

# Hyperspectral Image Super-Resolution via Subspace-Based Low Tensor Multi-Rank Regularization

Renwei Dian<sup>ID</sup>, *Student Member, IEEE*, and Shutao Li<sup>ID</sup>, *Fellow, IEEE*

**Abstract**—Recently, combining a low spatial resolution hyperspectral image (LR-HSI) with a high spatial resolution multispectral image (HR-MSI) into an HR-HSI has become a popular scheme to enhance the spatial resolution of HSI. We propose a novel subspace-based low tensor multi-rank regularization method for the fusion, which fully exploits the spectral correlations and non-local similarities in the HR-HSI. To make use of high spectral correlations, the HR-HSI is approximated by spectral subspace and coefficients. We first learn the spectral subspace from the LR-HSI via singular value decomposition, and then estimate the coefficients via the low tensor multi-rank prior. More specifically, based on the learned cluster structure in the HR-MSI, the patches in coefficients are grouped. We collect the coefficients in the same cluster into a three-dimensional tensor and impose the low tensor multi-rank prior on these collected tensors, which fully model the non-local self-similarities in the HR-HSI. The coefficients optimization is solved by the alternating direction method of multipliers. Experiments on two public HSI datasets demonstrate the advantages of our method.

**Index Terms**—Super-resolution, hyperspectral imaging, low tensor multi-rank, image fusion.

## I. INTRODUCTION

**H**YPERSPECTRAL images (HSIs) with hundreds of spectral bands contain abundant spectral information, which contributes the accurate identification of the objects. Therefore, HSIs have shown improved performance in many computer vision [1], [2] and earth remote sensing applications [3]–[8]. However, there is the inevitable trade-off between the spatial resolution and spectral resolution due to the limited sun irradiance [9]. Hence, HSIs often suffer from low spatial resolution. In contrast, compared with HSIs,

multispectral images (MSIs) is acquired with lower spectral resolution and higher spatial resolution. Therefore, MSI and HSI fusion has become a popular way to improve the spatial resolution of the HSI.

HSIs can be naturally represented by a three-dimensional tensor, and therefore it is meaningful to analyze HSIs from tensor perspective. Recently, reference [10] proposes a novel factorization for three-dimensional data, called as the tensor singular value decomposition (t-SVD). The t-SVD is based on a novel tensor product scheme, and it has optimal low-rank approximation (measured by the Frobenius norm) of a tensor [11]. Based on t-SVD, a novel tensor rank, referred to as tensor multi-rank, is in [10] and [12] to measure information complexity of three-dimensional data, which has shown superior performance in solving tensor recovery problems [12]–[15].

We propose a subspace based low tensor multi-rank (LTMR) regularization method for HSI super-resolution method by the HR-MSI and LR-HSI fusion. The LTMR mainly exploits two prior information of HR-HSI: high correlations among spectral bands and non-local self-similarities. The HR-HSI bands are highly correlated, and spectral vectors often live in a low-dimensional subspace. Therefore, we factorize the target HR-HSI as subspace multiplied by coefficients. Since LR-HSI preserves most of the spectral information, the spectral subspace is learned from it via SVD. Besides, the HR-HSI is often self-similar and therefore it has many similar patches. We firstly partition the HR-MSI as many full band patches (FBPs), and then similar FBPs are clustered together. The coefficients are also grouped based on the learned cluster from HR-MSI. We collect coefficients in the same cluster as the three-dimensional tensor. Since the three-dimensional tensor is constituted by similar coefficients, we impose the low tensor multi-rank prior on these tensors to make use of the non-local similarities. The optimization problem for coefficients estimation is solved via the alternating direction method of multipliers (ADMM) [16].

We organize the remaining parts of this paper as follows. In Section II, we review the representative HSI super-resolution literature. Section III gives notations on tensor. The proposed LTMR approach is presented in Section IV. In Section V, experimental results on simulated and real datasets are presented. Section VI gives the conclusion.

Manuscript received September 13, 2018; revised March 4, 2019; accepted April 9, 2019. Date of publication May 20, 2019; date of current version August 14, 2019. This work was supported in part by the National Natural Science Fund of China under Grant 61890962 and Grant 61520106001, in part by the Science and Technology Plan Project Fund of Hunan Province under Grant CX2018B171, Grant 2017RS3024, and Grant 2018TP1013, in part by the Science and Technology Talents Program of Hunan Association for Science and Technology under Grant 2017TJ-Q09, in part by the Hunan Provincial Innovation Foundation for Postgraduate, and in part by the China Scholarship Council. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Xin Li. (*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).

Digital Object Identifier 10.1109/TIP.2019.2916734

1057-7149 © 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.  
See [http://www.ieee.org/publications\\_standards/publications/rights/index.html](http://www.ieee.org/publications_standards/publications/rights/index.html) for more information.

## II. RELATED WORKS

Recently, HSI and MSI fusion based HSI spatial resolution improvement has received more and more attention. Recent HSI super-resolution approaches can be roughly categorized as three classes: matrix factorization based approaches [17]–[31], tensor factorization based approaches [32]–[37] and non-factorization based approaches [38]–[41].

Matrix factorization based HSI super-resolution approaches assume that each spectral pixel can be represented a linear combination of several spectral atoms, they factorize the target HR-HSI as the spectral basis and the coefficients. Kawakami *et al.* [17] solves the fusion problem with a sparse prior, which assumes the coefficients are sparse. Furthermore, references [18]–[21] solve the fusion problem by coupled non-negative matrix factorization scheme, which alternately updates spectral basis and coefficients for several iterations. Works [22]–[27] also make use of the spatial prior of images to solve the fusion problem. Zhou *et al.* [22] and Veganzones *et al.* [23] argue that HSI is locally low rank, and they learn the spectral basis for each local patch and solve the fusion problem in a patch-by-patch manner. Simoões *et al.* [24] use the subspace representation and exploit the total variation to obtain the spatial smoothness. Wei *et al.* [25] designs a sparsity regularizer to promote spatial similarities. Dong *et al.* [26] utilize the non-local sparse representation method to promote the non-local self-similarities of the HR-HSI, which achieves good fusion results. Han *et al.* [27] use a self-similarity constrained sparse representation to model the local and global similarities in the HR-HSI. References [26] and [27] also utilize the non-local similarities for the fusion, however our method differs from them in two aspects. Firstly, our method is based on low-dimensional subspace representation to model the spectral correlations. What is more, we design a tensor multi-rank based regularization term to utilize the non-local similarities.

Tensor factorization based approaches are another class of HSI super-resolution methods. Work [32] firstly proposes a non-local sparse tensor factorization method for HSI super-resolution, where they approximate the HR-HSI by dictionaries of three modes and core tensor. Besides, the similar FBPs in the HR-HSI are sparsely decomposed on the same dictionaries to better model the non-local similarities. Later, Li *et al.* [33] propose the coupled sparse tensor factorization algorithm to solve the fusion problem, and conduct sparse tensor factorization for HSI and MSI, simultaneously. In addition, Zhang *et al.* [34] propose a graph-regularized low-rank Tucker decomposition approach for HSI super-resolution, which promotes the spatial consistency and spectral smoothness from the HR-MSI and LR-HSI, respectively. Chang *et al.* [35] propose weighted low-rank tensor recovery method to solve the fusion problem, which gives the core tensor values different sparsity regularization parameters. Dian *et al.* [36] use the newly proposed tensor-train rank to regularize the fusion problem, which imposes low tensor-train rank constraint on the grouped 4-D tensors. Based on the

t-SVD, Xu *et al.* [37] propose a non-local tensor sparse representation model for the fusion, and the main difference between [32] and [37] is that they use different tensor sparse representation model.

Rather than decomposing the target HR-HSI as several factors, the non-factorization based fusion methods directly estimate the HR-HSI with appropriate priors. Considering the common imaging mechanism of commercial RGB cameras Fu *et al.* [38] proposes to combine an LR-HSI with a mosaic RGB based on non-local low-rank regularization. Zhang *et al.* [39] exploit the underlying manifold structure learned from HR-MSI to solve the fusion problem. Instead of imposing specific priors of HR-MSI, Qu *et al.* [40] design an unsupervised deep CNN for the fusion, which promotes the representations from two imaging sensors to admit a sparse Dirichlet distribution.

## III. PRELIMINARIES AND NOTATIONS

We denote an  $N$ -dimensional tensor as  $\mathcal{M} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ . The matrix  $\mathbf{M}_{(n)} \in \mathbb{R}^{I_n \times I_1 I_2 \dots I_{n-1} I_{n+1} \dots I_N}$  is mode- $n$  matricization of tensor  $\mathcal{M}$ , and its columns are  $n$ -mode vectors. The  $n$ -mode product of the matrix  $\mathbf{E} \in \mathbb{R}^{J_n \times I_n}$  and tensor  $\mathcal{N} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$  can be represented as

$$\mathcal{N} = \mathcal{M} \times_n \mathbf{E} \quad (1)$$

where  $\mathcal{N} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times J_n \times \dots \times I_N}$ , and its elements are calculated as

$$\mathcal{N}_{i_1 \dots i_{n-1} j_n i_{n+1} \dots i_N} = \sum_{i_n} \mathbf{E}_{j_n i_n} \mathcal{M}_{i_1 \dots i_{n-1} i_n i_{n+1} \dots i_N}. \quad (2)$$

The equation (1) is equivalent as matrix multiplication  $\mathbf{N}_{(n)} = \mathbf{E} \mathbf{M}_{(n)}$ .

