} = \{\mathcal{Z}^{(k,j)} \in \mathbb{R}^{d_W \times d_H \times N_S}\}_{j=1}^{N_k}$, where the spatial locations of $\mathcal{Z}^{(k,j)}$ are the same as $\mathcal{Y}^{(k,j)}$. The 3-D HR-HSI cubes in k^{th} group can constitute a 4-D tensor $\mathcal{Z}^k \in \mathbb{R}^{d_W \times d_H \times N_S \times N_k}$, which has four modes, i.e., width mode, height mode, spectral mode, and nonlocal mode. The TT rank of tensor \mathcal{Z}^k is defined as

$$\|\mathcal{Z}^k\|_{\text{TT}} = \sum_{t=1}^3 \alpha_t \text{LS}(\mathbf{Z}_{(t)}^k). \quad (20)$$

The weight α_t is defined as follows:

$$\alpha_t = \frac{\sqrt{\beta_t}}{\sum_{t=1}^3 \sqrt{\beta_t}}, \quad \text{with } \beta_t = \min \left(\prod_{j=1}^t I_j, \prod_{j=t}^3 I_j \right), \quad t = 1, 2, 3 \quad (21)$$

where I_j denotes the dimension of j^{th} mode of tensor \mathcal{Z}^k . The TT rank $\|\mathcal{Z}^k\|_{\text{TT}}$ contains the rank of matrices $\mathbf{Z}_{(1)}^k \in \mathbb{R}^{d_W \times d_H \times N_S \times N_k}$, $\mathbf{Z}_{(2)}^k \in \mathbb{R}^{d_W \times d_H \times N_S \times N_k}$, and $\mathbf{Z}_{(3)}^k \in \mathbb{R}^{d_W \times d_H \times N_S \times N_k}$. Especially, the rank of $\mathbf{Z}_{(2)}^k$ can capture the correlations between the two spatial modes and the last two modes (spectral and nonlocal modes), and the rank of $\mathbf{Z}_{(3)}^k$ can model the correlations between the first three modes (spatial and spectral modes) and the nonlocal mode. Since the strong nonlocal similarities and spatial-spectral correlations exist in the HR-HSI, the four modes of tensors $\mathcal{Z}^k \in \mathbb{R}^{d_W \times d_H \times N_S \times N_k}$, $k = 1, \dots, K$ are highly correlated, and thus the 4-D tensors have the LTTR properties. To incorporate the LTTR prior into the model (19), we obtain the following optimization problem:

$$\min_{\mathbf{Z}_{(3)}} \|\mathbf{X}_{(3)} - \mathbf{Z}_{(3)}\mathbf{BS}\|_F^2 + \|\mathbf{Y}_{(3)} - \mathbf{RZ}_{(3)}\|_F^2 + \lambda \sum_{k=1}^K \|\mathcal{Z}^k\|_{\text{TT}} \quad (22)$$

where λ is the TT rank regularization parameter. Combining with (20), optimization problem (22) is equivalent as

$$\min_{\mathbf{Z}_{(3)}} \|\mathbf{X}_{(3)} - \mathbf{Z}_{(3)}\mathbf{BS}\|_F^2 + \|\mathbf{Y}_{(3)} - \mathbf{RZ}_{(3)}\|_F^2 + \lambda \sum_{k=1}^K \sum_{t=1}^3 \alpha_t \text{LS}(\mathbf{Z}_{(t)}^k). \quad (23)$$

D. Optimization

The problem (23) is unconstrained optimization problem, and we use the framework of the ADMM [25] to solve it. By introducing three variables $\mathcal{U} = \mathcal{Z}$, $\mathcal{V} = \mathcal{Z}$, and $\mathcal{W} = \mathcal{Z}$,

the following augmented Lagrangian function is acquired:

$$\begin{aligned} L(\mathcal{U}, \mathcal{V}, \mathcal{W}, \mathcal{Z}, \mathcal{O}, \mathcal{P}, \mathcal{Q}) &= \|\mathbf{X}_{(3)} - \mathbf{Z}_{(3)}\mathbf{BS}\|_F^2 + \|\mathbf{Y}_{(3)} - \mathbf{RZ}_{(3)}\|_F^2 \\ &+ \mu \|\mathcal{U} - \mathcal{Z}\|_F^2 + \frac{\mathcal{O}}{2\mu} \|\mathcal{U}\|_F^2 + \lambda \sum_{k=1}^K \alpha_1 \text{LS}(\mathbf{U}_{(1)}^k) \\ &+ \mu \|\mathcal{V} - \mathcal{Z}\|_F^2 + \frac{\mathcal{P}}{2\mu} \|\mathcal{V}\|_F^2 + \lambda \sum_{k=1}^K \alpha_2 \text{LS}(\mathbf{V}_{(2)}^k) \\ &+ \mu \|\mathcal{W} - \mathcal{Z}\|_F^2 + \frac{\mathcal{Q}}{2\mu} \|\mathcal{W}\|_F^2 + \lambda \sum_{k=1}^K \alpha_3 \text{LS}(\mathbf{W}_{(3)}^k) \end{aligned} \quad (24)$$

where \mathcal{O} , \mathcal{P} , and \mathcal{Q} are the Lagrangian multipliers and μ is the penalty parameter. The augmented Lagrangian function (24) can be minimized by iteratively solving the following subproblems:

(a) Update \mathcal{Z} :

$$\begin{aligned} \mathcal{Z} &\in \underset{\mathcal{Z}}{\text{argmin}} L(\mathcal{U}, \mathcal{V}, \mathcal{W}, \mathcal{Z}, \mathcal{O}, \mathcal{P}, \mathcal{Q}) \\ &= \underset{\mathcal{Z}}{\text{argmin}} \|\mathbf{Y}_{(3)} - \mathbf{RZ}_{(3)}\|_F^2 + \|\mathbf{X}_{(3)} - \mathbf{Z}_{(3)}\mathbf{BS}\|_F^2 \\ &+ \mu \left\| \mathcal{U} - \mathcal{Z} + \frac{\mathcal{O}}{2\mu} \right\|_F^2 + \mu \left\| \mathcal{V} - \mathcal{Z} + \frac{\mathcal{P}}{2\mu} \right\|_F^2 \\ &+ \mu \left\| \mathcal{W} - \mathcal{Z} + \frac{\mathcal{Q}}{2\mu} \right\|_F^2. \end{aligned} \quad (25)$$

Problem (25) is strong convex. Therefore, we force the derivative of (25) with regard to $\mathbf{Z}_{(3)}$ to be zero, and obtain the following equation:

$$\mathbf{H}_1 \mathbf{Z}_{(3)} + \mathbf{Z}_{(3)} \mathbf{H}_2 = \mathbf{H}_3 \quad (26)$$

where

$$\begin{aligned} \mathbf{H}_1 &= \mathbf{R}^T \mathbf{R} + 3\mu \mathbf{I} \\ \mathbf{H}_2 &= (\mathbf{BS})(\mathbf{BS})^T \\ \mathbf{H}_3 &= \mathbf{R}^T \mathbf{Y}_{(3)} + \mathbf{X}_{(3)}(\mathbf{BS})^T + \mu \left(\mathbf{U}_{(3)} + \frac{\mathbf{O}_{(3)}}{2\mu} \right) \\ &+ \mu \left(\mathbf{V}_{(3)} + \frac{\mathbf{P}_{(3)}}{2\mu} \right) + \mu \left(\mathbf{W}_{(3)} + \frac{\mathbf{Q}_{(3)}}{2\mu} \right) \end{aligned} \quad (27)$$

