

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/333131087>

Hyperspectral and Multispectral Image Fusion via Nonlocal Low-Rank Tensor Decomposition and Spectral Unmixing

Preprint · May 2019

DOI: 10.13140/RG.2.2.18987.28968

CITATIONS

0

READS

294

6 authors, including:



[Yao Wang](#)

Xi'an Jiaotong University

58 PUBLICATIONS 700 CITATIONS

[SEE PROFILE](#)



[Xi-Le Zhao](#)

University of Electronic Science and Technology of China

109 PUBLICATIONS 910 CITATIONS

[SEE PROFILE](#)



[Deyu Meng](#)

Xi'an Jiaotong University

114 PUBLICATIONS 2,422 CITATIONS

[SEE PROFILE](#)



[Zongben Xu](#)

Xi'an Jiaotong University

416 PUBLICATIONS 9,781 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Deep Learning [View project](#)



machine learning [View project](#)

Hyperspectral and Multispectral Image Fusion via Nonlocal Low-Rank Tensor Decomposition and Spectral Unmixing

Kaidong Wang, Yao Wang, Xi-Le Zhao, Deyu Meng, Jonathan Cheung-Wai Chan and Zongben Xu

Abstract—Hyperspectral (HS) imaging has shown its superiority in many real applications. However, it is usually difficult to obtain high-resolution (HR) HS images through existing imaging techniques, due to the hardware limitations. To improve the spatial resolution of HS images, this paper proposes an effective hyperspectral-multispectral (HS-MS) image fusion method by combining the ideas of nonlocal low-rank tensor modeling and spectral unmixing. To be more precise, instead of unfolding the HS image into a matrix as done in the literature, we directly represent it as a tensor, then a designed nonlocal Tucker decomposition is used to model its underlying spatial-spectral correlation and the spatial self-similarity. The MS image serves mainly as a data constraint to maintain spatial consistency. To further reduce the spectral distortions in spatial enhancement, endmembers and abundances from the spectral are used for spectral regularization. An efficient algorithm based on the alternating direction method of multipliers (ADMM) is developed to solve the resulting model. Extensive experiments on both simulated and real data sets demonstrate the superiority of the proposed method over several state-of-the-art HS-MS image fusion methods.

Index Terms—Hyperspectral (HS) image, spatial enhancement, image fusion, nonlocal tensor decomposition, spectral unmixing.

I. INTRODUCTION

HYPERSPECTRAL imagery consists of images of one scene in tens to hundreds of discrete spectral bands at specific frequencies. The very high spectral resolution and coverage of HS images enable a precise identification of the materials present in the scene, which tends to significantly benefit the characterization of the imaged scene and greatly enhance performances in many applications including remote sensing [1], [2], [3], [4], [5], monitoring [6], object detection [7], military [8] and tracking [9]. However, the spatial resolution of HS images is usually poor due to the limited amount of incident energy, which limits its extensive applications to some extent. On the other hand, multispectral (MS) imaging sensors are capable of capturing images with increasingly higher spatial resolution. Therefore, a low-resolution HS image can be fused with the high-resolution MS image captured on

K. Wang, D. Meng and Z. Xu are with the School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China.

Y. Wang is with the School of Management and the Research Center for Statistics and Big Data, Xi'an Jiaotong University, Xi'an 710049, China.

X.-L. Zhao is with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 610054, China.

J. C.-W. Chan is with the Department of Electronics and Informatics, Vrije Universiteit Brussel, 1050 Brussel, Belgium.

Y. Wang is the corresponding author. Email: yao.s.wang@gmail.com

the same scene to reconstruct a high-resolution HS image [10], [11]. This procedure is known as HS and MS image fusion and has attracted great attention. Actually, new HS imager has been planned which carries also a MS image sensor [12]. This means high-spectral-resolution HS image and high-spatial-resolution MS image are simultaneously captured, more data sources for HS and MS image fusion are expected.

Pan-sharpening has been widely used to enhance the spatial resolution of MS images by the fusion of a MS image and the corresponding higher-resolution panchromatic image [13], [14], [15], [16]. There are mainly three categories of representative pan-sharpening techniques: component substitution (CS) [17], [18], multiresolution analysis (MRA) [19], [20] and sparse representation (SR) [16], [21]. These methods often suffer from serious spectral distortion because of the lack of spectral information in the panchromatic image. Obviously, pan-sharpening can be considered as a special case of the HS and MS image fusion problem, and thus those pan-sharpening methods can often be extended to the fusion of HS and MS images [22], [23]. For an in-depth review of pan-sharpening methods, see [10].

As for the HS and MS image fusion problem, there has been many fruitful researches in the last few years. The Bayesian framework is a frequently-used methodology to fuse a low-resolution HS image with a high-resolution MS image. This class of methods usually builds the posterior distribution of the expected high-resolution HS images based on the prior knowledge and observation model. In [24], a Bayesian method based on a maximum a posterior (MAP) estimation and a stochastic mixing model (SMM) is proposed for the HS and MS image fusion problem. In [25], the authors solve a Sylvester equation with closed-form solution to maximize the likelihoods of the observations and meanwhile incorporate prior information for the fusion problem, which leads to a Bayesian estimator and an effective fusion algorithm with significantly decreased computational complexity. A spectral linear mixing model based on non-parametric Bayesian dictionary learning and Bayesian sparse coding is employed in [26] to solve the fusion problem.

Matrix factorization is also a commonly-used methodology in HS-MS image fusion. The matrix factorization based fusion methods usually first unfold the targeted high-resolution HS image as a matrix, and then factorize it into a basis matrix and a coefficient matrix, where the basis and coefficient matrices are respectively extracted from the low-resolution HS image and high-resolution MS image. The coupled nonnegative ma-

trix factorization (CNMF) method proposed by Yokoya *et al.* [27] alternately unmixes the HS and MS image into endmember and abundance matrices by the CNMF algorithm based on a linear spectral mixture model. In [28], the authors apply an unmixing algorithm to the HS image to estimate a basis representing reflectance spectra, and then solve the high-resolution sparse coefficient using orthogonal matching pursuit. In [29], Huang *et al.* proposed a sparse matrix factorization method which learns from the low-resolution HS image a spectral dictionary containing pure signatures using the k -singular value decomposition (SVD) algorithm. Lanaras' work [30] jointly unmixes the HS and MS images into the pure reflectance spectra and the associated mixing coefficients, which leads to a coupled matrix factorization problem with some useful constraints imposed by elementary physical properties of spectral mixing. Dong *et al.* [31] proposed a nonnegative structured sparse representation (NSSR) approach which formulates the estimation of the high-resolution HS image as a joint estimation of the HS dictionary and the sparse codes based on the prior knowledge of the spatial-spectral sparsity of the HS image.

Although the matrix factorization based methods have achieved impressive performances in the HS and MS fusion problem, There are inherent drawbacks. All the matrix factorization methods need to unfold the three-dimensional data structures into matrices, and the unfolding procedure could break the spatial structure of the data, which makes it hard to fully exploit the spatial-spectral correlations of the HS images.

To maintain the inherent spatial structure of the HS images, Li *et al.* [32] proposed a coupled sparse tensor factorization (CSTF) based approach for the HS and MS image fusion problem. Different from the conventional matrix factorization based methods, the targeted high-resolution HS image is directly considered as a three-order tensor which can then be estimated by a core tensor and dictionaries of the three modes. Meanwhile, a sparsity prior of the core tensor is incorporated as a regularizer to model the high spatial-spectral correlations in the high-resolution HS image. Although achieving good performances, CSTF only considers the high spatial-spectral correlations and fails to thoroughly exploit other effective prior knowledge in the HS image. Xu's work [33] extracts the non-local similar patches to form a nonlocal patch tensor and use a tensor-tensor product based tensor sparse representation to model the patch tensors. Then the authors built the relationship between the HS and MS images by tensor-tensor product and design a unified objective function which then is solved by alternating direction method of multipliers.

In this paper, we proposed a nonlocal low-rank tensor decomposition and spectral unmixing based approach for dealing with the HS and MS fusion problem. To fully exploit the high spatial-spectral correlations and the nonlocal self-similarity of the high-resolution HS image, we segment it into many overlapped 3D patches and cluster those patches into some clusters, then every cluster is represented as a low-rank four-order tensor. We use Tucker decomposition to formulate the low-rank property of those tensors. Motivated by our previous works [34], [35], we then use a 3D total variation (3D-TV) as a regularization to formulate the spatial-spectral continuity

of the HS image. Finally, a nonconvex regularization based linear spectral unmixing is applied as a significant spectral regularization to reduce spectral distortions. Based on the above prior knowledge, we propose a unified optimization model to formulate the HS and MS fusion problem and develop an efficient algorithm using alternating direction method of multipliers (ADMM) to solve the proposed model. Experimental results demonstrate that the proposed procedure substantially outperforms several state-of-the-art HS and MS fusion methods.

The rest of this paper is organized as follows. In Section II, we give a basic introduction on tensors and formulate the fusion problem in tensor perspective. The proposed unified optimization model for HS and MS image fusion is then introduced in Section III. Section IV gives the ADMM based optimization algorithm to solve the proposed model. In Section V, experimental results on three image datasets are presented. And the conclusion is drawn in Section VI.

II. PROBLEM FORMULATION

Throughout this paper, both HS and MS images are denoted as three-order tensors which belong to $\mathbf{R}^{w \times h \times s}$, where w , h and s are the indexes of the width, height and spectral modes, respectively. In this section, we shall first review the basics about tensor and multi-linear algebra, and then formulate the fusion problem in tensor perspective.

A. Notions and Preliminaries on Tensors

A tensor can be seen as a multi-index numerical array, and the number of its modes is called the order of this tensor. An N -order real-valued tensor can be denoted as $\mathcal{X} \in \mathbf{R}^{I_1 \times I_2 \dots \times I_N}$ and its elements as x_{i_1, i_2, \dots, i_N} , where $1 \leq i_n \leq I_n$. Then a vector can be seen as a 1-order tensor and a matrix can be seen as a 2-order tensor. A subtensor is created by only using a fixed subset of the indices of the original tensor. Vector-valued subtensors are called fibers which are obtained by fixing all but one index. Matrix valued subtensors are called slices which are defined by fixing every index but two. The n -mode unfolding matrix $\mathbf{X}_{(n)} =: \text{unfold}_n(\mathcal{X}) \in \mathbf{R}^{I_n \times (I_1 \dots I_{n-1}, I_{n+1}, \dots, I_N)}$ is defined by taking all the mode- n fibers as its columns, and conversely, it can also be transformed back to the original tensor by $\mathcal{X} =: \text{fold}_n(\mathbf{X}_{(n)})$. The n -rank of \mathcal{X} , denoted as r_n , is defined by the rank of its n -mode unfolding matrix $\mathbf{X}_{(n)}$, and its multi-linear rank, denoted as (r_1, r_2, \dots, r_N) , is the tuple of all the n -rank, $n = 1, 2, \dots, N$.

The production between matrices can be generalized to the production of a tensor and a matrix. The mode- n product of a tensor $\mathcal{X} \in \mathbf{R}^{I_1 \times I_2 \dots \times I_N}$ by a matrix $\mathbf{A} \in \mathbf{R}^{J_n \times I_n}$, denoted by $\mathcal{X} \times_n \mathbf{A}$, is an N -order tensor $\mathcal{C} \in \mathbf{R}^{I_1 \times \dots \times J_n \times \dots \times I_N}$, whose elements are computed by

$$c_{i_1 \times \dots \times i_{n-1} \times j_n \times i_{n+1} \dots \times i_N} = \sum_{i_n} x_{i_1 \dots i_n \dots i_N} a_{j_n i_n}. \quad (1)$$

The mode- n production $\mathcal{C} = \mathcal{X} \times_n \mathbf{A}$ can also be calculated by the matrix multiplication $\mathbf{C}_{(n)} = \mathbf{B} \mathbf{A}_{(n)}$, then tensor \mathcal{C} can be computed by $\mathcal{C} = \text{fold}_n(\mathbf{C}_{(n)})$. It's worth mentioning that

TABLE I
NOTATIONS

Notations	Explanations
$\mathcal{X}, \mathbf{X}, \mathbf{x}, x$	tensor, matrix, vector, scalar.
$\mathbf{x}(:, i_2, i_3, \dots, i_N)$	fiber of tensor \mathcal{X} .
$\mathbf{X}(:, :, i_3, \dots, i_N)$	slice of tensor \mathcal{X} .
$\mathbf{X}_{(n)}$ or $\text{unfold}_n(\mathcal{X})$	mode- n matricization of tensor \mathcal{X} .
$\text{fold}_n(\mathbf{X}_{(n)})$	the inverse operation of $\text{unfold}_n(\mathcal{X})$.
$\text{Vec}(\mathcal{X})$	vectorization of tensor \mathcal{X} .
$\text{Ten}(\mathbf{x})$	tensorization of vector \mathbf{x} .
(r_1, r_2, \dots, r_N)	multi-linear rank, where $r_n = \text{Rank}(\mathbf{X}_{(n)})$.
$\ \mathcal{X}\ _0$	ℓ_0 norm of tensor \mathcal{X} .
$\ \mathcal{X}\ _1$	ℓ_1 norm of tensor \mathcal{X} .
$\ \mathcal{X}\ _F$	Frobenius norm of tensor \mathcal{X} .
$\mathcal{C} = \mathcal{X} \times_n \mathbf{A}$	mode- n multiplication of \mathcal{X} and \mathbf{A} .

for distinct modes in a series of multiplications, the order of the multiplications is irrelevant, i.e.,

$$\mathcal{X} \times_n \mathbf{A} \times_m \mathbf{B} = \mathcal{X} \times_m \mathbf{B} \times_n \mathbf{A} (n \neq m), \quad (2)$$

and for the same modes, we have

$$\mathcal{X} \times_n \mathbf{A} \times_n \mathbf{B} = \mathcal{X} \times_n (\mathbf{B}\mathbf{A}). \quad (3)$$

The ℓ_0 norm of a tensor \mathcal{X} , i.e., $\|\mathcal{X}\|_0$, is defined by the number of non-zero elements of \mathcal{X} . The ℓ_1 norm of \mathcal{X} is defined by $\|\mathcal{X}\|_1 = \sum_{i_1, \dots, i_N} |x_{i_1 \dots i_N}|$, and its Frobenius norm is $\|\mathcal{X}\|_F = \sqrt{\sum_{i_1, \dots, i_N} |x_{i_1 \dots i_N}|^2}$. All the tensor notations used in this paper are summarized in Table I. For more details about tensor and multi-linear algebra, please see [36], [37].

B. Problem Formulation in Tensor Perspective

The fusion problem of HS and MS images is aimed at reconstructing a both high-spatial-resolution (HR) and high-spectral-resolution image from a low-spatial-resolution (LR) HS image and a low-spectral-resolution MS image from the same scene. This problem is usually considered in matrix form, that is, unfolding the three-order tensor data into matrices, which can break the inherent structures of the images. To remedy this issue, we direct tackle the HS-MS image fusion problem in tensor perspectives, which could be capable of fully exploit the spatial-spectral correlations of the HS images.