For a three-dimensional tensor  $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ , we use  $\mathcal{A}(i, :, :)$ ,  $\mathcal{A}(:, i, :)$  and  $\mathcal{A}(:, :, i)$  to denote its  $i$ -th horizontal, lateral and frontal slice, respectively. We use  $\mathbf{A}^{(i)}$  to denote  $i$ -th frontal slice of tensor  $\mathcal{A}$ , i.e.,  $\mathbf{A}^{(i)} = \mathcal{A}(:, :, i)$ . The product of FFT on  $\mathcal{A}$  along the 3-rd mode is denoted by  $\tilde{\mathcal{A}} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ . We use the Matlab command `fft` to represent the FFT, and therefore we can obtain  $\tilde{\mathcal{A}} = \text{fft}(\mathcal{A}, [], 3)$ . Besides,  $\mathcal{A}$  can be computed by  $\tilde{\mathcal{A}}$  by using the inverse FFT, that is,  $\mathcal{A} = \text{ifft}(\tilde{\mathcal{A}}, [], 3)$ . The block diagonal matrix  $\text{bdiag}(\mathcal{A})$  for tensor  $\mathcal{A}$  is represented as

$$\text{bdiag}(\mathcal{A}) = \begin{bmatrix} \tilde{\mathbf{A}}^{(1)} & & & \\ & \tilde{\mathbf{A}}^{(2)} & & \\ & & \ddots & \\ & & & \tilde{\mathbf{A}}^{(I_3)} \end{bmatrix} \quad (3)$$

For tensor  $\mathcal{A}$ , its block circulant matrix  $\text{bcirc}(\mathcal{A}) \in \mathbb{R}^{I_1 I_3 \times I_2 I_3}$  is defined as

$$\text{bcirc}(\mathcal{A}) = \begin{bmatrix} \mathbf{A}^{(1)} & \mathbf{A}^{(I_3)} & \dots & \mathbf{A}^{(2)} \\ \mathbf{A}^{(2)} & \mathbf{A}^{(1)} & \dots & \mathbf{A}^{(3)} \\ \vdots & & \ddots & \vdots \\ \mathbf{A}^{(I_3)} & \mathbf{A}^{(I_3-1)} & \dots & \mathbf{A}^{(1)} \end{bmatrix} \quad (4)$$

We can diagonalize the block circulant matrix by the FFT, i.e.,

$$\text{bdiag}(\mathcal{A}) = (\mathbf{F} \otimes \mathbf{I}_{I_1}) \text{bcirc}(\mathcal{A}) (\mathbf{F}^{-1} \otimes \mathbf{I}_{I_2}) \quad (5)$$

where  $\mathbf{I}_{I_1}$  denotes identity matrix of size  $I_1 \times I_1$ . Based on equation (5), we can draw a conclusion that the rank of matrix  $\text{bdiag}(\mathcal{A})$  is equivalent to that of  $\text{bcirc}(\mathcal{A})$ . The matrix  $\text{bcirc}(\mathcal{A})$  is a new way of matricization of  $\mathcal{A}$ , and the frontal slices of  $\mathcal{A}$  are arranged in a circulant way, and therefore it can effectively model relationships among frontal slices. Therefore, the rank of  $\text{bcirc}(\mathcal{A})$  can measure the complexities of frontal slices and correlations between the frontal slices.

For  $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ , we define  $\text{Unfold}_F$  function as

$$\text{Unfold}_F(\mathcal{A}) = \begin{bmatrix} \mathbf{A}^{(1)} \\ \mathbf{A}^{(2)} \\ \vdots \\ \mathbf{A}^{(I_3)} \end{bmatrix} \quad (6)$$

Conversely, the matrix  $\text{Unfold}_F(\mathcal{A})$  can be transformed back to the tensor  $\mathcal{A}$  by  $\mathcal{A} = \text{fold}_F(\text{Unfold}_F(\mathcal{A}))$ .

Before the introduction of t-SVD, we first give some preliminaries. Let  $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$  and  $\mathcal{B} \in \mathbb{R}^{I_2 \times I \times I_3}$ , the t-product [10]  $\mathcal{A} * \mathcal{B}$  is defined to be a tensor of size  $I_1 \times I \times I_3$ ,

$$\mathcal{A} * \mathcal{B} = \text{Fold}_F(\text{bcirc}(\mathcal{A}) * \text{Unfold}_F(\mathcal{B})), \quad (7)$$

The conjugate transpose of a tensor  $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$  is  $\mathcal{A}^T \in \mathbb{R}^{I_2 \times I_1 \times I_3}$ , defined in [10]. The tensor [10]  $\mathcal{I} \in \mathbb{R}^{I_1 \times I_1 \times I_3}$  is identity tensor, and first frontal slice and other frontal slices of  $\mathcal{I}$  are identity matrix and zero, respectively. A tensor is called f-diagonal [10] if each frontal slice of the tensor is a diagonal matrix. A tensor  $\mathcal{U} \in \mathbb{R}^{I_1 \times I_1 \times I_3}$  is orthogonal if it stratifies

$$\mathcal{U} * \mathcal{U}^T = \mathcal{U}^T * \mathcal{U} = \mathcal{I}, \quad (8)$$

**Definition 3.1 (t-SVD [10]):** For  $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ , the T-SVD factors the tensor  $\mathcal{A}$  as

$$\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T, \quad (9)$$

where  $\mathcal{U} \in \mathbb{R}^{I_1 \times I_1 \times I_3}$ ,  $\mathcal{V} \in \mathbb{R}^{I_2 \times I_2 \times I_3}$  are orthogonal, and  $\mathcal{S} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$  is an f-diagonal tensor.

**Definition 3.2 (Tensor multi-rank [12]):** For tensor  $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ , the tensor multi-rank is a vector  $\mathbf{p} \in \mathbb{R}^{I_3}$ , where its  $i$ -th element is equal to the  $i$ -th frontal slice of tensor  $\tilde{\mathcal{A}}$ , i.e.,  $\mathbf{p}_i = \text{rank}(\tilde{\mathcal{A}}^{(i)})$ .

Based on the definition of tensor multi-rank,  $\ell_1$  norm of tensor multi-rank is equal to the rank of matrices  $\text{bdiag}(\mathcal{A})$  and  $\text{bcirc}(\mathcal{A})$ . As in the matrix case, we need convex relaxations for the rank measure. Therefore, For tensor multi-rank, we have the following convex relaxation

**Definition 3.3 (Tensor nuclear norm [12]):** For tensor  $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ , the tensor nuclear norm  $\|\mathcal{A}\|_{\text{TN}}$  is defined as the sum of the singular values of all the frontal slices of  $\tilde{\mathcal{A}}$ , i.e.,

$$\begin{aligned} \|\mathcal{A}\|_{\text{TN}} &= \sum_{i=1}^{I_3} \sum_j \sigma_j(\tilde{\mathcal{A}}^{(i)}) \\ &= \sum_{i=1}^{I_3} \sum_j \tilde{S}(i, i, j) \\ &= \|\text{bcirc}(\mathcal{A})\|_* \\ &= \|\text{bdiag}(\mathcal{A})\|_* \end{aligned} \quad (10)$$

where  $\sigma_j(\tilde{\mathcal{A}}^{(i)})$  denotes the  $j$ -th singular value of  $\tilde{\mathcal{A}}^{(i)}$ , and  $\|\cdot\|_*$  denotes the nuclear norm of matrix.  $\tilde{S} = \text{fft}(S)$ , where  $S$  is obtained by t-SVD, i.e.,  $\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T$ .

Compared with the sum of singular values, the log-sum (LS) of singular values has shown improved performance in approximating the rank of matrix [42]. For a matrix  $\mathbf{A}$ , the log-sum (LS) of its singular values is defined as

$$\text{LS}(\mathbf{A}) = \sum_i \log(\sigma_i(\mathbf{A}) + \varepsilon) \quad (11)$$

where  $\varepsilon > 0$  is a very small positive number. Replacing the sum of singular values as log-sum of singular values, we redefine the tensor multi-rank as

$$\|\mathcal{A}\|_{\text{TMR}} = \frac{1}{I_3} \sum_{i=1}^{I_3} \text{LS}(\tilde{\mathcal{A}}^{(i)}), \quad (12)$$

where  $I_3$  is the dimension of the third mode of  $\mathcal{A}$ , and the scalar is convenient for the optimization (31). In this paper, we use the proposed  $\|\mathcal{A}\|_{\text{TMR}}$  to depict the low-rank structure of a tensor.

#### IV. PROPOSED LTMR APPROACH

We present the proposed LTMR approach for HSI super-resolution in this section.

##### A. Observation Model

We denote the target HR-HSI as  $\mathcal{Z} \in \mathbb{R}^{W \times H \times S}$ , which has  $W \times H$  pixels and  $S$  bands.

$\mathcal{X} \in \mathbb{R}^{w \times h \times S}$  denotes the observed LR-HSI, which has  $S$  bands and  $w \times h$  pixels.  $\mathcal{X}$  can be seen as the spatially downsampled version of  $\mathcal{Z}$ , i.e.,

$$\mathbf{X}_{(3)} = \mathbf{Z}_{(3)} \mathbf{B} \mathbf{S}, \quad (13)$$

where  $\mathbf{Z}_{(3)} \in \mathbb{R}^{S \times WH}$  and  $\mathbf{X}_{(3)} \in \mathbb{R}^{S \times wh}$  are obtained by unfolding  $\mathcal{Z}$  and  $\mathcal{X}$  with the third mode, respectively. The matrix  $\mathbf{B} \in \mathbb{R}^{WH \times WH}$  represents the convolution blur, and it can be diagonalized, i.e.,

$$\mathbf{B} = \mathbf{F} \mathbf{K} \mathbf{F}^{-1}, \quad (14)$$

where  $\mathbf{F}$  represents fast Fourier transform (FFT), and  $\mathbf{F}^{-1}$  denotes the inverse FFT. The diagonal matrix  $\mathbf{K}$  holds eigenvalues of  $\mathbf{B}$ .  $\mathbf{S} \in \mathbb{R}^{WH \times wh}$  represents the downsampling matrix.

$\mathcal{Y} \in \mathbb{R}^{W \times H \times s}$  represents the HR-MSI of the same scene, which has  $s$  spectral bands and  $W \times H$  pixels.  $\mathcal{Y}$  can be modelled as spectrally downsampled version of  $\mathcal{Z}$ , i.e.,

$$\mathbf{Y}_{(3)} = \mathbf{R} \mathbf{Z}_{(3)}, \quad (15)$$

where  $\mathbf{Y}_{(3)} \in \mathbb{R}^{s \times WH}$  denotes the matrix acquired by unfolding  $\mathcal{Y}$  with third mode, and  $\mathbf{R} \in \mathbb{R}^{s \times S}$  represents the spectral response matrix.

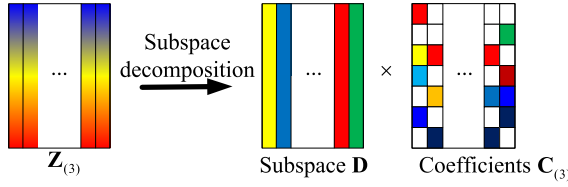


Fig. 1. The subspace decomposition of the HR-HSI.

