

A Truncated Matrix Decomposition for Hyperspectral Image Super-Resolution

Jianjun Liu[✉], Member, IEEE, Zebin Wu[✉], Senior Member, IEEE, Liang Xiao[✉], Member, IEEE,
Jun Sun[✉], Member, IEEE, and Hong Yan[✉], Fellow, IEEE

Abstract—Hyperspectral image super-resolution addresses the problem of fusing a low-resolution hyperspectral image (LR-HSI) and a high-resolution multispectral image (HR-MSI) to produce a high-resolution hyperspectral image (HR-HSI). In this paper, we propose a novel fusion approach for hyperspectral image super-resolution by exploiting the specific properties of matrix decomposition, which consists of four main steps. First, an endmember extraction algorithm is used to extract an initial spectral matrix from LR-HSI. Then, with the initial spectral matrix, we estimate the spatial matrix, i.e., the spatial-contextual information, from the degraded observations of HR-HSI. Third, the spatial matrix is further utilized to estimate the spectral matrix from LR-HSI by solving a least squares (LS)-based problem. Finally, the target HR-HSI is constructed by combining the estimated spectral and spatial matrixes. In particular, two models are proposed to estimate the spatial matrix. One is a simple case that involves a LS-based problem, and the other is an elaborate case that consists of two fidelity terms and a spatial regularizer, where the spatial regularizer aiming to restrain the range of solutions is achieved by exploiting the superpixel-level low-rank characteristics of HR-HSI. Experiment results conducted on both synthetic and real data sets demonstrate the effectiveness of the proposed approach as compared to other hyperspectral image super-resolution methods.

Index Terms—Super-resolution, hyperspectral image, matrix decomposition, low rank, superpixel.

I. INTRODUCTION

HYPERSPECTRAL images (HSIs) contain hundreds of contiguous and narrow spectral bands, with spectral

Manuscript received September 2, 2019; revised March 15, 2020 and May 29, 2020; accepted July 12, 2020. Date of current version July 28, 2020. This work was supported in part by the National Natural Science Foundation of China under Grant 61601201, Grant 61772274, and Grant 61871226, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20160188 and Grant BK20180018, in part by the Jiangsu Provincial Social Developing Project under Grant BE2018727, in part by the Hong Kong Scholars Program under Grant XJ2018113, in part by the Hong Kong Research Grants Council under Project C1007-15G, and in part by the City University of Hong Kong under Project 9610460. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Denis Kouame. (*Corresponding author: Jianjun Liu*)

Jianjun Liu is with the Jiangsu Provincial Engineering Laboratory for Pattern Recognition and Computational Intelligence, Jiangnan University, Wuxi 214122, China, and also with the Department of Electrical Engineering, City University of Hong Kong, Hong Kong (e-mail: liuofficial@163.com).

Zebin Wu and Liang Xiao are with the School of Computer Science, Nanjing University of Science and Technology, Nanjing 210094, China (e-mail: zebin.wu@gmail.com; xiaoliang@mail.njust.edu.cn).

Jun Sun is with the Jiangsu Provincial Engineering Laboratory for Pattern Recognition and Computational Intelligence, Jiangnan University, Wuxi 214122, China (e-mail: junsun@jiangnan.edu.cn).

Hong Yan is with the Department of Electrical Engineering, City University of Hong Kong, Hong Kong (e-mail: h.yan@cityu.edu.hk).

Digital Object Identifier 10.1109/TIP.2020.3009830

coverage ranging from the visible to infrared spectrum. This high spectral resolution of HSIs has proved useful for object detection [1], tracking [2], face recognition [3], [4] and land-cover classification [5]–[8]. Due to the limit of imaging sensors, there is always a tradeoff between the spectral resolution and spatial resolution of images [9]. For HSIs, the spectral resolution is required to be high, and thus the spatial resolution is sacrificed. Conversely, with the loss of much spectral information, conventional panchromatic or multispectral images (MSIs) achieve high spatial resolution. To obtain a high-resolution HSI (HR-HSI), an economical solution is to instead record a low-resolution HSI (LR-HSI) and a conventional high-resolution image, and to fuse them into a product that should be high in both spectral and spatial resolutions.

A lot of researches have well studied the fusion of LR-HSI and panchromatic image as the well-known hypersharpening problem [9], [10]. This fusion problem can be treated as an extension of the Pansharpening that fuses the low-resolution MSI with a corresponding panchromatic image [11], [12]. Typically, the hypersharpening approaches can be divided into four categories: component substitution (CS) [13], multiresolution analysis [14], model-based approaches [15]–[18] and deep learning [19]–[22]. Among these categories, the research of model-based approaches is the most classic one, where various techniques are introduced, including regularization [15], sparse representation [16], [17], spectral and spatial priors [18] and so on.

Compared with panchromatic images, MSIs have more spectral bands and a wider spectral coverage, thus having potential advantages in maintaining the spectral and spatial structures. Literarily, the fusion of LR-HSI and high-resolution MSI (HR-MSI) is often referred to as HSI super-resolution. Generally, HSI super-resolution process is more complex than that of hypersharpening due to the spatial-spectral information contained in MSIs. Thus, the approaches proposed for hypersharpening may be unsuitable for HSI super-resolution. Except the approaches extending from the hypersharpening techniques, HSI super-resolution approaches can be roughly divided into two categories, i.e., model-based approaches [23]–[46] and learning-based approaches [47]–[51].

For the model-based approaches, they mainly consider building a regularization model to solve HSI super-resolution problem according to the mechanism of spectral and spatial degradations, and thereby focus on designing effective fidelity terms and exploiting efficient regularization terms to enforce the desired solutions. According to whether to

separate the spectral and spatial modes, these approaches can be divided into two categories further: nonfactorization-based approaches [23]–[26] and factorization-based approaches [27]–[46]. As for the learning-based approaches, they are developing with the rise of deep learning in other fields, such as image recognition [52], super-resolution [53], segmentation [54], scene classification [55] and so on. Compared with the model-based approaches, the learning-based approaches often require samples to train the deep neural networks, which limits their applications in specific scenarios.

In this paper, we propose a fusion framework for HSI super-resolution by exploiting the specific properties of matrix decomposition (MD). Different from the existing model-based HSI super-resolution approaches, this MD fusion (MDF) framework consists of four steps. In the first step, we use an endmember extraction algorithm to extract an initial spectral matrix from LR-HSI. In the second step, with the initial spectral matrix, we estimate the spatial matrix from the degraded observations of HR-HSI. In the third step, the spatial matrix is further utilized to estimate the spectral matrix from LR-HSI by solving a least squares (LS)-based problem. In the last step, the estimated spectral and spatial matrixes are combined to construct the target HR-HSI. For the second step, we propose two models to estimate the spatial matrix. The first one is a simple case that estimates the spatial matrix from HR-MSI by solving a LS-based problem, and the second one is an elaborate case that designs a region-based low-rank (RLR) regularizer to enforce the superpixel-level low-rank characteristics of HR-HSI. Compared with the existing HSI super-resolution approaches, some of the innovative characteristics of the proposed approach are highlighted as follows.

- 1) An MDF framework is proposed. The second step of MDF belongs to the factorization-based HSI super-resolution approaches, and the third step is utilized to improve the performance of the second step further. The entire process of MDF can be treated as a hybrid approach of the model-based and CS-based ones.
- 2) An MDF method using LS regularization (LS-MDF) is proposed. LS-MDF is a simple and fast HSI super-resolution method, and it can achieve competitive results.
- 3) An MDF method using RLR regularization (RLR-MDF) is proposed. In the MDF framework, the RLR regularizer captures the superpixel-level low-rank characteristics of HR-HSI, and shows its potential advantages in restraining the range of solutions. In particular, one can terminate the algorithm of RLR-MDF ahead of time.

The remainder of this paper is organized as follows. Section II reviews the related works of HSI super-resolution. Section III briefly introduces superpixel segmentation and HSI super-resolution. In Section IV, we first introduce the MDF framework and then present the proposed LS-MDF and RLR-MDF methods and the corresponding algorithms. The effectiveness of the proposed approach is demonstrated in Section V by conducting experiments on three synthetic data sets and one real data set. Finally, Section VI concludes this paper with some remarks.

II. RELATED WORKS

In this section, we review the aforementioned two categories of the HSI super-resolution approaches.

A. Learning-Based HSI Super-Resolution

The learning-based approaches often describe the fusion process by constructing a deep network, and then feed LR-HSI and HR-MSI into the network to produce the fused image. For instance, Palsson *et al.* [48] propose a three-dimensional convolutional neural network (CNN) for HSI super-resolution. In [26], Dian *et al.* consider learning the image prior via deep residual CNN in the fusion process. In [50], Yang *et al.* extract the spatial and spectral features separately, and thereby propose a deep two-branches CNN for HSI super-resolution. In [51], Xie *et al.* first propose a fusion model and solve it by the proximal gradient method, and then construct a deep network by unfolding the iterative algorithm. Generally, there is no significant difference between the learning-based Pansharpening [20]–[22] and HSI super-resolution approaches, and one can extend the learning-based Pansharpening approaches for HSI super-resolution purpose easily.