As mentioned before, both HS and MS images can be naturally represented as three-order tensors. Precisely, the HR-HS image to be reconstructed is denoted as a three-order tensor $\mathcal{X} \in \mathbf{R}^{W \times H \times S}$, where W , H and S are the dimensions of the width, height and spectral mode, respectively. The LR-HS image is denoted as $\mathcal{Y}_h \in \mathbf{R}^{w \times h \times S}$ where $w < W$ and $h < H$, and it can be written as

$$\mathcal{Y}_h = \mathcal{X} \times_1 \mathbf{P}_1 \times_2 \mathbf{P}_2, \quad (4)$$

where $\mathbf{P}_1 \in \mathbf{R}^{w \times W}$ and $\mathbf{P}_2 \in \mathbf{R}^{h \times H}$ are the downsampling matrices along the width mode and height mode, respectively, which describe the spatial response of the imaging sensors. Specifically, equation (4) is actually a formal description model, and in our method the downsampling matrices respectively along the width mode and height mode, \mathbf{P}_1 and \mathbf{P}_2 ,

are actually not needed. In the viewpoint of matricization, \mathcal{X} and \mathcal{Y}_h are usually unfolded as $\mathbf{X}_{(3)} \in \mathbf{R}^{(W \times H) \times S}$ and $(\mathbf{Y}_h)_{(3)} \in \mathbf{R}^{(w \times h) \times S}$, and then $(\mathbf{Y}_h)_{(3)}$ can be written as

$$(\mathbf{Y}_h)_{(3)} = \mathbf{X}_{(3)} \mathbf{M}, \quad (5)$$

where $\mathbf{M} \in \mathbf{R}^{S \times s}$ denotes the degradation operator of blurring and downsampling associated with the LR-HS image. Equation (4) means that the point spread function (PSF) of the hyperspectral sensor and the downsampling matrices of the width mode and height modes are separable, then the action of the spatial subsampling matrix \mathbf{M} is decoupled with respect to the two spatial modes of \mathcal{X} :

$$\mathbf{M} = (\mathbf{P}_2 \otimes \mathbf{P}_1)^T. \quad (6)$$

Separability assumption is sometimes necessary, e.g., in [32] both \mathbf{P}_1 and \mathbf{P}_2 are needed in their method. As for our method, such assumption is not necessary because only \mathbf{M} is needed which is consistent with most of the existing matrix based methods.

The HR-MS image is denoted as $\mathcal{Y}_m \in \mathbf{R}^{W \times H \times s}$ where $s < S$, and it can be written as

$$\mathcal{Y}_m = \mathcal{X} \times_3 \mathbf{P}_3, \quad (7)$$

where $\mathbf{P}_3 \in \mathbf{R}^{s \times S}$ is the downsampling matrix of the spectral mode.

Reconstructing \mathcal{X} from \mathcal{Y}_h and \mathcal{Y}_m is obviously an inverse problem, and to get satisfactory \mathcal{X} , it is necessary to fully exploit the prior information of \mathcal{X} . In the following section, we will fully analyze the useful priors and characterize them using tensor algebra before describing our final model.

III. PROPOSED MODEL FOR HS-MS FUSION

With the above discussions, we should reconstruct \mathcal{X} based on

$$\mathcal{Y}_h = \mathcal{X} \times_1 \mathbf{P}_1 \times_2 \mathbf{P}_2, \quad \mathcal{Y}_m = \mathcal{X} \times_3 \mathbf{P}_3. \quad (8)$$

A nature procedure is to solve the following model:

$$\min_{\mathcal{X}} \frac{1}{2} \|\mathcal{Y}_h - \mathcal{X} \times_1 \mathbf{P}_1 \times_2 \mathbf{P}_2\|_F^2 + \frac{\lambda_1}{2} \|\mathcal{Y}_m - \mathcal{X} \times_3 \mathbf{P}_3\|_F^2. \quad (9)$$

However, solving (9) cannot get satisfactory result, because it is too simplified and has not considered any prior information of \mathcal{X} at all. In the following, we shall analyze the inherent structures of the HR-HS image \mathcal{X} in detail and depict mathematically the regularization for the model in (9).

A. The Underlying Priors of \mathcal{X}

It is known that \mathcal{X} shows strong spectral correlation which implies that its matrix unfolding in the spectral mode $\mathbf{X}_{(3)}$ should be low rank. Meanwhile, we can consider its nonlocal self-similarity to further exploit the spatial correlation. Spatial nonlocal self-similarity is a patch-based powerful prior and has been widely used in the community of image processing [38]-[39]. This prior means that one local patch in one image has many similar structure patches, and the similarity of patches actually implies the correlation of those patches. Inspired by [40], we can segment \mathcal{X} into many overlapped 3D patches

of the size $d \times d \times S$ and then collect these 3D patches as a patch set \mathcal{G} : $\mathcal{G} = \{\mathcal{P}_i \in \mathbf{R}^{d \times d \times S} : i \in \Gamma\}$, where Γ indicates the index set and \mathcal{P}_i is the i -th 3D patch in the set. We cluster the patch set \mathcal{G} into K clusters and then collect each cluster as a 4-order tensor, denoted by $\mathcal{Q}_p(\mathcal{X}) \in \mathbf{R}^{d \times d \times S \times N}$, $p = 1, 2, \dots, K$, where N is the number of 3D patches in every cluster. Due to the similarity of 3D patches in every cluster, for every $\mathcal{Q}_p(\mathcal{X})$, its matrix unfolding in the patch mode $\mathcal{Q}_p(\mathcal{X})_{(4)}$ should also be low rank.

Tucker decomposition [41] is a well-known form of tensor decomposition, and has shown its advantages in many applications. Every low rank tensor $\mathcal{Q}_p(\mathcal{X})$ can be approximated by the following Tucker decomposition form:

$$\mathcal{Q}_p(\mathcal{X}) = \mathcal{C}_p \times_1 \mathbf{W}_p \times_2 \mathbf{H}_p \times_3 \mathbf{S}_p \times_4 \mathbf{N}_p + \mathcal{E}_p, \quad (10)$$

where factor matrices \mathbf{W}_p , \mathbf{H}_p , \mathbf{S}_p and \mathbf{N}_p are orthogonal in columns for height, width, spectral and patch modes, respectively, $\mathcal{C}_p \in \mathbf{R}^{r_1 \times r_2 \times r_3 \times r_4}$ is the core tensor interacts with these factors, and \mathcal{E}_p is the disturbance. Considering the low rank property in spectral and patch modes of $\mathcal{Q}_p(\mathcal{X})$, r_3 and r_4 should satisfy $r_3 \ll S$, $r_4 < N$. After that all the 4-order tensor $\mathcal{Q}_p(\mathcal{X})$ has been approximatively estimated by the low rank tensor \mathcal{L}_p :

$$\mathcal{L}_p = \mathcal{C}_p \times_1 \mathbf{W}_p \times_2 \mathbf{H}_p \times_3 \mathbf{S}_p \times_4 \mathbf{N}_p. \quad (11)$$

\mathcal{X} can then be reconstructed by arranging those 3D patches in the original order, that is, $\mathcal{X} = \mathcal{Q}^{-1}(\{\mathcal{L}_p\}_{p=1,2,\dots,N})$, where $\{\mathcal{L}_p\}_{p=1,2,\dots,N}$ is the set of \mathcal{L}_p and \mathcal{Q}^{-1} is the reconstruction operator.

Besides the spatial-spectral correlation and nonlocal self-similarity, \mathcal{X} also shows spatial-spectral continuity which can be formulated by total variation (TV) regularization. TV regularization has been widely used to explore the spatial piecewise smooth structure for tackling the HS image restoration task [42], [43], [44]. Considering that there is also a strong local smooth structure of a HS image along its spectral mode, we thus use a 3D total variation (3D-TV) to model the spatial-spectral continuity of \mathcal{X} . The 3D-TV of \mathcal{X} can be defined by the following form:

$$\begin{aligned} \|\mathcal{X}\|_{3D-TV} := \sum_{i,j,k} \Big\{ & |\mathcal{X}(i,j,k) - \mathcal{X}(i+1,j,k)| \\ & + |\mathcal{X}(i,j,k) - \mathcal{X}(i,j+1,k)| \\ & + |\mathcal{X}(i,j,k) - \mathcal{X}(i,j,k+1)| \Big\}, \end{aligned} \quad (12)$$

which can be rewritten as the following equivalent form:

$$\|\mathcal{X}\|_{3D-TV} = \|\mathbf{D}(\mathcal{X})\|_1, \quad (13)$$

where $\mathbf{D}(\cdot) = [\mathbf{D}_w(\cdot), \mathbf{D}_h(\cdot), \mathbf{D}_s(\cdot)]$ is the three-dimensional difference operator, and $\mathbf{D}_w(\cdot), \mathbf{D}_h(\cdot), \mathbf{D}_s(\cdot)$ are the first-order difference operators along the width, height, and spectral modes, respectively.

With the above discussions, the model (9) is regularized as

$$\begin{aligned} & \min_{\mathcal{X}, \mathcal{C}_p, \mathbf{W}_p, \mathbf{H}_p, \mathbf{S}_p, \mathbf{V}_p} \frac{1}{2} \|\mathcal{Y}_h - \mathcal{X} \times_1 \mathbf{P}_1 \times_2 \mathbf{P}_2\|_F^2 + \frac{\lambda_1}{2} \|\mathcal{Y}_m - \mathcal{X} \times_3 \mathbf{P}_3\|_F^2 \\ & \quad + \lambda_2 \|\mathcal{X}\|_{3D-TV} \\ \text{s.t. } & \mathcal{X} = \mathcal{Q}^{-1} \left(\{\mathcal{C}_p \times_1 \mathbf{W}_p \times_2 \mathbf{H}_p \times_3 \mathbf{S}_p \times_4 \mathbf{N}_p\}_{p=1,2,\dots,N} \right), \\ & \mathbf{W}_p^T \mathbf{W}_p = \mathbf{I}, \mathbf{H}_p^T \mathbf{H}_p = \mathbf{I}, \mathbf{S}_p^T \mathbf{S}_p = \mathbf{I}, \mathbf{N}_p^T \mathbf{N}_p = \mathbf{I}. \end{aligned} \quad (14)$$

B. Linear Spectral Unmixing Based on MCP Regularization

The above model (14) only considers resolution enhancement, and such processing scheme tends to suffer from spectral distortions. Spectral unmixing [45] has been used as a significant spectral regularization to reduce spectral distortions [46]. Once unfolding the HR-HS image \mathcal{X} as $\mathbf{X}_{(3)} \in \mathbf{R}^{(S \times (W \times H))}$, then a sparse spectral unmixing model is formulated as:

$$\min_{\mathbf{E}, \mathbf{U}} \frac{1}{2} \|\mathbf{X}_{(3)} - \mathbf{EU}\|_F^2 + \lambda \|\mathbf{U}\|_0, \text{s.t. } \mathbf{E} \geq 0, \mathbf{U} \geq 0, \quad (15)$$

where $\mathbf{E} \in \mathbf{R}^{S \times K}$ contains K endmembers, while $\mathbf{U} \in \mathbf{R}^{K \times (W \times H)}$ represents fraction abundances and is regularized using sparsity constraint ℓ_0 norm. Due to the non-convexity of ℓ_0 norm, the sparsity constraint is often replaced by its convex relaxation form, usually ℓ_1 norm:

$$\min_{\mathbf{E}, \mathbf{U}} \frac{1}{2} \|\mathbf{X}_{(3)} - \mathbf{EU}\|_F^2 + \lambda \|\mathbf{U}\|_1, \text{s.t. } \mathbf{E} \geq 0, \mathbf{U} \geq 0. \quad (16)$$

ℓ_1 norm is a frequently-used sparsity constraint, however, several studies [47], [48], [49] have shown that the ℓ_1 norm (or LASSO) penalty over-penalizes large entries of vectors, and usually cannot avoid modelling bias. Fortunately, the folded-concave penalization such as SCAD (Smoothly Clipped Absolute Deviation) and MCP (Minmax Concave Plus) have been proposed and they are being described as nearly unbiased and superior to the convex penalization approach through numerous theoretical and numerical studies [49], [50], [51], [52], [53], [54], [55]. In this work, we use MCP rather than ℓ_1 norm as the sparsity constraint to achieve better unmixing performance.

MCP is a typical folded-concave penalty function and its form is as follows:

$$\psi_{a,\lambda}(t) := \begin{cases} \frac{a\lambda^2}{2}, & \text{if } |t| \geq a\lambda; \\ \lambda|t| - \frac{t^2}{2a}, & \text{otherwise.} \end{cases} \quad (17)$$

Its derivative is given by

$$\psi'_{a,\lambda}(t) = (\lambda - \frac{t}{a})_+, \text{ for some } a > 1, \quad (18)$$

where $(x)_+ = \max(x, 0)$. Using MCP penalty as the sparsity constraint, the above sparse unmixing model is formulated as:

$$\min_{\mathbf{E}, \mathbf{U}} \frac{1}{2} \|\mathbf{X}_{(3)} - \mathbf{EU}\|_F^2 + \lambda \|\mathbf{U}\|_{MCP}, \text{s.t. } \mathbf{E} \geq 0, \mathbf{U} \geq 0. \quad (19)$$

Except for the different spatial resolution, both the HR-HS image \mathcal{X} and LR-HS image \mathcal{Y}_h actually capture the same scene, and thus their endmembers should also be the

same [56]. Based on this, we can simultaneously unmix $\mathbf{X}_{(3)}$ and $(\mathbf{Y}_h)_{(3)}$, then they should have the same endmembers \mathbf{E} and different abundances \mathbf{U}_1 and \mathbf{U}_2 . We can use this sparse unmixing approach as a significant spectral regularization in model (14) to reduce spectral distortions, which leads to our final model:

$$\begin{aligned} & \min_{\mathcal{X}, \mathcal{C}_p, \mathbf{W}_p, \mathbf{V}_p} \frac{1}{2} \|\mathcal{Y}_h - \mathcal{X} \times_1 \mathbf{P}_1 \times_2 \mathbf{P}_2\|_F^2 + \frac{\lambda_1}{2} \|\mathcal{Y}_m - \mathcal{X} \times_3 \mathbf{P}_3\|_F^2 \\ & \quad + \lambda_2 \|\mathcal{X}\|_{3D-TV} + \frac{\lambda_3}{2} \|\mathbf{X}_{(3)} - \mathbf{E}\mathbf{U}_1\|_F^2 + \lambda_4 \|\mathbf{U}_1\|_{MCP} \\ & \quad + \frac{\lambda_5}{2} \|(\mathbf{Y}_h)_{(3)} - \mathbf{E}\mathbf{U}_2\|_F^2 + \lambda_6 \|\mathbf{U}_2\|_{MCP} \\ \text{s.t. } & \mathcal{X} = \mathcal{Q}^{-1} \left(\{\mathcal{C}_p \times_1 \mathbf{W}_p \times_2 \mathbf{H}_p \times_3 \mathbf{S}_p \times_4 \mathbf{N}_p\}_{p=1,2,\dots,N} \right), \\ & \mathbf{W}_p^T \mathbf{W}_p = \mathbf{I}, \mathbf{H}_p^T \mathbf{H}_p = \mathbf{I}, \mathbf{S}_p^T \mathbf{S}_p = \mathbf{I}, \mathbf{N}_p^T \mathbf{N}_p = \mathbf{I}, \\ & \mathbf{E} \geq 0, \mathbf{U}_1 \geq 0, \mathbf{U}_2 \geq 0. \end{aligned} \tag{20}$$