### B. Subspace Learning

HSI usually have the large correlation among bands, and spectral vectors often live in a low-dimension subspace [24], [43]. Therefore, the HR-HSI can be represented as

$$\mathcal{Z} = \mathcal{C} \times_3 \mathbf{D}, \quad (16)$$

where  $\mathbf{D} \in \mathbb{R}^{S \times L}$  is the subspace. The tensor  $\mathcal{C} \in \mathbb{R}^{W \times H \times L}$  is the coefficients, where  $\mathcal{C}(i, j, :)$  denotes the coefficient of the spectral pixel  $\mathcal{Z}(i, j, :)$ . As illustrated as in Fig. 1, equation (16) is equivalent as the following representation

$$\mathbf{Z}_{(3)} = \mathbf{D} \mathbf{C}_{(3)}, \quad (17)$$

where  $\mathbf{C}_{(3)} \in \mathbb{R}^{L \times WH}$  is obtained by unfolding  $\mathcal{C}$  along the third mode. Regarding to the subspace  $\mathbf{D}$ , we make two remarks: a) the values of  $L$  is small, i.e.,  $L < S$  means that spectral vectors live in a low-dimensional subspace, where the dimensionality reduction makes computationally efficient; b) the columns of the subspace are orthonormal, which makes the each spectral vector  $\mathcal{Z}(i, j, :)$  has the same Frobenius norm with corresponding coefficient  $\mathcal{C}(i, j, :)$ . In this way, the self-similarities in the image domain can be transformed to the subspace coefficients domain.

In this way, the fusion problem is reformulated as the estimation of subspace  $\mathbf{D}$  and coefficients  $\mathcal{C}$  from the HR-MSI and LR-HSI. The LR-HSI preserves the most of spectral information of the HR-HSI, and therefore LR-HSI and HR-HSI are assumed to live in the same spectral subspace. Hence, we estimate the low-dimensional spectral subspace from the LR-HSI via singular value decomposition (SVD), that is,

$$\mathbf{X}_{(3)} = \mathbf{U}_1 \Sigma_1 \mathbf{V}_1^T \quad (18)$$

Here,  $\mathbf{U}_1$  and  $\mathbf{V}_1$  are column-orthogonal matrices, and diagonal matrix  $\Sigma_1$  holds the singular values, which are arranged in non-increasing order. By reserving  $L$  largest singular and discarding the rest singular values from  $\Sigma_1$ , a low-dimensional approximation of  $\mathbf{X}_{(3)}$  is given by

$$\hat{\mathbf{X}}_{(3)} = \hat{\mathbf{U}}_1 \hat{\Sigma}_1 \hat{\mathbf{V}}_1^T \quad (19)$$

where  $\hat{\mathbf{U}}_1 = \mathbf{U}_1(:, 1:L)$  and  $\hat{\mathbf{V}} = \mathbf{V}(:, 1:L)$ . The learned subspace  $\mathbf{D}$  is equivalent as

$$\mathbf{D} = \hat{\mathbf{U}}_1 = \mathbf{U}_1(:, 1:L). \quad (20)$$

### C. Coefficients Estimation

Based on the subspace representation model (16), the target of the fusion is estimating the subspace  $\mathbf{D}$  and coefficients  $\mathcal{C}$  from  $\mathcal{X}$  and  $\mathcal{Y}$ . With the subspace  $\mathbf{D}$  known and combining

equations (13), (15), and (17), the coefficients  $\mathbf{C}_{(3)}$  is estimated by solving the following problem:

$$\min_{\mathbf{C}_{(3)}} \|\mathbf{X}_{(3)} - \mathbf{D} \mathbf{C}_{(3)} \mathbf{B} \mathbf{S}\|_F^2 + \|\mathbf{Y}_{(3)} - \mathbf{R} \mathbf{D} \mathbf{C}_{(3)}\|_F^2 + \phi(\mathcal{C}), \quad (21)$$

where  $\phi(\mathcal{C})$  stands for the regularization term with regard to coefficients  $\mathcal{C}$ . Note that, both  $\mathbf{C}_{(3)}$  and  $\mathcal{C}$  denote coefficients, and they are matrix representation and tensor representation of coefficients respectively. Both of the HR-MSI and LR-HSI are downsampled versions of HR-HSI, and therefore the solution of  $\mathcal{C}$  unique. In this way, some prior information of  $\mathcal{C}$  is needed to solve the problem.

We design a low tensor multi-rank prior to regularize the estimation of  $\mathcal{C}$ . The HR-HSIs have are many redundant and similar spatial-spectral structures. More specifically, there are many similar FBPs across the spatial domain. The HR-MSI preserves most of the spatial information, and therefore we can learn the non-local similarities from it. The HR-MSI is firstly spatially partitioned into several FBPs with the spatial size of  $\sqrt{q} \times \sqrt{q}$  and spectral size  $s$ , and then we cluster FBPs into  $K$  groups  $\mathcal{Y}^{(k)} = \{\mathcal{Y}^{(k,j)}\}_{j=1}^{N_k}$ ,  $k = 1, 2, \dots, K$ , where  $N_k$  denotes the number of FBPs in the  $k$ -th group. The efficient K-means++ method [44] is used for the clustering process. Since the columns of the subspace are orthonormal, which makes each spectral vector  $\mathcal{Z}(i, j, :)$  has the same Frobenius norm with corresponding coefficient  $\mathcal{C}(i, j, :)$ . In this way, the self-similarities in the image domain can be transformed into the subspace coefficients domain. We obtain the clusters of coefficients FBPs based on the learned cluster structure, that is,  $\mathcal{C}^{(k)} = \{\mathcal{C}^{(k,j)} \in \mathbb{R}^{\sqrt{q} \times \sqrt{q} \times L}\}_{j=1}^{N_k}$ , where the location of  $\mathcal{C}^{(k,j)}$  is the same as  $\mathcal{Y}^{(k,j)}$ . We collect the coefficients in  $k$ -th group  $\{\mathcal{C}^{(k,j)} \in \mathbb{R}^{\sqrt{q} \times \sqrt{q} \times L}\}_{j=1}^{N_k}$  into a new tensor  $\mathcal{C}^k \in \mathbb{R}^{N_k \times L \times q}$ . The tensor  $\mathcal{C}^k \in \mathbb{R}^{N_k \times L \times q}$  is constituted by similar FBPs, and hence it has low tensor multi-rank. To impose the low tensor multi-rank on the grouped coefficients, the coefficients are estimated by solving the following problem

$$\min_{\mathbf{C}_{(3)}} \|\mathbf{X}_{(3)} - \mathbf{D} \mathbf{C}_{(3)} \mathbf{B} \mathbf{S}\|_F^2 + \|\mathbf{Y}_{(3)} - \mathbf{R} \mathbf{D} \mathbf{C}_{(3)}\|_F^2 + \lambda \sum_{k=1}^K \|\mathcal{C}^k\|_{\text{TMR}}, \quad (22)$$

where  $\|\mathcal{C}^k\|_{\text{TMR}}$ , defined in (12), is the tensor multi-rank regularization term, and  $\lambda$  is the regularization parameter.

We solve the variational optimization problem (22) with the framework of the ADMM [16]. By introducing the variable  $\mathcal{V} = \mathcal{C}$ , the following augmented Lagrangian function is achieved:

$$L(\mathcal{C}, \mathcal{V}, \mathcal{G}) = \|\mathbf{X}_{(3)} - \mathbf{D} \mathbf{C}_{(3)} \mathbf{B} \mathbf{S}\|_F^2 + \|\mathbf{Y}_{(3)} - \mathbf{R} \mathbf{D} \mathbf{C}_{(3)}\|_F^2 + \mu \|\mathcal{V} - \mathcal{C}\|_F^2 + \frac{\mathcal{G}}{2\mu} \|\mathcal{V}\|_F^2 + \lambda \sum_{k=1}^K \|\mathcal{V}^k\|_{\text{TMR}} \quad (23)$$

where  $\mu$  is penalty parameter, and  $\mathcal{G}$  is the Lagrangian multiplier. We can decompose the minimization of function (23)



**Algorithm 1** A Closed-Form Solution of (25) for  $\mathbf{C}_{(3)}$ **Input:**  $\mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3, \mathbf{B}, \mathbf{S}$ 

- 1: Eigen-decomposition of  $\mathbf{B}$ :  $\mathbf{B} = \mathbf{F}\mathbf{K}\mathbf{F}^H$ ;
  - 2:  $\tilde{\mathbf{K}} = \mathbf{K}(\mathbf{1}_d \otimes \mathbf{1}_{wh})$ ;
  - 3: Eigen-decomposition of  $\mathbf{H}_1$ :  $\mathbf{H}_1 = \mathbf{Q}_1\mathbf{\Lambda}\mathbf{Q}_1^{-1}$ ;
  - 4:  $\tilde{\mathbf{H}}_3 = \mathbf{Q}_1^{-1}\mathbf{H}_3\mathbf{F}$ ;
  - 5: **for**  $i = 1$  to  $L$  **do**
  - 6:    $\mathbf{o}_l = \lambda_l^{-1}(\tilde{\mathbf{H}}_3)_l - \lambda_l^{-1}(\tilde{\mathbf{H}}_3)_l \tilde{\mathbf{K}}(\lambda_l d \mathbf{I}_n + \sum_{t=1}^d \mathbf{K}_t^2) \tilde{\mathbf{K}}^H$ ;
  - 7: **end for**
  - 8: Set  $\mathbf{C}_{(3)} = \mathbf{Q}_1 \mathbf{O} \mathbf{F}^H$ ;
- Output:**  $\mathbf{C}_{(3)}$ .