where $\mathbf{U}_{(3)}$, $\mathbf{V}_{(3)}$, $\mathbf{W}_{(3)}$, $\mathbf{O}_{(3)}$, $\mathbf{P}_{(3)}$, and $\mathbf{Q}_{(3)}$ are three-mode unfolding matrices of tensors \mathcal{U} , \mathcal{V} , \mathcal{W} , \mathcal{O} , \mathcal{P} , and \mathcal{Q} , respectively. $\mathbf{I} \in \mathbb{R}^{N_S \times N_S}$ is the identity matrix. If and only if the matrices \mathbf{H}_1 and $-\mathbf{H}_2$ do not have common eigenvalues [44], the solution of Sylvester equation (26) is unique. \mathbf{H}_1 is the positive matrix, and thus its eigenvalues are positive values. Since \mathbf{H}_2 is semipositive, the eigenvalues of $-\mathbf{H}_2$ are nonpositive. Therefore, \mathbf{H}_1 and $-\mathbf{H}_2$ has no common eigenvalue, the Sylvester equation (26) has unique solution. By exploiting the properties of \mathbf{B} and \mathbf{S} , we can solve Sylvester Equation (26) analytically and efficiently as demonstrated in [45]. We summarize the fast method for solving the equation (26) in **Algorithm 1**+

In **Algorithm 1**, we diagonalize the matrix \mathbf{B} as illustrated in (4). The diagonal matrix $\mathbf{D} \in \mathbb{C}^{N_W N_H \times N_W N_H}$ can be

Algorithm 1 An Analytical Solution of (26) w.r.t. $\mathbf{Z}_{(3)}$ **Input:** $\mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3, \mathbf{B}, \mathbf{S}$ **Output:** $\mathbf{Z}_{(3)}$

- (a) Eigen-decomposition of $\mathbf{B} = \mathbf{F}\mathbf{D}\mathbf{F}^H$ of \mathbf{B} :
- (b) $\tilde{\mathbf{D}} = \mathbf{D}(\mathbf{1}_d \otimes \mathbf{1}_{N_w N_h})$
- (c) Eigen-decomposition of \mathbf{H}_1 :
 $\mathbf{H}_1 = \mathbf{Q}_1 \Lambda \mathbf{Q}_1^{-1}$
- (d) $\tilde{\mathbf{H}}_3 = \mathbf{Q}_1^{-1} \mathbf{H}_3 \mathbf{F}$
- (e) Calculate auxiliary matrix $\tilde{\mathbf{Z}}_{(3)}$ row by row
For $l = 1$ **to** N_S **do**
 $\tilde{\mathbf{z}}_l = \lambda_l^{-1}(\tilde{\mathbf{H}}_3)_l - \lambda_l^{-1}(\tilde{\mathbf{H}}_3)_l \tilde{\mathbf{D}}(\lambda_l d \mathbf{I}_n + \sum_{t=1}^d \mathbf{D}_t^2) \tilde{\mathbf{D}}^H$
End
- (f) Set $\mathbf{Z}_{(3)} = \mathbf{Q}_1 \tilde{\mathbf{Z}}_{(3)} \mathbf{F}^H$

represented as

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{D}_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{D}_d \end{bmatrix} \quad (28)$$

where $\mathbf{D}_i \in \mathbb{C}^{N_w N_h \times N_w N_h}$ and d is the spatial downsampling factor. The diagonal matrix Λ and unitary matrix \mathbf{Q}_1 are acquired by the eigen-decomposition of \mathbf{H}_1 , i.e., $\mathbf{H}_1 = \mathbf{Q}_1 \Lambda \mathbf{Q}_1^{-1}$, where Λ can be written as

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{N_S} \end{bmatrix} \quad (29)$$

and the columns of \mathbf{Q}_1 hold eigenvectors of \mathbf{H}_1 . $\tilde{\mathbf{z}}_l$ represents the l^{th} row of $\tilde{\mathbf{Z}}_{(3)} = \mathbf{Q}_1^{-1} \mathbf{Z}_{(3)} \mathbf{F}$, that is, $\tilde{\mathbf{Z}}_{(3)} = [\tilde{\mathbf{z}}_1^T, \tilde{\mathbf{z}}_2^T, \dots, \tilde{\mathbf{z}}_{N_S}^T]^T$. $\mathbf{1}_d \in \mathbb{R}^d$ denotes vector, whose elements are all ones.

(b) Update \mathcal{U} :

$$\begin{aligned} \mathcal{U} &= \underset{\mathcal{U}}{\operatorname{argmin}} L(\mathcal{U}, \mathcal{V}, \mathcal{W}, \mathcal{Z}, \mathcal{O}, \mathcal{P}, \mathcal{Q}) \\ &= \underset{\mathcal{U}}{\operatorname{argmin}} \mu \left\| \mathcal{U} - \mathcal{Z} + \frac{\mathcal{O}}{2\mu} \right\|_F^2 + \lambda \sum_{k=1}^K \alpha_1 \operatorname{LS}(\mathbf{U}_{(1)}^k). \end{aligned} \quad (30)$$

The problem (30) can be separately solved for each cluster, i.e.,

$$\underset{\mathcal{U}}{\operatorname{argmin}} \sum_{k=1}^K \mu \left\| \mathbf{U}_{(1)}^k - \mathbf{Z}_{(1)}^k + \frac{\mathbf{O}_{(1)}^k}{2\mu} \right\|_F^2 + \lambda \sum_{k=1}^K \alpha_1 \operatorname{LS}(\mathbf{U}_{(1)}^k) \quad (31)$$

where $\mathbf{U}_{(1)}^k, \mathbf{Z}_{(1)}^k$, and $\mathbf{O}_{(1)}^k$ are the mode-1 canonical matrices of tensors $\mathcal{U}^k, \mathcal{Z}^k$, and \mathcal{O}^k . The 4-D tensors $\mathcal{U}^k, \mathcal{Z}^k$, and \mathcal{O}^k are constituted by k^{th} cube cluster in $\mathcal{U}^k, \mathcal{Z}^k$, and \mathcal{O}^k , respectively. The optimization problem (31) has the following solution:

$$\mathbf{U}_{(1)}^k = \mathbf{U}_1^k \tilde{\Sigma}_1^k \mathbf{V}_1^k, \quad 1 \leq k \leq K \quad (32)$$

where \mathbf{U}_1^k and \mathbf{V}_1^k are acquired by the SVD, i.e., $\mathbf{Z}_{(1)}^k - (\mathbf{O}_{(1)}^k/2\mu) = \mathbf{U}_1^k \tilde{\Sigma}_1^k \mathbf{V}_1^k$. $\tilde{\Sigma}_1^k$ is the diagonal matrix, whose diagonal elements are computed via $\tilde{\Sigma}_1^k(i, i) = D_{(\lambda \alpha_1/2\mu), \varepsilon}[\Sigma_1^k(i, i)]$. $D_{\alpha, \varepsilon}$ is the operator defined in Section IV-B. After obtaining mode-1 canonical matrix $\mathbf{U}_{(1)}^k$, we can acquire tensor \mathcal{U}^k via $\mathcal{U}^k = \operatorname{fold}_1(\mathbf{U}_{(1)}^k)$, and \mathcal{U} is obtained by placing the cube clusters $\mathcal{U}^k, 1 \leq k \leq K$, into the corresponding positions.