B. Model-Based HSI Super-Resolution

The nonfactorization based approaches build the models on the original images, and then exploit some spectral and spatial priors to enforce the desired solutions. For instance, Zhang *et al.* [23] propose a group spectral embedding-based fusion approach by exploring the multiple manifold structures of spectral bands and the low-rank structure of data. In [24], Zhang *et al.* exploit clustering manifold structure to regularize the result in the spatial domain. Xu *et al.* [25] formulate nonlocal similar patches as tensors and then incorporate tensor sparse representation into the model to regularize the final result. Dian *et al.* [26] exploit a low tensor-train rank prior to measure the relationship of nonlocal similar patches.

The factorization based approaches consider separating the spectral and spatial modes before building the models. References [27]–[29] assume the original HR-HSI lives in a low-dimensional subspace and thereby separate it into two parts. In [27], Simões *et al.* propose a subspace-based HSI super-resolution method called HySure, where a vector total variation (VTV) regularizer is introduced to promote piecewise-smooth solutions with discontinuities aligned across the hyperspectral bands. In [28], Wei *et al.* exploit a prior deriving from dictionary learning to enforce the solution. In [29], Dian *et al.* introduce a tensor multi-rank regularizer to measure the correlation of nonlocal similar patches. References [30]–[37] assume the original HR-HSI can be sparsely represented by an over-complete spectral dictionary, and then learn the spectral dictionary and update the coefficients accordingly. In these methods, various priors are introduced to enforce the solutions, such as nonnegativity [31]–[35], sparse [30]–[37], joint sparse [31], [35], spatial correlation [33], [36], [37], superpixel segmentation [35] and low-rank [34], [37]. References [38]–[44] assume the original HR-HSI can be

represented by some pure spectral signatures linearly, and incorporate nonnegative matrix factorization (NMF) [56] into the HSI super-resolution process. Besides the nonnegative constraint, some constraints or regularization terms are introduced in these methods, such as sum-to-one constraint [40], [41], [43], sparse constraint [39], [43], [44], simplex volume-based regularization [43] and spatial correlation regularization [42], [44]. Compared with the aforementioned subspace-based and dictionary-based methods, the two parts, i.e., endmember and abundance matrixes, of these NMF-based methods are updated iteratively. In addition, there are some works that utilize tensor decomposition [57] to capture the spectral-spatial characteristics of HSIs [45], [46]. In [45], Li *et al.* use tucker decomposition to factorize the HR-HSI into four parts and then solve the subproblems with respect to each part, just like the schema of dictionary learning. In [46], kanatsoulis *et al.* factorize the HR-HSI into three parts by using CANDECOMP/PARAFAC (CP) decomposition and then update each part iteratively.

III. PRELIMINARIES

A. Superpixel Segmentation

Superpixel segmentation partitions an image into several homogeneous regions, i.e., superpixels, whose shapes and sizes can be adaptively adjusted according to different spatial structures, thus providing a good delineation of the image boundaries and facilitating the subsequent processing. In this work, superpixel segmentation is extended to divide a MSI into several superpixels [58], [59]. Let $\mathbf{Z} \in \mathbb{R}^{N_b \times N_w N_h}$ be a HR-MSI, where N_b , N_w and N_h are the dimensions of the spectral band number, spatial width and spatial height, respectively. The goal of MSI superpixel segmentation is to partition \mathbf{Z} into K local similar regions $\{\mathbf{Z}_{(k)} \in \mathbb{R}^{N_b \times N_k} | k = 1, \dots, K\}$, where N_k is the number of pixels in k th superpixels and $N_w N_h = \sum_k N_k$. The adopted segmentation algorithm is the so-called entropy rate superpixel (ERS) [60]. Before the segmentation, an undirected graph is defined on the multispectral pixels, where vertexes are the multispectral pixels and every two spatially adjacent pixels are connected by an edge. For each edge, it is associated with a weight to measure the similarity between the two connected pixels,

$$\exp \left\{ -\frac{\omega^2(i, j) \cdot \|\mathbf{z}_i - \mathbf{z}_j\|_2^2}{2\sigma^2} \right\}, \quad (1)$$

where $\sigma > 0$ is the bandwidth, \mathbf{z}_i and \mathbf{z}_j are two pixels with spatial coordinates i and j , respectively, and $\omega(i, j)$ is their spatial Euclid distance. By using the defined graph, one can perform superpixel segmentation of MSIs via the ERS algorithm.

B. Hyperspectral Image Super-Resolution

HSI super-resolution aims to estimate a HR-HSI $\mathbf{X} \in \mathbb{R}^{N_B \times N_w N_h}$ by fusing an acquired LR-HSI $\mathbf{Y} \in \mathbb{R}^{N_B \times N_w N_h}$ and a corresponding HR-MSI $\mathbf{Z} \in \mathbb{R}^{N_b \times N_w N_h}$ of the same scene, where N_B ($N_B > N_b$) is the spectral band number, and N_w and N_h ($N_w > N_w$ and $N_h > N_h$) are the spatial width and height, respectively. The observations \mathbf{Y}

and \mathbf{Z} can be modeled as spatially degraded and spectrally degraded versions of the desired \mathbf{X} . Specifically, these two processes can be written as:

$$\mathbf{Y} = \mathbf{XBH} + \mathbf{E}_Y \quad (2)$$

$$\mathbf{Z} = \mathbf{RX} + \mathbf{E}_Z \quad (3)$$

where $\mathbf{B} \in \mathbb{R}^{N_w N_h \times N_w N_h}$ represents the spatial blurring operator, $\mathbf{H} \in \mathbb{R}^{N_w N_h \times N_w N_h}$ represents the spatial down-sampling operator, $\mathbf{R} \in \mathbb{R}^{N_b \times N_B}$ represents the spectral response of the multispectral sensor, and \mathbf{E}_Y and \mathbf{E}_Z are the residuals. Some approaches [23]–[26] consider building a regularization model consisting of two fidelity terms directly, which can be written as:

$$\min_{\mathbf{X}} \frac{1}{2} \|\mathbf{Y} - \mathbf{XBH}\|_F^2 + \frac{1}{2} \|\mathbf{Z} - \mathbf{RX}\|_F^2, \quad (4)$$

where $\|\cdot\|_F$ is the Frobenius norm. Since the problem (4) is ill-posed, different regularization terms are proposed to regulate \mathbf{X} .

Linear spectral mixing model (LSMM) [61], [62] assumes that the collected vectors of a HSI can be represented as a linear combination of some distinct spectral signatures, i.e., endmembers. Mathematically, given a HSI $\mathbf{X} \in \mathbb{R}^{N_B \times N_w N_h}$ containing J endmembers, LSMM can be written as

$$\mathbf{X} = \mathbf{AS} + \mathbf{E}, \quad (5)$$

where $\mathbf{A} \in \mathbb{R}^{N_B \times J}$ ($N_B \gg J$) represents the endmember matrix, $\mathbf{S} \in \mathbb{R}^{J \times N_w N_h}$ ($N_w N_h \gg J$) represents the abundance matrix and $\mathbf{E} \in \mathbb{R}^{N_B \times N_w N_h}$ is the residual. To make sense physically, both \mathbf{A} and \mathbf{S} are supposed to be nonnegative. Moreover, in some simplex-based methods \mathbf{S} needs to satisfy the sum-to-one constraint [63], [64]. In HSI super-resolution, some approaches [38]–[44] consider incorporating LSMM into the super-resolution process. For (4), it can be rewritten as:

$$\min_{\mathbf{A}, \mathbf{S}} \frac{1}{2} \|\mathbf{Y} - \mathbf{ASBH}\|_F^2 + \frac{1}{2} \|\mathbf{Z} - \mathbf{RAS}\|_F^2, \quad (6)$$

where the nonnegative constraints of \mathbf{A} and \mathbf{S} and the sum-to-one constraint of \mathbf{S} can be imposed on the model for specific purposes.

IV. PROPOSED APPROACH

A. MDF Framework

For LSMM, it often formulates as a NMF problem [38], [65]. Different from LSMM, i.e., NMF, we consider factoring a HSI $\mathbf{X} \in \mathbb{R}^{N_B \times N_w N_h}$ into two unconstrained matrixes, i.e., $\mathbf{X} \approx \mathbf{AS}$, and solving the following optimization problem

$$\min_{\mathbf{A}, \mathbf{S}} \|\mathbf{X} - \mathbf{AS}\|_F^2, \quad (7)$$

since we just use \mathbf{AS} to estimate \mathbf{X} and do not intend to finish the two tasks of fusion and unmixing simultaneously. Suppose \mathbf{X} is known, to obtain \mathbf{S} and \mathbf{A} one can update one matrix by fixing the other matrix, and the corresponding optimization problems can be written as

$$\mathbf{S}^{t+1} = \arg \min_{\mathbf{S}} \|\mathbf{X} - \mathbf{A}^t \mathbf{S}\|_F^2 \quad (8)$$

$$\mathbf{A}^{t+1} = \arg \min_{\mathbf{A}} \|\mathbf{X} - \mathbf{AS}^{t+1}\|_F^2 \quad (9)$$