as solving several manageable subproblems: (a) Update  $\mathbf{C}_{(3)}$ :

$$\begin{aligned} \mathbf{C}_{(3)} &\in \underset{\mathbf{C}_{(3)}}{\operatorname{argmin}} L(\mathcal{C}, \mathcal{V}, \mathcal{G}) \\ &= \underset{\mathbf{C}_{(3)}}{\operatorname{argmin}} \|\mathbf{X}_{(3)} - \mathbf{D}\mathbf{C}_{(3)}\mathbf{B}\mathbf{S}\|_F^2 \\ &\quad + \|\mathbf{Y}_{(3)} - \mathbf{R}\mathbf{D}\mathbf{C}_{(3)}\|_F^2 + \mu \|\mathbf{V}_{(3)} - \mathbf{C}_{(3)} + \frac{\mathbf{G}_{(3)}}{2\mu}\|_F^2 \end{aligned} \quad (24)$$

where  $\mathbf{V}_{(3)}$  and  $\mathbf{G}_{(3)}$  are matrices obtained by unfolding tensors  $\mathcal{V}$  and  $\mathcal{G}$  with the third mode, respectively. Problem (24) is strongly convex, which has unique solution. Hence, the derivative of (24) for  $\mathbf{C}_{(3)}$  is forced to be zero, and thus obtaining the Sylvester equation

$$\mathbf{H}_1 \mathbf{C}_{(3)} + \mathbf{C}_{(3)} \mathbf{H}_2 = \mathbf{H}_3, \quad (25)$$

The spectral subspace is obtained by SVD, and therefore it satisfies  $\mathbf{D}^T \mathbf{D} = \mathbf{I}_L$ . Hence, we can acquire the following equation:

$$\begin{aligned} \mathbf{H}_1 &= [\mathbf{R}\mathbf{D}]^T \mathbf{R}\mathbf{D} + \mu \mathbf{I}_L \\ \mathbf{H}_2 &= (\mathbf{B}\mathbf{S})(\mathbf{B}\mathbf{S})^T \\ \mathbf{H}_3 &= (\mathbf{R}\mathbf{D})^T \mathbf{Y}_{(3)} + \mu (\mathbf{V}_{(3)} + \frac{\mathbf{G}_{(3)}}{2\mu}) + \mathbf{D}^T \mathbf{X}_{(3)} (\mathbf{B}\mathbf{S})^T, \end{aligned} \quad (26)$$

where  $\mathbf{I}_L \in \mathbb{R}^{L \times L}$  is the identity matrix. By making use of the properties of  $\mathbf{B}$  and  $\mathbf{S}$  [45], equation (25) can be solved analytically and efficiently, and it does not need any iteration for  $\mathbf{C}_{(3)}$ . The fast approach for solving the equation (25) is presented in Algorithm 1.

In Algorithm 1, the blur matrix  $\mathbf{B}$  is decomposed as illustrated in (14), where the diagonal matrix  $\mathbf{K} \in \mathbb{C}^{WH \times WH}$  can be written as

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{K}_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{K}_d \end{bmatrix} \quad (27)$$

where  $\mathbf{K}_i \in \mathbb{C}^{wh \times wh}$ , and  $d$  is the spatial degradation factor. The unitary matrix  $\mathbf{Q}_1$  and diagonal matrix  $\mathbf{\Lambda}$  are obtained by the Eigen-decomposition, i.e.,  $\mathbf{H}_1 = \mathbf{Q}_1 \mathbf{\Lambda} \mathbf{Q}_1^{-1}$ , where  $\mathbf{\Lambda}$  is

represented as

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_L \end{bmatrix}, \quad (28)$$

and  $\mathbf{Q}_1$  holds eigenvectors of  $\mathbf{H}_1$  in its columns.  $\mathbf{o}_l$  represents  $l$ -th row of  $\mathbf{O} = \mathbf{Q}_1^{-1} \mathbf{C}_{(3)} \mathbf{F}$ , that is,  $\mathbf{O} = [\mathbf{o}_1^T, \mathbf{o}_2^T, \dots, \mathbf{o}_L^T]^T$ .  $\mathbf{1}_d \in \mathbb{R}^d$  is a vector, and its elements are all ones.

(b) Update  $\mathcal{V}$ :

$$\begin{aligned} \mathcal{V} &\in \underset{\mathcal{V}}{\operatorname{argmin}} L(\mathcal{C}, \mathcal{V}, \mathcal{G}) \\ &= \underset{\mathcal{V}}{\operatorname{argmin}} \mu \|\mathcal{V} - \mathcal{C} + \frac{\mathcal{G}}{2\mu}\|_F^2 + \lambda \sum_{k=1}^K \|\mathcal{V}^k\|_{\text{TMR}} \end{aligned} \quad (29)$$

We can solve the problem (29) independently for each cluster, i.e.,

$$\underset{\mathcal{V}^k}{\operatorname{argmin}} \sum_{k=1}^K \mu \|\mathcal{V}^k - \mathcal{C}^k + \frac{\mathcal{G}^k}{2\mu}\|_F^2 + \lambda \sum_{k=1}^K \|\mathcal{V}^k\|_{\text{TMR}} \quad (30)$$

where  $\mathcal{V}^k$  and  $\mathcal{G}^k$  are tensors formed by  $k$ -th cluster patches in  $\mathcal{V}$  and  $\mathcal{G}$ , respectively. Based on the definition of  $\|\mathcal{A}\|_{\text{TMR}}$  in equation (12), the  $k$ -th cluster  $\mathcal{V}^k$  can be estimated by solving the following problem:

$$\underset{\tilde{\mathcal{V}}^k}{\operatorname{argmin}} \frac{\mu}{q} \|\tilde{\mathcal{V}}^k - \tilde{\mathcal{H}}^k\|_F^2 + \lambda \sum_{i=1}^q \text{LS}(\tilde{\mathcal{V}}^k(:, :, i)) \quad (31)$$

where  $\tilde{\mathcal{V}}^k = \text{fft}(\mathcal{V}^k, [ ], 3)$  and  $\tilde{\mathcal{H}}^k = \text{fft}(\mathcal{C}^k - \frac{\mathcal{G}^k}{2\mu}, [ ], 3)$  are tensors obtained by conducting FFT on  $\mathcal{V}^k$  and  $\mathcal{C}^k - \frac{\mathcal{G}^k}{2\mu}$  with the third dimension, respectively. The tensor  $\tilde{\mathcal{V}}^k$  is estimated for each frontal slice respectively, and the LS based low-rank recovery problem [42] for each slice  $\tilde{\mathcal{V}}^k(:, :, i)$  has the closed-form solution:

$$\tilde{\mathcal{V}}^k(:, :, i) = \mathbf{U}_2 \tilde{\Sigma}_2 \mathbf{V}_2, \quad (32)$$

where  $\mathbf{U}_2$  and  $\mathbf{V}_2$  are acquired by the SVD on matrix  $\tilde{\mathcal{H}}^k(:, :, i)$ , i.e.,  $\tilde{\mathcal{H}}^k(:, :, i) = \mathbf{U}_2 \Sigma_2 \mathbf{V}_2$ .  $\tilde{\Sigma}_2$  is a diagonal matrix, and its diagonal elements are obtained by  $\tilde{\Sigma}_2(i, i) = E_{\frac{\lambda}{2\mu}, \varepsilon}[\Sigma_2(i, i)]$ . The thresholding operator  $E_{\lambda, \varepsilon}(\cdot)$  is given as

$$E_{\lambda, \varepsilon}(x) = \begin{cases} 0 & c_2 \leq 0 \\ \frac{c_1 + \sqrt{c_2}}{2} & c_2 > 0 \end{cases} \quad (33)$$

where  $c_1 = |x| - \varepsilon$  and  $c_2 = c_1^2 - 4(\alpha - \varepsilon|x|)$ . After obtaining  $\tilde{\mathcal{V}}^k$ ,  $\mathcal{V}^k$  can be computed the inverse FFT along the third dimension, i.e.,  $\mathcal{V}^k = \text{ifft}(\tilde{\mathcal{V}}^k, [ ], 3)$ . The algorithm for estimating  $\mathcal{V}^k$  is summarized in Algorithm 2. Finally,  $\mathcal{V}$  is acquired by aggregating FBP clusters  $\mathcal{V}^k, 1 \leq k \leq K$ .

(c) Update Lagrangian multipliers  $\mathcal{G}$ :

$$\mathcal{G} = \mathcal{G} + 2\mu(\mathcal{V} - \mathcal{C}) \quad (34)$$

The proposed LTMR method is outlined in Algorithm 3. The Algorithm 3 is stopped when the maximum number of iterations  $T$  is reached. We set  $T = 100$  in our experiments.

**Algorithm 2** A Closed-Form Solution of (30) w.r.t.  $\mathcal{V}^k$ 


---

**Input:**  $\mathcal{H}^k = \mathcal{C}^k - \frac{\mathcal{G}^k}{2\mu}$   
1:  $\tilde{\mathcal{H}}^k = \text{fft}(\mathcal{H}^k, [], 3)$ ;  
2: **for**  $i = 1$  to  $q$  **do**  
3:  $[\mathbf{U}_2, \mathbf{\Sigma}_2, \mathbf{V}_2] = \text{svd}(\tilde{\mathcal{H}}^k(:, :, i))$ ;  
4:  $\tilde{\mathbf{\Sigma}}_2 = \mathbf{D}_{\frac{\lambda}{2\mu}, \epsilon}[\mathbf{\Sigma}_2]$ ;  
5:  $\tilde{\mathcal{V}}^k(:, :, i) = \mathbf{U}_2 \tilde{\mathbf{\Sigma}}_2 \mathbf{V}_2$ ;  
6: **end for**  
7:  $\mathcal{V}^k = \text{ifft}(\tilde{\mathcal{V}}^k, [], 3)$ ;  
**Output:**  $\mathcal{V}^k$ .

---

**Algorithm 3** LTMR for HSI Super-Resolution

---

**Input:**  $\mathcal{X}, \mathcal{Y}, \lambda, K, L$ , and  $\sqrt{q}$ ,  
1: learn the subspace  $\mathbf{D}$  from  $\mathbf{X}_{(3)}$  via SVD;  
2: learn the cluster structure from the HR-MSI;  
3: **while** not converged **do**  
4: Update  $\mathbf{C}_{(3)}$  via solving equation (24);  
5: Update  $\mathcal{V}$  via solving equation (29);  
6: Update Lagrangian multipliers  $\mathcal{G}$  via equation (34);  
7: **end while**  
8:  $\mathcal{Z} = \mathcal{C} \times_3 \mathbf{D}$ ;  
**Output:**  $\mathcal{Z}$ .