(c) Update \mathcal{V} :

$$\begin{aligned} \mathcal{V} &= \underset{\mathcal{V}}{\operatorname{argmin}} L(\mathcal{U}, \mathcal{V}, \mathcal{W}, \mathcal{Z}, \mathcal{O}, \mathcal{P}, \mathcal{Q}) \\ &= \underset{\mathcal{V}}{\operatorname{argmin}} \mu \left\| \mathcal{V} - \mathcal{Z} + \frac{\mathcal{P}}{2\mu} \right\|_F^2 + \lambda \sum_{k=1}^K \alpha_2 \operatorname{LS}(\mathbf{V}_{(2)}^k). \end{aligned} \quad (33)$$

The problem (33) can be solved for each cluster separately, i.e.,

$$\underset{\mathcal{V}}{\operatorname{argmin}} \sum_{k=1}^K \mu \left\| \mathbf{V}_{(2)}^k - \mathbf{Z}_{(2)}^k + \frac{\mathbf{P}_{(2)}^k}{2\mu} \right\|_F^2 + \lambda \sum_{k=1}^K \alpha_2 \operatorname{LS}(\mathbf{V}_{(2)}^k) \quad (34)$$

where $\mathbf{V}_{(2)}^k, \mathbf{Z}_{(2)}^k$, and $\mathbf{P}_{(2)}^k$ are the mode-2 canonical matrices of tensors $\mathcal{V}^k, \mathcal{Z}^k$, and \mathcal{P}^k , respectively. The 4-D tensors $\mathcal{V}^k, \mathcal{Z}^k$, and \mathcal{P}^k are constituted by the k^{th} cube cluster in $\mathcal{V}^k, \mathcal{Z}^k$, and \mathcal{P}^k , respectively. The optimization problem (34) has the following solution:

$$\mathbf{V}_{(2)}^k = \mathbf{U}_2^k \tilde{\Sigma}_2^k \mathbf{V}_2^k, \quad 1 \leq k \leq K \quad (35)$$

where \mathbf{U}_2^k and \mathbf{V}_2^k are acquired by the SVD, i.e., $\mathbf{Z}_{(2)}^k - (\mathbf{P}_{(2)}^k/2\mu) = \mathbf{U}_2^k \tilde{\Sigma}_2^k \mathbf{V}_2^k$. $\tilde{\Sigma}_2^k$ is the diagonal matrix, whose diagonal elements are computed via $\tilde{\Sigma}_2^k(i, i) = D_{(\lambda \alpha_2/2\mu), \varepsilon}[\Sigma_2^k(i, i)]$. The tensor \mathcal{V}^k via $\mathcal{V}^k = \operatorname{fold}_2(\mathbf{V}_{(2)}^k)$, and \mathcal{V} is obtained by placing the cube clusters $\mathcal{V}^k, 1 \leq k \leq K$, into the corresponding positions.

(d) Update \mathcal{W} :

$$\begin{aligned} \mathcal{W} &= \underset{\mathcal{W}}{\operatorname{argmin}} L(\mathcal{U}, \mathcal{V}, \mathcal{W}, \mathcal{Z}, \mathcal{O}, \mathcal{P}, \mathcal{Q}) \\ &= \underset{\mathcal{W}}{\operatorname{argmin}} \mu \left\| \mathcal{W} - \mathcal{Z} + \frac{\mathcal{Q}}{2\mu} \right\|_F^2 + \lambda \sum_{k=1}^K \alpha_3 \operatorname{LS}(\mathbf{W}_{(3)}^k). \end{aligned} \quad (36)$$

The problem (36) can be solved for each cluster separately, i.e.,

$$\underset{\mathcal{W}}{\operatorname{argmin}} \sum_{k=1}^K \mu \left\| \mathbf{W}_{(3)}^k - \mathbf{Z}_{(3)}^k + \frac{\mathbf{Q}_{(3)}^k}{2\mu} \right\|_F^2 + \lambda \sum_{k=1}^K \alpha_3 \operatorname{LS}(\mathbf{W}_{(3)}^k) \quad (37)$$

where $\mathbf{W}_{(3)}^k, \mathbf{Z}_{(3)}^k$, and $\mathbf{Q}_{(3)}^k$ are the mode-3 canonical matrices of tensors $\mathcal{W}^k, \mathcal{Z}^k$, and \mathcal{Q}^k , respectively. The 4-D tensors $\mathcal{W}^k, \mathcal{Z}^k$, and \mathcal{Q}^k are constituted by the k^{th} cube cluster in $\mathcal{W}^k, \mathcal{Z}^k$, and \mathcal{Q}^k , respectively. The optimization problem (37) has the following solution:

$$\mathbf{W}_{(3)}^k = \mathbf{U}_3^k \tilde{\Sigma}_3^k \mathbf{V}_3^k, \quad 1 \leq k \leq K \quad (38)$$

where \mathbf{U}_3^k and \mathbf{V}_3^k are acquired by the SVD, i.e., $\mathbf{Z}_{(3)}^k - (\mathbf{Q}_{(3)}^k/2\mu) = \mathbf{U}_3^k \tilde{\Sigma}_3^k \mathbf{V}_3^k$. $\tilde{\Sigma}_3^k$ is the diagonal matrix,

Algorithm 2 LTTR-Based HSI Super-Resolution**Input:** $\mathcal{X}, \mathcal{Y}, K, \lambda$ **Output:** \mathcal{Z} **Non-local Clustering**(a) Group HR-MSI cubes into K clusters via, K-means++(b) Group HR-HSI cubes into K clusters according to the cluster

structure in the HR-MSI cubes

Solving (23) via ADMM:**while** not converged **do**(a) Update \mathcal{Z} via **Algorithm 1**(b) Update $\mathbf{U}_{(1)}^k (k = 1, \dots, K)$ via (32)Fold matrix $\mathbf{U}_{(1)}^k$ as tensor \mathcal{U}^k Acquire \mathcal{U} by replacing the cube clusters $\mathcal{U}^k (k = 1, \dots, K)$ (c) Update $\mathbf{V}_{(2)}^k (k = 1, \dots, K)$ via (35)Fold matrix $\mathbf{V}_{(2)}^k$ as tensor \mathcal{V}^k Acquire \mathcal{V} by replacing the cube clusters $\mathcal{V}^k (k = 1, \dots, K)$ (d) Update $\mathbf{W}_{(3)}^k (k = 1, \dots, K)$ via (38)Fold matrix $\mathbf{W}_{(3)}^k$ as tensor \mathcal{W}^k Acquire \mathcal{W} by replacing the cube clusters $\mathcal{W}^k (k = 1, \dots, K)$ (e) Update Lagrangian multipliers \mathcal{O}, \mathcal{P} , and \mathcal{Q} via (39)**End while**

whose diagonal elements are computed by $\tilde{\Sigma}_3^k(i, i) = D_{(\lambda\alpha_3/2\mu), \varepsilon}[\Sigma_3^k(i, i)]$. The tensor \mathcal{W}^k can be acquired via $\mathcal{W}^k = \text{fold}_3(\mathbf{W}_{(3)}^k)$, and \mathcal{W} is obtained by replacing the cube clusters $\mathcal{W}^k, 1 \leq k \leq K$ into the corresponding positions.

(e) Update Lagrangian multipliers \mathcal{O}, \mathcal{P} , and \mathcal{Q} :

$$\begin{aligned}\mathcal{O} &= \mathcal{O} + 2\mu(\mathcal{U} - \mathcal{Z}) \\ \mathcal{P} &= \mathcal{P} + 2\mu(\mathcal{V} - \mathcal{Z}) \\ \mathcal{Q} &= \mathcal{Q} + 2\mu(\mathcal{W} - \mathcal{Z}).\end{aligned}\quad (39)$$

The proposed LTTR method for HSI superresolution is summarized in **Algorithm 2**. The convergence of the ADMM (**Algorithm 2**) is hard to analyze because of the nonconvexity of logarithm function in the definition of relaxed TT rank. However, recent works [47] proved that the ADMM algorithm converges on some classes of nonconvex functions, e.g., the Lipschitz differentiable function. (The function is differentiable and its gradient is Lipschitz continuous.) We can verify that the logarithm function $\log(x + \varepsilon)$ is indeed Lipschitz differentiable, and thus, the ADMM algorithm for solving (23) can converge. We stop **Algorithm 2** when the maximum number of iterations T is attained.

