Nonnegative and Nonlocal Sparse Tensor Factorization-Based Hyperspectral Image Super-Resolution

Wei Wan, Weihong Guo, *Associate Member, IEEE*, Haiyang Huang, and Jun Liu

***Abstract*— Hyperspectral image (HSI) super-resolution refers to enhancing the spatial resolution of a 3-D image with many spectral bands (slices). It is a seriously ill-posed problem when the low-resolution (LR) HSI is the only input. It is better solved by fusing the LR HSI with a high-resolution (HR) multispectral image (MSI) for a 3-D image with both high spectral and spatial resolution. In this article, we propose a novel nonnegative and nonlocal 4-D tensor dictionary learning-based HSI super-resolution model using group-block sparsity. By grouping similar 3-D image cubes into clusters and then conduct super-resolution cluster by cluster using 4-D tensor structure, we not only preserve the structure but also achieve sparsity within the cluster due to the collection of similar cubes. We use 4-D tensor Tucker decomposition and impose nonnegative constraints on the dictionaries and group-block sparsity. Numerous experiments demonstrate that the proposed model outperforms many state-of-the-art HSI super-resolution methods.**

***Index Terms*— Hyperspectral imaging (HSI), nonlocal sparse tensor factorization (NLSTF), nonnegative tensor dictionary learning, super-resolution.**

I. Introduction

HYPERSPECTRAL images (HSIs) consist of a number of wavelength bands providing abundant spectral information. Different objects with different colors have a different spectral response. Even for the same object, it shows different spectral response under different status (e.g., a good nut versus a bad nut under the shell). This enables spectral imaging a strong modality to distinguish one object from another. It plays an important role in the fields of remote sensing and various computer vision tasks, such as object classification, tracking, and recognition. In practice, however, HSIs are severely limited in spatial resolution due to hardware constraints. Conversely, multispectral images (MSIs) usually have a high spatial resolution but low spectral resolution. The purpose of HSI super-resolution is to raise the spatial resolution of HSI. It is often done by fusing a high-resolution (HR) MSI and a low-resolution (LR) HSI.

The HSI super-resolution techniques which have been proposed in recent years mainly include the matrix factorization based methods and the tensor factorization based methods. Matrix factorization based methods rewrite an HSI as a 2-D matrix with each row corresponding to a spectral band and treat it as the product of a dictionary and a coefficient matrix. Kawakami *et al.* [1] introduced a sparse matrix factorization method, which reconstructs the HSI by multiplying the learned dictionary from the HR-MSI and sparse coefficients from the LR-HSI. Yokoya *et al.* [2] proposed a coupled nonnegative matrix factorization (CNMF) method for HSI super-resolution by alternately applying nonnegative matrix factorization (NMF) [3] unmixing to LR-HSI and HR-MSI. Meanwhile, a sparse NMF method was proposed by Wycoff *et al.* [4] to exploit both nonnegativity and sparsity on HSI. Simões *et al.* [5] presented a convex formulation for HSI super-resolution (known as HySure) based on vector total variation regularization, which promotes piecewise-smooth solutions with discontinuity along spectral direction. A nonparametric Bayesian sparse representation (BSR) method for the fusion of HSI and MSI was proposed by Akhtar *et al.* [6]. Dong *et al.* [7] proposed a nonnegative structured sparse representation (NSSR) approach based on the prior knowledge about spatio-spectral sparsity of the HSI. Lei *et al.* [8] developed a novel clustering manifold structure-based HSI super-resolution framework. The learned manifold structure from the HR-MSI can well capture the spatial correlation of the target HR-HSI. By organizing 3-D tensors into 2-D images, image factorization-based approaches gain computational simplicity and are often faster but lose high-dimensional structures. Just like when one collapses a 3-D paper origami into a 2-D piece of paper, the 3-D structure is lost even though the creases are still there. In comparison, tensor factorization-based approaches keep data structure better but pay the price of heavier computation.

Spectral data comes with a 3-D tensor structure. Tensor representation and modeling are better in preserving structures and have gained attention recently. They have provided impressive results in multidimensional HSI restoration. Qi *et al.* [9] proposed a tensor factorization-based denoising approach using an intrinsic tensor sparsity (ITS) measure. Their approach considers global correlation along the spectral

Manuscript received August 16, 2019; revised December 7, 2019 and March 7, 2020; accepted March 24, 2020. The work of Haiyang Huang and Jun Liu was supported by the National Key Research and Development Program of China under Grant 2017YFA0604903. The work of Wei Wan was supported by the China Scholarship Council under Grant 201706040141. The work of Weihong Guo was supported by the National Science Foundation under Grant DMS-1521582. *(Corresponding authors: Weihong Guo; Jun Liu.)*

Wei Wan, Haiyang Huang, and Jun Liu are with the Laboratory of Mathematics and Complex Systems (Ministry of Education of China), School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China (e-mail: weiwan@mail.bnu.edu.cn; hhywsg@bnu.edu.cn; jliu@bnu.edu.cn).

Weihong Guo is with the Department of Mathematics, Case Western Reserve University, Cleveland, OH 44106 USA (e-mail: wxg49@case.edu).

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/TGRS.2020.2987530

direction and nonlocal self-similarity along the spatial direction of the underlying HSI. Wei and Fu [10] presented a low-rank Bayesian tensor factorization (LBTF) approach for HSI denoising, which describes the inherent spatial–spectral correlation of HSIs. A tensor robust principle component analysis method based on tensor singular value decomposition (t-SVD) and tensor nuclear norm was proposed by Yang *et al.* [11]. They built a framework for hyperspectral compressive sensing with anomaly detection, which reconstructs the HSI and detects the anomalies simultaneously. For HSI super-resolution, Dian *et al.* [12] proposed a nonlocal sparse tensor factorization (NLSTF) model that induces spatial dictionaries and core tensors from HR-MSI and spectral dictionaries from LR-HSI, respectively. Dian *et al.* [13] further proposed a semiblind fusion method based on NLSTF. However, this method does not combine LR-HSI and HR-MSI to estimate the dictionary and the core tensor for each cluster. Chang *et al.* [14] presented a weighted low-rank tensor recovery model by using high-order singular value decomposition (HOSVD). It updates core tensors iteratively while fixing the dictionaries along every mode. A coupled canonical polyadic decomposition model was proposed by Kanatsoulis *et al.* [15]. This approach, however, did not consider the spatial nonlocal self-similarity and tensor sparsity prior. Very recently, Li *et al.* [16] proposed a coupled sparse tensor factorization (CSTF)-based HSI super-resolution model. The high spatial–spectral correlations in the HR-HSI are modeled by the sparsity of the core tensors. A coupled tensor factorization of the LR-HSI and of the HR-MSI is adopted. However, the nonlocal spatial similarity in HSIs was not modeled. Furthermore, a spatial–spectral sparse representation-based method (SSSR) was presented by Dian *et al.* [17] to exploit the nonlocal spatial similarities, prior to spectral unmixing and a sparse prior to the fusion problem. Dian *et al.* [18] proposed a low tensor-train rank (LTTR)-based HSI super-resolution approach, where an LTTR prior is designed to learn the correlations among the spatial, spectral, and nonlocal modes of each 4-D similar cluster. In addition, a subspace-based low tensor multirank regularization method for HSI super-resolution was proposed by Dian *et al.* [19]. They approximated the target HR-HSI with low-dimensional spectral subspace learned from the LR-HSI and corresponding coefficients estimated with low tensor multirank prior. Xu *et al.* [20] applied a tensor-tensor product-based tensor sparse representation that can preserve the spectral and spatial similarities between nonlocal similar patches in HSIs. Recently, some deep learning-based methods have been proposed for HSI super-resolution and remote sensing restoration. A deep residual pan-sharpening neural network (DRPNN) for the fusion of MSI and panchromatic image was proposed by Wei *et al.* [21]. They employed the nonlinearity of a convolutional neural network to achieve robust and high accuracy fusion. In addition, Dian *et al.* [22] proposed to learn the map between initialized HR-HSI and ground truth via a deep residual convolutional neural network. Then, the deep HSI sharpening method (DHSIS) incorporates the learned deep priors into the LR-HSI and HR-MSI fusion framework. In this article, we mainly study the tensor factorization-based methods in HSI super-resolution. Peng *et al.* [23] extended 2-D image group sparsity to HSI group-block sparsity and developed a nonlocal tensor dictionary learning model for HSI denoising problem. They employed the group-block sparsity prior to model the spatial nonlocal self-similarity and spectral global correlation simultaneously. By using the Tucker decomposition technique on each cluster of the noisy HSI directly, the quality of the restoration can be greatly improved.

Motivated by [23], we extend the group-block sparsity prior to HSI super-resolution. However, the HSI super-resolution problem is different from HSI denoising since it is more ill-posed. Denoising problem only has one input and the proposed fusion-based super-resolution work has two inputs, each one has a convolution relation with the underlying HR image along the spatial direction and the other one along the spectral direction. The data fidelity term thus contains two terms and due to the convolution, the computation is much harder than the denoising problem. Each of the dictionaries and the core tensor are estimated by using the information from both HR-MSI and LR-HSI. The group-block sparsity was based on the clustering of cubes. In the HSI denoising problem, there is only one clustering to be done, i.e., on the noisy input. In the HSI super-resolution problem, the clustering has to be done on two inputs and the clustering has to be matched up during the computation. In the HSI denoising work, the rank was predetermined using AIC/MDL proposed by Wax and Kailath [24] in 1985. The rank of the dictionary matrices are computed from the given noisy HSI beforehand and fixed during the computation. Especially, the dictionary matrices are narrow as they have less columns than rows. The results thus rely heavily on the accuracy of the rank precomputed. In our work, we do not fix the rank but use a more data-driven/adaptive scheme under a redundant scheme: our dictionary matrices are wide with more columns than rows and the group block sparsity is involved in the energy functional to be minimized to automatically determine the rank.