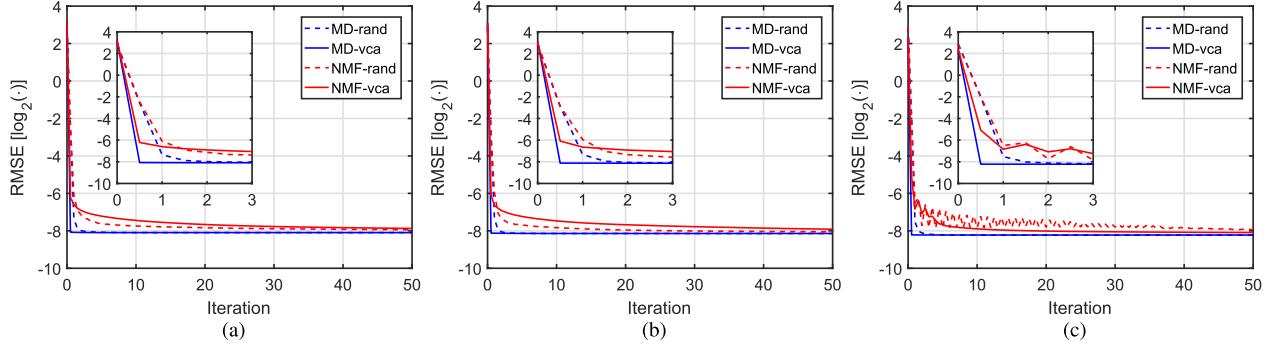


Fig. 1. RMSE as a function of the number of iterations for both MD and NMF methods using different initialization strategies, i.e., randomization and VCA. (a) University of Pavia data set. (b) Moffett field data set. (c) Washington DC Mall data set.

where $t = 0, 1, 2, \dots$. Both (8) and (9) have a close-form solution, and the iteration steps can be written as

$$\mathbf{S}^{t+1} = ((\mathbf{A}^t)^T \mathbf{A}^t)^{-1} (\mathbf{A}^t)^T \mathbf{X} \quad (10)$$

$$\mathbf{A}^{t+1} = \mathbf{X} (\mathbf{S}^{t+1})^T (\mathbf{S}^{t+1} (\mathbf{S}^{t+1})^T)^{-1} \quad (11)$$

Notably, the factorization process of (7) is exactly the CP decomposition [57] of a matrix. Thus, the above process is treated as MD. To distinguish with LSMM, in the MD process \mathbf{A} is called spectral matrix and \mathbf{S} is called spatial matrix.

To measure the performance of AS, we use the root mean squared error (RMSE), which is defined as

$$\|\mathbf{X} - \mathbf{AS}\|_F / \sqrt{N_B N_W N_H}. \quad (12)$$

Fig. 1 illustrates the RMSE results of MD and NMF as a function of the number of iterations, where two initialization strategies are used for both methods. The methods that initialize \mathbf{A} and \mathbf{S} randomly are suffixed by “-rand”, and the methods that initialize \mathbf{A} by vertex component analysis (VCA) [63] and \mathbf{S} randomly are suffixed by “-vca”. NMF is implemented by the alternating direction method of multipliers (ADMM) [64], [66]. Since both MD and NMF include two subproblems, each iteration in Fig. 1 is composed of two steps. The description of the three data sets can be seen in Section V-A. From Fig. 1, it can be seen that MD outperforms NMF. MD-vca only needs one step to obtain the optimal value, and MD-rand with one iteration can obtain a suboptimal value.

With the above observations, we design a MDF framework for HSI super-resolution by truncating the iterations.¹ The MDF can be divided into four steps: 1) initializing the spectral matrix \mathbf{A}^0 ; 2) estimating the spatial matrix \mathbf{S}^1 from the degraded observations of \mathbf{X} according to (8); 3) estimating the spectral matrix \mathbf{A}^1 from the degraded observations of \mathbf{X} according to (9); and 4) reconstructing \mathbf{X} by computing $\mathbf{A}^1 \mathbf{S}^1$. In MDF, steps 2 and 3 correspond to one iteration of MD, and the iterative update of \mathbf{S} and \mathbf{A} is not recommended, since the estimation of \mathbf{S} and \mathbf{A} is usually coarse and thus cumulative error may be introduced during the iteration process as shown in Fig. 2. Some specific properties of the proposed MDF framework are described as follows.

¹Since we truncate the iterations of MD, the related technique is called truncated matrix decomposition as shown in the title of this paper. ■

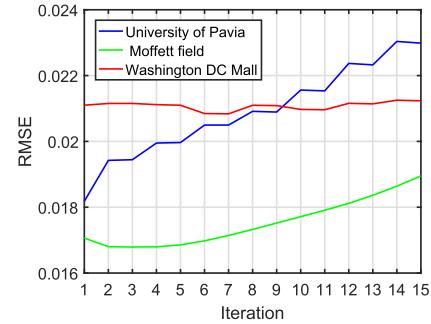


Fig. 2. RMSE as a function of the number of iterations for RLR-MDF when applied to the given synthetic data sets.

- 1) As demonstrated in Proposition 1, in MDF step-3 decreases the RMSE result of step-2 further, no matter whether \mathbf{A}^0 and \mathbf{S}^1 are accurate or not.
- 2) MDF with a coarse \mathbf{A}^0 can obtain a suboptimal value, since Fig. 1 illustrates that MD-rand with one iteration can obtain a suboptimal value.
- 3) Although in MDF the result of step-2 may be coarse, we can obtain an excellent result by step-3 if Proposition 2 is satisfied.

Proposition 1: For any spectral matrix \mathbf{A}^0 , we have

$$\|\mathbf{X} - \mathbf{A}^1 \mathbf{S}^1\|_F^2 \leq \|\mathbf{X} - \mathbf{A}^0 \mathbf{S}^1\|_F^2. \quad (13)$$

Proof: Let $t = 0$, then \mathbf{A}^1 is the solution of the optimization problem (9). Thus, $\|\mathbf{X} - \mathbf{A}^1 \mathbf{S}^1\|_F^2 \leq \|\mathbf{X} - \mathbf{A}^0 \mathbf{S}^1\|_F^2$. ■

Proposition 2: Suppose that $\bar{\mathbf{S}}^1$ is a coarse solution of (8) and $\bar{\mathbf{A}}^1$ is the corresponding solution of (9). If $\bar{\mathbf{S}}^1$ can be expressed as $\mathbf{T} \mathbf{S}^1$ with $\mathbf{T} \in \mathbb{R}^{N_B \times N_B}$ being any invertible matrix, we have

$$\|\mathbf{X} - \bar{\mathbf{A}}^1 \bar{\mathbf{S}}^1\|_F^2 = \|\mathbf{X} - \mathbf{A}^1 \mathbf{S}^1\|_F^2. \quad (14)$$

Proof: Since $\bar{\mathbf{A}}^1$ is the solution of (9) under $\bar{\mathbf{S}}^1$, we have $\bar{\mathbf{A}}^1 = \mathbf{X} (\bar{\mathbf{S}}^1)^T (\bar{\mathbf{S}}^1 (\bar{\mathbf{S}}^1)^T)^{-1}$. Then, we have

$$\begin{aligned} \bar{\mathbf{A}}^1 \bar{\mathbf{S}}^1 &= \mathbf{X} (\bar{\mathbf{S}}^1)^T (\bar{\mathbf{S}}^1 (\bar{\mathbf{S}}^1)^T)^{-1} \bar{\mathbf{S}}^1 \\ &= \mathbf{X} (\mathbf{T} \mathbf{S}^1)^T (\mathbf{T} \mathbf{S}^1 (\mathbf{T} \mathbf{S}^1)^T)^{-1} \mathbf{T} \mathbf{S}^1 \\ &= \mathbf{X} (\mathbf{S}^1)^T (\mathbf{S}^1 (\mathbf{S}^1)^T)^{-1} \mathbf{S}^1 \\ &= \mathbf{A}^1 \mathbf{S}^1 \end{aligned} \quad (15)$$

Thus, we have $\|\mathbf{X} - \bar{\mathbf{A}}^1 \bar{\mathbf{S}}^1\|_F^2 = \|\mathbf{X} - \mathbf{A}^1 \mathbf{S}^1\|_F^2$. ■

B. LS-MDF and RLR-MDF

In this subsection, we propose two fusion methods (i.e., LS-MDF and RLR-MDF) based on the MDF framework described in Section IV-A. For simplicity, the iteration index t of \mathbf{A} and \mathbf{S} is removed.

1) *Step 1 of MDF*: The first step is to initialize the spectral matrix \mathbf{A} and can be done by performing VCA on the spatially degraded observation $\mathbf{Y} \in \mathbb{R}^{N_B \times N_w N_h}$ of $\mathbf{X} \in \mathbb{R}^{N_B \times N_w N_H}$, since the spectral information is mainly contained in \mathbf{Y} . For the case when N_B is relatively small, one can perform singular value decomposition on \mathbf{Y} to initialize \mathbf{A} . In this step, one doesn't need to search an optimal spectral matrix and a suboptimal one is also satisfactory due to the specific properties of the MDF framework.

2) *Step 2 of MDF*: The second step estimates the spatial matrix \mathbf{S} . This step is equivalent to capture the spatial information of \mathbf{X} from its degraded observations. Since the spatial information is mainly contained in the spectrally degraded observation $\mathbf{Z} \in \mathbb{R}^{N_b \times N_w N_H}$, a simple and straightforward method is to solve the following LS problem according to (8)

$$\min_{\mathbf{S}} \|\mathbf{Z} - \mathbf{RAS}\|_F^2 + \lambda \|\mathbf{S}\|_F^2, \quad (16)$$

where $\lambda > 0$ is the regularization parameter. The problem (16) has a close-form solution, which can be written as

$$\mathbf{S} = (\mathbf{A}^T \mathbf{R}^T \mathbf{R} \mathbf{A} + \lambda \mathbf{I})^{-1} \mathbf{A}^T \mathbf{R}^T \mathbf{Z}, \quad (17)$$

where \mathbf{I} denotes the identity matrix. In LS-MDF, (16) is used as the second step, and the estimated spatial matrix is relatively coarse, since a simple Frobenius norm to regulate \mathbf{S} .