V. EXPERIMENTS

This section presents the experiments for HSI superresolution. The code of our method is released in the author's homepage <https://sites.google.com/view/renweidian>.

A. Experimental Database

To evaluate the performance of proposed LTTR method, experiments are conducted on both remotely sensed HSI data set and ground-based HSI data set. For ground-based HSI data set, we choose CAVE database [46],¹ which is widely used in HSI super-resolution literatures [28], [30], [33], [35], [48], [49]. The CAVE database [46] has 32 indoor HSIs captured by generalized assorted pixel camera. HSIs of the CAVE database have the size of $512 \times 512 \times 31$, and 512×512 is the dimension of spatial modes, and 31 is the dimension of the spectral mode. The spectral wavelength has interval 10 nm in the range of 400–700 nm. Some RGB images of the CAVE database are shown in Fig. 3. We use the HSIs of CAVE database for ground truths. We simulate the LR-HSI by using a 7×7 Gaussian blur (zero mean and standard deviation 2) and then downsample the blurred image every eight spectral pixels in two spatial modes for each reference image, that is, the downsampling factor is 8×8 . We simulate the HR-MSI (RGB image) of the same scene by downsampling the reference HSI along the spectral mode. The spectral downsampling filter comes from the response of a Nikon D700 camera.² The cube size is set as $8 \times 8 \times 31$ for CAVE database.

For the remotely sensed hyperspectral data, we utilize the Pavia University [50], acquired over the urban area of the University of Pavia, Italy. The used Pavia University is of size $128 \times 128 \times 115$, which has the 115 spectral bands and 128×128 spectral pixels. Since some bands contain water vapor absorption bands, we remove these bands and reduce the HSI as 93 bands. To generate the LR-HSI, we first blur the HR-HSI by using a 7×7 Gaussian blur (with zero mean and standard deviation 2) and then downsample every four pixels in both width and height modes of the HR-HSI. To simulate the HR-MSI of the same scene, the IKONOS-like reflectance spectral response filter [32] is used to downsample HR-HSI along the spectral mode. The cube size is set as $8 \times 8 \times 93$ for Pavia University.

B. Compared Methods

We compare the LTTR method with three current state-of-the-art HSI super-resolution approaches, including non-negative structured sparse representation (NSSR) [33], coupled spectral unmixing (CSU) [28], and nonlocal sparse tensor factorization method (NLSTF) [35]. The parameters for the compared methods are set for the best performance. Specifically, for the CSU, the number of iterations is set as 1500. For the NSSR, the number of dictionary atoms is set as 75, η_1 is set as 0.015 and 0.00015 for CAVE database and Pavia University, respectively, and η_2 is set as 10^{-4} . For the NLSTF, we rewrite the tensor sparse coding stage with ℓ_1 norm-based version, which makes the algorithm much faster and produces slightly better results.

¹<http://www.cs.columbia.edu/CAVE/databases/multispectral/>²https://www.maxmax.com/spectral_response.htm

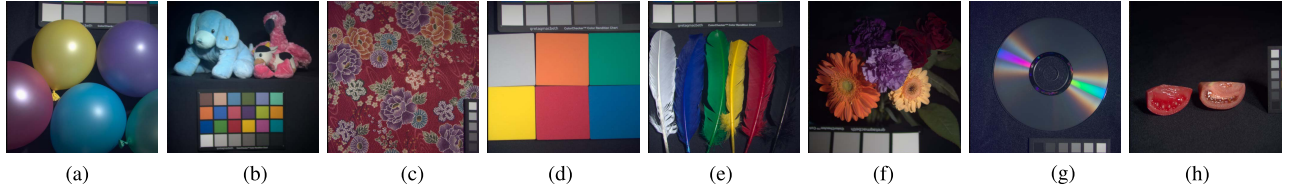


Fig. 3. RGB images from the CAVE database [46]. (a) Balloons. (b) Toys. (c) Cloth. (d) Sponges. (e) Feathers. (f) Flowers. (g) CD. (h) Tomatoes.

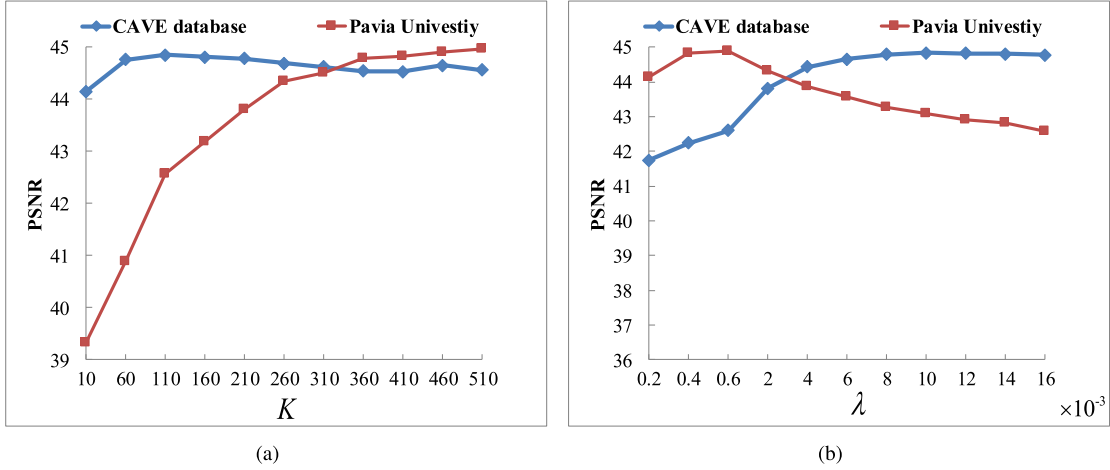


Fig. 4. PSNR curves as a function of parameters K and λ for the proposed LTTR method. (a) Number of clusters K . (b) Regularization parameter λ .

C. Quantitative Metrics

Four quantitative metrics are used to evaluate the quality of the reconstructed HSIs in our study.

The first index is the peak signal-to-noise ratio (PSNR). The PSNR for HSI is defined as the average value of all bands, i.e.,

$$\text{PSNR}(\mathcal{Z}, \hat{\mathcal{Z}}) = \frac{1}{N_S} \sum_{i=1}^{N_S} \text{PSNR}(\mathbf{Z}_i, \hat{\mathbf{Z}}_i) \quad (40)$$

where \mathbf{Z}_i and $\hat{\mathbf{Z}}_i$ denote the i^{th} band images of ground truth $\mathcal{Z} \in \mathbb{R}^{N_W \times N_H \times N_S}$ and estimated HR-HSI $\hat{\mathcal{Z}} \in \mathbb{R}^{N_W \times N_H \times N_S}$, respectively.

The second metric is the spectral angle mapper (SAM) [51], defined as the average angle between ground truth pixel \mathbf{z}_j and estimated pixel $\hat{\mathbf{z}}_j$, i.e.,

$$\text{SAM}(\mathcal{Z}, \hat{\mathcal{Z}}) = \frac{1}{N_W N_H} \sum_{j=1}^{N_W N_H} \arccos \frac{\hat{\mathbf{z}}_j^T \mathbf{z}_j}{\|\hat{\mathbf{z}}_j\|_2 \|\mathbf{z}_j\|_2}. \quad (41)$$

The SAM is given in degrees.