In this article, in order to keep the high-dimensional data structure, we propose a nonnegative coupled HSI super-resolution model from the tensor perspective using a nonlocal characteristic described by group-block sparsity. We first group similar 3-D cubes into clusters using K-means++. Similar 3-D cubes in each cluster are organized as a 4-D tensor instead of collapsing into a 3-D tensor or 2-D matrix. To the best of our knowledge, this is the first effort in using nonlocal sparsity on the 4-D tensor data structure for HSI super-resolution. The spectral correlation and similarity among 3-D cubes bring tensor sparsity under Tucker decomposition. Compared with the state-of-the-art HSI super-resolution methods, the contributions of the proposed method are as follows.

1) To the best of our knowledge, this is probably the first 4-D Tucker decomposition and group-block sparsity-based HSI super-resolution work. We propose to combine nonnegative and nonlocal sparsity for the task of HSI super-resolution. The nonnegativity constraint ensures the nonnegativity of the HR HSI intensity. The nonlocal sparsity denoted by group block sparsity is produced from grouping similar 3-D cubes into clusters. Tucker decomposition approximation is employed on

each cluster represented by a 4-D tensor. The computation is done on each cluster. Both clusters of HR-MSI and the corresponding ones in LR-HSI are coupled for estimating the dictionaries and the core tensor of the HR-HSI.

2) Different than what is done in the literature, the group-block sparsity is implemented in a more automatic data-driven manner. The rank of the dictionary matrices is not computed from the given noisy HSI beforehand and fixed as done in, for instance, [23]. In our work, we use a more data-driven/adaptive scheme under a redundant framework which better promotes sparsity than narrow dictionaries used in some works in the literature.
3) An efficient splitting approach is carefully designed for the proposed HSI super-resolution model and the $L_{1,2}$ norm of group sparsity in 2-D images is used to approximate the group-block sparsity in the tensor sparse coding process.

The rest of this article is organized as follows. In Section II, we introduce basis notations and review some preliminaries on tensors. We introduce the proposed HSI super-resolution model in Section III. Experimental results are laid out in Section IV, and conclusion follows in Section V.

II. Notations and Preliminaries

An Nth-order real-valued tensor is an N-dimensional data array denoted as $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \cdots \times I_N}$. Let $x_{i_1 i_2 \cdots i_N}$ be the $i_1 i_2 \cdots i_N$ entry. The Frobenius norm of $\mathcal{X}$ is here defined by $\|\mathcal{X}\|_F = (\sum_{i_1 i_2 \cdots i_N} x^2_{i_1 i_2 \cdots i_N})^{1/2}$. The mode-$n$ vectors are obtained by fixing every index but the one in the mode n. The mode-n matricization, also known as the unfolding or flattening, of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \cdots \times I_N}$ is denoted as $\mathbf{X}_{(n)} \in \mathbb{R}^{I_n \times (I_1 \cdots I_{n-1} I_{n+1} \cdots I_N)}$ and arranges the mode-n vectors to be the columns of the matrix. The n-mode product of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \cdots \times I_N}$ with a matrix $\mathbf{U} \in \mathbb{R}^{J \times I_n}$ is given by $\mathcal{X} \times_n \mathbf{U} \in \mathbb{R}^{I_1 \times \cdots \times I_{n-1} \times J \times I_{n+1} \times \cdots \times I_N}$ [25].

Definition 1 (Tucker Decomposition): A Tucker decomposition of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \cdots \times I_N}$ is a decomposition of $\mathcal{X}$ as a core tensor multiplied by a matrix along each mode

$$\mathcal{X} = \mathcal{Y} \times_1 \mathbf{A}_1 \times_2 \mathbf{A}_2 \cdots \times_N \mathbf{A}_N. \tag{1}$$

Here, $\mathbf{A}_n \in \mathbb{R}^{I_n \times M_n}$ are the factor matrices (also known as the dictionaries along each mode) and $\mathcal{Y} \in \mathbb{R}^{M_1 \times M_2 \cdots \times M_N}$ is the core tensor, i.e., the coefficient tensor of $\mathcal{X}$ over the N dictionaries.

In particular, the second-order tensor (image) $\mathbf{X}$ can be expressed as

$$\mathbf{X} = \mathbf{A}\mathbf{Y}\mathbf{B}^T$$

where $\mathbf{A}$ and $\mathbf{B}$ are the dictionaries related to mode-1 vector (column) and mode-2 vector (row), respectively, and $\mathbf{Y}$ is the coefficient matrix.

Let the vectorization of a tensor $\mathcal{X}$ be defined as $\text{vec}(\mathcal{X}) \equiv \text{vec}(\mathbf{X}_{(1)}) \in \mathbb{R}^{I_1 I_2 \cdots I_N}$, i.e., the long vector obtained by stacking all the one-mode vectors vertically. The mode-n unfolding of the Tucker decomposition is

$$\mathbf{X}_{(n)} = \mathbf{A}_n \mathbf{Y}_{(n)} (\mathbf{A}_N \otimes \cdots \otimes \mathbf{A}_{n+1} \otimes \mathbf{A}_{n-1} \otimes \cdots \otimes \mathbf{A}_1)^T \tag{2}$$

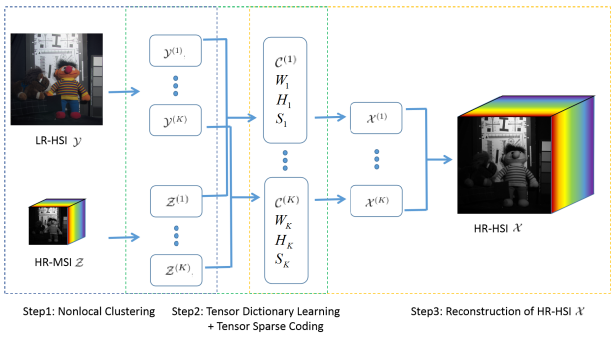


Fig. 1. Illustration of the proposed HSI super-resolution method.

and its vector representation is

$$\mathbf{x} = (\mathbf{A}_N \otimes \mathbf{A}_{N-1} \otimes \cdots \otimes \mathbf{A}_1)\mathbf{y} \tag{3}$$

where $\mathbf{x} = \text{vec}(\mathcal{X})$ and $\mathbf{y} = \text{vec}(\mathcal{Y})$.

III. Proposed Method

The purpose of HSI super-resolution is to estimate the HR-HSI $\mathcal{X} \in \mathbb{R}^{W \times H \times S}$ by fusing the LR-HSI $\mathcal{Y} \in \mathbb{R}^{w \times h \times S}$ with the HR-MSI $\mathcal{Z} \in \mathbb{R}^{W \times H \times s}$, where $w < W, h < H$, and $s < S$. The proposed HSI super-resolution method is shown in Fig. 1.

First, we construct a group of cubes $\{\mathcal{Z}_i\}_{i=1}^n \subset \mathbb{R}^{d_W \times d_H \times s}$ from HR-MSI $\mathcal{Z} \in \mathbb{R}^{W \times H \times s}$ by sweeping all across it with overlaps $l_w = d_W/2$ and $l_h = d_H/2$, where l_w, l_h are divisible into W, H, respectively, and $n = n_w n_h$ is the cube number and $n_w = (W - l_w)/(d_W - l_w), n_h = (H - l_h)/(d_H - l_h)$. Similarly, we can build a group of cubes $\{\mathcal{Y}_i\}_{i=1}^n \subset \mathbb{R}^{d_w \times d_h \times S}$ from LR-HSI $\mathcal{Y}$ and $d_w = d_W/a, d_h = d_H/a$, a is the scaling factor (e.g., $a = 4, d_W = d_H = 8, d_w = d_h = 2$). Note that each cube has full bands. We group the similar cubes of the HR-MSI $\mathcal{Z}$ into K clusters $\{\mathcal{Z}^{(k,j)}\}_{j=1}^{n_k}, k = 1, 2, \ldots, K$ by K-means++, where K is the number of clusters, and n_k is the number of cubes in the kth cluster. $\mathcal{Z}^{(k,j)} \in \mathbb{R}^{d_W \times d_H \times s}$ denotes the jth cube of the kth cluster. For convenience, we combine the similar cubes in the kth cluster together to formulate a fourth-order tensor: $\mathcal{Z}^{(k)} \in \mathbb{R}^{d_W \times d_H \times s \times n_k}$. Correspondingly, based on spatial alignment, the cubes of LR-HSI $\mathcal{Y} \in \mathbb{R}^{w \times h \times S}$ are grouped into K clusters the same way as $\mathcal{Z}$: $\{\mathcal{Y}^{(k,j)}\}_{j=1}^{n_k} \subset \mathbb{R}^{d_w \times d_h \times S}, k = 1, 2, \ldots, K$. Similarly, we organize 3-D tensors $\{\mathcal{Y}^{(k,j)}\}_{j=1}^{n_k}$ into 4-D tensors $\mathcal{Y}^{(k)} \in \mathbb{R}^{d_w \times d_h \times S \times n_k}$ so that $\mathcal{Y}^{(k)}(:,:,:,j) = \mathcal{Y}^{(k,j)}$. The HR-HSI $\mathcal{X}$ will be reconstructed cluster by cluster and each cluster is denoted as $\mathcal{X}^{(k)} \in \mathbb{R}^{d_W \times d_H \times S \times n_k}, k = 1, 2, \ldots, K$.

A. Reconstruction Constraints

As pointed out in [12] and [16], the acquired LR-HSI $\mathcal{Y}$ is the spatially (along first and second modes) downgraded version of $\mathcal{X}$

$$\mathcal{Y} = \mathcal{X} \times_1 \mathbf{P}_1 \times_2 \mathbf{P}_2 \tag{4}$$

where $\mathbf{P}_1 \in \mathbb{R}^{w \times W}$ and $\mathbf{P}_2 \in \mathbb{R}^{h \times H}$ are the blurring and downsampling matrices along the first (width) mode and second (height) modes, respectively. In addition, the HR-MSI $\mathcal{Z}$

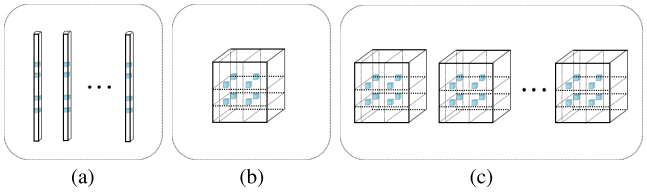


Fig. 2. Illustration of group sparsity, block sparsity, and group-block sparsity. Blue cubes represent nonzero entries. (a) Group sparsity. (b) Block sparsity. (c) Group-block sparsity.