For RLR-MDF, we consider capturing the spatial information from both \mathbf{Z} and \mathbf{Y} and design a RLR regularizer to restrain the solution. Specifically, we solve the following optimization problem according to (8)

$$\min_{\mathbf{S}} \frac{1}{2} \|\mathbf{Y} - \mathbf{ASBH}\|_F^2 + \frac{1}{2} \|\mathbf{Z} - \mathbf{RAS}\|_F^2 + \lambda \mathcal{S}(\mathbf{S}), \quad (18)$$

where $\mathcal{S}(\cdot)$ is the proposed RLR regularizer. Since \mathbf{Z} reflects the spatial information of \mathbf{X} preferably, it makes sense to transfer the spatial relationship of \mathbf{Z} to \mathbf{S} . To capture the image boundaries and exploit the low-rank attributes of adjacent pixels, a superpixel segmentation is performed on \mathbf{Z} , as shown in Section III-A. Then, the spatial regularizer $\mathcal{S}(\cdot)$ is defined as

$$\mathcal{S}(\mathbf{S}) = \sum_{k=1}^K \|\mathbf{S}_{(k)}\|_*, \quad (19)$$

where $\|\mathbf{S}_{(k)}\|_*$ denotes the nuclear norm (sum of the singular values) of $\mathbf{S}_{(k)}$. The problem (18) uses an objective function similar to (6), but it only updates \mathbf{S} and doesn't include the nonnegative and sum-to-one constraints. Moreover, (18) is similar to HySure [27], and the difference between them is the proposed spatial regularizer. That is to say, HySure can be used as the second step of MDF. Notably, in the second step of MDF we don't expect RLR to outperform the existing techniques, such as the VTV regularizer used in HySure, and just intend to exploit its potential advantages in the entire framework as shown in Section V.

Algorithm 1 LS-MDF for HSI Super-Resolution

- 1: **Input:** LR-HSI \mathbf{Y} , HR-MSI \mathbf{Z} , \mathbf{B} , \mathbf{H} , \mathbf{R} , λ , J .
 - 2: **Step 1:** Initialize \mathbf{A} by performing VCA on \mathbf{Y} .
 - 3: **Step 2:** Estimate \mathbf{S} by (17).
 - 4: **Step 3:** Update \mathbf{A} by (21).
 - 5: **Step 4:** $\mathbf{X} \leftarrow \mathbf{AS}$.
 - 6: **Output:** HR-HSI \mathbf{X} .
-

Algorithm 2 RLR-MDF for HSI Super-Resolution

- 1: **Input:** LR-HSI \mathbf{Y} , HR-MSI \mathbf{Z} , \mathbf{B} , \mathbf{H} , \mathbf{R} , λ , J , K , σ .
 - 2: Perform superpixel segmentation on \mathbf{Z} with parameters K and σ , and obtain the labels of each pixel.
 - 3: **Step 1:** Initialize \mathbf{A} by performing VCA on \mathbf{Y} .
 - 4: **Step 2:** Estimate \mathbf{S} by solving (18).
 - 5: **Step 3:** Update \mathbf{A} by (21).
 - 6: **Step 4:** $\mathbf{X} \leftarrow \mathbf{AS}$.
 - 7: **Output:** HR-HSI \mathbf{X} .
-

3) *Step 3 of MDF*: The third step updates the spectral matrix \mathbf{A} . Since the spectral information is mainly contained in \mathbf{Y} , we consider solving the following LS problem according to (9)

$$\min_{\mathbf{A}} \|\mathbf{Y} - \mathbf{AS}\|_F^2 + \lambda \|\mathbf{A}\|_F^2, \quad (20)$$

where $\bar{\mathbf{S}} = \mathbf{SBH}$ can be computed in advance. Specifically, one can obtain $\bar{\mathbf{S}}$ by blurring and downsampling the spatial matrix \mathbf{S} directly or by taking advantage of the particular properties of \mathbf{B} and \mathbf{H} in the frequency domain [67]. The close-form solution of (20) can be written as

$$\mathbf{A} = \mathbf{YS}^T (\bar{\mathbf{S}}\bar{\mathbf{S}}^T + \lambda \mathbf{I})^{-1}. \quad (21)$$

The problem (9) involves an over-determined system. For (20), its fidelity term also involves an over-determined system since we assume $N_w N_h \gg J$. Thus, one don't need to design a complex regularizer, and a simple Frobenius norm can be used to regulate \mathbf{A} and make the result robust. That is to say, using (21) to update the spectral matrix \mathbf{A} is feasible.

4) *Step 4 of MDF*: The last step are straightforward. After introducing the four steps of the MDF framework, the proposed LS-MDF and RLR-MDF methods can be implemented by performing the four steps successively. The complete algorithms of LS-MDF and RLR-MDF for HSI super-resolution are summarized in Algorithms 1 and 2, respectively.

Notably, the MDF framework is somewhat similar to the CS-based fusion approaches. The first step can be thought of as finding a suitable transformation, the second step can be thought of as a detail injection process, the third step is utilized to adjust the transformation adaptively and the last step is a reconstruction process.

C. Optimization Algorithm for (18)

For the proposed RLR-MDF method, there is an optimization problem (18), and ADMM can be used to solved it. By introducing three variables $\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3 \in \mathbb{R}^{J \times N_w N_H}$,

the optimization problem (18) can be rewritten as follows:

$$\begin{aligned} & \min_{\mathbf{S}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3} f(\mathbf{S}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3) \\ & \text{s.t. } \mathbf{V}_1 = \mathbf{S}\mathbf{B}, \mathbf{V}_2 = \mathbf{S}, \mathbf{V}_3 = \mathbf{S} \end{aligned} \quad (22)$$

where

$$\begin{aligned} f(\mathbf{S}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3) &= \frac{1}{2} \|\mathbf{Y} - \mathbf{A}\mathbf{V}_1\mathbf{H}\|_F^2 \\ &\quad + \frac{1}{2} \|\mathbf{Z} - \mathbf{R}\mathbf{A}\mathbf{V}_2\|_F^2 + \lambda\mathcal{S}(\mathbf{V}_3). \end{aligned} \quad (23)$$

The augmented Lagrangian function of (22) can be written as follows:

$$\begin{aligned} \mathcal{L}(\mathbf{S}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3, \mathbf{D}_1, \mathbf{D}_2, \mathbf{D}_3) \\ = f(\mathbf{S}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3) + \frac{\mu}{2} \|\mathbf{S}\mathbf{B} - \mathbf{V}_1 - \mathbf{D}_1\|_F^2 \\ + \frac{\mu}{2} \|\mathbf{S} - \mathbf{V}_2 - \mathbf{D}_2\|_F^2 + \frac{\mu}{2} \|\mathbf{S} - \mathbf{V}_3 - \mathbf{D}_3\|_F^2, \end{aligned} \quad (24)$$

where $\mu > 0$ is the penalty parameter and $\mathbf{D}_1, \mathbf{D}_2, \mathbf{D}_3 \in \mathbb{R}^{J \times N_W N_H}$ are auxiliary variables. The corresponding iteration procedure can be written as follows:

1) Solving \mathbf{S} Subproblem: Optimizing \mathcal{L} with respect to \mathbf{S} can be written as

$$\begin{aligned} \mathbf{S}^{(t+1)} &= \arg \min_{\mathbf{S}} \frac{\mu}{2} \|\mathbf{S}\mathbf{B} - \mathbf{V}_1^{(t)} - \mathbf{D}_1^{(t)}\|_F^2 \\ &\quad + \frac{\mu}{2} \|\mathbf{S} - \mathbf{V}_2^{(t)} - \mathbf{D}_2^{(t)}\|_F^2 + \frac{\mu}{2} \|\mathbf{S} - \mathbf{V}_3^{(t)} - \mathbf{D}_3^{(t)}\|_F^2 \end{aligned} \quad (25)$$

which has the solution

$$\begin{aligned} \mathbf{S}^{(t+1)} &= (\mathbf{B}\mathbf{B}^T + 2\mathbf{I})^{-1} ((\mathbf{V}_1^{(t)} + \mathbf{D}_1^{(t)})\mathbf{B}^T \\ &\quad + \mathbf{V}_2^{(t)} + \mathbf{D}_2^{(t)} + \mathbf{V}_3^{(t)} + \mathbf{D}_3^{(t)}). \end{aligned} \quad (26)$$

The computation can be efficiently performed by using the fast Fourier transform (FFT), and the first term $(\mathbf{B}\mathbf{B}^T + 2\mathbf{I})^{-1}$ can be computed in advance. The complexity of this subproblem under the assumption that $N_W > N_H$ is $\mathcal{O}(J N_W N_H \log N_W)$.