The third metric is the universal image quality Index (UIQI) [52]. We compute the UIQI on a sliding window of size 32×32 pixels and average on all window positions. Let $\mathbf{Z}_{i,j}$ and $\hat{\mathbf{Z}}_{i,j}$ denote the j^{th} window of i^{th} band ground truth image and estimated image, respectively. The UIQI between i^{th} band images \mathbf{Z}_i and $\hat{\mathbf{Z}}_i$ is given by

$$\text{UIQI}(\mathbf{Z}_i, \hat{\mathbf{Z}}_i) = \frac{1}{M} \sum_{j=1}^M \frac{4\sigma_{\mathbf{Z}_{i,j}\hat{\mathbf{Z}}_{i,j}}^2}{\sigma_{\mathbf{Z}_{i,j}}^2 + \sigma_{\hat{\mathbf{Z}}_{i,j}}^2} \frac{\mu_{\mathbf{Z}_{i,j}} \mu_{\hat{\mathbf{Z}}_{i,j}}}{\mu_{\mathbf{Z}_{i,j}}^2 + \mu_{\hat{\mathbf{Z}}_{i,j}}^2} \quad (42)$$

where M is the number of window positions, $\mu_{\mathbf{Z}_{i,j}}$ and $\sigma_{\mathbf{Z}_{i,j}}$ are the mean value and standard deviation of $\mathbf{Z}_{i,j}$, respectively, and $\sigma_{\mathbf{Z}_{i,j}\hat{\mathbf{Z}}_{i,j}}^2$ is the sample covariance between $\mathbf{Z}_{i,j}$ and $\hat{\mathbf{Z}}_{i,j}$. The UIQI of HSI is defined as the average value over all bands.

The fourth index is the relative dimensionless global error in synthesis (ERGAS) [53], defined as

$$\text{ERGAS}(\mathbf{Z}, \hat{\mathbf{Z}}) = \frac{100}{d} \sqrt{\frac{1}{N_S} \sum_{i=1}^{N_S} \frac{\text{MSE}(\mathbf{Z}_i, \hat{\mathbf{Z}}_i)}{\mu_{\mathbf{Z}_i}^2}} \quad (43)$$

where d is the spatial downsampling factor, $\text{MSE}(\mathbf{Z}_i, \hat{\mathbf{Z}}_i)$ is the mean square error between \mathbf{Z}_i and $\hat{\mathbf{Z}}_i$, and $\mu_{\mathbf{Z}_i}$ is the mean value of \mathbf{Z}_i .

D. Parameters Selection

For the LTTR method, we need to set two key parameters K and λ . The number of clusters K has an important influence on the cubes clustering process. Fig. 4(a) plots the PSNR of the reconstructed Toys (an HSI in the CAVE database) and of Pavia University as a function of K . As we can see from Fig. 4(a), the PSNR for CAVE database first rises as K varies from 10 to 110, and then, the PSNR decreases as K is bigger than 160. Hence, we set $K = 120$ for CAVE database. For Pavia University, the PSNR rises rapidly as K varies from 10 to 460, and then, it reaches a stable level. Therefore, K is set as 460 for Pavia University in our experiments. The selections of parameter K are very different for CAVE database and Pavia University, which is caused by different scenes captured by ground-based HSI (CAVE database) and remotely sensed HSI (Pavia University). Pavia University captures the scene

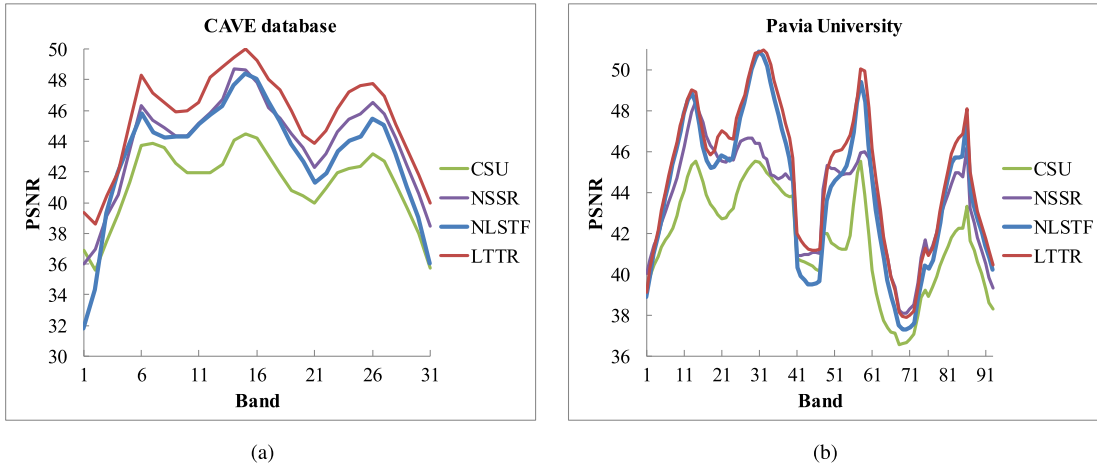


Fig. 5. Average PSNR as a function of the spectral bands for the compared methods. (a) CAVE database. (b) Pavia University.