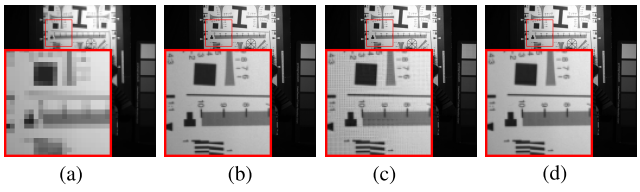


Fig. 3. Reconstruction results of $\mathcal{C}^{(k)}$ (c) with nonnegative constrains and those (d) without nonnegative constrains. (a) LR-HSI. (b) Ground truth. (c) PSNR = 40.28. (d) PSNR = 46.84.

contains less bands than the HR-HSI $\mathcal{X}$ and is the spectrally downgraded version of $\mathcal{X}$

$$\mathcal{Z} = \mathcal{X} \times_3 \mathbf{P}_3 \tag{5}$$

where $\mathbf{P}_3 \in \mathbb{R}^{s\times S}$ is the blurring and downsampling matrix of the third (spectral) mode. Noise was not mentioned in (4) and (5) for notation simplicity. It is considered in the models.

Thus, for the kth cluster, we have

$$\mathcal{Y}^{(k)} = \mathcal{X}^{(k)} \times_1 \mathbf{P}_1^* \times_2 \mathbf{P}_2^* \tag{6}$$

$$\mathcal{Z}^{(k)} = \mathcal{X}^{(k)} \times_3 \mathbf{P}_3 \tag{7}$$

where $\mathbf{P}_1^* \in \mathbb{R}^{d_w\times d_W}$ and $\mathbf{P}_2^* \in \mathbb{R}^{d_h\times d_H}$ have the same blurring and downsampling manner with $\mathbf{P}_1 \in \mathbb{R}^{w\times W}$ and $\mathbf{P}_2 \in \mathbb{R}^{h\times H}$, respectively. They differ only by size. Thus, we drop the notation $*$ in $\mathbf{P}_1$ and $\mathbf{P}_2$ for convenience.

B. Group-Block Sparsity Prior

We reconstruct the HR-HSI $\mathcal{X}$ cluster by cluster. Since the cubes in the same cluster are similar and should share certain characteristics, we attempt to enforce each cluster share the similar atoms from the spatial and spectral dictionaries and then use group-block sparsity to describe.

Definition 2 (Group-Block Sparsity [23]): For a higher-order core tensor $\mathcal{C} \in \mathbb{R}^{I_1\times I_2\cdots\times I_{N'}}$, its group-block sparsity with respect to the factors $\mathbf{A}_n \in \mathbb{R}^{M_n\times I_n}, n = 1, 2, \ldots, N (N \le N')$ is $\|\mathcal{C}\|_B = (m_1, m_2, \ldots, m_N)$ if and only if the smallest index subsets $\mathcal{I}_1, \mathcal{I}_2, \ldots, \mathcal{I}_N$ satisfying $c_{i_1 i_2\cdots i_{N'}} = 0$ for all $(i_1, i_2, \ldots, i_N) \notin \mathcal{I}_1 \times \mathcal{I}_2 \times \cdots \times \mathcal{I}_N$ contain $m_1, m_2, \ldots, m_N$ elements, respectively.

To help understand group-block sparsity, in Fig. 2, we illustrate the relation/difference among group-block sparsity, block sparsity, and group sparsity using a 4-D tensor, a 3-D tensor, and a 2-D matrix, respectively. As shown in Fig. 2(a), group sparsity of a 2-D matrix refers to all columns share zeros at the same locations, i.e., the matrix contains many zero rows. Block sparsity [see Fig. 2(b)] of a 3-D core tensor refers to nonzero entries of it are located within a block, leading to slices of zeros. It is motivated by the fact that block sparse signals are more likely to occur than totally random sparse ones [26]. A 4-D tensor can be considered as an array of 3-D tensors as shown in Fig. 2(c). Group-block sparsity of a 4-D tensor can be seen as the extended case of the group sparsity if you treat each 3-D tensor in the 4-D tensor as a column vector in a 2-D matrix. The 3-D tensors share locations of nonzero blocks as shown in Fig. 2(c). Moreover, when there is only one core tensor in each cluster, i.e., there is only one 3-D tensor in the array, the 4-D tensor is actually a 3-D tensor, its group-block sparsity is the same as the block sparsity.

Then, based on group-block sparsity prior and the Tucker decomposition, the kth cluster $\mathcal{X}^{(k)}$ of the HR-HSI $\mathcal{X}$ can be formulated as

$$\mathcal{X}^{(k)} = \mathcal{C}^{(k)} \times_1 \mathbf{W}_k \times_2 \mathbf{H}_k \times_3 \mathbf{S}_k, k = 1, 2, \ldots, K \quad \text{with } \|\mathcal{C}^{(k)}\|_B \preceq \left(m_k^W, m_k^H, m_k^S\right) \tag{8}$$

where the 2-D matrices $\mathbf{W}_k \in \mathbb{R}^{d_W\times r_k^W}$, $\mathbf{H}_k \in \mathbb{R}^{d_H\times r_k^H}$, and $\mathbf{S}_k \in \mathbb{R}^{S\times r_k^S}$ denote the dictionaries of the width mode with r_k^W atoms, the height mode with r_k^H atoms, and the spectral mode with r_k^S atoms, respectively. The 4-D tensor $\mathcal{C}^{(k)} \in \mathbb{R}^{r_k^W\times r_k^H\times r_k^S\times n_k}$ is the coefficient of $\mathcal{X}^{(k)}$ over the three dictionaries. Here, $m_k^W < r_k^W$, $m_k^H < r_k^H$, and $m_k^S < r_k^S$.

Combining the relationship (10) and (11) and the group-block sparsity, we propose the following model:

$$\begin{aligned}\min_{\mathbf{W}_k\ge 0,\mathbf{H}_k\ge 0,\mathbf{S}_k\ge 0,\mathcal{C}^{(k)}} & \times\frac{1}{2}\sum_{k=1}^{K}\|\mathcal{Y}^{(k)} - \mathcal{C}^{(k)} \times_1 (\mathbf{P}_1\mathbf{W}_k) \times_2 (\mathbf{P}_2\mathbf{H}_k) \times_3 \mathbf{S}_k\|_F^2 \\ & +\frac{\lambda}{2}\sum_{k=1}^{K}\|\mathcal{Z}^{(k)} - \mathcal{C}^{(k)} \times_1 \mathbf{W}_k \times_2 \mathbf{H}_k \times_3 (\mathbf{P}_3\mathbf{S}_k)\|_F^2 \\ & \text{s.t. } \|\mathcal{C}^{(k)}\|_B \preceq \left(m_k^W, m_k^H, m_k^S\right).\end{aligned} \tag{9}$$

We enforce nonnegativity of the dictionary matrices as image intensity is nonnegative. We, however, do not force $\mathcal{C}^{(k)}$, the core tensor, to be nonnegative to avoid artifacts due to strong constraints. After all, we do not need every component to be nonnegative to achieve overall nonnegativity. In Fig. 3, we compare the reconstruction results with and without nonnegative constraint on $\mathcal{C}^{(k)}$. We observe artifacts in Fig. 3(c) (with constraint on $\mathcal{C}^{(k)}$) while the one without constraint [in Fig. 3(d)] leads to better results in both PSNR and visual effects. Since (9) is separable with respect to k, the cluster index, the original nonlocal HSI super-resolution problem can be decomposed into K independent problems

$$\begin{aligned}\min_{\mathbf{W}_k\ge 0,\mathbf{H}_k\ge 0,\mathbf{S}_k\ge 0,\mathcal{C}^{(k)}} & \times\frac{1}{2}\|\mathcal{Y}^{(k)} - \mathcal{C}^{(k)} \times_1 (\mathbf{P}_1\mathbf{W}_k) \times_2 (\mathbf{P}_2\mathbf{H}_k) \times_3 \mathbf{S}_k\|_F^2 \\ & +\frac{\lambda}{2}\|\mathcal{Z}^{(k)} - \mathcal{C}^{(k)} \times_1 \mathbf{W}_k \times_2 \mathbf{H}_k \times_3 (\mathbf{P}_3\mathbf{S}_k)\|_F^2 \\ & \text{s.t. } \|\mathcal{C}^{(k)}\|_B \preceq \left(m_k^W, m_k^H, m_k^S\right), \quad \text{where } k = 1, 2, \ldots, K.\end{aligned} \tag{10}$$

Furthermore, the abovementioned problem can be equivalently reformulated as

$$\begin{aligned}\min_{\mathbf{W}_k\geq 0,\mathbf{H}_k\geq 0,\mathbf{S}_k\geq 0,\mathcal{C}^{(k,j)}} & \times\frac{1}{2}\sum_{j=1}^{n_k}\|\mathcal{Y}^{(k,j)}-\mathcal{C}^{(k,j)}\times_1(\mathbf{P}_1\mathbf{W}_k)\times_2(\mathbf{P}_2\mathbf{H}_k)\times_3\mathbf{S}_k\|_F^2\\ &+\frac{\lambda}{2}\sum_{j=1}^{n_k}\|\mathcal{Z}^{(k,j)}-\mathcal{C}^{(k,j)}\times_1\mathbf{W}_k\times_2\mathbf{H}_k\times_3(\mathbf{P}_3\mathbf{S}_k)\|_F^2\\ \text{s.t. } & \|\mathcal{C}^{(k,j)}\|_B\preceq\left(m_k^W,m_k^H,m_k^S\right),\\ & \text{for each } k=1,2,\ldots,K.\end{aligned}\tag{11}$$

We use an alternating minimization method to solve (11). We alternate some subproblems iteratively until convergence. In what follows, we explain how to solve each subproblem.