2) Solving \mathbf{V}_1 Subproblem: Optimizing \mathcal{L} with respect to \mathbf{V}_1 can be written as

$$\begin{aligned} \mathbf{V}_1^{(t+1)} &= \arg \min_{\mathbf{V}_1} \frac{1}{2} \|\mathbf{Y} - \mathbf{A}\mathbf{V}_1\mathbf{H}\|_F^2 \\ &\quad + \frac{\mu}{2} \|\mathbf{S}^{(t+1)}\mathbf{B} - \mathbf{V}_1 - \mathbf{D}_1^{(t)}\|_F^2. \end{aligned} \quad (27)$$

Similar to [27], we take advantage of the masking matrix \mathbf{H} to separate \mathbf{V}_1 into $\mathbf{V}_1\mathbf{H}$ and $\mathbf{V}_1\bar{\mathbf{H}}$ with $\bar{\mathbf{H}}$ being the complement matrix of \mathbf{H} . The solution can be written as

$$\begin{aligned} \mathbf{V}_1^{(t+1)}\mathbf{H} &= (\mathbf{A}^T\mathbf{A} + \mu\mathbf{I})^{-1} (\mathbf{A}^T\mathbf{Y} \\ &\quad + \mu(\mathbf{S}^{(t+1)}\mathbf{B} - \mathbf{D}_1^{(t)}))\mathbf{H} \end{aligned} \quad (28)$$

$$\mathbf{V}_1^{(t+1)}\bar{\mathbf{H}} = (\mathbf{S}^{(t+1)}\mathbf{B} - \mathbf{D}_1^{(t)})\bar{\mathbf{H}} \quad (29)$$

The computation can be efficiently performed via the FFT, and the terms $(\mathbf{A}^T\mathbf{A} + \mu\mathbf{I})^{-1}$ and $\mathbf{A}^T\mathbf{Y}$ can be precomputed. The complexity of this subproblem is $\mathcal{O}(J N_W N_H \log N_W + J^2 N_W N_H)$.

3) Solving \mathbf{V}_2 Subproblem: Optimizing \mathcal{L} with respect to \mathbf{V}_2 can be written as

$$\mathbf{V}_2^{(t+1)} = \arg \min_{\mathbf{V}_2} \frac{1}{2} \|\mathbf{Z} - \mathbf{R}\mathbf{A}\mathbf{V}_2\|_F^2 + \frac{\mu}{2} \|\mathbf{S}^{(t+1)} - \mathbf{V}_2 - \mathbf{D}_2^{(t)}\|_F^2, \quad (30)$$

Algorithm 3 Solving (18)

-
- 1: **Input:** LR-HSI \mathbf{Y} , HR-MSI \mathbf{Z} , \mathbf{B} , \mathbf{H} , \mathbf{R} , λ , J , K , spectral matrix \mathbf{A} and superpixel-level labels of \mathbf{Z} 's pixels.
 - 2: Initialize $\mathbf{S}^{(0)}$, $\mathbf{V}_1^{(0)}$, $\mathbf{V}_2^{(0)}$, $\mathbf{V}_3^{(0)}$, $\mathbf{D}_1^{(0)}$, $\mathbf{D}_2^{(0)}$, $\mathbf{D}_3^{(0)}$ by zero matrix, $\mu = 0.05$, $maxIter = 200$, $\epsilon = 10^{-3}$.
 - 3: **for** $t = 1 : maxIter$ **do**
 - 4: Update $\mathbf{S}^{(t)}$ by (26)
 - 5: Update $\mathbf{V}_1^{(t)}$ by (28) and (29)
 - 6: Update $\mathbf{V}_2^{(t)}$ by (31)
 - 7: Update $\mathbf{V}_3^{(t)}$ by (34)
 - 8: Update $\mathbf{D}_1^{(t)}$, $\mathbf{D}_2^{(t)}$ and $\mathbf{D}_3^{(t)}$ by (35), (36) and (37)
 - 9: **if** $\|\mathbf{AS}^{(t-1)} - \mathbf{AS}^{(t)}\|_F / \|\mathbf{AS}^{(t-1)}\|_F < \epsilon$ **then**
 - 10: **break**
 - 11: **end if**
 - 12: **end for**
 - 13: **Output:** The spectral matrix \mathbf{S} .
-

which has the solution

$$\mathbf{V}_2^{(t+1)} = (\mathbf{A}^T\mathbf{R}^T\mathbf{RA} + \mu\mathbf{I})^{-1} (\mathbf{A}^T\mathbf{R}^T\mathbf{Z} + \mu(\mathbf{S}^{(t+1)} - \mathbf{D}_2^{(t)})). \quad (31)$$

Only the term $(\mathbf{S}^{(t+1)} - \mathbf{D}_2^{(t)})$ cannot be precomputed, and the complexity of this subproblem is $\mathcal{O}(J^2 N_W N_H)$.

4) Solving \mathbf{V}_3 Subproblem: Optimizing \mathcal{L} with respect to \mathbf{V}_3 can be written as

$$\mathbf{V}_3^{(t+1)} = \arg \min_{\mathbf{V}_3} \lambda \sum_{k=1}^K \|\mathbf{V}_{3(k)}\|_* + \frac{\mu}{2} \|\mathbf{S}^{(t+1)} - \mathbf{V}_3 - \mathbf{D}_3^{(t)}\|_F^2. \quad (32)$$

Since each superpixel $\mathbf{V}_{3(k)}$ is nonoverlapping, we can solve it independently. The above problem of each $\mathbf{V}_{3(k)}$ can be written as follows:

$$\min_{\mathbf{V}_{3(k)}} \lambda \|\mathbf{V}_{3(k)}\|_* + \frac{\mu}{2} \|\mathbf{S}_{(k)}^{(t+1)} - \mathbf{V}_{3(k)} - \mathbf{D}_{3(k)}^{(t)}\|_F^2, \quad (33)$$

which has the solution

$$\mathbf{V}_{3(k)}^{(t+1)} = \mathcal{D}_{\lambda/\mu}(\mathbf{S}_{(k)}^{(t+1)} - \mathbf{D}_{3(k)}^{(t)}), \quad (34)$$

where $\mathcal{D}_{\lambda/\mu}(\cdot)$ is the singular value threshold function [68]. The complexity of this subproblem under the assumption that $J < N_k$ is $\mathcal{O}(J^2 N_W N_H)$.

5) Updating Multipliers: The multipliers associated with \mathcal{L} is updated as follows:

$$\mathbf{D}_1^{(t+1)} = \mathbf{D}_1^{(t)} - (\mathbf{S}^{(t+1)}\mathbf{B} - \mathbf{V}_1^{(t+1)}) \quad (35)$$

$$\mathbf{D}_2^{(t+1)} = \mathbf{D}_2^{(t)} - (\mathbf{S}^{(t+1)} - \mathbf{V}_2^{(t+1)}) \quad (36)$$

$$\mathbf{D}_3^{(t+1)} = \mathbf{D}_3^{(t)} - (\mathbf{S}^{(t+1)} - \mathbf{V}_3^{(t+1)}) \quad (37)$$

Ideally, the solution \mathbf{S} of (18) should satisfy Proposition 2. However, it is difficult to demonstrate this in practice. Experiment shows that the above procedure can obtain a satisfactory result for MDF. The complete algorithm of (18) is summarized in Algorithm 3. Generally, most algorithms use the difference

of results between the current two iterations to check the convergence condition, i.e.,

$$\|\mathbf{S}^{(t-1)} - \mathbf{S}^{(t)}\|_F / \|\mathbf{S}^{(t-1)}\|_F. \quad (38)$$

Different from them, in this work we use the following formula to terminate the algorithm ahead of time

$$\|\mathbf{AS}^{(t-1)} - \mathbf{AS}^{(t)}\|_F / \|\mathbf{AS}^{(t-1)}\|_F, \quad (39)$$

since our goal is to estimate \mathbf{X} by \mathbf{AS} . For the proposed RLR-MDF method, the most time-consuming part is solving (18), and the overall complexity of solving (18) is $\mathcal{O}(JN_W N_H \log N_W + J^2 N_W N_H)$ per iteration.

V. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we conduct experiments on both synthetic and real data sets to demonstrate the performance of the proposed approach. Before the following experiments, all data sets are scaled to the range [0, 1]. For the synthetic data sets, the quality of the fused images are assessed with the RMSE, the peak signal-noise-ratio (PSNR), the spectral angle mapper (SAM), the relative dimensionless global error in synthesis (ERGAS), and the universal image quality index (UIQI) [9], [69]. Notably, for the extended ERS segmentation, the input images are scaled to the range [0, 255]. All experiments are carried out using a 64-b quad-core CPU 3.40-GHz processor with 12-GB memory.

A. Synthetic Data Sets

Given a real-life HSI as reference image, we generate the two observation images, i.e., LR-HSI and HR-MSI, according to the Wald's protocol [70]. Details of the three reference images are presented as follows.

- 1) University of Pavia data set. This data set is an urban image acquired by the Reflective Optics System Imaging Spectrometer (ROSIS). The ROSIS sensor is characterized by 115 spectral bands covering the spectral range 0.43–0.86 μm . This image, with size of 610 × 340 pixels, has spatial resolution of 1.3 m per pixel. There are 103 spectral bands remained after removal of noisy bands. We select the up-left 580 × 340-pixel-size image as the reference image.
- 2) Moffett field data set. This data set is a mixed urban rural image acquired by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). The AVIRIS sensor is characterized by 224 spectral bands covering the spectral range 0.4–2.5 μm . The size of this image is 395 × 185 with a spatial resolution of 20 m. The number of spectral bands is reduced to 176 by removing water absorption bands. We select the up-left 380 × 180-pixel-size image as the reference image.
- 3) Washington DC Mall data set. This data set is the Hyperspectral digital imagery collection experiment (HYDICE) image taken over the Washington DC Mall. The HYDICE sensor is characterized by 210 spectral bands covering the spectral range 0.4–2.4 μm . This image contains 1208 × 307 pixels with a spatial resolution of about 2.8 m. Bands in the region where

the atmosphere is opaque have been omitted, leaving 191 bands. We select a 580 × 300-pixel-size image as the reference image.