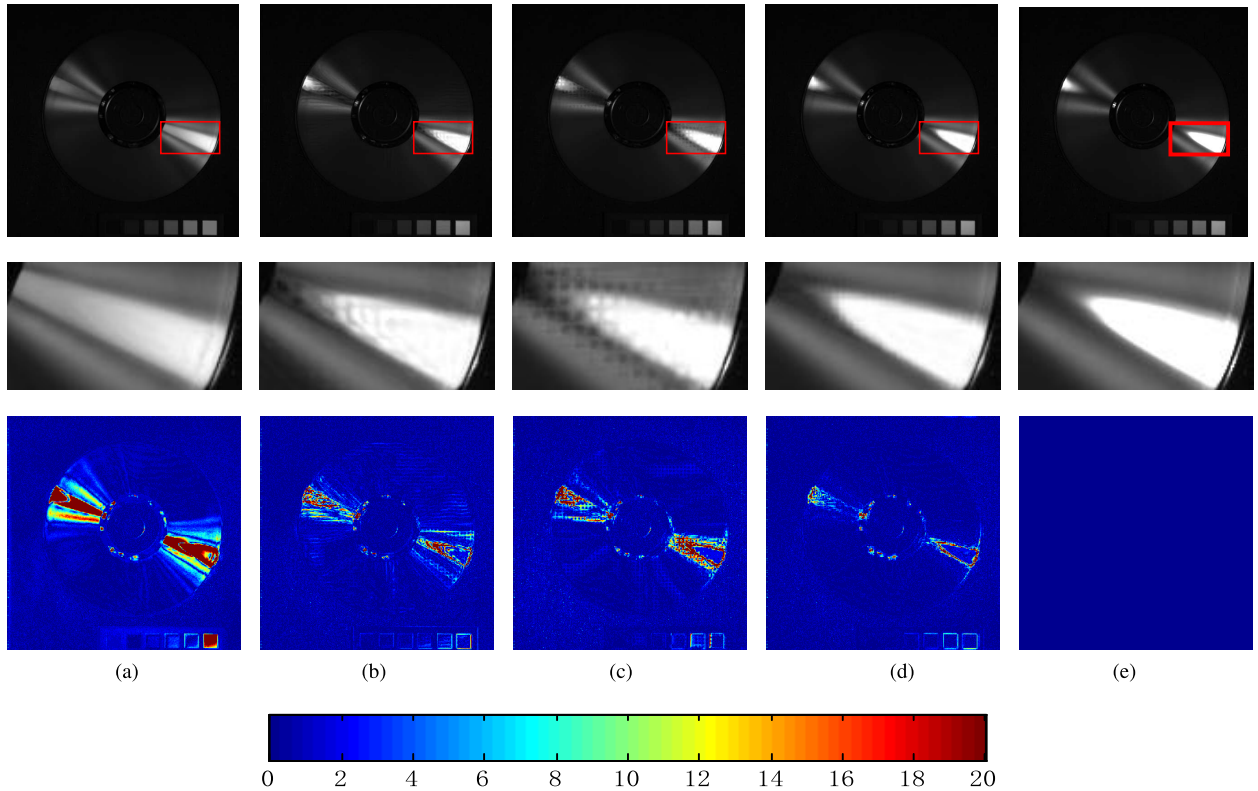


Fig. 6. First row: reconstructed images of the compared methods for *CD* (an HSI in CAVE database) at the ninth band. Second row: magnified images of the marked region. Third row: corresponding error images of the compared methods for *CD* (an HSI in CAVE database) at the ninth band. (a) CSU [28]. (b) NSSR [33]. (c) NLSTF [35]. (d) LTTR. (e) Ground truth.

over the urban area of the University of Pavia, Italy, which contains many different objects. Compared with Pavia University, the HSIs in CAVE database capture much smaller scenes, which contain much fewer objects. Therefore, the parameter K is set much bigger in Pavia University than that in CAVE database.

The parameter λ is the weight of the LTTR term, which has an importance on the solution. Fig. 4(b) plots the PSNR curves of the reconstructed Toys (an HSI in the CAVE database) and of Pavia University as a function of λ . As we can see from Fig. 4(b), the PSNR for CAVE database rises, as λ varies

from 2×10^{-4} to 0.01, and then, it does not have obvious changes as λ grows further. For Pavia University, the PSNR rises as λ grows from 2×10^{-4} to 4×10^{-4} , and then, it decreases quickly as λ is bigger than 6×10^{-4} . Therefore, we set $\lambda = 0.01$ for CAVE database and $\lambda = 4 \times 10^{-4}$ for Pavia University. The parameter λ is set bigger for the CAVE database than that for Pavia University. The reason is that HSIs in CAVE database has much higher spatial resolution than that of Pavia University. In this way, the nonlocal spatial-spectral similarities in the CAVE database are stronger, and we need the bigger regularization parameter.

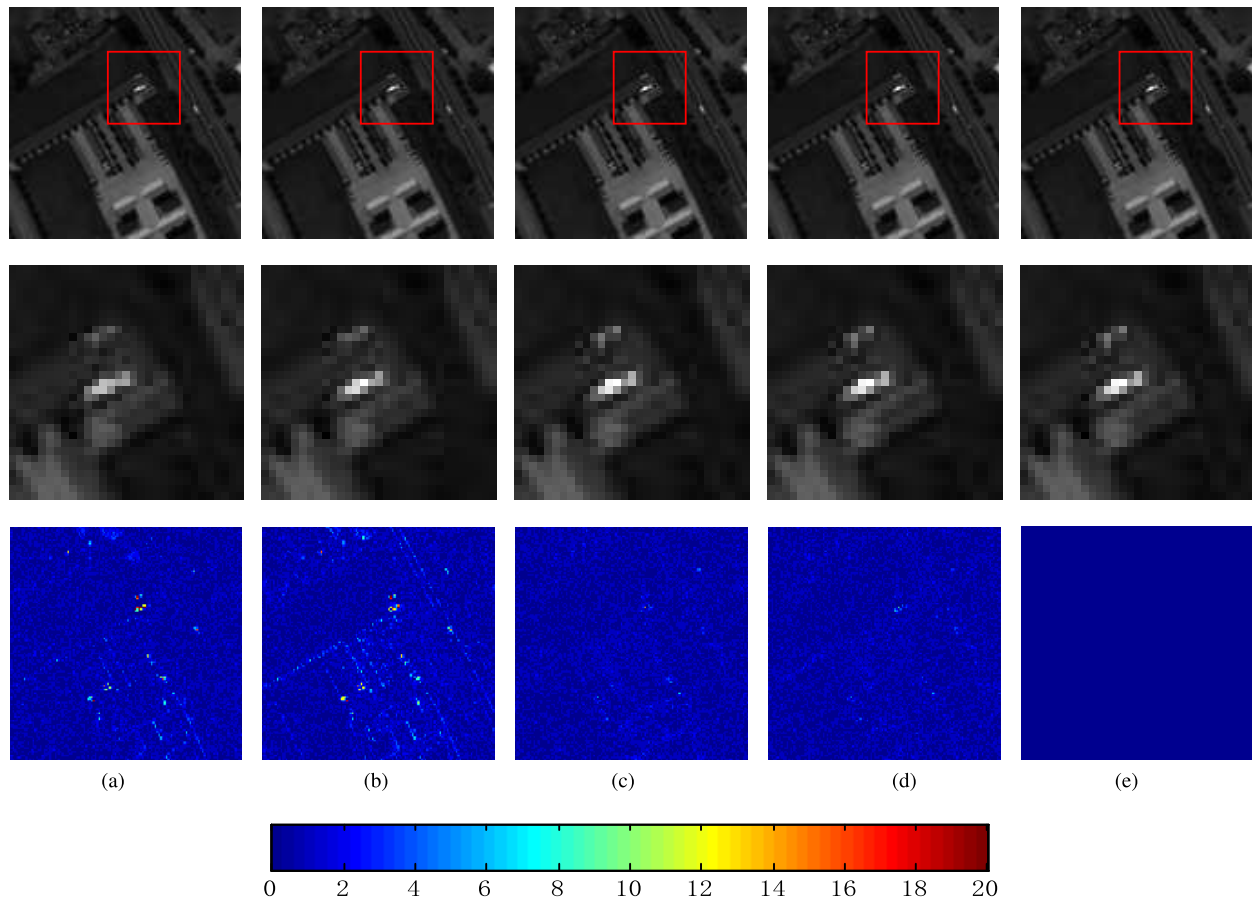


Fig. 7. First row: reconstructed images of the compared methods for Pavia University at 30th band. Second row: magnified images of the marked region. Third row: corresponding error images of the compared methods for the Pavia University at 30th band. (a) CSU [28]. (b) NSSR [33]. (c) NLSTF [35]. (d) LTTR. (e) Ground truth.

TABLE I

AVERAGE QUANTITATIVE RESULTS OF THE COMPARED METHODS ON THE CAVE DATA SET [46]

Method	CAVE database [47]			
	PSNR	SAM	UIQI	ERGAS
Best Values	$+\infty$	0	1	0
CSU [28]	41.282	7.078	0.858	1.679
NSSR [33]	43.978	5.230	0.885	1.188
NLSTF [35]	43.183	5.991	0.880	1.452
LTTR	46.161	4.324	0.902	0.925

TABLE II

QUANTITATIVE RESULTS OF THE COMPARED METHODS ON THE PAVIA UNIVERSITY [50]

Method	Pavia University [51]			
	PSNR	SAM	UIQI	ERGAS
Best Values	$+\infty$	0	1	0
CSU [28]	41.688	1.979	0.991	1.264
NSSR [33]	43.635	1.726	0.994	1.049
NLSTF [35]	44.244	1.770	0.994	1.030
LTTR	44.995	1.677	0.995	0.937

E. Experimental Results

The average quality metrics of the compared approaches on the CAVE database is shown in Table I. We highlight the best results in bold for clarity. As we can see from Table I, the proposed LTTR method performs the best among the compared approach in terms of all quality metrics, which demonstrates the effectiveness of LTTR prior. The NSSR takes the second place in terms of quality metrics, which also uses the nonlocal similarities for reconstruction. To compare the performance of the compared approaches in each spectral band, Fig. 5(a) shows the average PSNR as a function of the spectral bands over CAVE database. As we can see from Fig. 5(a), the LTTR performs constantly better than the other methods in each band. For visual comparison, we show the reconstructed images and corresponding error images for CD

(an HSI in CAVE database) at the ninth band in Fig. 6. We magnify a meaningful region of each reconstructed image for better visual comparison. As we can see from the magnified regions, the HR images recovered by CSU, NSSR, and NLSTF have obvious artifacts, and the HR images recovered by LTTR is more close to the ground truth.