IV. OPTIMIZATION ALGORITHM

We now optimize the model (20) using a multi-block version of ADMM [57]-[58]. Mathematically, (20) has the following equivalent form:

$$\begin{aligned} & \min_{\mathcal{A}, \mathcal{B}, \mathcal{F}, \mathcal{Z}, \mathcal{C}, \mathbf{W}, \mathbf{H}, \mathbf{S}} \frac{1}{2} \|\mathcal{Y}_h - \mathcal{A} \times_1 \mathbf{P}_1 \times_2 \mathbf{P}_2\|_F^2 + \frac{\lambda_1}{2} \|\mathcal{Y}_m - \mathcal{B} \times_3 \mathbf{P}_3\|_F^2 \\ & \quad + \lambda_2 \|\mathcal{F}\|_1 + \frac{\lambda_3}{2} \|\mathbf{G}_{(3)} - \mathbf{E}\mathbf{U}_1\|_F^2 + \lambda_4 \|\mathbf{V}_1\|_{MCP} \\ & \quad + \frac{\lambda_5}{2} \|(\mathbf{Y}_h)_{(3)} - \mathbf{E}\mathbf{U}_2\|_F^2 + \lambda_6 \|\mathbf{V}_2\|_{MCP} \\ \text{s.t. } & \mathcal{X} = \mathcal{Q}^{-1} \left(\{\mathcal{C}_p \times_1 \mathbf{W}_p \times_2 \mathbf{H}_p \times_3 \mathbf{S}_p \times_4 \mathbf{N}_p\}_{p=1,2,\dots,N} \right), \\ & \mathbf{W}_p^T \mathbf{W}_p = \mathbf{I}, \mathbf{H}_p^T \mathbf{H}_p = \mathbf{I}, \mathbf{S}_p^T \mathbf{S}_p = \mathbf{I}, \mathbf{N}_p^T \mathbf{N}_p = \mathbf{I}, \\ & \mathbf{E} \geq 0, \mathbf{V}_1 \geq 0, \mathbf{V}_2 \geq 0, \\ & \mathcal{A} = \mathcal{X}, \mathcal{B} = \mathcal{X}, \mathcal{F} = \mathbf{D}(\mathcal{Z}), \\ & \mathcal{Z} = \mathcal{X}, \mathcal{G} = \mathcal{X}, \mathbf{U}_1 = \mathbf{V}_1, \mathbf{U}_2 = \mathbf{V}_2, \end{aligned} \tag{21}$$

where $\mathbf{G}_{(3)}$ is the mode-3 matricization of tensor \mathcal{G} and matrix \mathbf{D} is three-dimensional difference operator. The augmented Lagrangian function for (21) is:

$$\begin{aligned} & L_A(\mathcal{X}, \mathcal{A}, \mathcal{B}, \mathcal{F}, \mathcal{Z}, \mathcal{C}, \mathbf{W}, \mathbf{H}, \mathbf{S}) \\ & = \frac{1}{2} \|\mathcal{Y}_h - \mathcal{A} \times_1 \mathbf{P}_1 \times_2 \mathbf{P}_2\|_F^2 + \frac{\lambda_1}{2} \|\mathcal{Y}_m - \mathcal{B} \times_3 \mathbf{P}_3\|_F^2 \\ & \quad + \lambda_2 \|\mathcal{F}\|_1 + \frac{\lambda_3}{2} \|\mathbf{G}_{(3)} - \mathbf{E}\mathbf{U}_1\|_F^2 + \lambda_4 \|\mathbf{V}_1\|_{MCP} \\ & \quad + \frac{\lambda_5}{2} \|(\mathbf{Y}_h)_{(3)} - \mathbf{E}\mathbf{U}_2\|_F^2 + \lambda_6 \|\mathbf{V}_2\|_{MCP} \\ & \quad + \frac{\beta}{2} \|\mathcal{A} - \mathcal{X} - \frac{\boldsymbol{\lambda}^{\mathcal{A}}}{\beta}\|^2 + \frac{\beta}{2} \|\mathcal{B} - \mathcal{X} - \frac{\boldsymbol{\lambda}^{\mathcal{B}}}{\beta}\|^2 \\ & \quad + \frac{\beta}{2} \|\mathcal{F} - \mathbf{D}(\mathcal{Z}) - \frac{\boldsymbol{\lambda}^{\mathcal{F}}}{\beta}\|^2 + \frac{\beta}{2} \|\mathcal{Z} - \mathcal{X} - \frac{\boldsymbol{\lambda}^{\mathcal{Z}}}{\beta}\|^2 \\ & \quad + \frac{\beta}{2} \|\mathcal{G} - \mathcal{X} - \frac{\boldsymbol{\lambda}^{\mathcal{G}}}{\beta}\|^2 + \frac{\beta}{2} \|\mathbf{U}_1 - \mathbf{V}_1 - \frac{\boldsymbol{\lambda}^{\mathbf{U}_1}}{\beta}\|^2 \\ & \quad + \frac{\beta}{2} \|\mathbf{U}_2 - \mathbf{V}_2 - \frac{\boldsymbol{\lambda}^{\mathbf{U}_2}}{\beta}\|^2, \end{aligned}$$

where $\boldsymbol{\lambda}^{\mathcal{A}}, \boldsymbol{\lambda}^{\mathcal{B}}, \boldsymbol{\lambda}^{\mathcal{F}}, \boldsymbol{\lambda}^{\mathcal{Z}}, \boldsymbol{\lambda}^{\mathcal{G}}, \boldsymbol{\lambda}^{\mathbf{U}_1}$ and $\boldsymbol{\lambda}^{\mathbf{U}_2}$ are the Lagrange multiplier vectors, and β is the positive penalty scalar. Based on multi-block ADMM, we can approximately solve this optimization problem by alternatively minimizing one variable with the others fixed. The optimization problem of L_A with respect to each variable can be solved by the following sub-problems:

1) $\mathcal{X}, \mathcal{C}_p, \mathbf{W}_p, \mathbf{H}_p, \mathbf{S}_p, \mathbf{N}_p$ sub-problem: The sub-problem to optimize L_A with respect to $\mathcal{X}, \mathcal{C}_p, \mathbf{W}_p, \mathbf{H}_p, \mathbf{S}_p, \mathbf{N}_p$ is:

$$\begin{aligned} & \min_{\mathcal{X}, \mathcal{C}_p, \mathbf{W}_p, \mathbf{H}_p, \mathbf{S}_p, \mathbf{N}_p} \frac{\beta}{2} \|\mathcal{A} - \mathcal{X} - \frac{\boldsymbol{\lambda}^{\mathcal{A}}}{\beta}\|^2 + \frac{\beta}{2} \|\mathcal{B} - \mathcal{X} - \frac{\boldsymbol{\lambda}^{\mathcal{B}}}{\beta}\|^2 \\ & \quad + \frac{\beta}{2} \|\mathcal{Z} - \mathcal{X} - \frac{\boldsymbol{\lambda}^{\mathcal{Z}}}{\beta}\|^2 + \frac{\beta}{2} \|\mathcal{G} - \mathcal{X} - \frac{\boldsymbol{\lambda}^{\mathcal{G}}}{\beta}\|^2 \\ \text{s.t. } & \mathcal{X} = \mathcal{Q}^{-1} \left(\{\mathcal{C}_p \times_1 \mathbf{W}_p \times_2 \mathbf{H}_p \times_3 \mathbf{S}_p \times_4 \mathbf{N}_p\}_{p=1,2,\dots,N} \right), \\ & \mathbf{W}_p^T \mathbf{W}_p = \mathbf{I}, \mathbf{H}_p^T \mathbf{H}_p = \mathbf{I}, \mathbf{S}_p^T \mathbf{S}_p = \mathbf{I}, \mathbf{N}_p^T \mathbf{N}_p = \mathbf{I}. \end{aligned} \tag{22}$$

The above equation can be easily transformed into the following equivalent problem:

$$\begin{aligned} & \min_{\mathcal{C}_p, \mathbf{W}_p, \mathbf{H}_p, \mathbf{S}_p, \mathbf{V}_p} 2\beta \|\mathcal{Q}^{-1} \left(\{\mathcal{C}_p \times_1 \mathbf{W}_p \times_2 \mathbf{H}_p \times_3 \mathbf{S}_p \times_4 \mathbf{N}_p\}_{p=1,2,\dots,N} \right) - \mathcal{L}\|^2 \\ \text{s.t. } & \mathbf{W}_p^T \mathbf{W}_p = \mathbf{I}, \mathbf{H}_p^T \mathbf{H}_p = \mathbf{I}, \mathbf{S}_p^T \mathbf{S}_p = \mathbf{I}, \mathbf{V}_p^T \mathbf{V}_p = \mathbf{I}, \end{aligned} \tag{23}$$

where $\mathcal{L} = \frac{1}{4\beta} ((\beta\mathcal{A} - \boldsymbol{\lambda}^{\mathcal{A}}) + (\beta\mathcal{B} - \boldsymbol{\lambda}^{\mathcal{B}}) + (\beta\mathcal{Z} - \boldsymbol{\lambda}^{\mathcal{Z}})) + (\beta\mathcal{G} - \boldsymbol{\lambda}^{\mathcal{G}})$. This problem can be rewritten as

$$\begin{aligned} & \min_{\mathcal{C}_p, \mathbf{W}_p, \mathbf{H}_p, \mathbf{S}_p, \mathbf{V}_p} \sum_{p=1}^K \frac{1}{2} \|\mathcal{C}_p \times_1 \mathbf{W}_p \times_2 \mathbf{H}_p \times_3 \mathbf{S}_p \times_4 \mathbf{V}_p - \mathbf{Q}_p(\mathcal{L})\|^2 \\ \text{s.t. } & \mathbf{W}_p^T \mathbf{W}_p = \mathbf{I}, \mathbf{H}_p^T \mathbf{H}_p = \mathbf{I}, \mathbf{S}_p^T \mathbf{S}_p = \mathbf{I}, \mathbf{V}_p^T \mathbf{V}_p = \mathbf{I}. \end{aligned} \tag{24}$$

Then similar to the Joint HOOI Algorithm in [40], this problem can be approximately solved by alternatively updating the following formulas:

$$\mathcal{C}_p = \mathbf{Q}_p(\mathcal{L}) \times_1 \mathbf{W}_p^T \times_2 \mathbf{H}_p^T \times_3 \mathbf{S}_p^T \times_4 \mathbf{N}_p^T, \tag{25}$$

$$\mathbf{W}_p = \text{SVD}((\mathbf{Q}_p(\mathcal{L}) \times_2 \mathbf{H}_p^T \times_3 \mathbf{S}_p^T \times_4 \mathbf{N}_p^T)_{(1)}, r_1) \tag{26}$$

$$\mathbf{H}_p = \text{SVD}((\mathbf{Q}_p(\mathcal{L}) \times_1 \mathbf{W}_p^T \times_3 \mathbf{S}_p^T \times_4 \mathbf{N}_p^T)_{(2)}, r_2) \tag{27}$$

$$\mathbf{S}_p = \text{SVD}((\mathbf{Q}_p(\mathcal{L}) \times_1 \mathbf{W}_p^T \times_2 \mathbf{H}_p^T \times_4 \mathbf{N}_p^T)_{(3)}, r_3) \tag{28}$$

$$\mathbf{N}_p = \text{SVD}((\mathbf{Q}_p(\mathcal{L}) \times_1 \mathbf{W}_p^T \times_2 \mathbf{H}_p^T \times_3 \mathbf{S}_p^T)_{(4)}, r_4), \tag{29}$$

where $\text{SVD}(\mathbf{A}, r)$ indicates top r singular vectors of matrix \mathbf{A} . Then \mathcal{X} can be updated by $\mathcal{X} = \mathcal{Q}^{-1} \left(\{\mathcal{C}_p \times_1 \mathbf{W}_p \times_2 \mathbf{H}_p \times_3 \mathbf{S}_p \times_4 \mathbf{N}_p\}_{p=1,2,\dots,N} \right)$.

2) \mathcal{A} sub-problem: Optimizing L_A with respect to \mathcal{A} can be written as:

$$\min_{\mathcal{A}} \frac{1}{2} \|\mathcal{Y}_h - \mathcal{A} \times_1 \mathbf{P}_1 \times_2 \mathbf{P}_2\|_F^2 + \frac{\beta}{2} \|\mathcal{A} - \mathcal{X} - \frac{\boldsymbol{\lambda}^{\mathcal{A}}}{\beta}\|^2 \tag{30}$$

which can be formalized as the following matrix format:

$$\min_{\mathbf{A}_{(3)}} \frac{1}{2} \|(\mathbf{Y}_h)_{(3)} - \mathbf{A}_{(3)} \mathbf{M}\|_F^2 + \frac{\beta}{2} \|\mathbf{A}_{(3)} - \mathbf{X}_{(3)} - \frac{\boldsymbol{\lambda}_{(3)}^{\mathcal{A}}}{\beta}\|^2 \quad (31)$$

where $\mathbf{A}_{(3)}, (\mathbf{Y}_h)_{(3)}, \boldsymbol{\lambda}_{(3)}^{\mathcal{A}}, \mathbf{X}_{(3)}$ are the 3-mode unfolding matrices of $\mathcal{A}, \mathcal{Y}_h, \boldsymbol{\lambda}^{\mathcal{A}}, \mathcal{X}$, respectively, and \mathbf{M} is the spatial downsampling matrix. The problem in Eq. (31) can be treated as solving the following linear system:

$$(\mathbf{M}\mathbf{M}^T + \beta\mathbf{I})\mathbf{A}_{(3)}^T = \beta\mathbf{X}_{(3)}^T + (\boldsymbol{\lambda}_{(3)}^{\mathcal{A}})^T + \mathbf{M}((\mathbf{Y}_h)_{(3)})^T. \quad (32)$$

Obviously, this linear system can be solved by off-the-shelf conjugate gradient techniques.

3) \mathcal{B} sub-problem: Optimizing L_A with respect to \mathcal{B} can be formalized as :

$$\min_{\mathcal{B}} \frac{\lambda_1}{2} \|\mathcal{Y}_m - \mathcal{B} \times_3 \mathbf{P}_3\|_F^2 + \frac{\beta}{2} \|\mathcal{B} - \mathcal{X} - \frac{\boldsymbol{\lambda}^{\mathcal{B}}}{\beta}\|^2 \quad (33)$$

which can be matriculated as follows:

$$\min_{\mathbf{B}_{(3)}} \frac{\lambda_1}{2} \|(\mathbf{Y}_m)_{(3)} - \mathbf{P}_3 \mathbf{B}_{(3)}\|_F^2 + \frac{\beta}{2} \|\mathbf{B}_{(3)} - \mathbf{X}_{(3)} - \frac{\boldsymbol{\lambda}_{(3)}^{\mathcal{B}}}{\beta}\|^2$$

where $\mathbf{B}_{(3)}, (\mathbf{Y}_m)_{(3)}, \boldsymbol{\lambda}_{(3)}^{\mathcal{B}}, \mathbf{X}_{(3)}$ are the mode-3 unfolding matrices of $\mathcal{B}, \mathcal{Y}_m, \boldsymbol{\lambda}^{\mathcal{B}}, \mathcal{X}$, respectively, and \mathbf{P}_3 is the spectral downsampling matrix. The problem in Eq. (IV-3) can be treated as solving the following linear system:

$$(\lambda_1 \mathbf{P}_3^* \mathbf{P}_3 + \beta \mathbf{I}) \mathbf{B}_{(3)} = \beta \mathbf{X}_{(3)} + \boldsymbol{\lambda}_{(3)}^{\mathcal{B}} + \lambda_1 \mathbf{P}_3^* (\mathbf{Y}_m)_{(3)} \quad (34)$$

where the superscript * denotes the adjoint matrix. Obviously, this linear system can be solved by off-the-shelf conjugate gradient techniques.