To create the LR-HSI, we spatially blur the corresponding reference image and then downsample the blurred image by a factor of 5 in each direction. For the blurring process, a Gaussian blurring of size 9 × 9 with standard deviation 2.12 is applied to each band of the reference image. To create the HR-MSI, the spectral response of the IKONOS satellite is used, and thus a 4-band image is obtained by averaging the bands of the reference image according to the spectral response profiles of the RGB and NIR bands. Unless otherwise specified, Gaussian noise is added to the LR-HSI (SNR = 30 dB) and to the HR-MSI (SNR = 40 dB).

B. Influence of Parameters

There are four free parameters in the proposed RLR-MDF method: 1) the regularization parameter λ , 2) the rank of the spectral and spatial matrixes J , 3) the number of superpixels K , and 4) the bandwidth of superpixel segmentation σ . In this subsection, we investigate them and show how they impact quality measures of the proposed method. Unless otherwise specified, in the experiments, the spatial blur \mathbf{B} and the spectral responses of the sensor \mathbf{R} are estimated by the method proposed in [27]. Fig. 3 illustrates the RMSE results of RLR-MDF as a function of λ and J . It can be seen that, for all data sets, there is a wide optimal range for the choice of λ and J . A small λ can obtain moderate results and $\lambda = 10^{-2}$ is the best. J should not be too large, and a moderate value is better. Thus, λ and J are eventually set as 10^{-2} and 30 respectively for all data sets. Then we evaluate the RMSE performance of RLR-MDF over K and σ as shown in Fig. 4. It can be seen that the values of RMSE are less affected by the choice of K and σ . The choice of K depends on the image size and the number of homogeneous regions. K is eventually set as 200 for University of Pavia data set and 400 for the two remaining data sets. For the choice of σ , there is a wide optimal range, and we can get suboptimal results when σ is chosen in the range [15, 25].

C. Influence of Spectral Matrix Update

In this section, we investigate the third step of MDF and show how it improves relative to the second step of MDF. Specifically, we show the improvement of RLR-MDF over RLR-MDF without spectral matrix update (denoted by RLR). Fig. 5 illustrates the RMSE results of RLR and RLR-MDF as a function of the number of iterations. It can be seen that, for all data sets, RLR-MDF outperforms RLR consistently, and the gains are significant. RLR-MDF can achieve a suboptimal result in the first few iterations. For RLR and RLR-MDF, their convergence properties depending on RMSE are similar, and thus the suggested stopping condition (39) is reasonable.

D. Experiment Results on Synthetic Data Sets

In order to evaluate the performance of the proposed approach, ten methods are compared in this subsection. The first method, used as a baseline one, is to solve the problem (4)

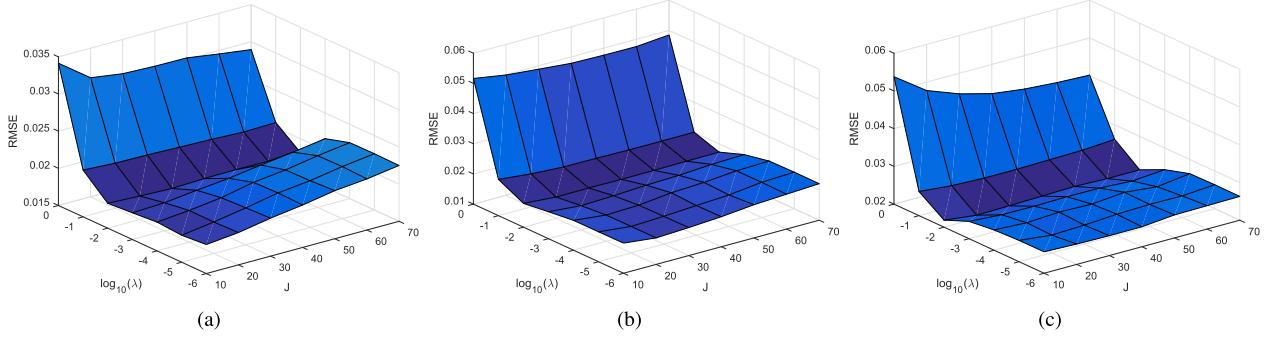


Fig. 3. The RMSE curve as a function of J and λ . (a) University of Pavia data set. (b) Moffett field data set. (c) Washington DC Mall data set.

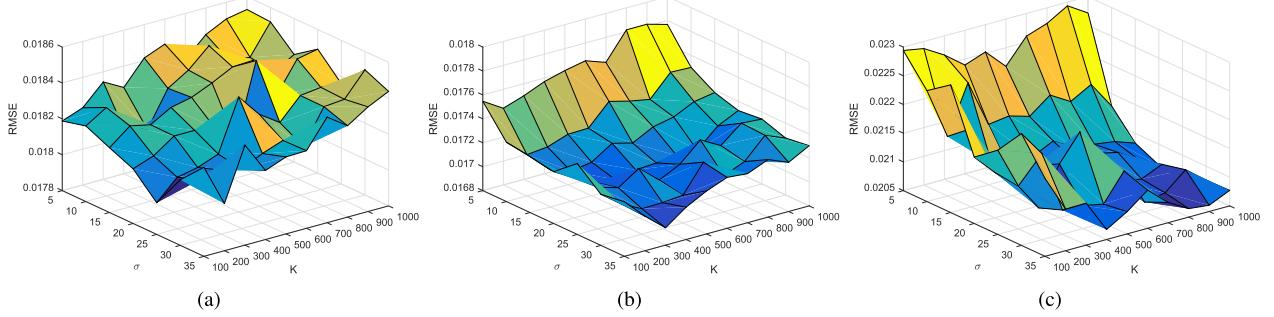


Fig. 4. The RMSE curve as a function of K and σ . (a) University of Pavia data set. (b) Moffett field data set. (c) Washington DC Mall data set.

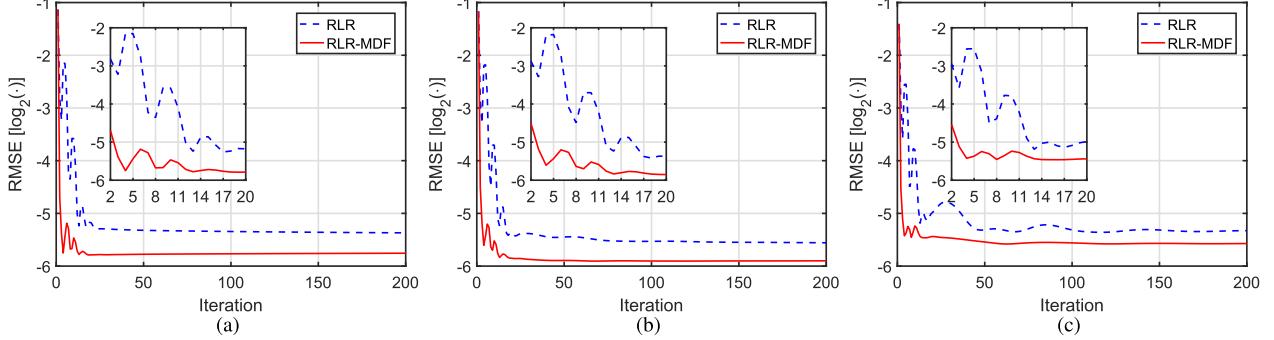


Fig. 5. RMSE as a function of the number of iterations for RLR and RLR-MDF. (a) University of Pavia data set. (b) Moffett field data set. (c) Washington DC Mall data set.

with Gaussian prior (denoted by SYLV), which can be done by solving a Sylvester equation [71]. The second method is one of the NMF-based super-resolution approaches, called coupled spectral unmixing (CSU) [40]. The next four methods are HySure [27], NSSR [33], CSTF [45] and NPTSR [25]. For the aforementioned six methods, their free parameters are tuned to be optimal with our data sets. The last four methods are those related to this work: 1) RLR, 2) the proposed LS-MDF, 3) the MDF using HySure as the second step (denoted by VTV-MDF), and 4) the proposed RLR-MDF method. For RLR, its free parameters are as same as those of RLR-MDF. For LS-MDF, λ and J are set as 10^{-2} and 30. As for VTV-MDF, its free parameters are fixed to be the same as those used in HySure. Two sets of experiments are performed in this subsection. The first one estimates \mathbf{B} and \mathbf{R} by using the method proposed in [27], and the second one uses the exact \mathbf{B} and \mathbf{R} . When using the estimated spatial response, the spatial blur kernel is first factorized by rank-1 MD and then fed into CSTF, since CSTF assumes the two modes of spatial response are separable.