C. Tensor Dictionary Learning

To update $\mathbf{W}_k$, the subproblem can be rewritten in mode-1 as follows:

$$\min_{\mathbf{W}_k\geq 0}\frac{1}{2}\|\mathbf{Y}_1-\mathbf{P}_1\mathbf{W}_k\mathbf{G}_1\|_F^2+\frac{\lambda}{2}\|\mathbf{Z}_1-\mathbf{W}_k\mathbf{Q}_1\|_F^2\tag{12}$$

where $\mathbf{Y}_1=[\mathbf{Y}_{(1)}^{(k,1)},\mathbf{Y}_{(1)}^{(k,2)},\ldots,\mathbf{Y}_{(1)}^{(k,n_k)}]$, $\mathbf{Z}_1=[\mathbf{Z}_{(1)}^{(k,1)},\mathbf{Z}_{(1)}^{(k,2)},\ldots,\mathbf{Z}_{(1)}^{(k,n_k)}]$, $\mathbf{G}_1=[\mathbf{C}_{(1)}^{(k,1)}\mathbf{G}_{(1)},\mathbf{C}_{(1)}^{(k,2)}\mathbf{G}_{(1)},\ldots,\mathbf{C}_{(1)}^{(k,n_k)}\mathbf{G}_{(1)}]$, and $\mathbf{Q}_1=[\mathbf{C}_{(1)}^{(k,1)}\mathbf{Q}_{(1)},\mathbf{C}_{(1)}^{(k,2)}\mathbf{Q}_{(1)},\ldots,\mathbf{C}_{(1)}^{(k,n_k)}\mathbf{Q}_{(1)}]$. In addition, $\mathbf{Y}_{(1)}^{(k,j)}$, $\mathbf{Z}_{(1)}^{(k,j)}$, and $\mathbf{C}_{(1)}^{(k,j)}$ are mode-1 unfolding matrices of tensors $\mathcal{Y}^{(k,j)}$, $\mathcal{Z}^{(k,j)}$, and $\mathcal{C}^{(k,j)}$, respectively, and $\mathbf{G}_{(1)}=(\mathbf{S}_k\otimes(\mathbf{P}_2\mathbf{H}_k))^T$, $\mathbf{Q}_{(1)}=((\mathbf{P}_3\mathbf{S}_k)\otimes\mathbf{H}_k)^T$.

By using $\text{vec}(\mathbf{AXB})=(\mathbf{B}^T\otimes\mathbf{A})\text{vec}(\mathbf{X})$, and by letting $\mathbf{w}_k=\text{vec}(\mathbf{W}_k)$, we can easily rewrite the abovementioned problem as

$$\min_{\mathbf{w}_k}\frac{1}{2}\|\mathbf{u}_1-\mathbf{D}_1\mathbf{w}_k\|_2^2+\mathcal{X}_+(\mathbf{w}_k)\tag{13}$$

where $\mathbf{u}_1=[\text{vec}(\mathbf{Y}_1)^T,\sqrt{\lambda}\text{vec}(\mathbf{Z}_1)^T]^T$, $\mathbf{D}_1=[(\mathbf{G}_1^T\otimes\mathbf{P}_1)^T,\sqrt{\lambda}(\mathbf{Q}_1^T\otimes\mathbf{I})^T]^T$, and

$$\mathcal{X}_+(\mathbf{w}_k)=\begin{cases}0, & \mathbf{w}_k\geq 0\\ \infty, & \text{otherwise.}\end{cases}$$

Since $\mathbf{D}_1^T\mathbf{D}_1$ is sometimes not of full rank, it is not accurate to solve the abovementioned minimization problem directly by the projection method. To better solve for $\mathbf{w}_k$, we use an auxiliary variable $\widetilde{\mathbf{w}}_k$ to split $\mathbf{w}_k$, and rewrite (13) as

$$\begin{aligned}\min_{\mathbf{w}_k,\widetilde{\mathbf{w}}_k}\ & \frac{1}{2}\|\mathbf{u}_1-\mathbf{D}_1\mathbf{w}_k\|_2^2+\mathcal{X}_+\left(\widetilde{\mathbf{w}}_k\right)\\ \text{s.t. }\ & \widetilde{\mathbf{w}}_k=\mathbf{w}_k.\end{aligned}\tag{14}$$

Then, we use standard ADMM to produce the following iteration scheme:

$$\begin{cases}\mathbf{w}_k^{t+1}=\arg\min_{\mathbf{w}_k}\left\{\frac{1}{2}\|\mathbf{u}_1-\mathbf{D}_1\mathbf{w}_k\|_2^2+\frac{\eta_w}{2}\|\widetilde{\mathbf{w}}_k^t-\mathbf{w}_k-\mathbf{v}_1^t\|_2^2\right\}\\ \widetilde{\mathbf{w}}_k^{t+1}=\arg\min_{\widetilde{\mathbf{w}}_k}\left\{\mathcal{X}_+\left(\widetilde{\mathbf{w}}_k\right)+\frac{\eta_w}{2}\|\widetilde{\mathbf{w}}_k-\mathbf{w}_k^{t+1}-\mathbf{v}_1^t\|_2^2\right\}\\ \mathbf{v}_1^{t+1}=\mathbf{v}_1^t+\eta_w\left(\mathbf{w}_k^{t+1}-\widetilde{\mathbf{w}}_k^{t+1}\right).\end{cases}\tag{15}$$

Each subproblem has a closed-form solution, which is written as

$$\mathbf{w}_k^{t+1}=\left(\mathbf{D}_1^T\mathbf{D}_1+\eta_w\mathbf{I}\right)^{-1}\left(\mathbf{D}_1^T\mathbf{u}_1+\eta_w\widetilde{\mathbf{w}}_k^t-\eta_w\mathbf{v}_1^t\right)\tag{16}$$

$$\widetilde{\mathbf{w}}_k^{t+1}=\max\left(\mathbf{w}_k^{t+1}+\mathbf{v}_1^t,0\right).\tag{17}$$

Equation (16) can be solved using QR or Choleskey factorization.

$\mathbf{H}_k$ and $\mathbf{S}_k$ can be updated similarly. For $\mathbf{H}_k$, use $\mathbf{h}_k=\text{vec}(\mathbf{H}_k)$ and rewrite (11) in mode-2 matricization as follows:

$$\min_{\mathbf{h}_k,\widetilde{\mathbf{h}}_k}\frac{1}{2}\|\mathbf{u}_2-\mathbf{D}_2\mathbf{h}_k\|_2^2+\mathcal{X}_+\left(\widetilde{\mathbf{h}}_k\right),\quad\text{s.t. }\widetilde{\mathbf{h}}_k=\mathbf{h}_k.\tag{18}$$

$\mathbf{h}_k$ can be calculated by the following iteration scheme:

$$\begin{cases}\mathbf{h}_k^{t+1}=\arg\min_{\mathbf{h}_k}\left\{\frac{1}{2}\|\mathbf{u}_2-\mathbf{D}_2\mathbf{h}_k\|_2^2+\frac{\eta_h}{2}\|\widetilde{\mathbf{h}}_k^t-\mathbf{h}_k-\mathbf{v}_2^t\|_2^2\right\}\\ \widetilde{\mathbf{h}}_k^{t+1}=\arg\min_{\widetilde{\mathbf{h}}_k}\left\{\mathcal{X}_+\left(\widetilde{\mathbf{h}}_k\right)+\frac{\eta_h}{2}\|\widetilde{\mathbf{h}}_k-\mathbf{h}_k^{t+1}-\mathbf{v}_2^t\|_2^2\right\}\\ \mathbf{v}_2^{t+1}=\mathbf{v}_2^t+\eta_h\left(\mathbf{h}_k^{t+1}-\widetilde{\mathbf{h}}_k^{t+1}\right)\end{cases}\tag{19}$$

where $\mathbf{u}_2=[\text{vec}(\mathbf{Y}_2)^T,\sqrt{\lambda}\text{vec}(\mathbf{Z}_2)^T]^T$, $\mathbf{Y}_2=[\mathbf{Y}_{(2)}^{(k,1)},\mathbf{Y}_{(2)}^{(k,2)},\ldots,\mathbf{Y}_{(2)}^{(k,n_k)}]$, and $\mathbf{Z}_2=[\mathbf{Z}_{(2)}^{(k,1)},\mathbf{Z}_{(2)}^{(k,2)},\ldots,\mathbf{Z}_{(2)}^{(k,n_k)}]$. In addition, $\mathbf{D}_2=[(\mathbf{G}_2^T\otimes\mathbf{P}_2)^T,\sqrt{\lambda}(\mathbf{Q}_2^T\otimes\mathbf{I})^T]^T$, $\mathbf{G}_2=[\mathbf{C}_{(2)}^{(k,1)}\mathbf{G}_{(2)},\mathbf{C}_{(2)}^{(k,2)}\mathbf{G}_{(2)},\ldots,\mathbf{C}_{(2)}^{(k,n_k)}\mathbf{G}_{(2)}]$, $\mathbf{G}_{(2)}=(\mathbf{S}_k\otimes(\mathbf{P}_1\mathbf{W}_k))^T$, $\mathbf{Q}_2=[\mathbf{C}_{(2)}^{(k,1)}\mathbf{Q}_{(2)},\mathbf{C}_{(2)}^{(k,2)}\mathbf{Q}_{(2)},\ldots,\mathbf{C}_{(2)}^{(k,n_k)}\mathbf{Q}_{(2)}]$, and $\mathbf{Q}_{(2)}=((\mathbf{P}_3\mathbf{S}_k)\otimes\mathbf{W}_k)^T$.