4) \mathcal{Z} sub-problem: The sub-problem to optimize L_A with respect to \mathcal{Z} has the following form:

$$\min_{\mathcal{Z}} \frac{\beta}{2} \|\mathcal{F} - \mathbf{D}(\mathcal{Z}) - \frac{\boldsymbol{\lambda}^{\mathcal{F}}}{\beta}\|^2 + \frac{\beta}{2} \|\mathcal{Z} - \mathcal{X} - \frac{\boldsymbol{\lambda}^{\mathcal{Z}}}{\beta}\|^2. \quad (35)$$

The problem in Eq. (35) can be treated as solving the following linear system:

$$(\beta \mathbf{D}^* \mathbf{D} + \beta \mathbf{I}) \mathcal{Z} = \beta \mathcal{X} + \boldsymbol{\lambda}^{\mathcal{Z}} + \mathbf{D}^* (\beta \mathcal{F} - \boldsymbol{\lambda}^{\mathcal{F}}) \quad (36)$$

Thanks to the block-circulant structure of the matrix corresponding to the operator $D^* D$, it can be diagonalized by the 3D FFT matrix. Therefore, \mathcal{Z} can be fast computed by

$$\text{ifftn} \left(\frac{\text{fftn}(\beta \mathcal{X} + \boldsymbol{\lambda}^{\mathcal{Z}} + \mathbf{D}^* (\beta \mathcal{F} - \boldsymbol{\lambda}^{\mathcal{F}}))}{\beta \mathbf{1} + \beta(|\text{fftn}(\mathbf{D}_h)|^2 + |\text{fftn}(\mathbf{D}_v)|^2 + |\text{fftn}(\mathbf{D}_t)|^2)} \right), \quad (37)$$

where fftn and ifftn respectively indicate fast 3D Fourier transform and its inverse transform, $|\cdot|^2$ is the element-wise square, and the division is also performed element-wisely. Note that the denominator in the equation can be pre-calculated outside the main loop, avoiding the extra computational cost.

5) \mathcal{F} sub-problem: Optimizing L_A with respect to \mathcal{F} can be written as :

$$\min_{\mathcal{F}} \frac{\beta}{2} \|\mathcal{F} - \mathbf{D}(\mathcal{Z}) - \frac{\boldsymbol{\lambda}^{\mathcal{F}}}{\beta}\|^2 + \lambda_2 \|\mathcal{F}\|_1. \quad (38)$$

This sub-problem can be solved by the well-known soft shrinkage operator as follows:

$$\mathcal{F} = \text{soft}(\mathbf{D}(\mathcal{Z}) + \frac{\boldsymbol{\lambda}^{\mathcal{F}}}{\beta}, \frac{\lambda_2}{\beta}), \quad (39)$$

where $\text{soft}(\mathbf{a}, \tau) := \text{sgn}(\mathbf{a}) \cdot \max(|\mathbf{a}| - \tau, 0)$.

6) \mathcal{G} sub-problem: The sub-problem to optimize L_A with respect to \mathcal{G} is:

$$\min_{\mathcal{G}} \frac{\lambda_3}{2} \|\mathbf{G}_{(3)} - \mathbf{E} \mathbf{U}_1\|_F^2 + \frac{\beta}{2} \|\mathcal{G} - \mathcal{X} - \frac{\boldsymbol{\lambda}^{\mathcal{G}}}{\beta}\|^2, \quad (40)$$

which is equivalent to

$$\min_{\mathcal{G}} \|\mathcal{G} - \frac{1}{\lambda_3 + \beta} (\lambda_3 \text{fold}_3(\mathbf{E} \mathbf{U}_1) + \beta(\mathcal{X} + \frac{\boldsymbol{\lambda}^{\mathcal{G}}}{\beta}))\|^2. \quad (41)$$

This problem has the following closed-form solution:

$$\mathcal{G} = \frac{1}{\lambda_3 + \beta} (\lambda_3 \text{fold}_3(\mathbf{E} \mathbf{U}_1) + \beta(\mathcal{X} + \frac{\boldsymbol{\lambda}^{\mathcal{G}}}{\beta})). \quad (42)$$

7) $\mathbf{U}_1, \mathbf{U}_2$ sub-problem: Optimizing L_A with respect to \mathbf{U}_1 can be formalized as :

$$\min_{\mathbf{U}_1} \frac{\lambda_3}{2} \|\mathbf{G}_{(3)} - \mathbf{E} \mathbf{U}_1\|_F^2 + \frac{\beta}{2} \|\mathbf{U}_1 - \mathbf{V}_1 - \frac{\boldsymbol{\lambda}^{\mathbf{U}_1}}{\beta}\|^2. \quad (43)$$

This problem can be treated as solving the following linear system:

$$(\lambda_3 \mathbf{E}^T \mathbf{E} + \beta \mathbf{I}) \mathbf{U}_1 = \beta \mathbf{V}_1 + \boldsymbol{\lambda}^{\mathbf{U}_1} + \lambda_3 \mathbf{E}^T \mathbf{G}_{(3)}. \quad (44)$$

where \mathbf{I} is identity matrix. This linear system can be solved by off-the-shelf conjugate gradient techniques.

The sub-problem to optimize L_A with respect to \mathbf{U}_2 has similar form:

$$\min_{\mathbf{U}_2} \frac{\lambda_5}{2} \|(\mathbf{Y}_h)_{(3)} - \mathbf{E} \mathbf{U}_2\|_F^2 + \frac{\beta}{2} \|\mathbf{U}_2 - \mathbf{V}_2 - \frac{\boldsymbol{\lambda}^{\mathbf{U}_2}}{\beta}\|^2. \quad (45)$$

This problem can also be treated as solving a linear system:

$$(\lambda_2 \mathbf{E}^T \mathbf{E} + \beta \mathbf{I}) \mathbf{U}_2 = \beta \mathbf{V}_2 + \boldsymbol{\lambda}^{\mathbf{U}_2} + \lambda_5 \mathbf{E}^T (\mathbf{Y}_h)_{(3)}. \quad (46)$$

8) $\mathbf{V}_1, \mathbf{V}_2$ sub-problem: The sub-problem to optimize L_A with respect to \mathbf{V}_1 is:

$$\begin{aligned} \min_{\mathbf{V}_1} \lambda_4 \|\mathbf{V}_1\|_{\text{MCP}} + \frac{\beta}{2} \|\mathbf{U}_1 - \mathbf{V}_1 - \frac{\boldsymbol{\lambda}^{\mathbf{U}_1}}{\beta}\|^2 \\ \text{s.t. } \mathbf{V}_1 \geq 0. \end{aligned} \quad (47)$$

This problem can be solved using the iteratively reweighted ℓ_1 norm method as follows:

$$\begin{aligned} \mathbf{V}_1 &= [\text{soft}(\mathbf{U}_1 - \frac{\boldsymbol{\lambda}^{\mathbf{U}_1}}{\beta}, \frac{\lambda_4}{\beta})]_+; \\ \boldsymbol{\Omega} &= \psi'_{a,\lambda}(\mathbf{V}_1); \\ \mathbf{V}_1 &= [\text{soft}(\mathbf{U}_1 - \frac{\boldsymbol{\lambda}^{\mathbf{U}_1}}{\beta}, \frac{\lambda_4}{\beta} \boldsymbol{\Omega})]_+, \end{aligned} \quad (48)$$

where $\psi'_{a,\lambda}(\cdot)$ is the derivative of MCP penalty function and computed element-wisely for the matrix \mathbf{V}_1 , $\text{soft}(\cdot, \cdot)$ is the soft shrinkage operator as in Eq.(39), and $[x]_+ = \max\{x, 0\}$.

Similarly, optimizing L_A with respect to \mathbf{V}_2 can be written as :

$$\begin{aligned} & \min_{\mathbf{V}_2} \lambda_6 \|\mathbf{V}_2\|_{\text{MCP}} + \frac{\beta}{2} \|\mathbf{U}_2 - \mathbf{V}_2 - \frac{\boldsymbol{\lambda}^{\mathbf{U}_2}}{\beta}\|^2 \\ & \text{s.t. } \mathbf{V}_2 \geq 0. \end{aligned} \quad (49)$$

This problem can also be solved as in Eq.(48):

$$\begin{aligned} \mathbf{V}_2 &= [\text{soft}(\mathbf{U}_2 - \frac{\boldsymbol{\lambda}^{\mathbf{U}_2}}{\beta}, \frac{\lambda_6}{\beta})]_+; \\ \boldsymbol{\Omega} &= \psi'_{a,\lambda}(\mathbf{V}_2); \\ \mathbf{V}_2 &= [\text{soft}(\mathbf{U}_2 - \frac{\boldsymbol{\lambda}^{\mathbf{U}_2}}{\beta}, \frac{\lambda_6}{\beta} \boldsymbol{\Omega})]_+. \end{aligned} \quad (50)$$

9) **E sub-problem:** Optimizing L_A with respect to \mathbf{E} can be written as :

$$\begin{aligned} & \min_{\mathbf{E}} \frac{\lambda_3}{2} \|\mathbf{G}_{(3)} - \mathbf{E}\mathbf{U}_1\|^2 + \frac{\lambda_5}{2} \|(\mathbf{Y}_h)_{(3)} - \mathbf{E}\mathbf{U}_2\|^2 \\ & \text{s.t. } \mathbf{E} \geq 0, \end{aligned} \quad (51)$$

which is equivalent to the following optimization problem:

$$\begin{aligned} \mathbf{E} &= \arg \min_{\mathbf{E}} \|\mathbf{G}_{(3)} - \mathbf{E}\mathbf{U}_1\|^2 + \frac{\lambda_5}{\lambda_3} \|(\mathbf{Y}_h)_{(3)} - \mathbf{E}\mathbf{U}_2\|^2 \\ & \text{s.t. } \mathbf{e}_k \geq 0, \end{aligned}$$

where \mathbf{e}_k is the k -th column of \mathbf{E} . To reduce the computational complexity, similar to ODL method [59] and the spectral dictionary learning method in [31], instead of updating the whole \mathbf{E} at once, we can solve Eq.(IV-9) by using block coordinate descent [60], i.e., during each iteration we update only one column of \mathbf{E} while keeping the others fixed under the non-negative constraint. Let $\mathbf{E}^{(t)}$ denotes the dictionary obtained after the t -th iteration, $\boldsymbol{\alpha}_k$ denotes the k -th row of \mathbf{U}_1 , $\boldsymbol{\beta}_k$ denotes the k -th row of \mathbf{U}_2 , and let $\mathbf{e}_k^{(t+1)} = \mathbf{e}_k^{(t)} + \Delta \mathbf{e}_k$. Then $\Delta \mathbf{e}_k$ can be obtained by solving the following optimization problem:

$$\begin{aligned} \Delta \mathbf{e}_k &= \arg \min_{\Delta \mathbf{e}_k} \|\mathbf{G}_{(3)} - \mathbf{E}^{(t)}\mathbf{U}_1 - \Delta \mathbf{e}_k \boldsymbol{\alpha}_k\|^2 + \\ & \quad \frac{\lambda_5}{\lambda_3} \|(\mathbf{Y}_h)_{(3)} - \mathbf{E}^{(t)}\mathbf{U}_2 - \Delta \mathbf{e}_k \boldsymbol{\beta}_k\|^2 \\ & \text{s.t. } \mathbf{e}_k^{(t)} + \Delta \mathbf{e}_k \geq 0. \end{aligned} \quad (52)$$

Let $\mathbf{H}_1^{(t)} = \mathbf{G}_{(3)} - \mathbf{E}^{(t)}\mathbf{U}_1$ and $\mathbf{H}_2^{(t)} = (\mathbf{Y}_h)_{(3)} - \mathbf{E}^{(t)}\mathbf{U}_2$ denote the residual matrices after the t -th iteration, then $\mathbf{H}_1^{(t)} \in \mathbb{R}^{S \times (W \times H)}$, $\mathbf{H}_2^{(t)} \in \mathbb{R}^{S \times (w \times h)}$. Let $n_1 = W \times H$, $n_2 = w \times h$, then problem (52) can be rewritten as follows:

$$\begin{aligned} \Delta \mathbf{e}_k &= \arg \min_{\Delta \mathbf{e}_k} \|\mathbf{H}_1^{(t)} - \Delta \mathbf{e}_k \boldsymbol{\alpha}_k\|^2 + \frac{\lambda_5}{\lambda_3} \|\mathbf{H}_2^{(t)} - \Delta \mathbf{e}_k \boldsymbol{\beta}_k\|^2 \\ & \text{s.t. } \mathbf{e}_k^{(t)} + \Delta \mathbf{e}_k \geq 0 \end{aligned}$$

$$\begin{aligned} &= \arg \min_{\Delta \mathbf{e}_k} \sum_{s=1}^S \left(\sum_{i=1}^{n_1} ((\mathbf{H}_1^{(t)})_{s,i} - \Delta \mathbf{e}_k(s) \boldsymbol{\alpha}_k(i))^2 + \right. \\ & \quad \left. \frac{\lambda_5}{\lambda_3} \sum_{j=1}^{n_2} ((\mathbf{H}_2^{(t)})_{s,j} - \Delta \mathbf{e}_k(s) \boldsymbol{\beta}_k(j))^2 \right) \\ & \text{s.t. } \mathbf{e}_k^{(t)} + \Delta \mathbf{e}_k \geq 0 \end{aligned}$$

$$\begin{aligned} &= \arg \min_{\Delta \mathbf{e}_k} \sum_{s=1}^S \left(\sum_{i=1}^{n_1} \boldsymbol{\alpha}_k(i)^2 \Delta \mathbf{e}_k(s)^2 + \frac{\lambda_5}{\lambda_3} \sum_{j=1}^{n_2} \boldsymbol{\beta}_k(j)^2 \Delta \mathbf{e}_k(s)^2 - \right. \\ & \quad \left. 2(\boldsymbol{\alpha}_k(\mathbf{H}_1^{(t)})^T + \frac{\lambda_5}{\lambda_3} \boldsymbol{\beta}_k(\mathbf{H}_2^{(t)})^T) \Delta \mathbf{e}_k(s) \right) + \text{Constant} \\ & \quad \text{s.t. } \mathbf{e}_k^{(t)} + \Delta \mathbf{e}_k \geq 0 \\ &= \arg \min_{\Delta \mathbf{e}_k} \sum_{s=1}^S (\Delta \mathbf{e}_k(s)) - \frac{\boldsymbol{\alpha}_k(\mathbf{H}_1^{(t)})^T + \frac{\lambda_5}{\lambda_3} \boldsymbol{\beta}_k(\mathbf{H}_2^{(t)})^T}{\sum_{i=1}^{n_1} \boldsymbol{\alpha}_k(i)^2 + \frac{\lambda_5}{\lambda_3} \sum_{j=1}^{n_2} \boldsymbol{\beta}_k(j)^2} \\ & \quad + \text{Constant} \\ & \quad \text{s.t. } \mathbf{e}_k^{(t)} + \Delta \mathbf{e}_k \geq 0 \\ &= \arg \min_{\Delta \mathbf{e}_k} \|\Delta \mathbf{e}_k - \frac{\mathbf{H}_1^{(t)} \boldsymbol{\alpha}_k^T + \frac{\lambda_5}{\lambda_3} \mathbf{H}_2^{(t)} \boldsymbol{\beta}_k^T}{\sum_{i=1}^{n_1} \boldsymbol{\alpha}_k(i)^2 + \frac{\lambda_5}{\lambda_3} \sum_{j=1}^{n_2} \boldsymbol{\beta}_k(j)^2}\|_2^2 \\ & \quad + \text{Constant} \\ & \quad \text{s.t. } \mathbf{e}_k^{(t)} + \Delta \mathbf{e}_k \geq 0 \end{aligned} \quad (53)$$

where $\mathbf{H}_{1,s,i}^{(t)}$ and $\mathbf{H}_{2,s,j}^{(t)}$ denote the (s, i) -th element of $\mathbf{H}_1^{(t)}$ and the (s, j) -th element of $\mathbf{H}_2^{(t)}$, respectively; $\boldsymbol{\alpha}_k(i)$, $\boldsymbol{\beta}_k(j)$ and $\Delta \mathbf{e}_k(s)$ denote the i -th element of $\boldsymbol{\alpha}_k$, the j -th element of $\boldsymbol{\beta}_k$ and the s -th element of $\Delta \mathbf{e}_k$, respectively; $\mathbf{H}_{1,s,:}^{(t)}$ and $\mathbf{H}_{2,s,:}^{(t)}$ denote the s -th row of $\mathbf{H}_1^{(t)}$ and the s -th row of $\mathbf{H}_2^{(t)}$, respectively. Based on Eq.(53), we can see that the solution of Eq.(52) can be given by

$$\mathbf{e}_k^{(t+1)} = [\mathbf{e}_k^{(t)} + \frac{\mathbf{H}_1^{(t)} \boldsymbol{\alpha}_k^T + \frac{\lambda_5}{\lambda_3} \mathbf{H}_2^{(t)} \boldsymbol{\beta}_k^T}{\sum_{i=1}^{n_1} \boldsymbol{\alpha}_k(i)^2 + \frac{\lambda_5}{\lambda_3} \sum_{j=1}^{n_2} \boldsymbol{\beta}_k(j)^2}]_+. \quad (54)$$