Table I shows the five quantitative results of the compared methods for University of Pavia data set. The running times in seconds are also included for reference. The best values of this table and the following are marked in bold for clarity. As seen from Table I, all methods outperform the baseline method SLYV, LS-MDF obtains competitive results when compared with the other methods, and the proposed RLR-MDF method gives the best quantitative results. When using the exact responses, the quantitative results of all methods are improved. In all cases, RLR-MDF performs better than RLR. For the estimated responses, the quantitative results of VTV-MDF are better than those of HySure, while for the exact responses, there are no significant gaps between them. The performance of RLR is similar to that of HySure, while under the MDF framework, RLR-MDF outperforms VTV-MDF consistently. For computation cost, LS-MDF is the fastest method, and the proposed RLR-MDF method is only inferior to LS-MDF and SLYV.

Fig. 6 illustrates the fusion results of the compared methods using the estimated responses in form of RGB images,

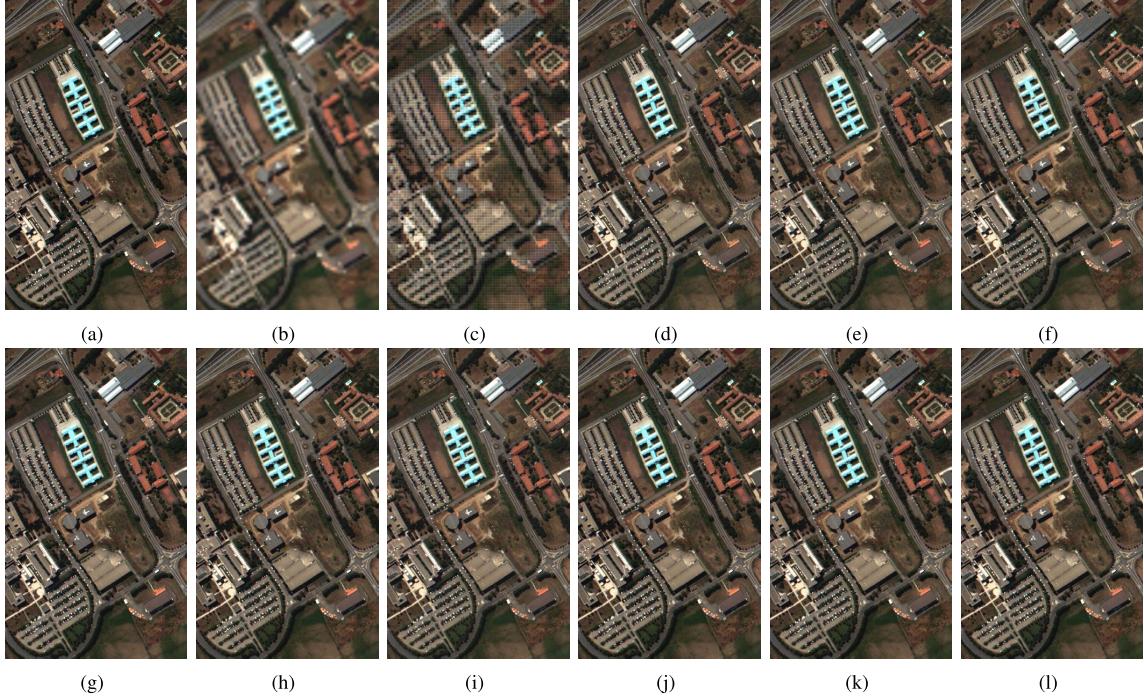


Fig. 6. RGB images of HSI super-resolution results using estimated spatial and spectral responses when applied to University of Pavia data set. (a) Reference image. (b) LR-HSI. (c) SLYV. (d) CSU. (e) HySure. (f) NSSR. (g) CSTF. (h) NPTSR. (i) RLR. (j) LS-MDF. (k) VTV-MDF. (l) RLR-MDF.

TABLE I

QUALITY MEASURES FOR UNIVERSITY OF PAVIA DATA SET
USING DIFFERENT SUPER-RESOLUTION METHODS

Method	RMSE	PSNR	SAM	ERGAS	UIQI	Time
Best Values	0	$+\infty$	0	0	1	0
Estimated \mathbf{B} and \mathbf{R}						
SLYV	0.1163	19.995	8.293	6.301	0.708	6
CSU	0.0253	32.528	2.988	1.554	0.976	187
HySure	0.0240	32.911	2.910	1.510	0.977	664
NSSR	0.0278	31.817	3.268	1.697	0.971	441
CSTF	0.0314	30.896	2.948	1.801	0.977	509
NPTSR	0.0202	35.198	2.790	1.292	0.982	989
RLR	0.0241	32.904	2.948	1.529	0.977	324
LS-MDF	0.0202	36.052	2.783	1.251	0.983	3
VTV-MDF	0.0194	35.497	2.638	1.240	0.982	727
RLR-MDF	0.0183	36.392	2.485	1.176	0.983	67
Exact \mathbf{B} and \mathbf{R}						
SLYV	0.0412	28.982	5.579	2.396	0.944	5
CSU	0.0189	36.142	2.556	1.233	0.982	234
HySure	0.0178	36.985	2.483	1.155	0.984	660
NSSR	0.0203	35.367	2.823	1.309	0.982	415
CSTF	0.0190	36.500	2.659	1.213	0.982	507
NPTSR	0.0176	37.405	2.476	1.135	0.984	1007
RLR	0.0180	37.019	2.527	1.152	0.984	289
LS-MDF	0.0199	36.556	2.755	1.229	0.983	3
VTV-MDF	0.0181	36.646	2.549	1.174	0.983	728
RLR-MDF	0.0174	37.415	2.416	1.126	0.984	66

TABLE II

QUALITY MEASURES FOR MOFFETT FIELD DATA SET
USING DIFFERENT SUPER-RESOLUTION METHODS

Method	RMSE	PSNR	SAM	ERGAS	UIQI	Time
Best Values	0	$+\infty$	0	0	1	0
Estimated \mathbf{B} and \mathbf{R}						
SLYV	0.0956	21.133	10.602	6.360	0.741	3
CSU	0.0280	31.819	5.223	2.000	0.968	79
HySure	0.0211	34.077	4.333	1.424	0.976	223
NSSR	0.0238	33.220	4.428	1.704	0.968	115
CSTF	0.0242	32.964	4.663	1.716	0.971	167
NPTSR	0.0233	33.447	4.528	1.682	0.976	625
RLR	0.0213	33.932	4.562	1.414	0.975	150
LS-MDF	0.0285	32.985	5.805	2.128	0.965	2
VTV-MDF	0.0175	36.566	3.473	1.267	0.982	237
RLR-MDF	0.0170	37.049	3.358	1.233	0.981	35
Exact \mathbf{B} and \mathbf{R}						
SLYV	0.0621	26.399	10.858	4.361	0.868	3
CSU	0.0216	34.688	3.814	1.589	0.978	97
HySure	0.0173	36.939	3.435	1.256	0.982	226
NSSR	0.0198	35.232	3.963	1.429	0.975	111
CSTF	0.0205	35.227	4.292	1.502	0.974	169
NPTSR	0.0206	35.589	3.937	1.521	0.979	624
RLR	0.0170	36.867	3.561	1.222	0.982	139
LS-MDF	0.0286	32.993	5.768	2.132	0.965	2
VTV-MDF	0.0174	36.809	3.466	1.262	0.982	240
RLR-MDF	0.0170	37.199	3.342	1.233	0.982	34

according to the spectral response of the IKONOS satellite. The reference image, LR-HSI and HR-MSI are also presented. Visually, it can be observed that none of these methods exhibit an obvious spectral distortion. SLYV has severe spatial distortion, and the results produced by the other methods are visually very similar. Fig. 9 (a) shows the PSNR as a function of spectral band for the aforementioned methods that use the estimated responses. It can be seen that the proposed RLR-MDF method performs best in most bands. LS-MDF achieves

highest PSNR results in many bands, but it performs worse in some bands. Fig. 10 (a) shows the SAM between the reference image and the results of the aforementioned methods for each pixel, with the pixels sorted in order of ascending error. As illustrated in this figure, RLR-MDF outperforms the others at the pixel level.