---

## V. EXPERIMENTS

## A. Datasets

To test the effectiveness of the proposed LTMR method, experiments are conducted on both simulated HSI data and real HSI data. The first HSI dataset is Columbia computer vision laboratory (CAVE) [46], which is ground-based. The CAVE dataset [46] has 32 indoor HSIs with high quality. Each HSI in the CAVE dataset has  $512 \times 512$  pixels and 31 spectral bands, where  $512 \times 512$  is the number of pixels, and 31 is the number of spectral bands. The HSIs has the spectral range of 400nm-700nm with a wavelength interval 10nm. The first two spectral bands of CAVE dataset has a blur, and therefore we remove the first two bands to avoid bias. We use the HSI from CAVE dataset as a reference image. The LR-HSI is simulated by applying a  $7 \times 7$  Gaussian blur (standard deviation 2) and then by downsampling every 32 pixels in two spatial dimensions. The response of a Nikon D700 camera is used to generate the three-band HR-MSI (RGB image).

The second dataset is Pavia University [47], acquired by the reflective optics system imaging spectrometer optical sensor over the area of Pavia University. The Pavia University have the 115 spectral bands and  $610 \times 340$  spectral pixels. By removing the bands of low SNR, we reduce the HSI as 93 bands. We select the up-left  $256 \times 256 \times 93$ -pixel-size image as the ground truth. To simulate the LR-HSI, we firstly filter each band of HR-HSI by a  $7 \times 7$  Gaussian blur (standard deviation 2) and then downsample every four pixels in two spatial modes. The HR-MSI of four bands is simulated by using IKONOS-like reflectance spectral response filter [25].

## B. Compared Approaches

Our method is compared with three recent state-of-the-art methods for HSI super-resolution, including subspace based method (HySure) [24],<sup>1</sup> non-negative structured sparse representation (NSSR) [26],<sup>2</sup> and coupled sparse tensor factorization (CSTF) [32].<sup>3</sup> Among the compared approaches, the CSTF is based on Tucker factorization, and the NSSR and HySure are based on matrix factorization. We set the parameters for the compared approaches for the best performance. Specifically, for the HySure, we set  $\lambda_m = 1$  and  $\lambda_\phi = 5 \times 10^{-5}$ . For the NSSR,  $\eta_1$  is set as  $10^{-4}$  and 0.015 for the Pavia University and CAVE dataset, respectively,  $\eta_2$  is set as  $10^{-4}$ , and  $K$  is set as 75,. For the CSTF, we set  $n_w = 260$ ,  $n_h = 260$ ,  $n_s = 15$ , and  $\lambda = 10^{-5}$ .

## C. Quantitative Metrics

Four indexes are utilized to measure the quality of the fusion results.

The peak signal to noise ratio (PSNR) is used as the first index. The PSNR for HSI defined as,

$$\text{PSNR}(\mathcal{Z}, \hat{\mathcal{Z}}) = \frac{1}{S} \sum_{i=1}^S \text{PSNR}(\mathbf{Z}_i, \hat{\mathbf{Z}}_i), \quad (35)$$

where  $\mathbf{Z}_i$  and  $\hat{\mathbf{Z}}_i$  represent  $i$ -th band of  $\mathcal{Z} \in \mathbb{R}^{W \times H \times S}$  and fused results  $\hat{\mathcal{Z}} \in \mathbb{R}^{W \times H \times S}$ , respectively.

We use the spectral angle mapper (SAM) [48] as the second quality metric, which reflects the spectral error between fusion result and ground truth.

The third index is the universal image quality index (UIQI) [49], which is calculated on a sliding window of size  $32 \times 32$  and averaged on all windows. The uqi for two windows  $\mathbf{a}$  and  $\mathbf{b}$  is given by

$$\text{UIQI}(\mathbf{a}, \mathbf{b}) = \frac{4\sigma_{\mathbf{ab}}^2}{\sigma_{\mathbf{a}}^2 + \sigma_{\mathbf{b}}^2} \frac{\mu_{\mathbf{a}}\mu_{\mathbf{b}}}{\mu_{\mathbf{a}}^2 + \mu_{\mathbf{b}}^2}, \quad (36)$$

where  $\sigma_{\mathbf{ab}}$  is the sample covariance between  $\mathbf{a}$  and  $\mathbf{b}$ , and  $\sigma_{\mathbf{a}}$  and  $\mu_{\mathbf{a}}$  denote the standard deviation and mean value of  $\mathbf{a}$ , respectively.

The relative dimensionless global error in synthesis (ERGAS) [50] is the fourth quality metric, which is calculated as

$$\text{ERGAS}(\mathcal{Z}, \hat{\mathcal{Z}}) = \frac{100}{d} \sqrt{\frac{1}{S} \sum_{i=1}^S \frac{\text{MSE}(\mathbf{Z}_i, \hat{\mathbf{Z}}_i)}{\mu_{\mathbf{Z}_i}^2}}, \quad (37)$$

where  $\text{MSE}(\mathbf{Z}_i, \hat{\mathbf{Z}}_i)$  denote the mean square error between  $\mathbf{Z}_i$  and  $\hat{\mathbf{Z}}_i$ , and  $d$  is spatial downsampling factor. The smaller ERGAS, the better fusion results.

## D. Parameters Discussion

In our method, four key parameters: the number of atoms  $L$ , the number of clusters  $K$ ,  $\lambda$ , and  $\sqrt{q}$  need to be discussed.

<sup>1</sup><https://github.com/alfaiate/HySure>

<sup>2</sup><http://see.xidian.edu.cn/faculty/wsdong>

<sup>3</sup><https://sites.google.com/view/renweidian>

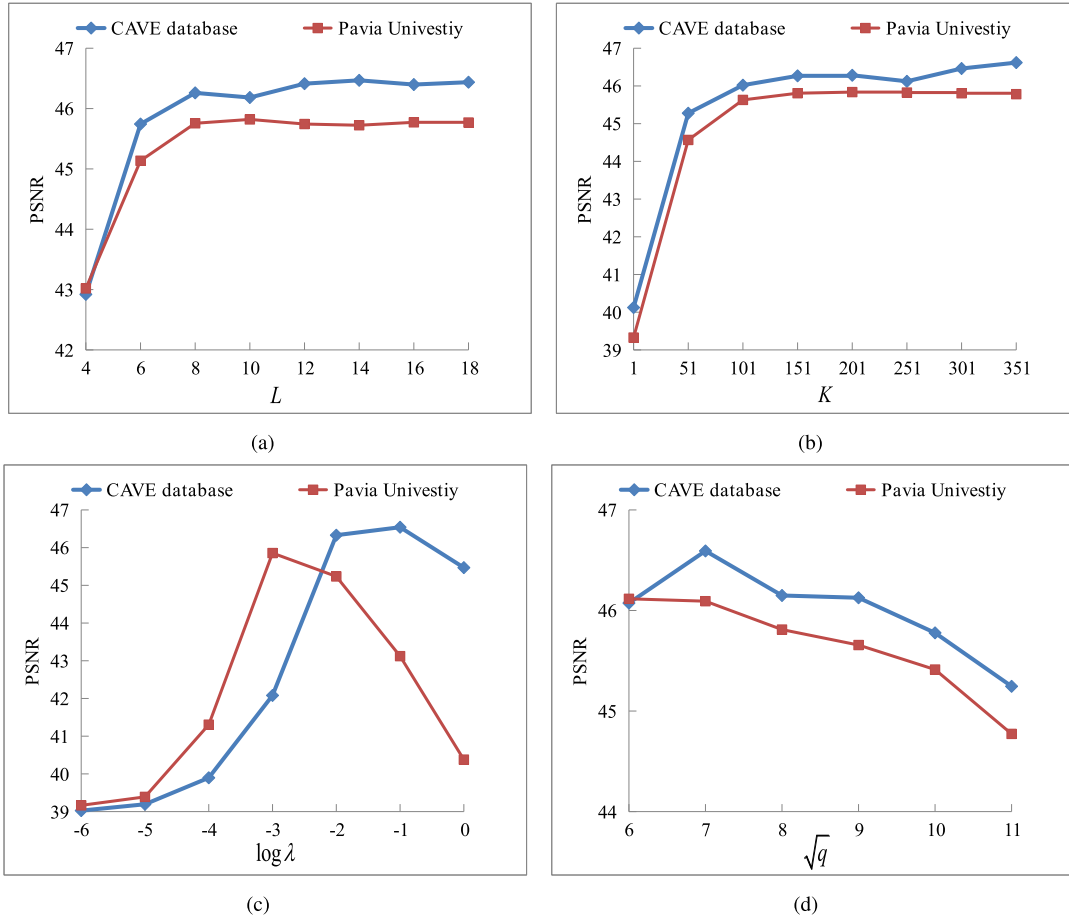


Fig. 2. The PSNR varies from parameters  $L$ ,  $K$ ,  $\log \lambda$ , and  $\sqrt{q}$ , respectively. (a) is for  $L$ . (b) is for  $K$ . (c) is for  $\log \lambda$ . (d) is for  $\sqrt{q}$ .

To discuss the effect of the subspace dimension  $L$ , we plot the PSNR curves of the fused results of *Balloons* (an HSI in the CAVE dataset) and Pavia University varying from  $L$  in Fig. 2 (a). As we can see from Fig. 2 (a) that the PSNR rises for both *Balloons* and Pavia University as  $L$  changed from 4 to 10, and then it keeps relatively stable as  $L$  increases further. Therefore, 10 atoms are enough to preserve the spectral information, which demonstrates that spectral vectors indeed live in a low-dimensional subspace, and we can reduce the dimension of spectral mode via subspace representation. Therefore, we set  $L = 10$  for both CAVE dataset and Pavia University.

The number of groups  $K$  is a very important parameter for FBPs clustering. Fig. 2 (b) shows the PSNR curves of the fused results of *Balloons* and of Pavia University varying from  $K$ . As can be seen from Fig. 2 (b), the figure starts from  $K = 1$ , which means that all FBPs belong to the same group and the non-local clustering operation is not conducted. The PSNR for *Balloons* dataset and Pavia University firstly rises as  $K$  varies from 1 to 201, which means that the non-local clustering operation really works. The PSNR for two images keeps stable as  $K$  is bigger than 201. Hence, we set  $K = 201$  for both Pavia University and CAVE dataset.