The PSNR, SAM, UIQI, and ERGAS of the test methods on Pavia University are shown in Table II. As we can see from Table II, the proposed LTTR approach still outperforms other compared approaches, and the NSSR and NLSTF produce comparable results. To compare the performance of the compared approaches in different spectral bands, Fig. 5(b) shows the PSNR curves as a function of the spectral bands over Pavia University for the compared approaches. As we can see from Fig. 5(b), the LTTR delivers the best results at most of the

spectral bands among the compared approaches. The performance gain in the CAVE database is much higher than that in the Pavia University. The main reason is that CAVE database is ground-based HSI, and it has the high spatial resolution. However, Pavia University is remotely sensed HSI, its spatial resolution is much lower than that of CAVE database. There are many similar spatial structures in the high spatial resolution image. The proposed LTTR uses similar cubes to constitute a 4-D tensor and forces LTTR prior on the 4-D tensor, which can effectively exploit the nonlocal similarities. Therefore, the LTTR is more effective for HSIs with higher spatial resolution (CAVE database). This claim can be supported by the discussion of LTTR regularization parameter λ . As we can see from Fig. 4(b), the optimal λ for CAVE database (0.01) is much larger than that for Pavia University (4×10^{-4}), which demonstrates that the LTTR prior is more effective for the CAVE database. Fig. 7 shows the recovered images and corresponding error images of the compared approaches for Pavia University at 30th band. Similarly, a specific region of each recovered image is also magnified for visual comparison. As we can see from the magnified regions, all compared methods deliver clear and sharp spatial details. The HR-HSI reconstructed by CSU has a few artifacts, and the NSSR, NLSTF, and LTTR perform well in reconstructing the detailed structures.

VI. CONCLUSION

This paper presents the LTTR-based HSI super-resolution method by fusing the LR-HSI with HR-MSI of the same scene to recover the HR-HSI. In the LTTR approach, we design a novel LTTR prior to regularize the super-resolution problem. Specifically, based on the learned clusters of HR-MSI cubes, we build clusters of HR-HSI cubes with the same spatial structure. HR-HSI cubes in each cluster can constitute a 4-D tensor, and we impose the LTTR prior on these 4-D tensors, which can effectively correlations among spatial, spectral, and nonlocal modes. The super-resolution problem is solved efficiently with the ADMM. Experimental results demonstrate the superiority of the LTTR approach.

For future works, the LTTR can be improved in two aspects. First, the spectral band can also be divided into several groups to better exploit the similarities between different spectral bands, which may further promote the performance. Second, the LTTR regularization parameter λ is the same for all 4-D tensors in the LTTR, and it may be better to choose different λ for different 4-D tensors. In addition, the proposed LTTR prior may also be effective for other HSI restoration problems, such as HSI denoising.

ACKNOWLEDGMENT

The authors would like to thank the editors and anonymous reviewers for their insightful comments and suggestions, which have significantly improved this paper.

REFERENCES

- [1] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.
- [2] Y. Gao, X. Wang, Y. Cheng, and Z. J. Wang, "Dimensionality reduction for hyperspectral data based on class-aware tensor neighborhood graph and patch alignment," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 8, pp. 1582–1593, Aug. 2015.
- [3] N. Akhtar and A. Mian, "Nonparametric coupled Bayesian dictionary and classifier learning for hyperspectral classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 9, pp. 4038–4050, Sep. 2018, doi: [10.1109/TNNLS.2017.2742528](https://doi.org/10.1109/TNNLS.2017.2742528).
- [4] Q. Wang, J. Lin, and Y. Yuan, "Salient band selection for hyperspectral image classification via manifold ranking," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 6, pp. 1279–1289, Jun. 2016.
- [5] P. Zhong and R. Wang, "Jointly learning the hybrid CRF and MLR model for simultaneous denoising and classification of hyperspectral imagery," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 7, pp. 1319–1334, Jul. 2014.
- [6] L. Fang, N. He, S. Li, P. Ghamisi, and J. A. Benediktsson, "Extinction profiles fusion for hyperspectral images classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1803–1815, Mar. 2018.
- [7] L. Fang, N. He, S. Li, A. J. Plaza, and J. Plaza, "A new spatial-spectral feature extraction method for hyperspectral images using local covariance matrix representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3534–3546, Jun. 2018.
- [8] 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.
- [9] L. Loncan *et al.*, "Hyperspectral pansharpening: A review," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 3, pp. 27–46, Sep. 2015.
- [10] S. Yang, M. Wang, Z. Feng, Z. Liu, and R. Li, "Deep sparse tensor filtering network for synthetic aperture radar images classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 8, pp. 3919–3924, Aug. 2018, doi: [10.1109/TNNLS.2017.2688466](https://doi.org/10.1109/TNNLS.2017.2688466).
- [11] Z. Chen, K. Batselier, J. A. K. Suykens, and N. Wong, "Parallelized tensor train learning of polynomial classifiers," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 10, pp. 4621–4632, Oct. 2018, doi: [10.1109/TNNLS.2017.2771264](https://doi.org/10.1109/TNNLS.2017.2771264).
- [12] T.-X. Jiang, T.-Z. Huang, X.-L. Zhao, T.-Y. Ji, and L.-J. Deng, "Matrix factorization for low-rank tensor completion using framelet prior," *Inf. Sci.*, vols. 436–437, pp. 403–417, Apr. 2018.
- [13] Q. Zhao, G. Zhou, L. Zhang, A. Cichocki, and S.-I. Amari, "Bayesian robust tensor factorization for incomplete multiway data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 4, pp. 736–748, Apr. 2016.
- [14] W. Hu, D. Tao, W. Zhang, Y. Xie, and Y. Yang, "The twist tensor nuclear norm for video completion," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 12, pp. 2961–2973, Dec. 2017.
- [15] T. Xie, S. Li, L. Fang, and L. Liu, "Tensor completion via nonlocal low-rank regularization," *IEEE Trans. Cybern.*, to be published, doi: [10.1109/TCYB.2018.2825598](https://doi.org/10.1109/TCYB.2018.2825598).
- [16] Y. Liu, F. Shang, W. Fan, J. Cheng, and H. Cheng, "Generalized higher order orthogonal iteration for tensor learning and decomposition," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 12, pp. 2551–2563, Dec. 2016.
- [17] X. Chen *et al.*, "A generalized model for robust tensor factorization with noise modeling by mixture of Gaussians," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 11, pp. 5380–5393, Nov. 2018, doi: [10.1109/TNNLS.2018.2796606](https://doi.org/10.1109/TNNLS.2018.2796606).
- [18] I. V. Oseledets, "Tensor-train decomposition," *SIAM J. Sci. Comput.*, vol. 33, no. 5, pp. 2295–2317, Sep. 2011.
- [19] J. A. Bengua, H. N. Phien, H. D. Tuan, and M. N. Do, "Efficient tensor completion for color image and video recovery: Low-rank tensor train," *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2466–2479, May 2017.
- [20] Q. Shi, H. Lu, and Y.-M. Cheung, "Rank-one matrix completion with automatic rank estimation via L1-norm regularization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 10, pp. 4744–4757, Oct. 2018.
- [21] S. Yang, K. Zhang, and M. Wang, "Learning low-rank decomposition for pan-sharpening with spatial-spectral offsets," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 8, pp. 3647–3657, Aug. 2018.
- [22] Y. Wang, L. Wu, X. Lin, and J. Gao, "Multiview spectral clustering via structured low-rank matrix factorization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 10, pp. 4833–4843, Oct. 2018.
- [23] P. Zhou, Z. Lin, and C. Zhang, "Integrated low-rank-based discriminative feature learning for recognition," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 5, pp. 1080–1093, May 2016.
- [24] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2009, pp. 2272–2279.