10) **Updating Multipliers:** According to the ADMM, the multipliers associated with L_A are updated by the following formulas:

$$\left\{ \begin{array}{l} \boldsymbol{\lambda}^A \leftarrow \boldsymbol{\lambda}^A - \gamma \beta (\mathcal{A} - \mathcal{X}) \\ \boldsymbol{\lambda}^B \leftarrow \boldsymbol{\lambda}^B - \gamma \beta (\mathcal{B} - \mathcal{X}) \\ \boldsymbol{\lambda}^Z \leftarrow \boldsymbol{\lambda}^Z - \gamma \beta (\mathcal{Z} - \mathcal{X}) \\ \boldsymbol{\lambda}^F \leftarrow \boldsymbol{\lambda}^F - \gamma \beta (\mathcal{F} - \mathbf{D}(\mathcal{Z})) \\ \boldsymbol{\lambda}^G \leftarrow \boldsymbol{\lambda}^G - \gamma \beta (\mathcal{G} - \mathcal{X}) \\ \boldsymbol{\lambda}^{\mathbf{U}_1} \leftarrow \boldsymbol{\lambda}^{\mathbf{U}_1} - \gamma \beta (\mathbf{U}_1 - \mathbf{V}_1) \\ \boldsymbol{\lambda}^{\mathbf{U}_2} \leftarrow \boldsymbol{\lambda}^{\mathbf{U}_2} - \gamma \beta (\mathbf{U}_2 - \mathbf{V}_2), \end{array} \right. \quad (55)$$

where γ is a parameter associated with convergence rate of the algorithm. The proposed algorithm for model (20) can now be summarized in Algorithm 1.

V. EXPERIMENTS

A. Compared Methods and Performance Evaluation Measures

In this section, the performances of our proposed algorithm are compared with several state-of-the-art HS and MS image fusion methods, including the CNMF (coupled nonnegative matrix factorization) method [27], the HSRCSU (HS coupled spectral unmixing) method [30], the NSSR (nonnegative structured sparse representation) method [31] and the CSTF

Algorithm 1 Optimization Algorithm for The Model (20)

Input: LR-HS image \mathcal{Y}_h ; HR-MS image \mathcal{Y}_m ; Spatial down-sampling matrix \mathbf{M} ; Spectral down-sampling matrix \mathbf{P}_3 ; Algorithm parameters $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, r_1, r_2, r_3, r_4$.

Initialization: $\mathcal{X}, \mathbf{U}_1, \mathbf{U}_2$ and \mathbf{E} are randomly initialized; $\mathcal{A}, \mathcal{B}, \mathcal{Z}, \mathcal{G}$ are all initialized to be equal to \mathcal{X} ; \mathbf{V}_1 and \mathbf{V}_2 are initialized to be equal to \mathbf{U}_1 and \mathbf{U}_2 , respectively; \mathcal{F} is initialized by 0.

while not converge **do**

1. Updating \mathcal{X} and $\mathcal{C}_p, \mathbf{W}_p, \mathbf{H}_p, \mathbf{S}_p, \mathbf{N}_p$ by the following algorithm:

while not converge **do**

- (1). Updating \mathcal{C}_p via Eq.(25);
- (2). Updating $\mathbf{W}_p, \mathbf{H}_p, \mathbf{S}_p$ and \mathbf{V}_p via Eqs. (26), (27), (28) and (29).

end while

2. Updating \mathcal{A} via Eq.(32);

3. Updating \mathcal{B} via Eq.(34);

4. Updating \mathcal{Z} via Eq.(37);

5. Updating \mathcal{F} via Eq.(39);

6. Updating \mathcal{G} via Eq.(42);

7. Updating $\mathbf{U}_1, \mathbf{U}_2$ via Eq.(44), (46), respectively;

8. Updating $\mathbf{V}_1, \mathbf{V}_2$ via Eqs.(48), (50), respectively;

9. Updating \mathbf{E} via Eq.(54);

10. Updating multipliers via Eqs.(55).

end while

Output: The reconstructed HR-HS image \mathcal{X} .

(coupled sparse tensor factorization) method [32]. The assessments used to evaluate the quality of reconstructed HS images contain the root mean square error (RMSE), the peak signal-to-noise ratio(PSNR), structural similarity index (SSIM), spectral angle mapper (SAM) [61], relative dimensionless global error in synthesis (ERGAS) [62], universal image quality index (UIQI) [63], the degree of distortion (DD) and the correlation coefficient (CC). It is known that the higher values of PSNR, SSIM, UIQI and CC and the smaller values of RMSE, SAM, ERGAS and DD show better performances.

B. Experimental Data Sets

Three different data sets of HS images are used to compare the performances of our proposed method and other fusion methods. The first dataset is the Pavia University image [64] acquired by the reflective opticssystem imaging spectrometer (ROSIS) optical sensor over the urban area of the University of Pavia, Italy. This image is of size $610 \times 340 \times 115$ with a spatial resolution of 1.3 m, and the number of bands has been reduced to 93 bands after removing the water vapor absorption bands (with spectral range from 0.43 to 0.86 μm). A subimage with the size of $256 \times 256 \times 93$ is used as the groundtruth HR-HS image. Then the corresponding LR-HS image of size $64 \times 64 \times 93$ is constructed by applying a 5×5 Gaussian spatial filter with standard deviation 2.5 on each band of the HR-HS image and downampling every four pixels in both height and width directions. A four-band HR-MS image of size $256 \times 256 \times 4$

is constructed by filtering the HR-HS image with a IKONOS like reflectance spectral responses [25]. And following [61], the LR-HS and HR-MS image are both contaminated by zero-mean additive Gaussian noises, where SNR=35 dB for the first 43 bands and SNR=30 dB for the remaining 50 bands of the LR-HS image while SNR=30 dB for all bands of the HR-MS image.

The second data set is the Pavia Centre image which is also acquired by the ROSIS sensor during a flight campaign over Pavia, Italy. After removing the water vapor absorption bands and some of the samples containing no information, Pavia Centre is of size $1096 \times 715 \times 92$. A subimage with the size of $256 \times 256 \times 92$ is used as the ground truth HR-HS image, and the corresponding contaminated LR-HS and HR-MS image are constructed as in the Pavia University image.

The third data set is the Salinas image which is collected by the NASA's Airborne Visible and Infrared Imaging Spectrometer (AVIRIS) [65] over Salinas Valley, California. This image is of size $512 \times 217 \times 224$ covering the wavelength range 400nm-2500nm with 10nm sampling interval. The number of bands is reduced to 194 bands after removing the bands covering heavy noise and water absorption, then a subimage with the size of $200 \times 200 \times 194$ is used as the ground truth HR-HS image. The corresponding LR-HS image of size $50 \times 50 \times 194$ is constructed by applying a 5×5 Gaussian kernel with standard deviation 2.5 following by a down-sampling factor of 4. The HR-MS image is generated by down-sampling the original images in the spectral domain using the spectral transform operator of the Landsat TM [27], where the spectral response is transformed according to Landsat TM bands 1-5 and 7, covering the spectral regions of 450-520, 520-600, 630-690, 760-900, 1550-1750, and 2080-2350 nm [66]. Then the LR-HS and HR-MS image are both contaminated by zero-mean additive Gaussian noises, where SNR=40 dB for the LR-HS image and SNR=35 dB for the HR-MS image.

C. Experimental Results

In this section, we will compare the fusion results of the five tested methods on the three data sets.

Table II shows the RMSE, PSNR, SSIM, SAM, ERGAS, UIQI, DD and CC of the reconstructed HR-HS image by CNMF, HSRCSU, NSSR, CSTF and our method for the Pavia University data set, where the best results are marked in bold for clarity. As we can see from Table II, our method can get the best results for all the assessment indices in this data set, which shows the superiority of our method to the other fusion methods in terms of reconstruction accuracy. This is because we have fully exploited the underlying priors of the targeted HR-HS image, including spatial-spectral correlation, nonlocal self-similarity and spatial-spectral continuity. In addition, the term of spectral unmixing based on MCP regularization has effectively reduced the spectral distortions to further improve the reconstruction performances.

In order to further intuitively compare the performances of those fusion methods, we show the 20th and 60th bands and corresponding error images reconstructed by NSSR, CSTF and our method in Fig. 1. From Fig. 1, we can see while all this

tested methods can well reconstruct the spatial structures of the HR-HS image, our method obviously outperforms the others in recovering the details of the original HR-HS image.

Table II shows the quantitative results of the Pavia Centre data set. We can see from Table II that our method gets the best scores for all the assessments. The 10th and 50th bands and corresponding error images reconstructed by NSSR, CSTF and our method are shown in Fig. 2, in which we can see that our method outperforms the other two methods in recovering image details.

The quantitative results of those tested methods on the Salinas data set are also shown in Table II. The corresponding error images for the 30th and 80th bands reconstructed by NSSR, CSTF and our method are shown in Fig. 3. We can see that our method can get the best reconstruction results both in terms of quantitative assessments and recovering image details.

Fig. 4 shows the PSNR, SSIM and UIQI curves as functions of the spectral bands over the (a) Pavia University, (b) Pavia Centre and (c) Salinas data set for the test methods. It can be seen that our method consistently outperforms other methods at most of the spectral bands for all the three data sets.

To verify the effectiveness of the spectral unmixing term in our model, we further compare our method with a simplified version of our model without the spectral unmixing term, i.e., let the parameters $\lambda_3 = \lambda_4 = \lambda_5 = \lambda_6 = 0$ in our model (20), and the quantitative results are shown in Table III. We can see from Table III that the spectral unmixing term in our model is assuredly able to improve the performances of our method. And as an example, Fig. 5 shows the 70th bands of Pavia Centre data set and corresponding error images reconstructed by the original version and simplified version (without the spectral unmixing term) of our method. We can see that compared with the simplified version, the original version method performs better in recovering the details in the image, which further proves the effectiveness of the spectral unmixing term in our model. In addition to this, we respectively select two pixels in every tested data set and plot their spectral signature curves estimated by all the compared methods in Fig. 6, and we can see that the original version method can recover more accurate spectral signature than the simplified version, which testifies the effectiveness of the spectral unmixing term in reducing spectral distortion.

D. Empirical Analysis for Algorithm Convergence

We provide an empirical analysis for the convergence of the proposed optimization algorithm on the three test data sets. We use the relative change $\text{relChgX} := \frac{\|\mathcal{X}^k - \mathcal{X}^{k-1}\|_F}{\|\mathcal{X}^{k-1}\|_F}$, the relative error $\text{relErrX} := \frac{\|\mathcal{X}^k - \mathcal{X}_0\|_F}{\|\mathcal{X}_0\|_F}$, the objective function value in model (20) and the PSNR value as the four assessment indices of algorithm convergence, where \mathcal{X}^k is the recovered result in k -th iteration and \mathcal{X}_0 is the ground truth HR-HS image.

In Fig. 7, we show the curves of the relative change, the relative error, the objective function value and the PSNR as functions of the iteration steps in the proposed optimization algorithm for the three test data sets. We observe from Fig. 7 that the curves of all assessment indexes come to a stable value when the algorithm reach a relatively high iteration number,

which suggests that the proposed algorithm converges well empirically.

E. Implementation Issues and Parameters Discussions

In Algorithm 1, those parameters can be divided into three groups based on their characteristics: (1) $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$ and λ_6 provide a trade-off between the reconstruction error term, the 3D-TV term and the spectral unmixing term in the model, in our experiments those parameters are empirically fixed, i.e., $\lambda_1 = 1, \lambda_2 = 10^{-4}, \lambda_3 = 0.2, \lambda_4 = 0.001 * \lambda_3, \lambda_5 = 0.1, \lambda_6 = 0.001 * \lambda_5$; (2) the size of 3D patch w_{patch} , the size of search window around one patch S_{patch} , the number of collected similar 3D patches N_{patch} and the sliding distance d_{patch} are the parameters to construct the nonlocal low rank 4D tensors, where $w_{patch}, S_{patch}, N_{patch}$ and d_{patch} are empirically set as 8, 36, 45, 7, respectively, [67], [68]; (3) r_1, r_2, r_3, r_4 control the complexity of spatial redundancy, spectral redundancy and similarity of 3D patches, respectively, where r_1, r_2, r_4 are empirically fixed as 8, 8, $\text{ceil}^1(N_{patch} * 0.45)$, respectively. Actually r_3 is the only parameter that needs to tune carefully in our method. r_3 controls the spectral correlation of the reconstructed HR-HS image and it may vary for different HS images obviously. We empirically found that r_3 can be tuned in the range [3,10] for various HS images, and for the three test data sets in our experiments $r_3 = 3$ is a proper choice. All the parameters in Algorithm 1 and their corresponding values in our experiments are summarized in Table IV.

VI. CONCLUSION

This paper presents an effective HS and MS image fusion method based on nonlocal low-rank tensor decomposition and spectral unmixing. To fully exploit the high spatial-spectral correlations and the nonlocal self-similarity of the HR-HS image, we first segment it into many overlapped 3D patches and cluster those patches into some clusters, which can be considered as four-order tensors. Then the Tucker decomposition is used to formulate the low-rank property of those tensors, and the corresponding MS image is used as a data consistency constraint to maintain the spatial information. Meanwhile, a 3D total variation (3D-TV) is used as a regularizer to exploit the spatial-spectral continuity of the HS image. Finally, a MCP regularization based linear spectral unmixing is applied as a significant spectral regularizer to reduce spectral distortions. Based on the aforementioned prior knowledge, a unified optimization model is constructed to formulate the HS and MS fusion problem and an efficient ADMM based algorithm is developed to solve the resulting model. Experimental results on three different data sets demonstrate that the proposed method can achieve smaller reconstruction errors and better visual quality than several state-of-the-art HS and MS fusion methods.