$\lambda$  is the low tensor multi-rank regularization parameter, which directly affects the solution of (30). Fig. 2 (c) plots the

PSNR curves of our method on *Balloons* and Pavia University varying from  $\log \lambda$  ( $\log$  is base 10.). From Fig. 2 (c), we can see that the PSNR for the Pavia University has the sharp rise as  $\log \lambda$  varies from  $-6$  to  $-3$ , and then it declines rapidly as  $\log \lambda$  grows further. The PSNR for *Balloons* rises, as  $\log \lambda$  varies from  $-6$  to  $-2$ , and then it declines as  $\log \lambda$  is greater than  $-1$ . Hence,  $\lambda$  is set as  $3 \times 10^{-2}$  and  $10^{-3}$  for the CAVE dataset and Pavia University, respectively. The optimal  $\lambda$  for the CAVE dataset is greater than that for the Pavia University. The potential reason is that HSIs in CAVE dataset has much higher resolution than that in the Pavia University, and therefore the HSIs in CAVE dataset have stronger non-local similarities, which results in greater optimal  $\lambda$ .

In our experiment, the spatial size of the patches is set as  $\sqrt{q} \times \sqrt{q}$ , which can also affect the performance of our method. To test the effect of spatial size, Fig. 2 (d) plots the PSNR curves of our method on *Balloons* and Pavia University as functions of  $\sqrt{q}$ . We can see from the figure that the PSNR for the CAVE dataset goes up as  $\sqrt{q}$  varies from 6 to 7, and then the PSNR declines. The PSNR for the Pavia University holds stable when  $\sqrt{q}$  changes from 6 to 7, and then it decreases. Therefore, we set  $\sqrt{q} = 7$  with overlap 4 for both two datasets.

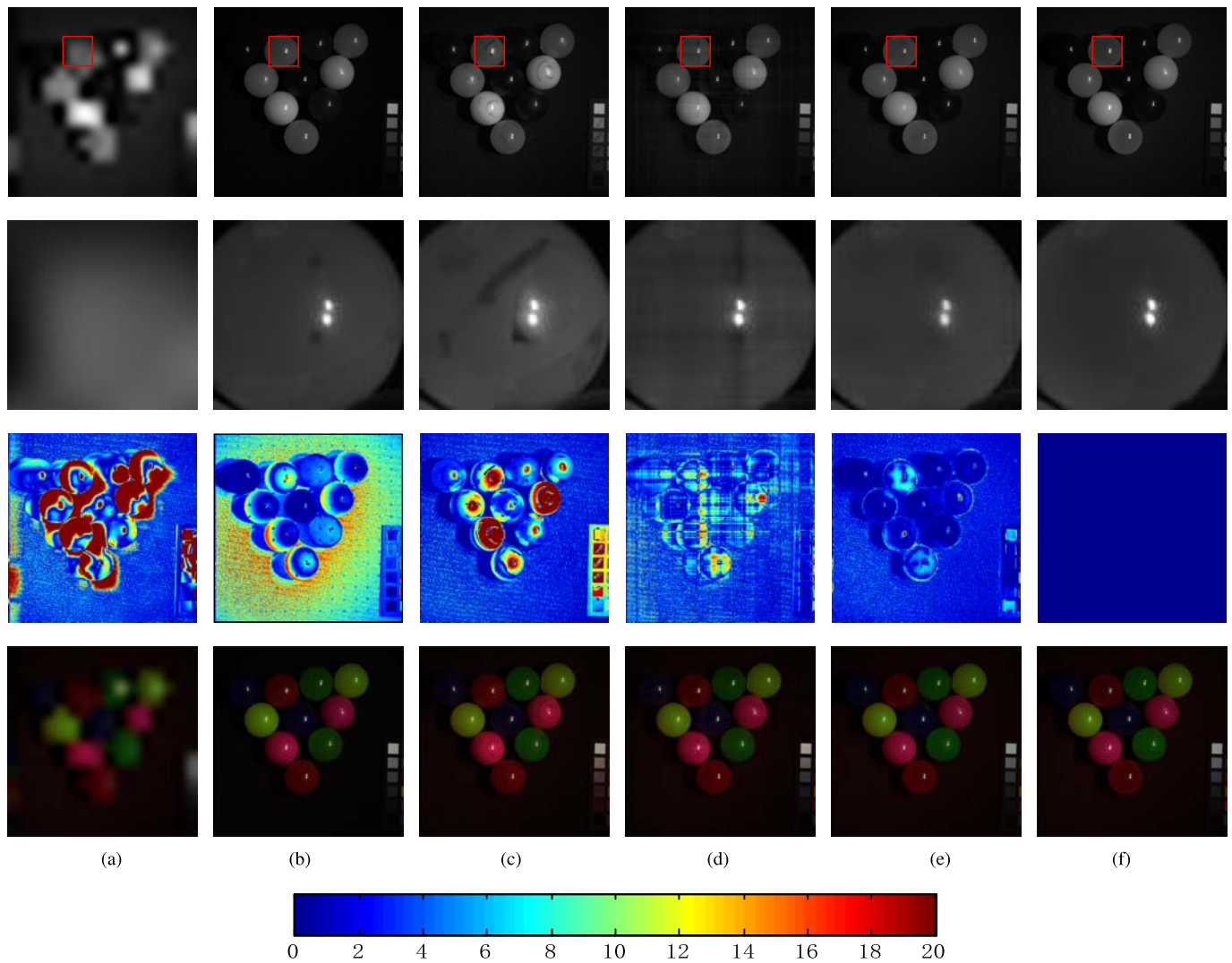


Fig. 3. First row shows the fused images by the competing approaches for *Superballs* (an HSI in CAVE dataset) at band 29. Second row shows magnified version of the marked area. Third row shows the error images of the test methods for *Superballs* at band 29. Fourth row shows false color images formed by bands 29, 10, and 5 for *Superballs*. (a) LR-HSI. (b) HySure [24]. (c) NSSR [26]. (d) CSTF [33]. (e) LTMR. (f) Ground truth.

TABLE I  
AVERAGE QUANTITATIVE INDEXES OF THE COMPETING  
APPROACHES ON THE CAVE DATASET [46]

Method	CAVE dataset [46]			
	PSNR	SAM	UIQI	ERGAS
Best Values	$+\infty$	0	1	0
HySure [24]	39.399	14.357	0.848	0.477
NSSR [26]	39.319	8.788	0.875	0.484
CSTF [33]	41.392	9.712	0.821	0.390
LTMR	<b>42.607</b>	<b>6.769</b>	<b>0.890</b>	<b>0.337</b>

#### E. Experimental Results on Simulated Data

Table I reports the average quality metrics of test methods on the CAVE dataset. We highlight the best results in bold. From I, we can see that our LTMR method performs consistently better than other testing methods according to all quality metrics, which indicates that HR-HSIs recovered by our method have better spatial and spectral qualities in

CAVE dataset. For example, the PSNR of our method is greater than that of the state-of-the-art CSTF method by 1.2dB, and the SAM is reduced by 2 degrees compared with NSSR. The improvement mainly comes from two aspects. Firstly, we exploit the prior that HSI lives in a low-dimensional subspace. What is more, the proposed LTMR prior can well depict the non-local self-similarities of the HR-HSI. In addition, the NSSR and CSTF perform better than HySure. It is because that HSIs in CAVE dataset have a very high spatial resolution, and the total variation used in HySure may not be a good regularizer for very high-resolution HSIs. Besides, the CSTF does not use the non-local similarities of the HR-HSI. For visual comparison, the recovered HR-HSIs, corresponding error images of the competing approaches at band 29 for *Superballs* (an HSI in CAVE dataset), and false color images formed by bands 29, 10, and 5 are shown in Fig. 3. A representative region of each reconstructed image is magnified. As can be seen from the magnified images and error images, the fused images by the HySure, the NSSR, and