For $\mathbf{S}_k$, mode-3 matricization is used

$$\min_{\mathbf{s}_k,\widetilde{\mathbf{s}}_k}\frac{1}{2}\|\mathbf{u}_3-\mathbf{D}_3\mathbf{s}_k\|_2^2+\mathcal{X}_+\left(\widetilde{\mathbf{s}}_k\right),\quad\text{s.t. }\widetilde{\mathbf{s}}_k=\mathbf{s}_k.\tag{20}$$

$\mathbf{s}_k$ can be found by solving

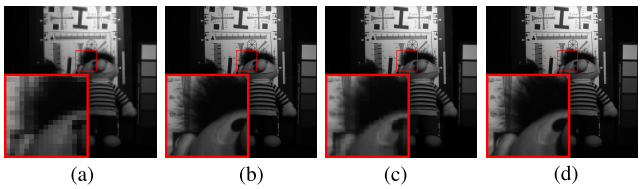
$$\begin{cases}\mathbf{s}_k^{t+1}=\arg\min_{\mathbf{s}_k}\left\{\frac{1}{2}\|\mathbf{u}_3-\mathbf{D}_3\mathbf{s}_k\|_2^2+\frac{\eta_s}{2}\|\widetilde{\mathbf{s}}_k^t-\mathbf{s}_k-\mathbf{v}_3^t\|_2^2\right\}\\ \widetilde{\mathbf{s}}_k^{t+1}=\arg\min_{\widetilde{\mathbf{s}}_k}\left\{\mathcal{X}_+\left(\widetilde{\mathbf{s}}_k\right)+\frac{\eta_s}{2}\|\widetilde{\mathbf{s}}_k-\mathbf{s}_k^{t+1}-\mathbf{v}_3^t\|_2^2\right\}\\ \mathbf{v}_3^{t+1}=\mathbf{v}_3^t+\eta_s\left(\mathbf{s}_k^{t+1}-\widetilde{\mathbf{s}}_k^{t+1}\right)\end{cases}\tag{21}$$

where $\mathbf{s}_k=\text{vec}(\mathbf{S}_k)$, $\mathbf{u}_3=[\text{vec}(\mathbf{Y}_3)^T,\sqrt{\lambda}\text{vec}(\mathbf{Z}_3)^T]^T$, $\mathbf{Y}_3=[\mathbf{Y}_{(3)}^{(k,1)},\mathbf{Y}_{(3)}^{(k,2)},\ldots,\mathbf{Y}_{(3)}^{(k,n_k)}]$, $\mathbf{Z}_3=[\mathbf{Z}_{(2)}^{(k,1)},\mathbf{Z}_{(3)}^{(k,2)},\ldots,\mathbf{Z}_{(2)}^{(k,n_k)}]$, $\mathbf{D}_3=[(\mathbf{G}_3^T\otimes\mathbf{I})^T,\sqrt{\lambda}(\mathbf{Q}_3^T\otimes\mathbf{P}_3)^T]^T$, $\mathbf{G}_3=[\mathbf{C}_{(3)}^{(k,1)}\mathbf{G}_{(3)},\mathbf{C}_{(3)}^{(k,2)}\mathbf{G}_{(3)},\ldots,\mathbf{C}_{(3)}^{(k,n_k)}\mathbf{G}_{(3)}]$, $\mathbf{G}_{(3)}=((\mathbf{P}_2\mathbf{H}_k)\otimes(\mathbf{P}_1\mathbf{W}_k))^T$, and $\mathbf{Q}_3=[\mathbf{C}_{(3)}^{(k,1)}\mathbf{Q}_{(3)},\mathbf{C}_{(3)}^{(k,2)}\mathbf{Q}_{(3)},\ldots,\mathbf{C}_{(3)}^{(k,n_k)}\mathbf{Q}_{(3)}]$, $\mathbf{Q}_{(3)}=(\mathbf{H}_k\otimes\mathbf{W}_k)^T$.

D. Tensor Sparse Coding

Once the dictionaries $\mathbf{W}_k$, $\mathbf{H}_k$, and $\mathbf{S}_k$ are updated, our next problem is to update $\mathcal{C}^{(k)}$. For fixed $\mathbf{W}_k$, $\mathbf{H}_k$, and $\mathbf{S}_k$, $C^{(k)}$ is updated by solving

$$\begin{aligned}\min_{\mathcal{C}^{(k,j)}}\ & \frac{1}{2}\sum_{j=1}^{n_k}\|\mathcal{Y}^{(k,j)}-\mathcal{C}^{(k,j)}\times_1(\mathbf{P}_1\mathbf{W}_k)\times_2(\mathbf{P}_2\mathbf{H}_k)\times_3\mathbf{S}_k\|_F^2\\ &+\frac{\lambda}{2}\sum_{j=1}^{n_k}\|\mathcal{Z}^{(k,j)}-\mathcal{C}^{(k,j)}\times_1\mathbf{W}_k\times_2\mathbf{H}_k\times_3(\mathbf{P}_3\mathbf{S}_k)\|_F^2\\ \text{s.t. }\ & \|\mathcal{C}^{(k,j)}\|_B\preceq\left(m_k^W,m_k^H,m_k^S\right),\quad\text{where } k=1,2,\ldots,K.\end{aligned}\tag{22}$$



(a) (b) (c) (d)

Fig. 4. Reconstruction results by using (c) $\ell_{0,\infty}$ norm and (d) $\ell_{1,2}$ norm. (a) LR-HSI. (b) Ground truth. (c) PSNR = 34.99. (d) PSNR = 46.84.

According to the relationship between the Tucker model and Kronecker representation, the abovementioned problem can be formulated as

$$\min_{\mathbf{C}^{(k)}} \frac{1}{2}\|\mathbf{Y}^{(k)} - \mathbf{D}_1\mathbf{C}^{(k)}\|_F^2 + \frac{\lambda}{2}\|\mathbf{Z}^{(k)} - \mathbf{D}_2\mathbf{C}^{(k)}\|_F^2 + \lambda_c\|\mathbf{C}^{(k)}\|_{p,q} \tag{23}$$

where $\mathbf{C}^{(k)} = [\mathbf{c}^{(k,1)}, \mathbf{c}^{(k,2)}, \ldots, \mathbf{c}^{(k,n_k)}]$, $\mathbf{Y}^{(k)} = [\mathbf{y}^{(k,1)}, \mathbf{y}^{(k,2)}, \ldots, \mathbf{y}^{(k,n_k)}]$, and $\mathbf{Z}^{(k)} = [\mathbf{z}^{(k,1)}, \mathbf{z}^{(k,2)}, \ldots, \mathbf{z}^{(k,n_k)}]$ are matrices, $\mathbf{c}^{(k,j)} = \text{vec}(\mathcal{C}^{(k,j)})$, $\mathbf{y}^{(k,j)} = \text{vec}(\mathcal{Y}^{(k,j)})$, and $\mathbf{z}^{(k,j)} = \text{vec}(\mathcal{Z}^{(k,j)})$ are vectors by stacking all the mode-1 vectors of tensor $\mathcal{C}^{(k,j)}$, $\mathcal{Y}^{(k,j)}$, and $\mathcal{Z}^{(k,j)}$, respectively, and the matrices $\mathbf{D}_1 = \mathbf{S}_k \otimes (\mathbf{P}_2\mathbf{H}_k) \otimes (\mathbf{P}_1\mathbf{W}_k) \in \mathbb{R}^{d_w d_h S \times r_k^W r_k^H r_k^S}$, $\mathbf{D}_2 = (\mathbf{P}_3\mathbf{S}_k) \otimes \mathbf{H}_k \otimes \mathbf{W}_k \in \mathbb{R}^{d_W d_H s \times r_k^W r_k^H r_k^S}$ are the dictionaries. Note the first two terms in (23) are identical to the energy function in (22). The only difference is the data structure. With the matricization of $\mathcal{C}^{(k)}$, the group-block sparsity is relaxed to the following group sparsity regularizer [27] in 2-D images:

$$\|\mathbf{A}\|_{p,q} := \sum_{i=1}^{N} \|\alpha_i\|_q^p \tag{24}$$

where α_i denotes the i-th row of $\mathbf{A}$. The pair (p, q) is usually set as (1, 2) or (0, ∞). In the $\ell_{0,\infty}$ case, this optimization problem is intractable, and could be solved by a greedy approach such as simultaneous orthogonal matching pursuit (SOMP). We have observed experimentally that the reconstructed HSI is in general of better quality when using the $\ell_{1,2}$ norm rather than the pseudo $\ell_{0,\infty}$ norm, as shown in Fig. 4. Here, we use the convex $\ell_{1,2}$ norm for updating $\mathbf{C}^{(k)}$, which can be solved efficiently with ADMM. First, (23) can be equivalently simplified as

$$\min_{\mathbf{C}^{(k)}} \frac{1}{2}\|\mathbf{U}^{(k)} - \mathbf{D}\mathbf{C}^{(k)}\|_F^2 + \lambda_c\|\mathbf{C}^{(k)}\|_{1,2} \tag{25}$$

where $\mathbf{U}^{(k)} = \begin{bmatrix} \mathbf{Y}^{(k)} \\ \sqrt{\lambda}\mathbf{Z}^{(k)} \end{bmatrix}$ and $\mathbf{D} = \begin{bmatrix} \mathbf{D}_1 \\ \sqrt{\lambda}\mathbf{D}_2 \end{bmatrix}$. To solve $\mathbf{C}^{(k)}$, we add an auxiliary variable $\mathbf{B}^{(k)}$ to substitute $\mathbf{C}^{(k)}$, and rewrite the abovementioned problem as

$$\min_{\mathbf{C}^{(k)}} \frac{1}{2}\|\mathbf{U}^{(k)} - \mathbf{D}\mathbf{B}^{(k)}\|_F^2 + \lambda_c\|\mathbf{C}^{(k)}\|_{1,2}, \text{ s.t. } \mathbf{B}^{(k)} = \mathbf{C}^{(k)}. \tag{26}$$

Then we use standard ADMM to produce the following scheme:

$$\begin{cases} \mathbf{B}^{(k)^{t+1}} = \arg\min_{\mathbf{B}^{(k)}} \left\{ \frac{1}{2}\|\mathbf{U}^{(k)} - \mathbf{D}\mathbf{B}^{(k)}\|_F^2 + \frac{\eta_c}{2}\|\mathbf{B}^{(k)} - \mathbf{C}^{(k)^t} - \mathbf{V}^t\|_F^2 \right\} \\ \mathbf{C}^{(k)^{t+1}} = \arg\min_{\mathbf{C}^{(k)}} \left\{ \lambda_c\|\mathbf{C}^{(k)}\|_{1,2} + \frac{\eta_c}{2}\|\mathbf{B}^{(k)^{t+1}} - \mathbf{C}^{(k)} - \mathbf{V}^t\|_F^2 \right\} \\ \mathbf{V}^{t+1} = \mathbf{V}^t + \eta_c\left(\mathbf{C}^{(k)^{t+1}} - \mathbf{B}^{(k)^{t+1}}\right). \end{cases} \tag{27}$$

Algorithm 1 Nonlocal Sparse Tensor Decomposition-Based HSI Super-Resolution

Input: LR-HSI $\mathcal{Y}$, HR-MSI $\mathcal{Z}$, $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$ and $\lambda > 0$.
Output: HR-HSI $\mathcal{X}$.
Group the similar cubes of $\mathcal{Z}$ and $\mathcal{Y}$ into cluster tensors $\mathcal{Z}^{(k)}$ and $\mathcal{Y}^{(k)}$ by K-means ++, and initialize $\mathbf{W}_k, \mathbf{H}_k$ and $\mathbf{S}_k$ by using Tensor factorization technique on $\mathcal{Z}^{(k)}$ and $\mathcal{Y}^{(k)}$, $k = 1, \ldots, K$. For each cluster k,
repeat
  step 1. Update core tensor $\mathcal{C}^{(k)}$ by solving (27).
  step 2. Update dictionary $\mathbf{W}_k$ by solving (15).
  step 3. Update dictionary $\mathbf{H}_k$ by solving (19).
  step 4. Update dictionary $\mathbf{S}_k$ by solving (21).
until stopping criterion is satisfied.
Estimate $\mathcal{X}^{(k)}$ by the Equation (8). Reformulate $\mathcal{X}^{(k)}$ to obtain $\mathcal{X}$.