¹ $\text{ceil}(a)$ indicates the smallest integer larger than a .

TABLE II
QUANTITATIVE RESULTS OF THE TEST METHODS ON THE PAVIA UNIVERSITY,
PAVIA CENTRE AND SALINAS DATA SETS.

Pavia University							
Methods	RMSE	PSNR	SSIM	SAM	ERGAS	UIQI	DD
CNMF	2.997	38.598	0.952	1.969	1.072	0.970	2.208
HSRCSU	3.192	38.050	0.958	1.774	1.155	0.967	2.164
NSSR	2.669	39.603	0.966	1.741	0.943	0.977	1.905
CSTF	2.549	40.003	0.961	1.687	0.921	0.976	1.899
OURS	2.114	41.627	0.978	1.298	0.756	0.985	1.454
Pavia Centre							
Methods	RMSE	PSNR	SSIM	SAM	ERGAS	UIQI	DD
CNMF	2.853	39.025	0.964	2.063	1.162	0.978	2.084
HSRCSU	2.985	38.633	0.973	1.807	1.229	0.978	1.913
NSSR	2.452	40.340	0.977	1.874	0.995	0.984	1.783
CSTF	2.230	41.164	0.975	1.748	0.909	0.985	1.682
OURS	2.039	41.943	0.982	1.369	0.827	0.989	1.451
Salinas							
Methods	RMSE	PSNR	SSIM	SAM	ERGAS	UIQI	DD
CNMF	3.312	37.729	0.941	1.669	0.933	0.953	2.115
HSRCSU	3.825	36.478	0.948	1.638	1.182	0.952	2.186
NSSR	2.935	38.776	0.954	1.498	0.823	0.963	1.770
CSTF	2.926	38.805	0.952	1.488	0.815	0.963	1.768
OURS	2.775	39.266	0.963	1.399	0.795	0.970	1.523

TABLE III
QUANTITATIVE RESULTS OF THE SIMPLIFIED AND ORIGINAL VERSION OF OUR METHOD
ON THE PAVIA UNIVERSITY, PAVIA CENTRE AND SALINAS DATA SETS.

Pavia University							
Methods	RMSE	PSNR	SSIM	SAM	ERGAS	UIQI	DD
Simplified Version	2.133	41.550	0.978	1.307	0.759	0.985	1.457
Original Version	2.114	41.627	0.978	1.298	0.756	0.985	1.454
Pavia Centre							
Methods	RMSE	PSNR	SSIM	SAM	ERGAS	UIQI	DD
Simplified Version	2.148	41.491	0.980	1.424	0.858	0.988	1.481
Original Version	2.039	41.943	0.982	1.369	0.827	0.989	1.451
Salinas							
Methods	RMSE	PSNR	SSIM	SAM	ERGAS	UIQI	DD
Simplified Version	2.834	39.083	0.960	1.408	0.798	0.967	1.574
Original Version	2.775	39.266	0.963	1.399	0.795	0.970	1.523

TABLE IV
PARAMETERS AND THEIR CORRESPONDING VALUES IN ALGORITHM 1.

Trade-off	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
	1	0.0001	0.2	0.0002	0.1	0.0001
3D patch	w_{patch}	S_{patch}	N_{patch}	d_{patch}		
	8	36	45	7		
Redundancy	r_1	r_2	r_3	r_4		
	8	8	3	21		

REFERENCES

- [1] F. D. Van der Meer, H. M. Van der Werff, F. J. Van Ruitenbeek, C. A. Hecker, W. H. Bakker, M. F. Noomen, M. Van Der Meijde, E. J. M. Carranza, J. B. De Smeth, and T. Woldai, “Multi-and hyperspectral geologic remote sensing: A review,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 14, no. 1, pp. 112–128, 2012.
- [2] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, “Hyperspectral remote sensing data analysis and future challenges,” *IEEE Geoscience and remote sensing magazine*, vol. 1, no. 2, pp. 6–36, 2013.
- [3] W. Song, S. Li, L. Fang, and T. Lu, “Hyperspectral image classification with deep feature fusion network,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 6, pp. 3173–3184, 2018.
- [4] L. Fang, N. He, S. Li, P. Ghamisi, and J. A. Benediktsson, “Extinction profiles fusion for hyperspectral images classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 3, pp. 1803–1815, 2018.
- [5] Y. Yuan, X. Zheng, and X. Lu, “Discovering diverse subset for unsupervised hyperspectral band selection,” *IEEE Transactions on Image Processing*, vol. 26, no. 1, pp. 51–64, 2017.
- [6] D. Lorente, N. Aleixos, J. Gómez-Sanchis, S. Cubero, O. L. García-Navarrete, and J. Blasco, “Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment,” *Food and Bioprocess Technology*, vol. 5, no. 4, pp. 1121–1142, 2012.
- [7] Y. Zhang, B. Du, L. Zhang, and T. Liu, “Joint sparse representation and multitask learning for hyperspectral target detection,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 2, pp. 894–906, 2017.
- [8] J. Yang, Y.-Q. Zhao, J. C.-W. Chan, and S. G. Kong, “Coupled sparse denoising and unmixing with low-rank constraint for hyperspectral image,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 3, pp. 1818–1833, 2016.
- [9] H. Van Nguyen, A. Banerjee, and R. Chellappa, “Tracking via object reflectance using a hyperspectral video camera,” in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*. IEEE, 2010, pp. 44–51.
- [10] L. Loncan, L. B. De Almeida, J. M. Bioucas-Dias, X. Briottet, J. Chanussot, N. Dobigeon, S. Fabre, W. Liao, G. A. Licciardi, M. Simoes *et al.*, “Hyperspectral pansharpening: A review,” *IEEE Geoscience and remote*

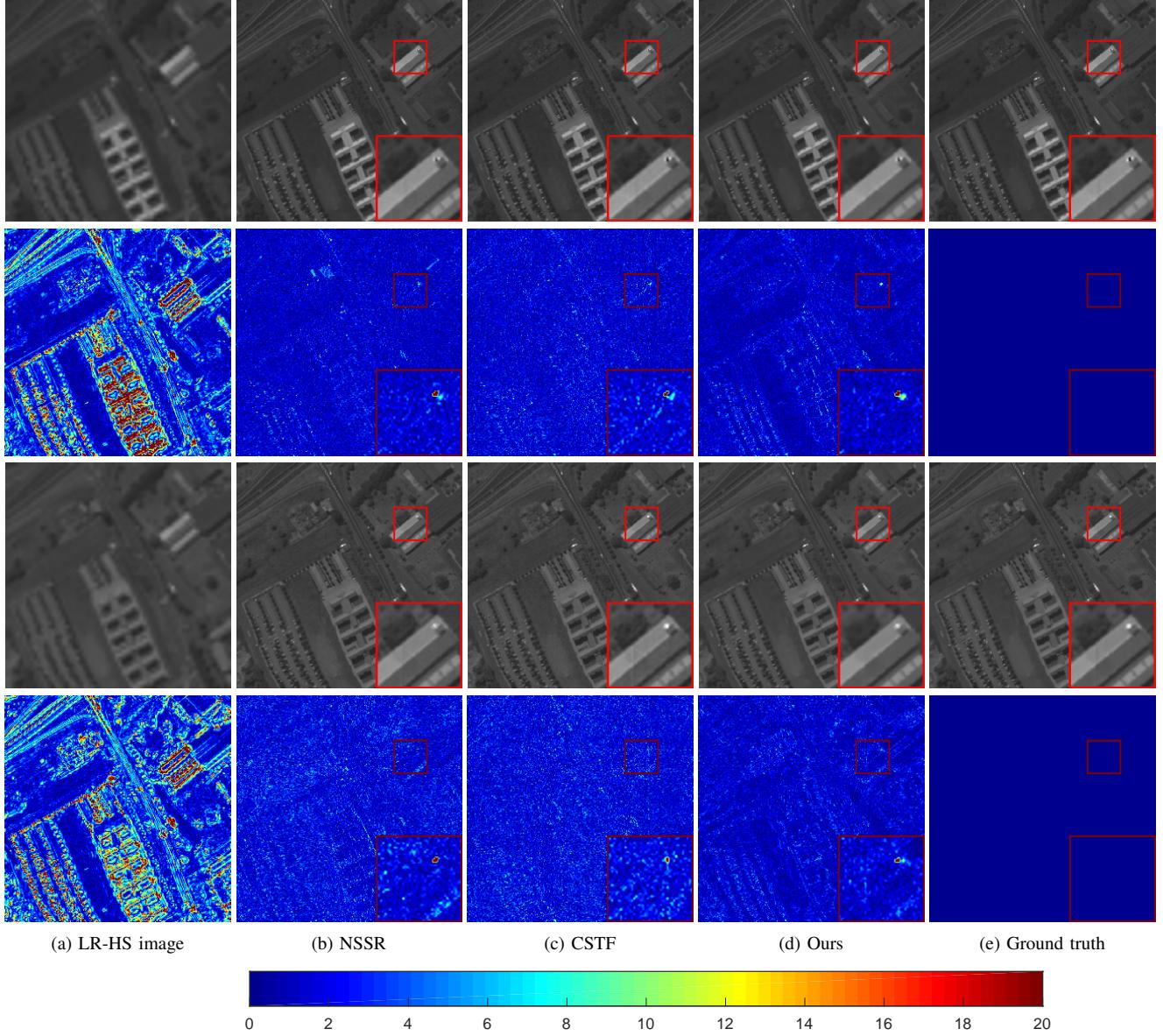


Fig. 1. Reconstructed images and corresponding error images of Pavia University for the 20th and 60th bands. The first and second rows show the reconstructed images for the 20th band and corresponding error images, respectively; The third and forth rows show the reconstructed images for the 60th band and corresponding error images, respectively; (a) LR-HS image; (b) the NSSR method; (c) the CSTF method ; (d) Our method; (e) Ground truth.

- sensing magazine*, vol. 3, no. 3, pp. 27–46, 2015.
- [11] N. Yokoya, C. Grohnfeldt, and J. Chanussot, “Hyperspectral and multispectral data fusion: A comparative review of the recent literature,” *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 2, pp. 29–56, 2017.
 - [12] H. Kaufmann, K. Segl, S. Chabrillat, S. Hofer, T. Stoffler, A. Mueller, R. Richter, G. Schreier, R. Haydn, and H. Bach, “Enmap a hyperspectral sensor for environmental mapping and analysis,” in *2006 IEEE International Symposium on Geoscience and Remote Sensing*. IEEE, 2006, pp. 1617–1619.
 - [13] X. X. Zhu, C. Grohnfeldt, and R. Bamler, “Exploiting joint sparsity for pansharpening: the j-sparsefi algorithm,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 5, pp. 2664–2681, 2016.
 - [14] G. Vivone, L. Alparone, J. Chanussot, M. Dalla Mura, A. Garzelli, G. A. Licciardi, R. Restaino, and L. Wald, “A critical comparison among pansharpening algorithms,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 5, pp. 2565–2586, 2015.
 - [15] Y. Wei, Q. Yuan, H. Shen, and L. Zhang, “Boosting the accuracy of multispectral image pansharpening by learning a deep residual network,” *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 10, pp. 1795–1799, 2017.
 - [16] X. X. Zhu and R. Bamler, “A sparse image fusion algorithm with application to pan-sharpening,” *IEEE transactions on geoscience and remote sensing*, vol. 51, no. 5, pp. 2827–2836, 2013.
 - [17] B. Aiazzi, S. Baronti, and M. Selva, “Improving component substitution pansharpening through multivariate regression of ms + pan data,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 10, pp. 3230–3239, 2007.
 - [18] W. CARPER, T. LILLESAND, and R. KIEFER, “The use of intensity-hue-saturation transformations for merging spot panchromatic and multispectral image data,” *Photogrammetric Engineering and remote sensing*, vol. 56, no. 4, pp. 459–467, 1990.
 - [19] J. Liu, “Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details,” *International Journal of Remote Sensing*, vol. 21, no. 18, pp. 3461–3472, 2000.
 - [20] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, “Mtf-tailored multiscale fusion of high-resolution ms and pan imagery,” *Photogrammetric Engineering & Remote Sensing*, vol. 72, no. 5, pp. 591–596, 2006.
 - [21] S. Li and B. Yang, “A new pan-sharpening method using a compressed

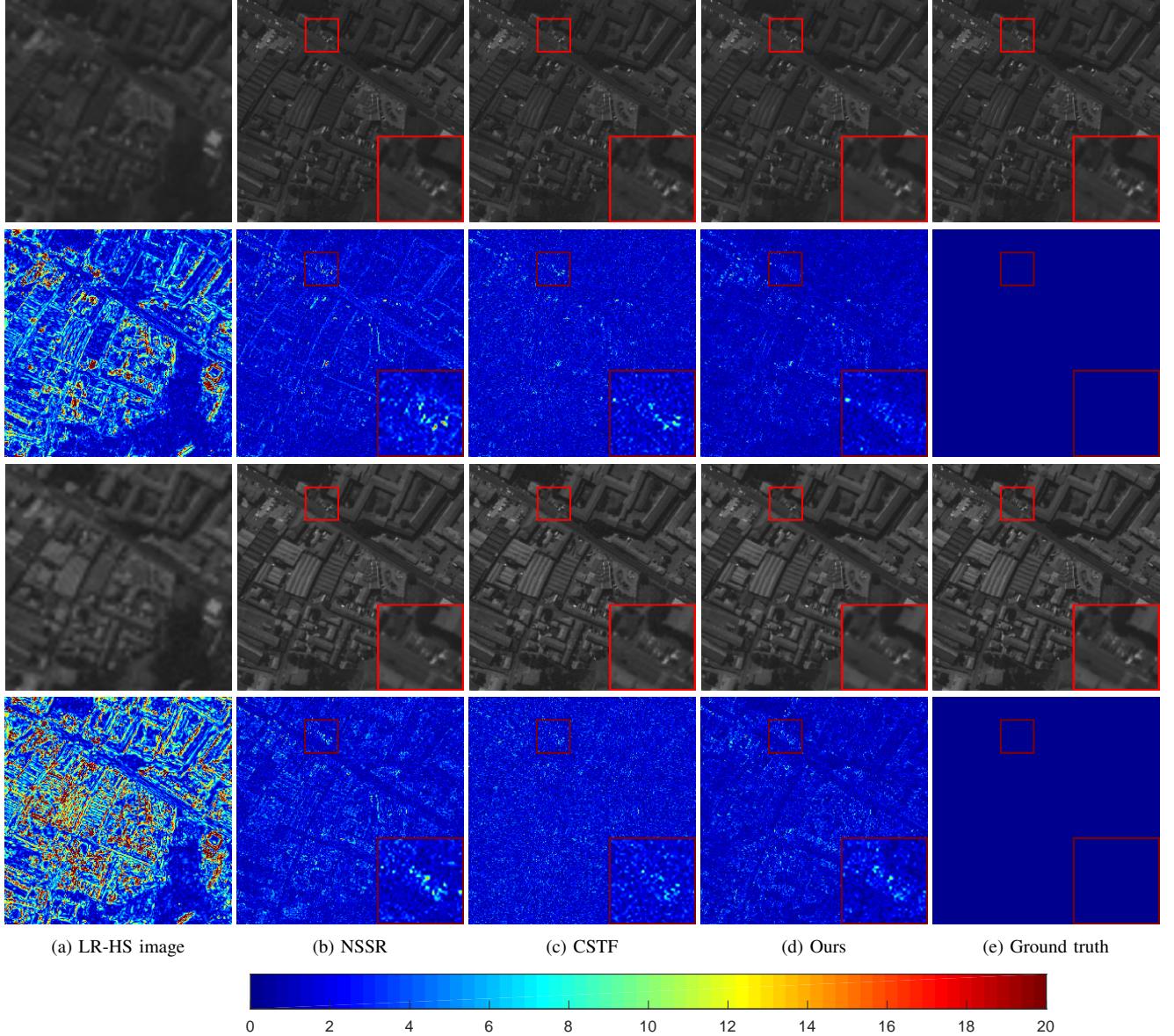


Fig. 2. Reconstructed images and corresponding error images of Pavia Centre for the 10th and 50th bands. The first and second rows show the reconstructed images for the 10th band and corresponding error images, respectively; The third and forth rows show the reconstructed images for the 50th band and corresponding error images, respectively; (a) LR-HS image; (b) the NSSR method; (c) the CSTF method ; (d) Our method; (e) Ground truth.