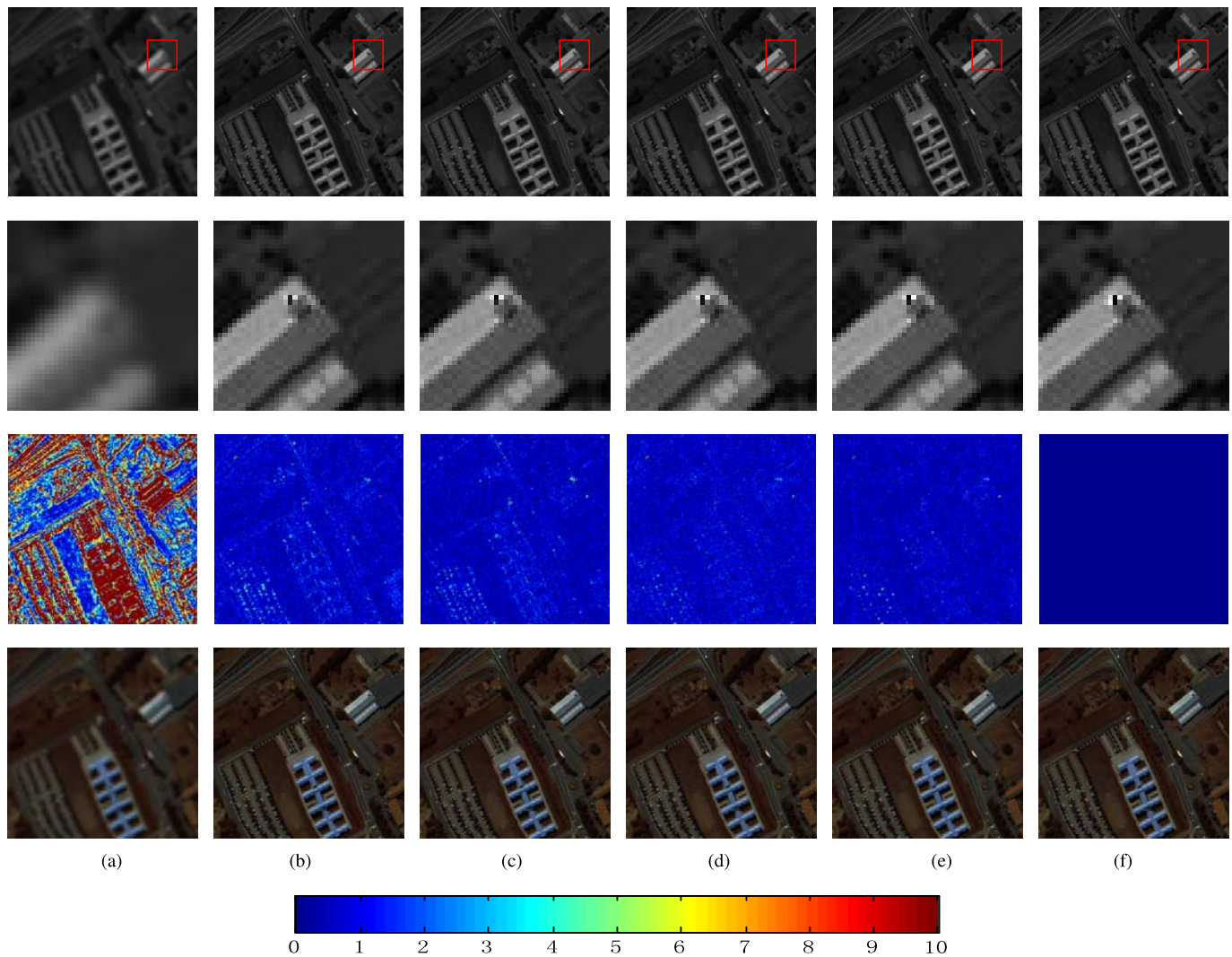


Fig. 4. First row: reconstructed images of the competing approaches for the Pavia University at 30-th band. Second row shows magnified version of the marked area. Third row: corresponding error images of the competing approaches for the Pavia University at 30-th band. Fourth row shows the false color images formed by bands 60, 30, and 5 for the Pavia University. (a) LR-HSI. (b) HySure [24]. (c) NSSR [26]. (d) CSTF [33]. (e) LTMR. (f) Ground truth.

TABLE II  
QUANTITATIVE INDEXES OF THE COMPETING APPROACHES  
ON THE PAVIA UNIVERSITY [47]

Method	Pavia University [47]			
	PSNR	SAM	UIQI	ERGAS
Best Values	$+\infty$	0	1	0
HySure [24]	44.619	1.738	0.995	0.923
NSSR [26]	44.709	1.657	0.995	0.927
CSTF [33]	44.450	1.779	0.995	0.952
LTMR	<b>45.820</b>	<b>1.493</b>	<b>0.996</b>	<b>0.805</b>

CSTF have obvious flaws, and the HR-HSI reconstructed by LTMR has fewer artifacts.

The quality metrics of the competing approaches on the Pavia University are shown in Table II. Our LTMR method still has obvious advantages over other testing approaches in terms of all quality indexes. Meanwhile, the HySure delivers comparable performance to the NSSR and CSTF. Fig. 4 shows the reconstructed images at band 30, error images at band 30, and false color images formed by bands 60, 30, and 5 of

the competing approaches for the Pavia University. From the magnified images, we can see that all testing approaches can provide clear and sharp details compared with the LR-HSI, which demonstrates the effectiveness of these super-resolution methods. As we can see from the error images, the HR-HSIs produced by LTMR and CSTF have fewer errors than that produced by HySure and NSSR. To compare the spectral reflectance of some certain materials, Fig. 5 (a) and (b) show the PSNR as functions of spectral band for *Asphalt* and *Trees*, respectively. The PSNR values measure the similarities of spectral reflectance between the fusion results and ground truth on the two regions. As we can see from the figure, the proposed LTMR method has the highest PSNR for most of the spectral bands of two areas, which further demonstrates the advantages of our method.

#### F. Experimental Results on Real Data

In order to further evaluate our method, we apply our method to real HSI and MSI fusion. The HSI dataset is taken by the Hyperion sensor loaded on Earth Observing-1 satellite,

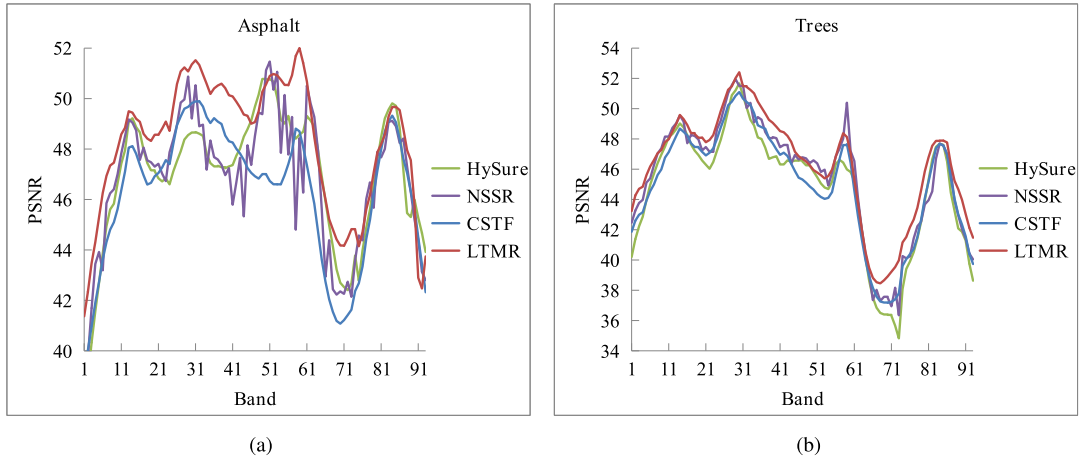


Fig. 5. PSNR of *Asphalt* and *Trees* as a function of spectral band. (a) *Asphalt*; (b) *Trees*.

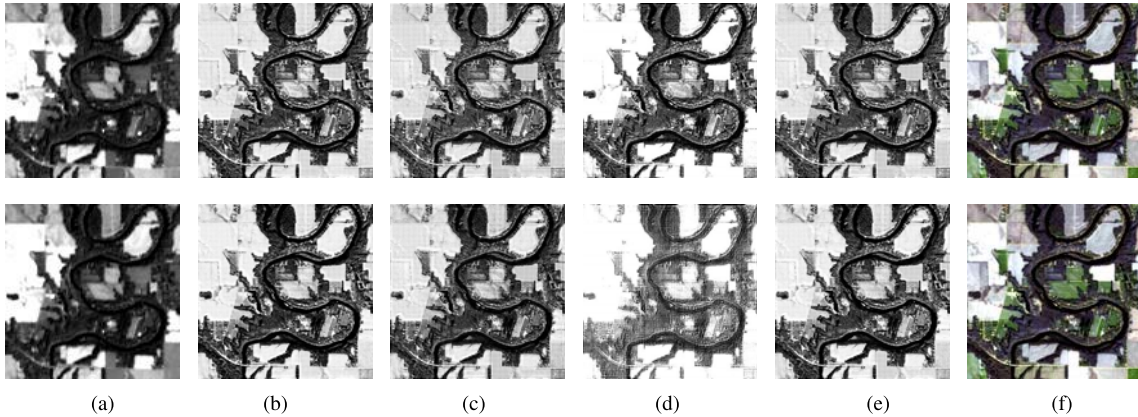


Fig. 6. The first and second rows show the recovered HR-HSIs at 4-th and 15-th bands for Hyperion data, respectively. (a) LR-HSI. (b) HySure [24]. (c) NSSR [26]. (d) CSTF [33]. (e) LTMR. (f) HR-MSI.

which has the spatial resolution of 30 m. The Hyperion HSI contains 220 spectral bands in the spectral range of 400-2500nm. Some bands of the Hyperion HSI suffers from noise, and 89 bands are remained after removing the noisy bands. The used Hyperion HSI is of size  $120 \times 120 \times 89$ . The MSI data with 13 bands is collected by the Sentinel-2A satellite. Four bands with 10m spatial resolution are used for the fusion. The spatial downsampling factor is 3, and the size of MSI data is  $360 \times 360 \times 4$ . The spectral response matrix  $\mathbf{R}$  and convolution blur matrix  $\mathbf{B}$  are estimated via the method proposed in [24]. The first and second rows of Fig. 6 show the fused HR-HSI at 4-th and 15-th band, respectively. As we can see from the figure that all the methods can obviously improve the spatial resolution, compared with the original LR-HSI. Besides, the HySure, NSSR, and LTMR perform better than the CSTF. The underlying reason is that the CSTF is designed for the case that the convolution blur can be decomposed in two spatial modes, however the estimated convolution blur is not decomposable.

#### G. The Effectiveness of the Low Tensor Multi-Rank Regularization

To demonstrate the effectiveness of the low tensor multi-rank regularization, we discuss the low tensor multi-rank

regularization parameter  $\lambda$ . The  $\lambda$  is the low tensor multi-rank regularization parameter, and bigger value means lower tensor multi-rank. Fig. 2 (c) plots the PSNR of our method on *Balloons* and of Pavia University as functions of  $\log \lambda$ . As we can see from Fig. 2 (c), the PSNR for both of CAVE dataset and Paiva university has a rising process, as  $\log \lambda$  increases, which demonstrates the proposed low tensor multi-rank term really works.

Besides, we add the experiments of  $\lambda = 0$ , where the low tensor multi-rank regularization term is not used. Table III shows the quantitative results of with/without the regularization term on the CAVE dataset and Pavia University. From the table, we can see that the low tensor multi-rank regularization can deliver obviously better results on both two datasets.