For subproblem $\mathbf{B}^{(k)^{t+1}}$, it has a closed-form solution

$$\mathbf{B}^{(k)^{t+1}} = \left(\mathbf{D}^T\mathbf{D} + \eta_c\mathbf{I}\right)^{-1}\left(\mathbf{D}^T\mathbf{U}^{(k)} + \eta_c\mathbf{C}^{(k)^t} + \eta_c\mathbf{V}^t\right). \tag{28}$$

Observe that the subproblem $\mathbf{C}^{(k)^{t+1}}$ can be separated with respect to each row of $\mathbf{C}^{(k)}$. Focusing on the ith row of $\mathbf{C}^{(k)}$, denoted as $\mathbf{C}_i^{(k)}$, it suffices to solve a problem in the following form:

$$\min_{\mathbf{C}_i^{(k)}} \lambda_c\|\mathbf{C}_i^{(k)}\|_2 + \frac{\eta_c}{2}\|\mathbf{B}_i^{(k)^{t+1}} - \mathbf{C}_i^{(k)} - \mathbf{V}_i^t\|_2^2. \tag{29}$$

It has a closed form solution in terms of shrinkage

$$\mathbf{C}_i^{(k)^{t+1}} = \mathcal{S}\left(\mathbf{B}_i^{(k)^{t+1}} - \mathbf{V}_i^t, \frac{\lambda_c}{\eta_c}\right) \tag{30}$$

where $\mathcal{S}$ is a shrinkage operator and $\mathcal{S}(f, \mu) = (f/\|f\|_2)\max\{\|f\|_2 - \mu, 0\}$.

Once the dictionaries $\mathbf{W}_k, \mathbf{H}_k, \mathbf{S}_k$, and core tensors $\{\mathcal{C}^{(k,j)}\}_{j=1}^{n_k}$ are known, all overlapping HR-HSI cubes can be estimated. Finally, the estimated cube sets can be returned to the original place to reconstruct the HR-HSI $\mathcal{X}$. Overlapping regions take the average as intensity.

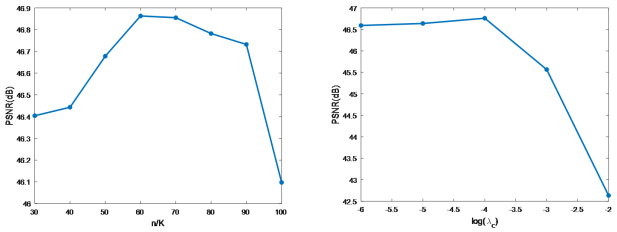
We now summarize the proposed algorithm in Algorithm 1.

In addition, we state the convergence of the ADMM algorithm for $\mathbf{W}_k$, $\mathbf{H}_k$, $\mathbf{S}_k$, and $\mathcal{C}^{(k)}$ in the following theorem. The basic convergence result can be found in several references, such as [28].

Theorem 1: Let the constant $\eta_w > 0, \eta_h > 0, \eta_s > 0$, and $\eta_c > 0$ be given and suppose that there exists a KKT point for problem (14), (18), (20), and (26), respectively. Then these sequences $(\mathbf{w}_k^t, \widetilde{\mathbf{w}}_k^t)$, $(\mathbf{h}_k^t, \widetilde{\mathbf{h}}_k^t)$, $(\mathbf{s}_k^t, \widetilde{\mathbf{s}}_k^t)$, and $(\mathbf{B}^{(k)^t}, \mathbf{C}^{(k)^t})$ generated by the iteration scheme (15), (19), (21), and (27) converge to some KKT point, respectively.

IV. Numerical Experiments

In this section, we numerically demonstrate the superior performance of the proposed model. We compare it with several related HSI super-resolution methods: subspace regularization method HySure [5], CNMF [2], and CSTF [16]. We test the effectiveness of our proposed method on the CAVE database [29], Harvard database [30], Indian Pines



(a) (b)

Fig. 5. PSNR curve as a function of parameters (a)n/K and (b)$\log(\lambda_c)$.

image [31], and the Pavia image [32]. The CAVE database contains 32 HSIs, each with 31 spectral bands, ranging from 400 to 700 nm in 10-nm steps, and a spatial resolution of 512 × 512. The Harvard database containing 50 HSIs of size 1392 × 1040 × 31 was captured with the wavelengths in the range of 420–720 nm at an interval of 10 nm. For the convenience of spatial sampling, we choose the upper left 1024×1024 pixels as the ground truth image. The Indian Pines image was obtained by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor, and the whole image contains 145 × 145 pixels and 224 spectral bands in the wavelength range 400–2500 nm. In our experiments, a subimage of size 120 × 120 × 224 is used as the ground truth image $\mathcal{X}$. The Pavia image was obtained by the reflective optics system imaging spectrometer (ROSIS), and the whole image contains 610 × 340 pixels and 115 spectral bands. In our experiments, the water vapor absorption bands were removed and only a subimage of size 256 × 256 × 93 was used to evaluate the performance of HSI super-resolution. All the key parameters are optimally assigned in each method, and the experiments are run under Windows 7 and MATLAB R2017a with Intel Core i5-5200U CPU at2.80 GHz and 8-GB memory.

A. Implementation Details

In the experimental tests, we run the proposed method for different values of key parameters. In general, the proposed method performs best when the number of clusters $K = (n/60)$, n is the total number of cubes, and the sparsity regularization parameter λ_c is set as 10^{-4}. For illustration intuitively, the PSNR curves of the reconstructed results of chart and stuffed toys in the CAVE dataset for $(n/K) \in [30, 100]$ and $\log(\lambda_c) \in [-6, -2]$ are plotted in Fig. 5. We can clearly see that the PSNR value reaches its maximum at $(n/K) = 60$ and $\log(\lambda_c) = -4$. Similar phenomena can be observed for other test images. In addition, we test the proposed method when the log operation of the penality parameters $\eta_w, \eta_h, \eta_s, \eta_c$ varies from 1 to -3, and then they are set as $\eta_w = \eta_h = 0.1, \eta_s = \eta_c = 0.01$ to get the best results. In addition, in order to speed up the algorithm, we adopt the Tucker decomposition on $\mathcal{Z}^{(k)}$ and extract the dictionaries of the width mode and height mode to initialize $\mathbf{W}_k$ and $\mathbf{H}_k$. We adopt the Tucker decomposition on $\mathcal{Y}^{(k)}$ and extract its dictionary of the spectral mode to initialize $\mathbf{S}_k$. Without special statements, the stopping criterion of the algorithm for each cluster is $(\|\mathcal{X}^{(k)^t} - \mathcal{X}^{(k)^{t-1}}\|^2 / \|\mathcal{X}^{(k)^{t-1}}\|^2) < 1 \times 10^{-6}$. As for the selection of the number of dictionary atoms, we set $r_k^W = r_k^H = 2a$ and $r_k^S = 8$, where a is the scale factor. In addition, the spatial size of HR-MSI cubes is empirically chosen as $2a \times 2a$ (i.e., $d_W = 2a, d_H = 2a$), then the size of the corresponding LR-HSI cubes is 2×2 (i.e., $d_w = 2, d_h = 2$).

TABLE I

COMPARISON ON THE AVERAGE VALUES OF FOUR QUANTITATIVE MEASURES ON 32 SCENES FROM CAVE DATASET WITH UNIFORM BLUR

Method	HySure [5]	CNMF [2]	CSTF [16]	Proposed
$a = 8$				
PSNR	42.121	43.718	44.762	**46.794**
RMSE	2.532	2.172	1.998	**1.544**
SAM	10.588	5.345	6.088	**4.369**
ERGAS	1.419	1.166	1.076	**0.859**
$a = 16$				
PSNR	39.820	43.344	43.546	**44.900**
RMSE	3.590	2.266	2.370	**1.939**
SAM	16.209	5.509	6.822	**5.226**
ERGAS	0.956	0.620	0.624	**0.531**
$a = 32$				
PSNR	38.505	42.301	42.226	**42.943**
RMSE	4.464	2.571	2.848	**2.437**
SAM	20.011	**5.959**	7.837	6.495
ERGAS	0.570	0.345	0.367	**0.326**

B. Experimental Results on Simulated Images

In this section, we first use the CAVE database as ground truth image $\mathcal{X}$, and simulate the LR-HSI $\mathcal{Y}$ by uniformly averaging over disjoint $a \times a$ blocks (e.g., $a = 8, 16, 32$). The HR-MSI $\mathcal{Z}$ is generated by degrading the HSI $\mathcal{X}$ using a spectral transform matrix $\mathbf{P}_3$ based on the response of a Nikon D700 camera. All of the four quantitative measures for all compared HSI super-resolution methods on all 32 scenes have been calculated and recorded. Table I lists their average values of all compared methods. The first and third rows of Fig. 6 show the constructed images of chart and stuffed toy at 630 nm with the scaling factor 16 and 32, respectively. Their corresponding absolute error images are shown in the second and fourth rows. Blue color corresponds to low error. We observe from the zoomed-in patches that the proposed method performs better in recovering details, especially for regions containing more texture information. The proposed method shows better performance than all the other three methods both quantitatively and qualitatively in most cases.