- sensing technique,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 2, pp. 738–746, 2011.
- [22] R. B. Gomez, A. Jazaeri, and M. Kafatos, “Wavelet-based hyperspectral and multispectral image fusion,” in *Geo-Spatial Image and Data Exploitation II*, vol. 4383. International Society for Optics and Photonics, 2001, pp. 36–43.
- [23] Y. Zhang and M. He, “Multi-spectral and hyperspectral image fusion using 3-d wavelet transform,” *Journal of Electronics (China)*, vol. 24, no. 2, pp. 218–224, 2007.
- [24] R. C. Hardie, M. T. Eismann, and G. L. Wilson, “Map estimation for hyperspectral image resolution enhancement using an auxiliary sensor,” *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1174–1184, 2004.
- [25] Q. Wei, N. Dobigeon, and J.-Y. Tourneret, “Fast fusion of multi-band images based on solving a sylvester equation,” *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 4109–4121, 2015.
- [26] N. Akhtar, F. Shafait, and A. Mian, “Bayesian sparse representation for hyperspectral image super resolution,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3631–3640.
- [27] N. Yokoya, T. Yairi, and A. Iwasaki, “Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 2, pp. 528–537, 2012.
- [28] R. Kawakami, Y. Matsushita, J. Wright, M. Ben-Ezra, Y.-W. Tai, and K. Ikeuchi, “High-resolution hyperspectral imaging via matrix factorization,” in *CVPR 2011*. IEEE, 2011, pp. 2329–2336.
- [29] B. Huang, H. Song, H. Cui, J. Peng, and Z. Xu, “Spatial and spectral image fusion using sparse matrix factorization,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 3, pp. 1693–1704, 2014.
- [30] C. Lanaras, E. Baltsavias, and K. Schindler, “Hyperspectral super-resolution by coupled spectral unmixing,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 3586–3594.
- [31] W. Dong, F. Fu, G. Shi, X. Cao, J. Wu, G. Li, and X. Li, “Hyperspectral image super-resolution via non-negative structured sparse representation,” *IEEE Transactions on Image Processing*, vol. 25, no. 5, pp. 2337–2352, 2016.
- [32] S. Li, R. Dian, L. Fang, and J. M. Bioucas-Dias, “Fusing hyperspectral and multispectral images via coupled sparse tensor factorization,” *IEEE Transactions on Image Processing*, vol. 27, no. 8, pp. 4118–4130, 2018.

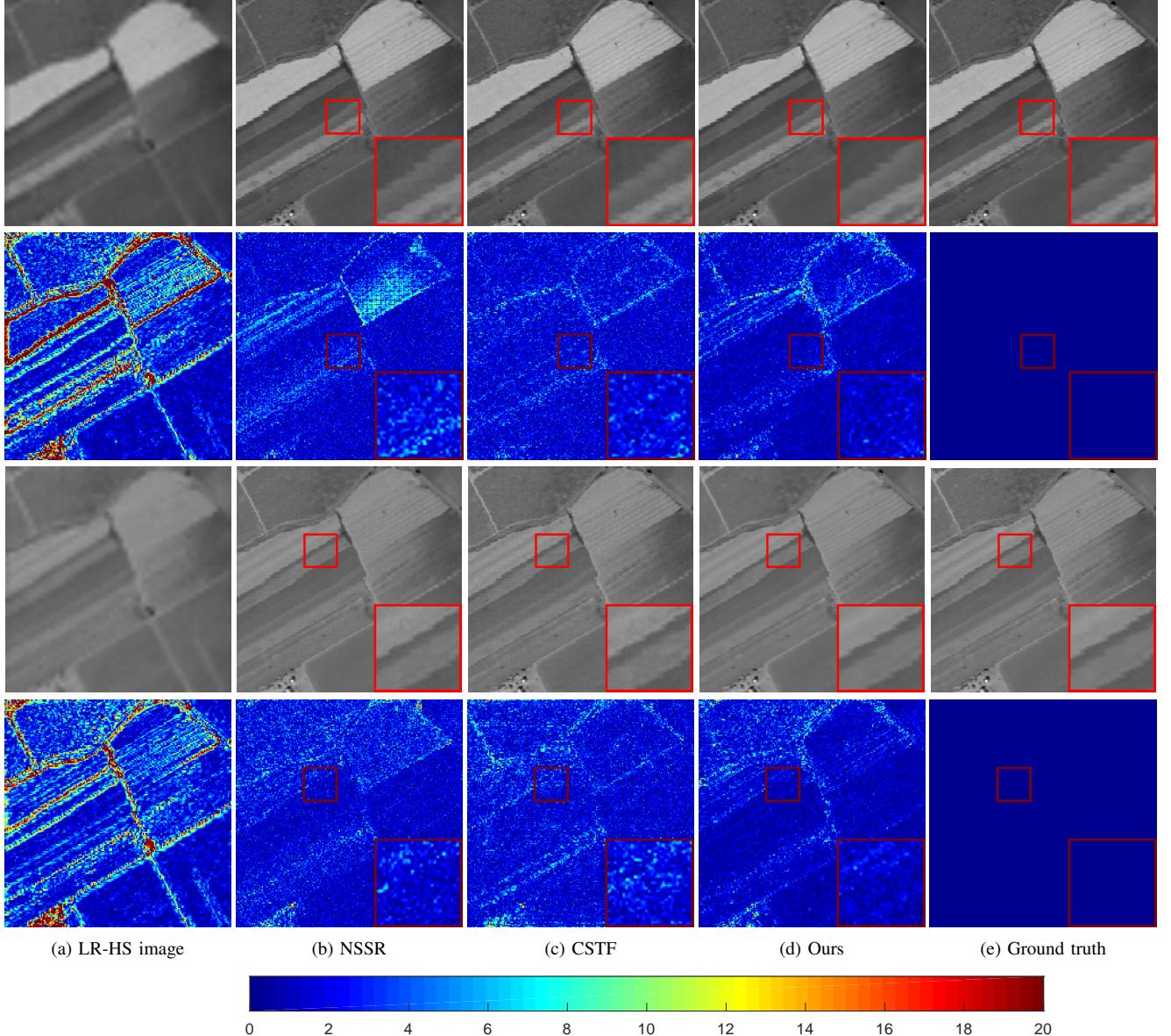


Fig. 3. Reconstructed images and corresponding error images of Salinas for the 30th and 80th bands. The first and second rows show the reconstructed images for the 30th band and corresponding error images, respectively; The third and forth rows show the reconstructed images for the 80th band and corresponding error images, respectively; (a) LR-HS image; (b) the NSSR method; (c) the CSTF method ; (d) Our method; (e) Ground truth.

- [33] Y. Xu, Z. Wu, J. Chanussot, and Z. Wei, "Nonlocal patch tensor sparse representation for hyperspectral image super-resolution," *IEEE Transactions on Image Processing*, vol. 28, no. 6, pp. 3034–3047, 2019.
- [34] Y. Wang, L. Lin, Q. Zhao, T. Yue, M. Deyu, and L. Yee, "Compressive sensing of hyperspectral images via joint tensor tucker decomposition and weighted total variation regularization," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 12, pp. 2457–2461, 2017.
- [35] Y. Wang, J. Peng, Q. Zhao, M. Deyu, L. Yee, and Z. Xi-Le, "Hyperspectral image restoration via total variation regularized low-rank tensor decomposition," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 4, pp. 1227–1243, 2018.
- [36] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.
- [37] A. Cichocki, R. Zdunek, A. H. Phan, and S.-i. Amari, *Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation*. John Wiley & Sons, 2009.
- [38] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, 2009, pp. 2272–2279.
- [39] Y. Peng, D. Meng, Z. Xu, C. Gao, Y. Yang, and B. Zhang, "De-composable nonlocal tensor dictionary learning for multispectral image denoising," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 2949–2956.
- [40] W. Cao, Y. Wang, J. Sun, D. Meng, C. Yang, A. Cichocki, and Z. Xu, "Total variation regularized tensor rpca for background subtraction from compressive measurements," *IEEE Transactions on Image Processing*, vol. 25, no. 9, pp. 4075–4090, 2016.
- [41] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966.
- [42] W. He, H. Zhang, L. Zhang, and H. Shen, "Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 1, pp. 178–188, 2016.
- [43] Z. Wu, Q. Wang, Z. Wu, and Y. Shen, "Total variation-regularized weighted nuclear norm minimization for hyperspectral image mixed denoising," *Journal of Electronic Imaging*, vol. 25, no. 1, p. 013037, 2016.
- [44] C. Jiang, H. Zhang, L. Zhang, H. Shen, and Q. Yuan, "Hyperspectral image denoising with a combined spatial and spectral weighted hyperspectral total variation model," *Canadian Journal of Remote Sensing*,

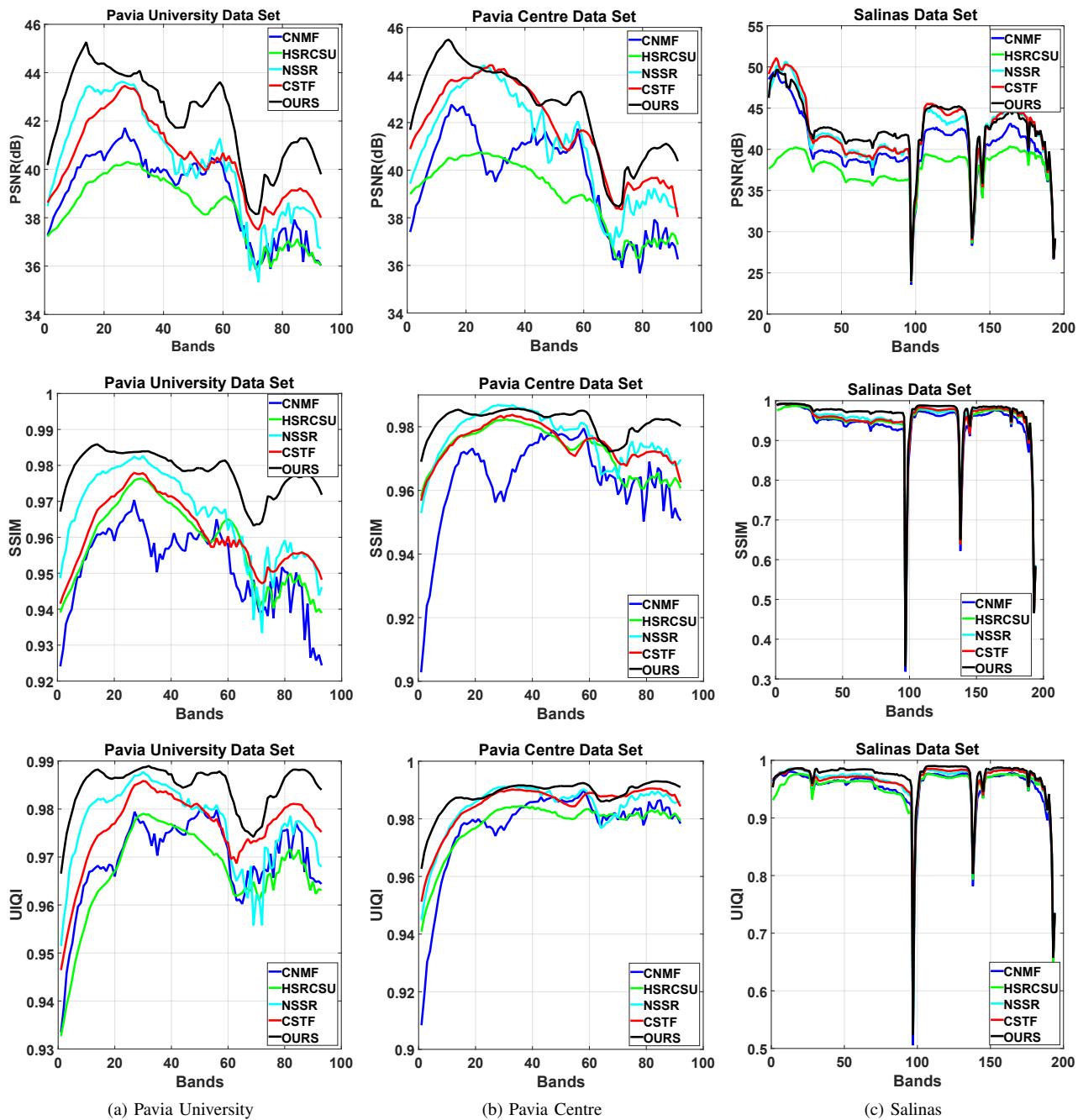


Fig. 4. PSNR, SSIM and UIQI values of each band in (a) Pavia University, (b) Pavia Centre and (c) Salinas data set.