- [25] 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, p. 122, 2011.
- [26] R. Kawakami, Y. Matsushita, J. Wright, M. Ben-Ezra, Y. Tai, and K. Ikeuchi, "High-resolution hyperspectral imaging via matrix factorization," in *Proc. IEEE CVPR*, Jun. 2011, pp. 2329–2336.
- [27] 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.
- [28] C. Lanaras, E. Baltsavias, and K. Schindler, "Hyperspectral super-resolution by coupled spectral unmixing," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 3586–3594.
- [29] Q. Wei, J. Bioucas-Dias, N. Dobigeon, J.-Y. Tourneret, M. Chen, and S. Godsill, "Multiband image fusion based on spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7236–7249, Dec. 2016.
- [30] N. Akhtar, F. Shafait, and A. Mian, "Sparse spatio-spectral representation for hyperspectral image super-resolution," in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2014, pp. 63–78.
- [31] 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.
- [32] 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.
- [33] 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.
- [34] 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.
- [35] R. Dian, L. Fang, and S. Li, "Hyperspectral image super-resolution via non-local sparse tensor factorization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jul. 2017, pp. 3862–3871.
- [36] 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.
- [37] K. Zhang, M. Wang, S. Yang, and L. Jiao, "Spatial-Spectral-graph-regularized low-rank tensor decomposition for multispectral and hyperspectral image fusion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1030–1040, Apr. 2018, doi: 10.1109/JSTARS.2017.2785411.
- [38] Y. Qu, H. Qi, and C. Kwan, "Unsupervised sparse dirichlet-net for hyperspectral image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2018, pp. 2511–2520.
- [39] L. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, Sep. 1966.
- [40] A. Cichocki, N. Lee, I. Oseledets, A.-H. Phan, Q. Zhao, and D. P. Mandic, "Tensor networks for dimensionality reduction and large-scale optimization: Part 1 low-rank tensor decompositions," *Found. Trends Mach. Learn.*, vol. 9, nos. 4–5, p. 249, 2016.
- [41] Q. Xie *et al.*, "Multispectral images denoising by intrinsic tensor sparsity regularization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2016, pp. 1692–1700.
- [42] K. Zhang, M. Wang, and S. Yang, "Multispectral and hyperspectral image fusion based on group spectral embedding and low-rank factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 3, pp. 1363–1371, Mar. 2017.
- [43] D. Arthur and S. Vassilvitskii, "K-means++: The advantages of careful seeding," in *Proc. Annu. ACM-SIAM Symp. Discrete Algorithms*, 2007, pp. 1027–1035.
- [44] R. H. Bartels and G. W. Stewart, "Solution of the matrix equation $AX + XB = C$ [F4]," *Commun. ACM*, vol. 15, no. 9, pp. 820–826, Sep. 1972.
- [45] Q. Wei, N. Dobigeon, and J.-Y. 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.
- [46] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, "Generalized assorted pixel camera: Postcapture control of resolution, dynamic range, and spectrum," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2241–2253, Sep. 2010.
- [47] Y. Wang, W. Yin, and J. Zeng. (2015). "Global convergence of ADMM in nonconvex nonsmooth optimization." [Online]. Available: <https://arxiv.org/abs/1511.06324>
- [48] H. Kwon and Y.-W. Tai, "RGB-guided hyperspectral image upsampling," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 307–315.
- [49] N. Akhtar, F. Shafait, and A. Mian, "Hierarchical beta process with Gaussian process prior for hyperspectral image super resolution," in *Proc. Eur. Conf. Comput. Vis.*, Oct. 2016, pp. 103–120.
- [50] F. Dell'Acqua, P. Gamba, A. Ferrari, J. A. Palmason, J. A. Benediktsson, and K. Arnason, "Exploiting spectral and spatial information in hyperspectral urban data with high resolution," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 322–326, Oct. 2004.
- [51] R. H. Yuhas, A. F. H. Goetz, and J. W. Boardman, "Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm," in *Proc. 3rd Annu. JPL Airborne Geosci. Workshop*, vol. 1, 1992, pp. 147–149.
- [52] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.
- [53] L. Wald, "Quality of high resolution synthesised images: Is there a simple criterion?" in *Proc. Int. Conf. Fusion Earth Data*, Jan. 2000, pp. 99–103.



Renwei Dian (S'16) received the B.S. degree from the Wuhan University of Science and Technology, Wuhan, China, in 2015. He is currently pursuing the Ph.D. degree with the Laboratory of Vision and Image Processing, Hunan University, Changsha, China.

Since 2017, he has been a Visiting Ph.D. Student with the University of Lisbon, Lisbon, Portugal, supported by the China Scholarship Council. His current research interests include hyperspectral image super-resolution, image fusion, tensor decomposition, and deep learning.

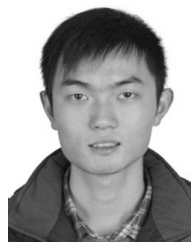


Shutao Li (M'07–SM'15–F'19) received the B.S., M.S., and Ph.D. degrees from Hunan University, Changsha, China, in 1995, 1997, and 2001, respectively.

In 2001, he joined the College of Electrical and Information Engineering, Hunan University, where he is currently a Full Professor. From 2002 to 2003, he was a Post-Doctoral Fellow with Prof. John Shawe-Taylor with Royal Holloway College, University of London, London, U.K. In 2005, he joined the Department of Computer Science, Hong Kong

University of Science and Technology, Hong Kong, as a Visiting Professor, where he became a Research Associate in 2011. He has authored or co-authored over 160 refereed papers. His current research interests include compressive sensing, sparse representation, image processing, and pattern recognition.

Dr. Li is currently an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and the IEEE textscTransactions on Instrumentation and Measurement. He is currently a member of the Editorial Board of the Information Fusion and the Sensing and Imaging. He was a recipient of two 2nd-Grade National Awards at the Science and Technology Progress of China in 2004 and 2006.



Leyuan Fang (S'10–M'14–SM'17) received the B.S. and Ph.D. degrees from the College of Electrical and Information Engineering, Hunan University, Changsha, China, in 2008 and 2015, respectively.

From 2011 to 2012, he was a Visiting Ph.D. Student with the Department of Ophthalmology, Duke University, Durham, NC, USA, supported by the China Scholarship Council. Since 2017, he has been an Associate Professor with the College of Electrical and Information Engineering, Hunan University. His current research interests include sparse representation and multiresolution analysis in remote sensing and medical image processing.

Dr. Fang was a recipient of the Scholarship Award for Excellent Doctoral Student granted by the Chinese Ministry of Education in 2011.