Next, we further demonstrate the performance of the proposed method to Gaussian blur on the CAVE database and the Harvard database. We simulate $\mathcal{Y}$ by employing a 8×8 Gaussian blur function of standard deviation 3 to $\mathcal{X}$ before downsampling with scaling factor 8. The quantitative results obtained with different test methods are compared in Table II. We can clearly see that the proposed method significantly outperforms other methods with respective to all the quantitative measures.

C. Experimental Results on Remote Sensing Data

In this section, we mainly demonstrate the performance of the proposed method on the Indian Pines image. The HR-MSI $\mathcal{Z}$ with six bands is degenerated by using Landsat sensor, and each spectral band captures the 450–520-, 520–600-,

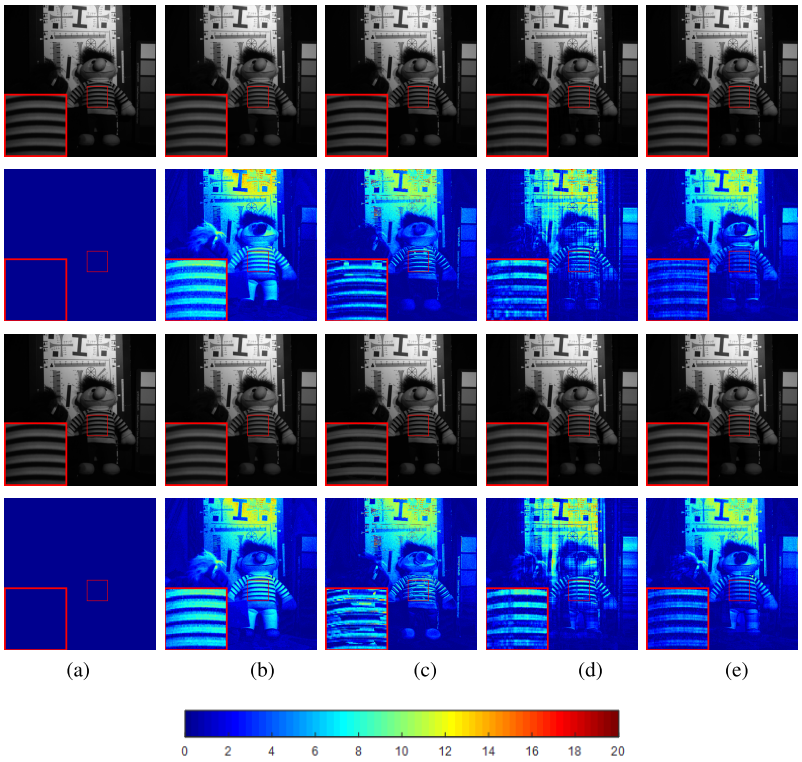


Fig. 6. Reconstructed images and the corresponding error images of chart and stuffed toy at 630 nm with uniform blur and scaling factor (Top two rows) $a = 16$ and (Bottom two rows) $a = 32$. (a) Ground truth. (b) HySure. (c) CNMF. (d) CSTF. (e) Proposed.

TABLE II
COMPARISON ON THE AVERAGE VALUES OF FOUR QUANTITATIVE MEASURES ON CAVE DATA SET AND HARVARD DATA SET WITH GAUSSIAN BLUR

Method	HySure [5]	CNMF [2]	CSTF [16]	Proposed
Cave dataset				
PSNR	42.051	43.625	43.873	**46.651**
RMSE	2.546	2.233	2.150	**1.578**
SAM	10.577	5.344	6.292	**4.391**
ERGAS	1.417	1.178	1.157	**0.879**
Harvard dataset				
PSNR	44.248	45.877	45.471	**46.526**
RMSE	1.901	1.629	1.758	**1.506**
SAM	3.616	2.981	3.289	**2.699**
ERGAS	1.277	0.998	1.002	**0.946**

630–690-, 760–900-, 1550–1750-, and 2080–2350-nm information of the HSI $\mathcal{X}$, respectively. The LR-HSI $\mathcal{Y}$ is generated as the spatial degradation manner for the CAVE data set and the Harvard data set. Both uniform blur and Gaussian blur are tested with scaling factor 8 (see Table III). We observe that the proposed method still leads to the best results among all test methods.

D. Experimental Results on Remote Sensing Data Corrupted by Gaussian Noise

In this section, we consider to verify the robustness of the proposed model to the noise by using on the Pavia data with Gaussian noise. We generate the noisy LR-HSI $\mathcal{Y}$ and HR-MSI $\mathcal{Z}$ in the same way as CSTF. The noisy LR-HSI $\mathcal{Y}$ is simulated by first uniformly averaging over disjoint 8×8 blocks of $\mathcal{X}$, and then adding Gaussian noise. We use SNRh to denote SNR of the simulated noisy LR-HSI. To generate the noisy HR-MSI $\mathcal{Z}$, first, $\mathcal{X}$ is downsampled by using the IKONOS-like reflectance spectral response filter [33], and we then add Gaussian noise. We use SNRm to represent SNR of the noisy HR-MSI. Fig. 7 shows the reconstructed images and the corresponding absolute error images from CSTF and the proposed method at the 45th and 60th bands. We only listed the comparison with CSTF as it outperforms others on these data. The quantitative results obtained with different test methods on the Pavia image are compared in Table IV. We observe that the proposed method still leads to the best results among all test methods.

E. Experimental Results on Real Data Set

In this section, we further evaluate the performance of the proposed method on real HSI and MSI data set captured by the Hyperion sensor and the Sentinel-2A satellite as described in [16]. The used sizes of HSI and MSI are $120 \times 120 \times 89$ and $360 \times 360 \times 4$, respectively. Therefore, the spatial downsampling factor is 3. For the real data set, the spatial degradation matrices $\mathbf{P}_1, \mathbf{P}_2$, and the spectral response matrix $\mathbf{P}_3$ are unknown. In this experiment, the convolution blur matrix and spectral response matrix are estimated according to the method developed in [5]. Since the proposed method and CSTF are designed for the case that the convolution blur matrix can be decomposed into two spatial modes, we use two separate blur responses to approximate the estimated convolution

TABLE III

PERFORMANCE COMPARISON OF THE METHODS ON THE INDIAN PINES IMAGE WITH SCALING FACTOR 8

Method	CNMF [2]	NLSTF [12]	NSSR [7]	HySure [5]	CSTF [16]	Proposed
Uniform blur						
PSNR	42.911	49.854	47.868	46.446	50.801	**51.399**
RMSE	2.533	1.193	1.426	1.566	0.900	**0.889**
SAM	2.376	1.087	1.133	1.457	0.895	**0.858**
ERGAS	9.406	9.375	9.376	9.382	**9.368**	**9.368**
Gaussian blur						
PSNR	42.736	46.265	46.708	45.780	48.583	**49.222**
RMSE	2.593	1.669	1.561	1.728	1.224	**1.140**
SAM	2.428	1.365	1.235	1.536	1.140	**1.021**
ERGAS	9.408	9.385	9.382	9.384	9.374	**9.370**

TABLE IV

PERFORMANCE COMPARISON OF NOISY CASES ON THE PAVIA DATA WITH SCALING FACTOR 8

Method	CNMF [2]	NLSTF [12]	NSSR [7]	HySure [5]	CSTF [16]	Proposed
SNRh=35dB, SNRm=40dB						
PSNR	38.728	42.548	42.578	42.635	43.073	**43.442**
RMSE	3.117	2.006	2.095	2.259	2.195	**1.913**
SAM	2.690	2.280	2.215	2.400	2.254	**1.984**
ERGAS	0.868	0.555	0.566	0.579	0.612	**0.531**
SNRh=30dB, SNRm=35dB						
PSNR	38.483	39.541	40.398	41.589	42.073	**42.214**
RMSE	3.189	2.748	2.581	2.319	2.149	**2.125**
SAM	2.761	3.262	2.794	2.533	2.279	**2.254**
ERGAS	0.891	0.791	0.726	0.639	0.606	**0.596**

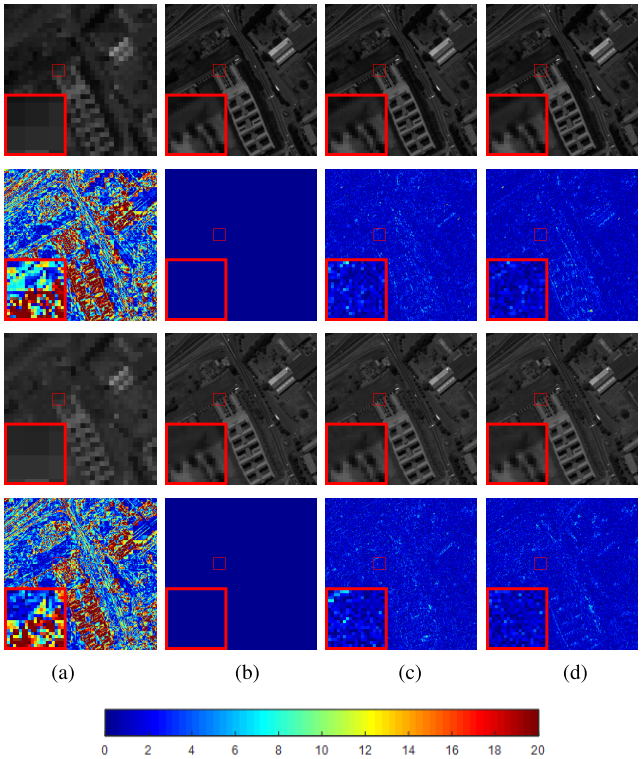


Fig. 7. Reconstructed images and the corresponding error images of the Pavia image at the 45th and 60th bands with scaling factor $a = 8$. (a) LR-HSI. (b) Ground truth. (c) CSTF. (d) Proposed.

blur matrix. Fig. 8 shows the reconstructed HR-HSIs from CSTF and the proposed method. We observe that the results of CSTF contain obvious artifacts near edges and our method gives better visual quality for this real data set.