- vol. 42, no. 1, pp. 53–72, 2016.
- [45] N. Keshava and J. F. Mustard, “Spectral unmixing,” *IEEE signal processing magazine*, vol. 19, no. 1, pp. 44–57, 2002.
- [46] C. Yi, Y.-Q. Zhao, J. Yang, J. C.-W. Chan, and S. G. Kong, “Joint hyperspectral superresolution and unmixing with interactive feedback,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 7, pp. 3823–3834, 2017.
- [47] H. Zou, “The adaptive lasso and its oracle properties,” *Journal of the American statistical association*, vol. 101, no. 476, pp. 1418–1429, 2006.
- [48] C. Leng, Y. Lin, and G. Wahba, “A note on the lasso and related procedures in model selection,” *Statistica Sinica*, pp. 1273–1284, 2006.
- [49] J. Fan and R. Li, “Variable selection via nonconcave penalized likelihood and its oracle properties,” *Journal of the American statistical Association*, vol. 96, no. 456, pp. 1348–1360, 2001.
- [50] C.-H. Zhang *et al.*, “Nearly unbiased variable selection under minimax concave penalty,” *The Annals of statistics*, vol. 38, no. 2, pp. 894–942, 2010.
- [51] C.-H. Zhang, T. Zhang *et al.*, “A general theory of concave regularization for high-dimensional sparse estimation problems,” *Statistical Science*, vol. 27, no. 4, pp. 576–593, 2012.
- [52] J. Fan, L. Xue, and H. Zou, “Strong oracle optimality of folded concave penalized estimation,” *Annals of statistics*, vol. 42, no. 3, p. 819, 2014.
- [53] S. Wang, D. Liu, and Z. Zhang, “Nonconvex relaxation approaches to robust matrix recovery,” in *IJCAI*, 2013, pp. 1764–1770.
- [54] D. Liu, T. Zhou, H. Qian, C. Xu, and Z. Zhang, “A nearly unbiased matrix completion approach,” in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 2013, pp. 210–225.
- [55] W. Cao, Y. Wang, C. Yang, X. Chang, Z. Han, and Z. Xu, “Folded-concave penalization approaches to tensor completion,” *Neurocomputing*, vol. 152, pp. 261–273, 2015.
- [56] M. Simões, J. Bioucas-Dias, L. B. Almeida, and J. Chanussot, “A convex formulation for hyperspectral image superresolution via subspace-based regularization,” *IEEE Transactions on Geoscience and Remote Sensing*,

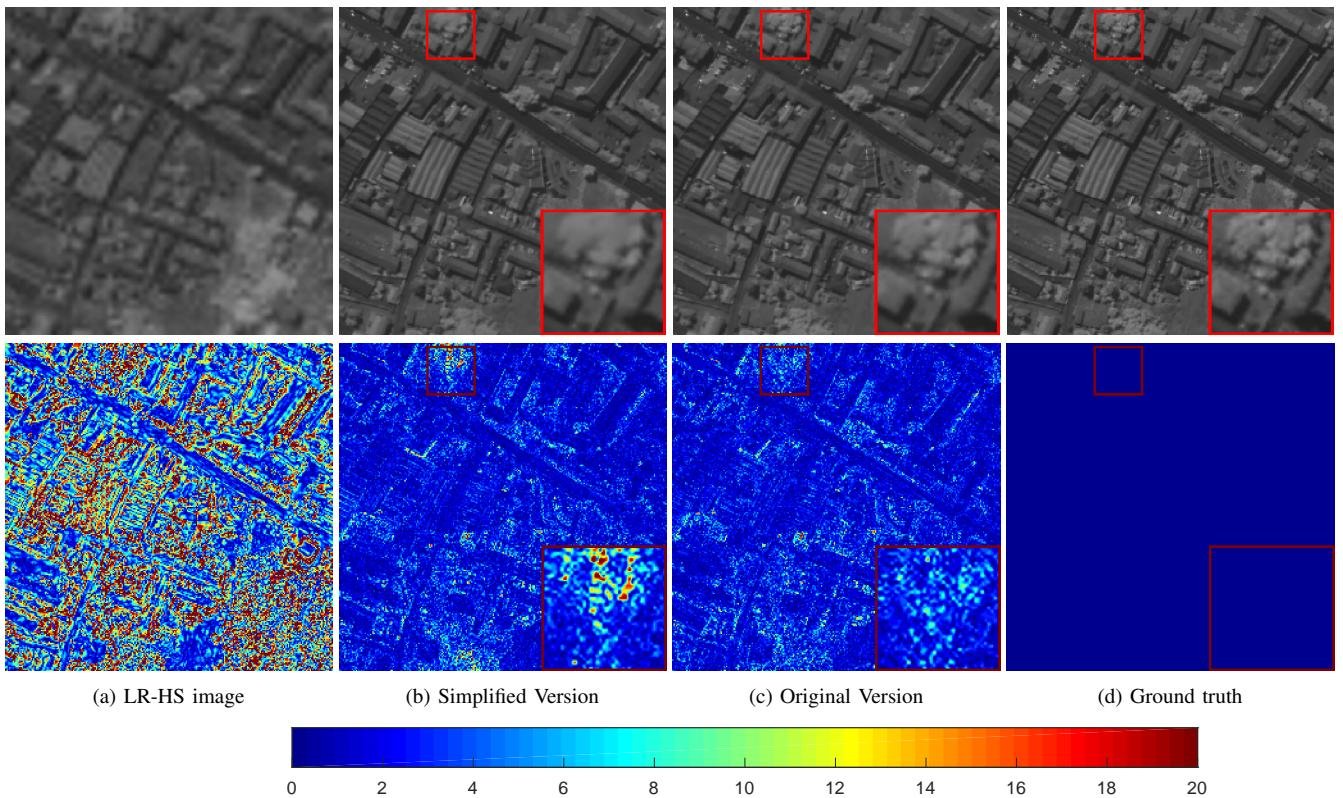


Fig. 5. Reconstructed images and corresponding error images of Pavia Centre for the 70th band. (a) LR-HS image; (b) the simplified version method; (c) the original version method ; (d) Ground truth.

- vol. 53, no. 6, pp. 3373–3388, 2015.
- [57] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.*, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.
 - [58] D. P. Robinson and R. Tappenden, “A flexible admm algorithm for big data applications,” *Journal of Scientific Computing*, vol. 71, no. 1, pp. 435–467, 2017.
 - [59] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, “Online learning for matrix factorization and sparse coding,” *Journal of Machine Learning Research*, vol. 11, no. Jan, pp. 19–60, 2010.
 - [60] J. Friedman, T. Hastie, H. Höfling, R. Tibshirani *et al.*, “Pathwise coordinate optimization,” *The Annals of Applied Statistics*, vol. 1, no. 2, pp. 302–332, 2007.
 - [61] Q. Wei, J. Bioucas-Dias, N. Dobigeon, and J.-Y. Tourneret, “Hyperspectral and multispectral image fusion based on a sparse representation,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 7, pp. 3658–3668, 2015.
 - [62] L. Wald, “Quality of high resolution synthesised images: Is there a simple criterion?” in *Third conference “Fusion of Earth data: merging point measurements, raster maps and remotely sensed images”*. SEE/URISCA, 2000, pp. 99–103.
 - [63] Z. Wang and A. C. Bovik, “A universal image quality index,” *IEEE signal processing letters*, vol. 9, no. 3, pp. 81–84, 2002.
 - [64] F. Dell’Acqua, P. Gamba, A. Ferrari, J. A. Palmason, J. A. Benediktsson, and K. Árnason, “Exploiting spectral and spatial information in hyperspectral urban data with high resolution,” *IEEE Geoscience and Remote Sensing Letters*, vol. 1, no. 4, pp. 322–326, 2004.
 - [65] R. O. Green, M. L. Eastwood, C. M. Sarture, T. G. Chrien, M. Aronsson, B. J. Chippendale, J. A. Faust, B. E. Pavri, C. J. Chovit, M. Solis *et al.*, “Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (aviris),” *Remote sensing of environment*, vol. 65, no. 3, pp. 227–248, 1998.
 - [66] M. T. Eismann and R. C. Hardie, “Hyperspectral resolution enhancement using high-resolution multispectral imagery with arbitrary response functions,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 455–465, 2005.
 - [67] W. Dong, G. Shi, and X. Li, “Nonlocal image restoration with bilateral variance estimation: a low-rank approach,” *IEEE transactions on image processing*, vol. 22, no. 2, pp. 700–711, 2013.
 - [68] W. Dong, G. Shi, X. Li, Y. Ma, and F. Huang, “Compressive sensing via nonlocal low-rank regularization,” *IEEE Transactions on Image Processing*, vol. 23, no. 8, pp. 3618–3632, 2014.

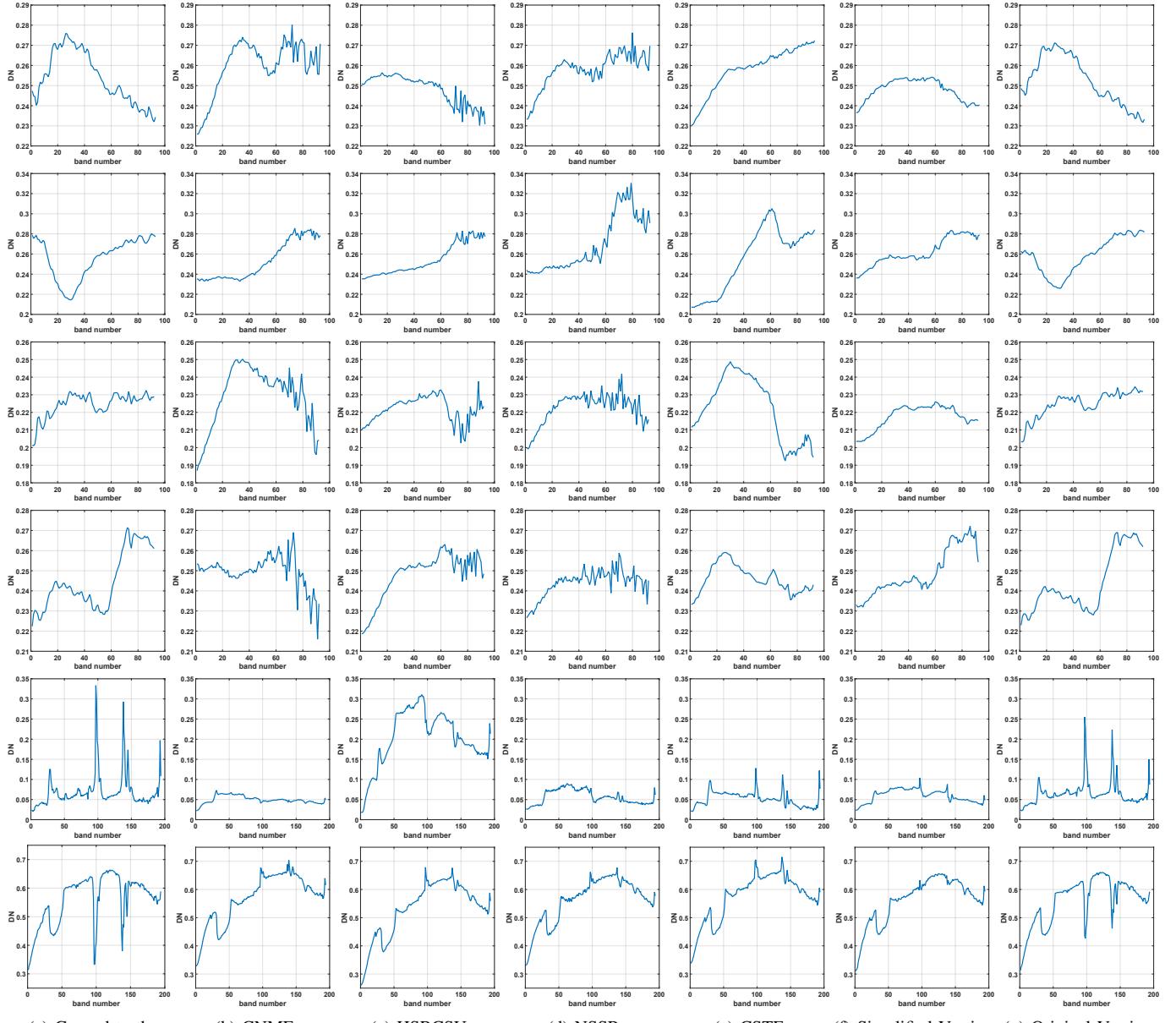


Fig. 6. Spectral signature curves estimated by all the compared methods: (a) the original spectral signatures; (b) CNMF; (c) HSRCSU; (d) NSSR; (e) CSTF; (f) the simplified version of our method; (g) the original version of our method. The first two rows: pixel (76,134) and (133,135) in Pavia University data set; the middle two rows: pixel (83,42) and (84,123) in Pavia Centre data set; the last two rows: pixel (127,128) and (186,44) in Salinas data set.

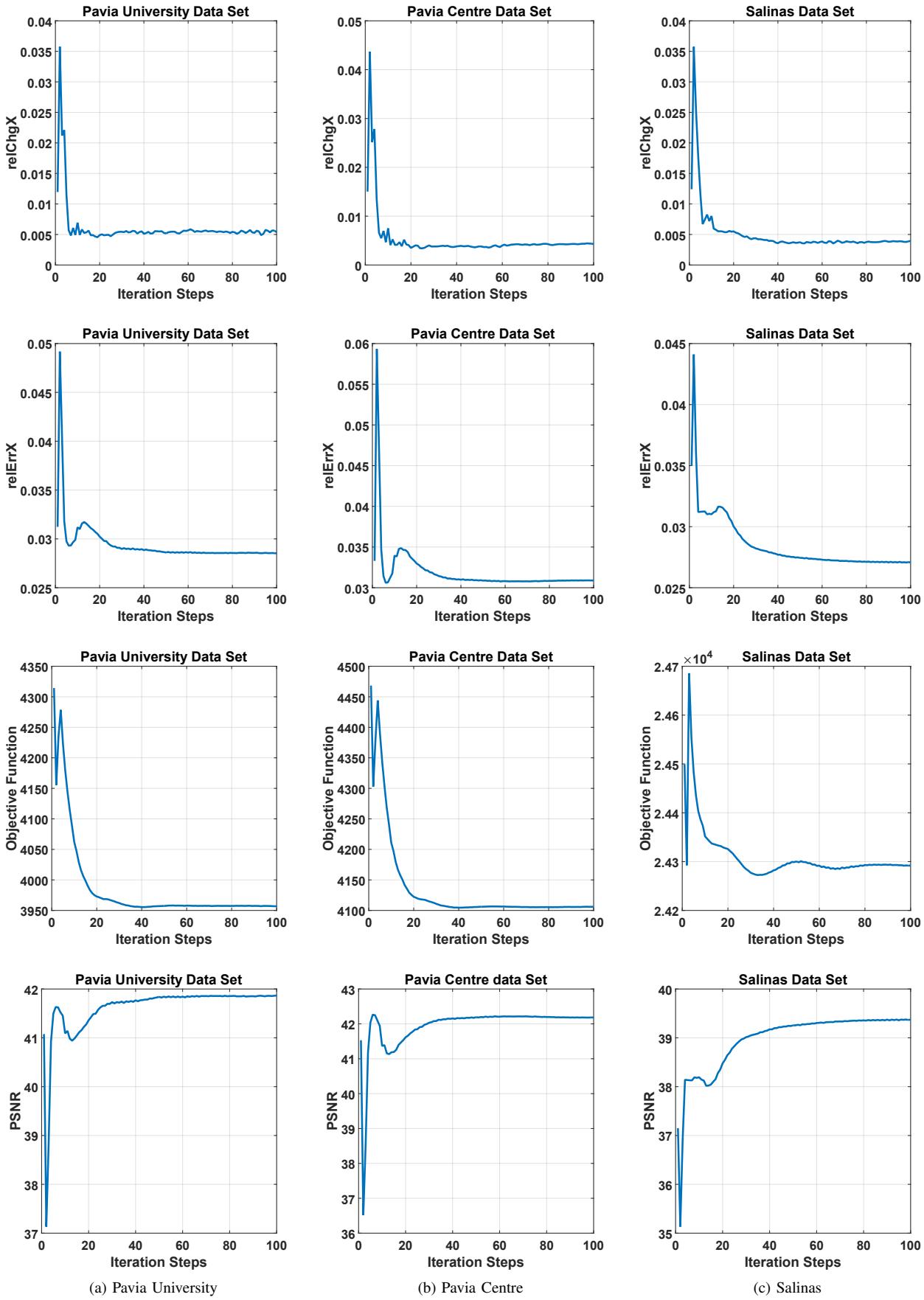


Fig. 7. The empirical analysis of algorithm convergence. The first, second, third and forth row are the relChgX , relErrX , objection function values and PSNR curves, respectively, as functions of the iteration steps in the proposed optimization algorithm in (a) Pavia University, (b) Pavia Centre and (c) Salinas data set.