#### H. Computation Efficiency

To show the computation efficiency of the compared approaches, Table IV reports the average computational time in seconds of the compared methods on the CAVE dataset, Pavia University, and Hyperion dataset. All experiments are implemented at Matlab R2018b, and the computer has Intel Core-E5-2603 CPU with 1.6-GHz and 96-GB random access memory. As can be seen from the table, the NSSR is the fastest methods on the CAVE dataset, and the HySure needs

TABLE III  
QUANTITATIVE RESULTS OF WITH/WITHOUT THE REGULARIZATION TERM ON THE CAVE DATASET AND PAVIA UNIVERSITY

Methods	CAVE dataset				Pavia University			
	PSNR	SAM	UIQI	ERGAS	PSNR	SAM	UIQI	ERGAS
Best Values	$+\infty$	0	1	0	$+\infty$	0	1	0
LTMR ( $\lambda=0$ )	36.933	11.416	0.826	0.644	39.142	3.709	0.975	1.991
LTMR	<b>42.607</b>	<b>6.769</b>	<b>0.890</b>	<b>0.337</b>	<b>45.820</b>	<b>1.493</b>	<b>0.996</b>	<b>0.805</b>

TABLE IV  
AVERAGE COMPUTATIONAL TIME IN SECONDS OF COMPARED APPROACHES OF TESTING METHODS

Dataset	Method			
	HySure	NSSR	CSTF	LTMR
CAVE dataset	341	<b>207</b>	311	620
Pavia University	<b>89</b>	100	282	207
Hyperion dataset	<b>188</b>	303	316	350

the least computational time on the Pavia University and Hyperion dataset. The main reason is that HySure is spectral subspace based, and therefore it has the speed advantage on the HSI dataset, which has more spectral bands. Our method needs relatively more computation time on the three datasets, and the main computation burden of our method comes from solving (34) of the manuscript, where we compute the low tensor multi-rank approximation for each cluster. The time complexity for solving (30) is  $O(qKL^2N_m)$ , where  $N_m = \max(N_1, \dots, N_k)$  is the maximum number of patches among all clusters.

## VI. CONCLUSIONS

In this paper, we propose a subspace-based low tensor multi-rank regularization approach to enhance the spatial resolution of the HSI. Based on the low-dimensional spectral subspace representation model, we approximate the target HR-HSI with low-dimensional spectral subspace and corresponding coefficients. Since the LR-HSI maintains most of the spectral information, we first learn the low-dimensional spectral subspace from the LR-HSI. With the subspace known, the coefficients are estimated with low tensor multi-rank prior. Specifically, According to the learned clusters on FBPs of HR-MSI, we build clusters of FBPs in the coefficients. We collect the coefficient FBPs in the same cluster as a three-dimensional tensor and impose the low tensor multi-rank prior on these tensors, which fully learn the non-local self-similarities of the HR-MSI. Experimental results on both simulated and real datasets indicates the advantages of our method.

## ACKNOWLEDGEMENT

We would like to appreciate the editors and anonymous reviewers for their careful and valuable suggestions and comments, which have highly improved this work.

## REFERENCES

- [1] M. Uzair, A. Mahmood, and A. Mian, "Hyperspectral face recognition with spatio-spectral information fusion and PLS regression," *IEEE Trans. Image Process.*, vol. 24, no. 3, pp. 1127–1137, Mar. 2015.
- [2] Y. Yuan, X. Zheng, and X. Lu, "Discovering diverse subset for unsupervised hyperspectral band selection," *IEEE Trans. Image Process.*, vol. 26, no. 1, pp. 51–64, Jan. 2017.
- [3] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. M. 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.
- [4] P. Zhong and R. Wang, "Learning conditional random fields for classification of hyperspectral images," *IEEE Trans. Image Process.*, vol. 19, no. 7, pp. 1890–1907, Jul. 2010.
- [5] B. Du, Y. Zhang, L. Zhang, and D. Tao, "Beyond the sparsity-based target detector: A hybrid sparsity and statistics-based detector for hyperspectral images," *IEEE Trans. Image Process.*, vol. 25, no. 11, pp. 5345–5357, Nov. 2016.
- [6] N. He, L. Fang, S. Li, A. Plaza, and J. Plaza, "Remote sensing scene classification using multilayer stacked covariance pooling," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 12, pp. 6899–6910, Dec. 2018.
- [7] N. He *et al.*, "Feature extraction with multiscale covariance maps for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 755–769, Feb. 2019.
- [8] Z. Huang, S. Li, L. Fang, H. Li, and J. A. Benediktsson, "Hyperspectral image denoising with group sparse and low-rank tensor decomposition," *IEEE Access*, vol. 6, pp. 1380–1390, Dec. 2017.
- [9] L. Loncan *et al.*, "Hyperspectral pansharpening: A review," *IEEE Trans. Geosci. Remote Sens.*, vol. 3, no. 3, pp. 27–46, Sep. 2015.
- [10] M. E. Kilmer, K. Braman, N. Hao, and R. C. Hoover, "Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 1, pp. 148–172, 2013.
- [11] O. Semerci, N. Hao, M. E. Kilmer, and E. L. Miller, "Tensor-based formulation and nuclear norm regularization for multienergy computed tomography," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1678–1693, Apr. 2014.
- [12] Z. Zhang, G. Ely, S. Aeron, N. Hao, and M. Kilmer, "Novel methods for multilinear data completion and de-noising based on tensor-SVD," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 3842–3849.
- [13] P. Zhou, C. Lu, Z. Lin, and C. Zhang, "Tensor factorization for low-rank tensor completion," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1152–1163, Mar. 2018.
- [14] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor robust principal component analysis: Exact recovery of corrupted low-rank tensors via convex optimization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 5249–5257.
- [15] T.-X. Jiang, T.-Z. Huang, X.-L. Zhao, and L.-J. Deng. (2017). "A novel nonconvex approach to recover the low-tubal-rank tensor data: When t-SVD meets PSSV." [Online]. Available: <https://arxiv.org/abs/1712.05870>
- [16] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.
- [17] R. Kawakami, Y. Matsushita, J. Wright, M. Ben-Ezra, Y.-W. Tai, and K. Ikeuchi, "High-resolution hyperspectral imaging via matrix factorization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 2329–2336.
- [18] 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.
- [19] 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.



- [20] 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.
- [21] 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.
- [22] Y. Zhou, L. Feng, C. Hou, and S.-Y. Kung, "Hyperspectral and multispectral image fusion based on local low rank and coupled spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 10, pp. 5997–6009, Oct. 2017.
- [23] M. A. Veganzones, M. Simões, G. Licciardi, N. Yokoya, J. M. Bioucas-Dias, and J. Chanussot, "Hyperspectral super-resolution of locally low rank images from complementary multisource data," *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 274–288, Jan. 2016.
- [24] M. Simoes, 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.
- [25] 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.
- [26] 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.
- [27] X.-H. Han, B. Shi, and Y. Zheng, "Self-similarity constrained sparse representation for hyperspectral image super-resolution," *IEEE Trans. Image Process.*, vol. 27, no. 11, pp. 5625–5637, Nov. 2018.
- [28] C. Yi, Y.-Q. Zhao, and J. C.-W. Chan, "Hyperspectral image super-resolution based on spatial and spectral correlation fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 7, pp. 4165–4177, Jul. 2018.
- [29] C. 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.
- [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] R. Dian, S. Li, L. Fang, and Q. Wei, "Multispectral and hyperspectral image fusion with spatial-spectral sparse representation," *Inf. Fusion*, vol. 49, pp. 262–270, Sep. 2019.
- [32] R. Dian, L. Fang, and S. Li, "Hyperspectral image super-resolution via non-local sparse tensor factorization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2017, pp. 5344–5353.
- [33] 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.
- [34] 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 Observat. Remote Sens.*, vol. 11, no. 4, pp. 1030–1040, Apr. 2018. doi: [10.1109/JSTARS.2017.2785411](https://doi.org/10.1109/JSTARS.2017.2785411).
- [35] Y. Chang, L. Yan, H. Fang, S. Zhong, and Z. Zhang. (2017). "Weighted low-rank tensor recovery for hyperspectral image restoration." [Online]. Available: <https://arxiv.org/abs/1709.00192>
- [36] R. Dian, S. Li, and L. Fang, "Learning a low tensor-train rank representation for hyperspectral image super-resolution," *IEEE Trans. Neural Netw. Learn. Syst.*, to be published. doi: [10.1109/TNNLS.2018.2885616](https://doi.org/10.1109/TNNLS.2018.2885616).
- [37] Y. Xu, Z. Wu, J. Chanussot, and Z. Wei, "Nonlocal patch tensor sparse representation for hyperspectral image super-resolution," *IEEE Trans. Image Process.*, vol. 28, no. 6, pp. 3034–3047, Jun. 2019.
- [38] Y. Fu, Y. Zheng, H. Huang, I. Sato, and Y. Sato, "Hyperspectral image super-resolution with a mosaic RGB image," *IEEE Trans. Image Process.*, vol. 27, no. 11, pp. 5539–5552, Nov. 2018.
- [39] L. Zhang, W. Wei, C. Bai, Y. Gao, and Y. Zhang, "Exploiting clustering manifold structure for hyperspectral imagery super-resolution," *IEEE Trans. Image Process.*, vol. 27, no. 12, pp. 5969–5982, Dec. 2018.
- [40] Y. Qu, H. Qi, and C. Kwan, "Unsupervised sparse Dirichlet-Net for hyperspectral image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 2511–2520.
- [41] 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.
- [42] 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.
- [43] L. Zhuang and J. M. Bioucas-Dias, "Fast hyperspectral image denoising and inpainting based on low-rank and sparse representations," *IEEE J. Sel. Topics Appl. Earth Observat. Remote Sens.*, vol. 11, no. 3, pp. 730–742, Mar. 2018.
- [44] D. Arthur and S. Vassilvitskii, "K-means++: The advantages of careful seeding," in *Proc. Annu. ACM-SIAM Symp. Discrete Algorithms*, 2007, pp. 1027–1035.
- [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] 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.
- [48] R. H. Yuhas, A. F. H. Goetz, and J. W. Boardman, "Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm," *JPL Airborne Geosci. Workshop*, vol. 1, 1992, pp. 147–149.
- [49] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.
- [50] 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.

From 2017 to 2018, he was a Visiting Ph.D. Student with the University of Lisbon, Lisbon, Portugal, supported by the China Scholarship Council. His 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. From May to October 2001, he was a Research Associate with the Department of Computer Science, The Hong Kong University of Science and Technology. From 2002 to 2003, he was a Post-Doctoral Fellow with the Royal Holloway College, University of London. From April to June 2005, he was a Visiting Professor with the Department of Computer Science, The Hong Kong University of Science and Technology. He is currently a Full Professor with the College of Electrical and Information Engineering, Hunan University. He has authored or coauthored more than 200 refereed papers. He received two 2nd-Grade State Scientific and Technological Progress Awards of China in 2004 and 2006. His current research interests include image processing, pattern recognition, and artificial intelligence.

He is a member of the Editorial Board of the *Information Fusion* and the *Sensing and Imaging*. He is an Associate Editor of the *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING* and the *IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT*.