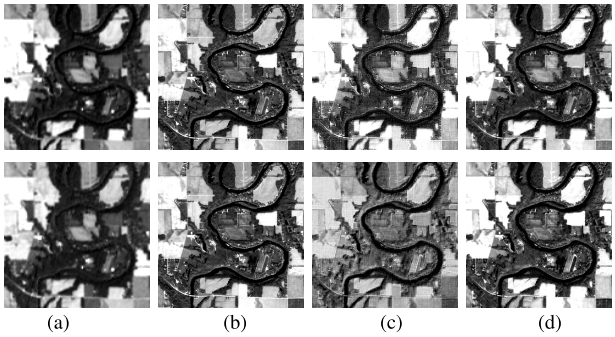


Fig. 8. Reconstructed images of the real data set at (Top row) 15th band and (Bottom row) 58th band. (a) LR-HSI. (b) HR-MSI. (c) CSTF. (d) Proposed.

TABLE V

NECESSITY OF USING BOTH NONNEGATIVE CONSTRAINT AND GROUP-BLOCK SPARSITY. CAVE DATA SET WITH UNIFORM BLUR AND SCALING FACTOR 8 ARE USED

Method	PSNR	RMSE	SAM	ERGAS
Group-block sparsity only	32.140	7.739	4.682	6.264
Non-negative only	44.048	2.359	5.220	1.171
Both constraints	**46.794**	**1.544**	**4.369**	**0.859**

F. Effectiveness of the Nonnegative Constraint and Group-Block Sparsity Constraint

In this section, we have also discussed the effectiveness of the nonnegative constraint on dictionaries and group-block sparsity constraints on core tensors. To demonstrate effectively, we add the experiments without the nonnegative constraint (i.e., group-block sparsity only, $\eta_w = \eta_h = \eta_s = 0$) and the experiments without group-block sparsity constraint (i.e., nonnegative only, $\lambda_c = 0$). Their quantitative results on the CAVE data set compared to the proposed approach with both constraints are listed in Table V. It should be noted that

TABLE VI
RUNNING TIME COMPARISON ON A TEST IMAGE OF SIZE $512 \times 512 \times 31$ WITH SCALING FACTOR 16

Method	CNMF [2]	NLSTF [12]	NSSR [7]	HySure [5]	CSTF [16]	Proposed
time(s)	**53.24**	143.92	163.56	332.52	1549.52	804.83

$\lambda_c = 0$ are set in (31) and core tensors $\mathbf{C}^{(k)}$ are still solved by using the ADMM structure instead of solving a quadratic problem directly by setting $\lambda_c = 0$ in (29). We can clearly observe that both the nonnegative constraint and group-block sparsity regularization are reasonable and effective, and the proposed approach produces obviously the best results.

G. Computation Cost

Actually, the computation of the proposed algorithm is more expensive than matrix-based methods due to unavoidable clustering and reshaping data between tensors and matrices, vectors. Generally speaking, for an image with a size of $512 \times 512 \times 31$, our proposed model takes about 41 s to perform one iteration with our unoptimized MATLAB codes and CPU implementation. However, the proposed algorithm runs much faster than the state-of-the-art tensor-based CSTF method [16]. The whole running time comparison is shown in Table VI. We will study how to employ block pro-linear method to accelerate the computation. We will also study how to handle the cases when the degradation operators $\mathbf{P}_1$, $\mathbf{P}_2$, and $\mathbf{P}_3$ are unknown. Inverse filters will be explored regarding this issue.

V. CONCLUSION

In this article we propose a novel HSI super-resolution approach based on nonnegative tensor dictionary learning and group-block sparsity. We first divide the LR-HSI $\mathcal{Y}$ and HR-MSI $\mathcal{Z}$ into 3-D cubes and cluster them into some clusters represented by 4-D tensors. By using the Tucker tensor factorization technique for each cluster, the problem of HSI super-resolution can be reformulated as a joint estimation of core tensors and dictionaries of three modes. The sparse representation ensured that similar cubes share the same atoms from the spatial and spectral dictionaries, and we can well reconstruct some texture information. Both quantitative and qualitative numerical experiments demonstrate that the proposed model can provide some state-of-the-art results in HSI super-resolution.

ACKNOWLEDGMENT

The authors would like to thank Dr. Xiaoxiang Zhu from the Technical University of Munich and German Aerospace Center (DLR) for the discussion on hyperspectral and multispectral data and results. They would also like to thank Dr. Renwei Dian for providing the real data set.

REFERENCES

[1] R. Kawakami, J. Wright, Y. W. Tai, Y. Matsushita, M. Ben-Ezra, and K. Ikeuchi, "High-resolution hyperspectral imaging via matrix factorization," in *Proc. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 2329–2336.

[2] 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.

[3] X. Liu, W. Xia, B. Wang, and L. Zhang, "An approach based on constrained nonnegative matrix factorization to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 757–772, Feb. 2011.

[4] E. Wycoff, T.-H. Chan, K. Jia, W.-K. Ma, and Y. Ma, "A non-negative sparse promoting algorithm for high resolution hyperspectral imaging," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 1409–1413.

[5] 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.

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

[7] 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.

[8] 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.

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

[10] K. Wei and Y. Fu, "Low-rank Bayesian tensor factorization for hyperspectral image denoising," *Neurocomputing*, vol. 331, pp. 412–423, Feb. 2019.

[11] Y. Xu, Z. Wu, J. Chanussot, and Z. Wei, "Joint reconstruction and anomaly detection from compressive hyperspectral images using mahalanobis distance-regularized tensor RPCA," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 5, pp. 2919–2930, May 2018.

[12] 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. (CVPR)*, Jul. 2017, pp. 3862–3871.

[13] R. Dian, S. Li, L. Fang, T. Lu, and J. M. Bioucas-Dias, "Nonlocal sparse tensor factorization for semiblind hyperspectral and multispectral image fusion," *IEEE Trans. Cybern.*, early access, Nov. 28, 2019, doi: 10.1109/TCYB.2019.2951572.

[14] Y. Chang, L. Yan, H. Fang, S. Zhong, and Z. Zhang, "Weighted low-rank tensor recovery for hyperspectral image restoration," 2017, *arXiv:1709.00192*. [Online]. Available: https://arxiv.org/abs/1709.00192

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

[16] 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.

[17] 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.

[18] R. Dian, S. Li, and L. Fang, "Learning a low tensor-train rank representation for hyperspectral image super-resolution," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 9, pp. 2672–2683, Sep. 2019.

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

[20] 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.

[21] Y. Wei, Q. Yuan, H. Shen, and L. Zhang, "Boosting the accuracy of multispectral image pansharpening by learning a deep residual network," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1795–1799, Oct. 2017.

[22] 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.
[23] Y. Peng, D. Meng, Z. Xu, C. Gao, Y. Yang, and B. Zhang, "Decomposable nonlocal tensor dictionary learning for multispectral image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2949–2956.
[24] M. Wax and T. Kailath, "Detection of signals by information criteria," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-33, no. 2, pp. 387–392, 2nd Quart., 1985.
[25] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, Aug. 2009.
[26] C. F. Caiafa and A. Cichocki, "Computing sparse representations of multidimensional signals using kronecker bases," *Neural Comput.*, vol. 25, no. 1, pp. 186–220, Jan. 2013.
[27] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep. 2009, pp. 2272–2279.
[28] J. Eckstein and W. Yao, "Understanding the convergence of the alternating direction method of multipliers: Theoretical and computational perspectives," *Pacific J. Optim.*, vol. 11, no. 4, pp. 619–644, Jun. 2015.
[29] 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.
[30] A. Chakrabarti and T. Zickler, "Statistics of real-world hyperspectral images," in *Proc. CVPR*, Jun. 2011, pp. 193–200.
[31] G. Vane, R. O. Green, T. G. Chrien, H. T. Enmark, E. G. Hansen, and W. M. Porter, "The airborne visible/infrared imaging spectrometer (AVIRIS)," *Remote Sens. Environ.*, vol. 44, nos. 2–3, pp. 127–143, 1993.
[32] 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.
[33] 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.

Wei Wan received the B.S. degree in mathematics from Ludong University, Yantai, China, in 2013, the M.S. degree in applied mathematics from the Minzu University of China, Beijing, China, in 2016, and the Ph.D. degree in applied mathematics from Beijing Normal University, Beijing, in 2019.

She is currently a Post-Doctoral Fellow with the Yau Mathematical Sciences Center, Tsinghua University, Beijing. Her research interests include image inpainting, hyperspectral image super-resolution, partial differential equations, and deep learning.

Weihong Guo (Associate Member, IEEE) received the B.S. degree in computational mathematics from the Minzu University of China, Beijing, China, in 1999, and the M.S. degree in statistics and the Ph.D. degree in applied mathematics from the University of Florida, Gainesville, FL, USA, in 2007.

She was a Mathematics Assistant Professor with The University of Alabama, Tuscaloosa, AL, USA, from 2007 to 2009. She is currently an Applied Mathematics Associate Professor with Case Western Reserve University, Cleveland, OH, USA. She has published over 40 articles in various journals such as, *SIAM Journal on Imaging Sciences*, *Inverse Problems and Imaging*, the IEEE TRANSACTIONS ON IMAGE PROCESSING, *Magnetic Resonance Imaging*, *Magnetic Resonance in Medicine*, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, and so on. Her current research interests include variational image reconstruction, image super-resolution, image segmentation, and scientific computing.

Dr. Guo serves as an editor for an international journal.

Haiyang Huang received the M.S. and Ph.D. degrees in mathematics from Beijing Normal University, Beijing, China, in 1981 and 1993, respectively.

She is currently a Professor with the School of Mathematical Science, Beijing Normal University. Her research interests include PDE theory, mathematical biology, mathematics modeling, and image processing.

Jun Liu received the B.S. degree in mathematics from Hunan Normal University, Changsha, China, in 2004, and the M.S. and Ph.D. degrees in computational mathematics from Beijing Normal University (BNU), Beijing, China, in 2008 and 2011, respectively.

He is currently an Associate Professor with BNU. His research interests include variational, optimal transport, and deep learning-based image processing algorithms and their applications.