Table II reports the RMSE, PSNR, SAM, ERGAS, UIQI and running time results of the compared methods for Moffett field data set. From this table, we can see that using the

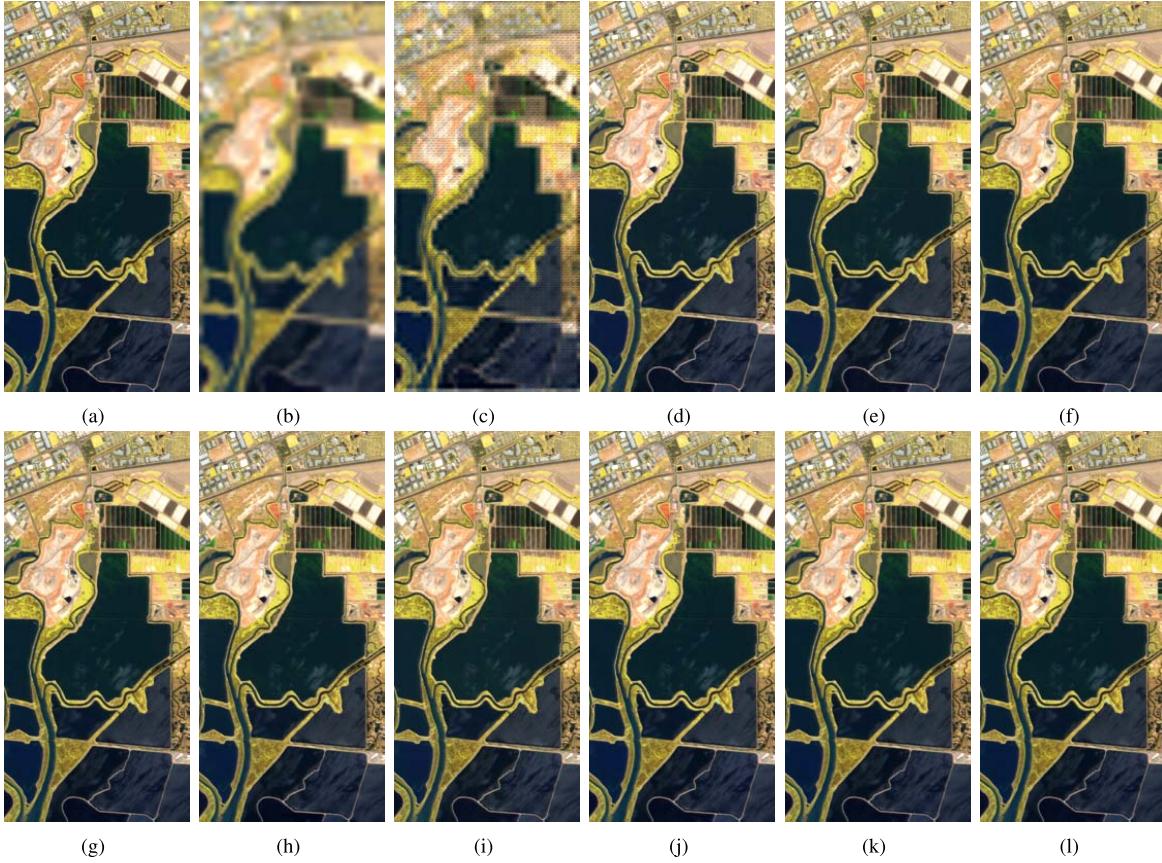


Fig. 7. RGB images of HSI super-resolution results using estimated spatial and spectral responses when applied to Moffett field data set. (a) Reference image. (b) LR-HSI. (c) SLYV. (d) CSU. (e) HySure. (f) NSSR. (g) CSTF. (h) NPTSR. (i) RLR. (j) LS-MDF. (k) VTV-MDF. (l) RLR-MDF.

exact responses are better for the final results, the baseline method SLYV performs the worst, and the proposed RLR-MDF method gives most of the best quantitative results. For the MDF framework, it can improve the performance of RLR in all cases and that of HySure using the estimated responses. As for the simplest method LS-MDF, it gives moderate results and outperforms SLYV in all quality measures and running time. In Fig. 7, we show the original images and the fusion results of the compared methods using the estimated responses in form of RGB images. It can be seen that there is no spectral distortion for all methods, and SLYV is the only one that exhibits an obvious spatial distortion. For these methods using the estimated responses, Fig. 9 (b) illustrates the PSNR as a function of spectral band. It is clear that RLR-MDF performs best followed by VTV-MDF. Moreover, Fig. 10 (b) illustrates the SAM between the reference image and their results for each pixel. It can be observed that RLR-MDF is the best.

Table III summarizes the five quality measures and running time results of the compared methods for Washington DC Mall data set. From this table, we can see that when using the estimated responses, RLR-MDF gives most of the best quantitative results and RLR-MDF and VTV-MDF outperform RLR and HySure respectively, for most quality measures. When using the exact responses, RLR-MDF outperforms RLR for most quantitative results and VTV-MDF and HySure show comparable results. As for LS-MDF, it is the fastest method and gives competitive results. Fig. 8 illustrates the RGB images of LR-HSI, HR-MSI, the reference HR-HSI and the

TABLE III
QUALITY MEASURES FOR WASHINGTON DC MALL DATA SET
USING DIFFERENT SUPER-RESOLUTION METHODS

Method	RMSE	PSNR	SAM	ERGAS	UIQI	Time
Best Values	0	$+\infty$	0	0	1	0
Estimated B and R						
SLYV	0.0919	19.064	8.383	58.030	0.498	11
CSU	0.0273	32.692	3.652	20.886	0.928	112
HySure	0.0257	33.495	3.398	19.473	0.934	566
NSSR	0.0266	31.998	3.570	19.443	0.918	345
CSTF	0.0302	30.245	3.640	21.968	0.884	522
NPTSR	0.0220	32.656	3.160	20.439	0.908	1766
RLR	0.0246	33.864	3.313	18.061	0.941	379
LS-MDF	0.0361	33.201	5.059	20.858	0.959	4
VTV-MDF	0.0224	34.411	3.075	19.647	0.929	652
RLR-MDF	0.0211	35.470	2.851	18.388	0.948	146
Exact B and R						
SLYV	0.0453	24.266	6.761	44.711	0.682	9
CSU	0.0177	35.741	2.296	20.971	0.951	252
HySure	0.0218	35.348	2.955	19.453	0.939	556
NSSR	0.0227	33.464	3.190	19.739	0.928	340
CSTF	0.0224	35.286	3.118	20.619	0.962	520
NPTSR	0.0195	34.901	2.517	19.952	0.946	1736
RLR	0.0203	35.761	2.850	17.719	0.945	444
LS-MDF	0.0360	33.272	5.010	20.852	0.959	4
VTV-MDF	0.0217	34.930	2.917	19.640	0.930	612
RLR-MDF	0.0201	36.132	2.704	17.779	0.952	222

fusion results of the compared methods using the estimated responses. Through visual inspection, we can see that no method exhibits an obvious spectral distortion, and all methods provide good spatial quality, except SLYV. Figs. 9 (c) and 10

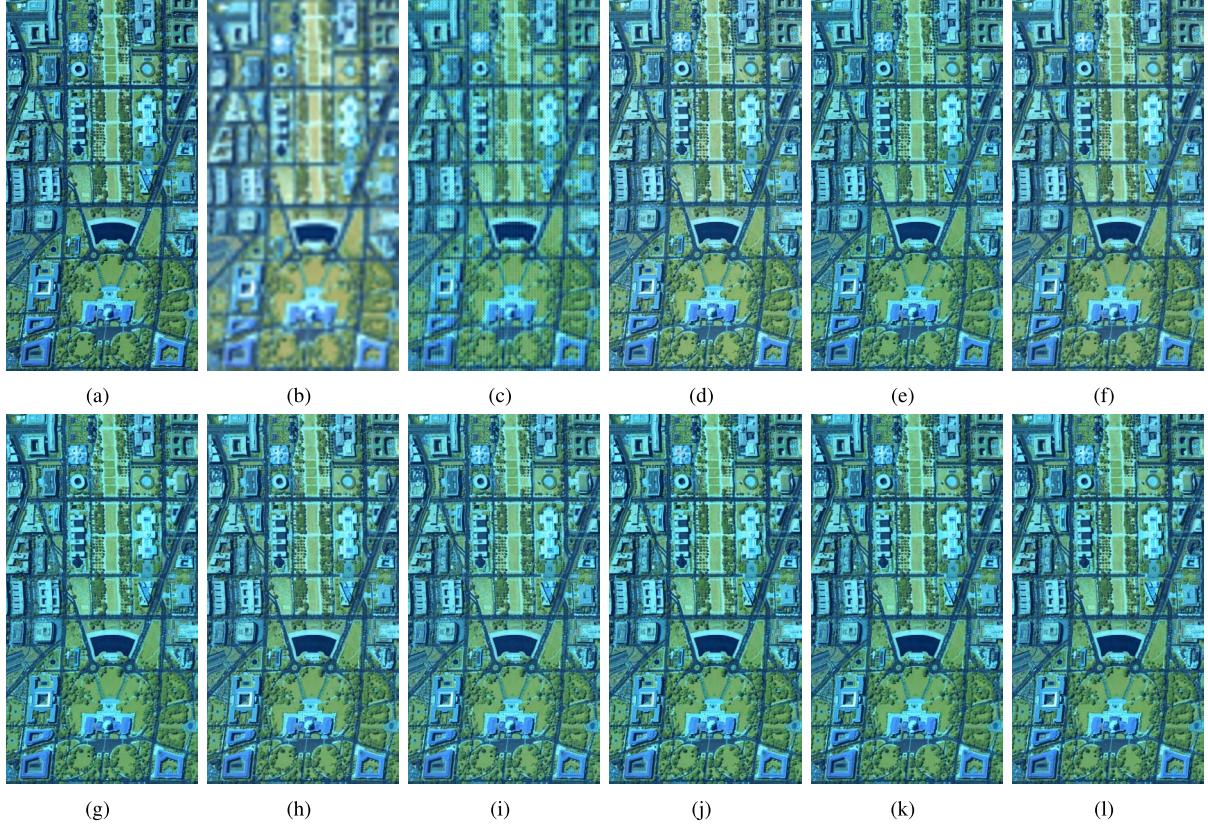


Fig. 8. RGB images of HSI super-resolution results using estimated spatial and spectral responses when applied to Washington DC Mall data set. (a) Reference image. (b) LR-HSI. (c) CSU. (d) HySure. (e) NSSR. (f) CSTF. (g) NPTSR. (h) RLR. (i) LS-MDF. (j) VTV-MDF. (k) RLR-MDF.

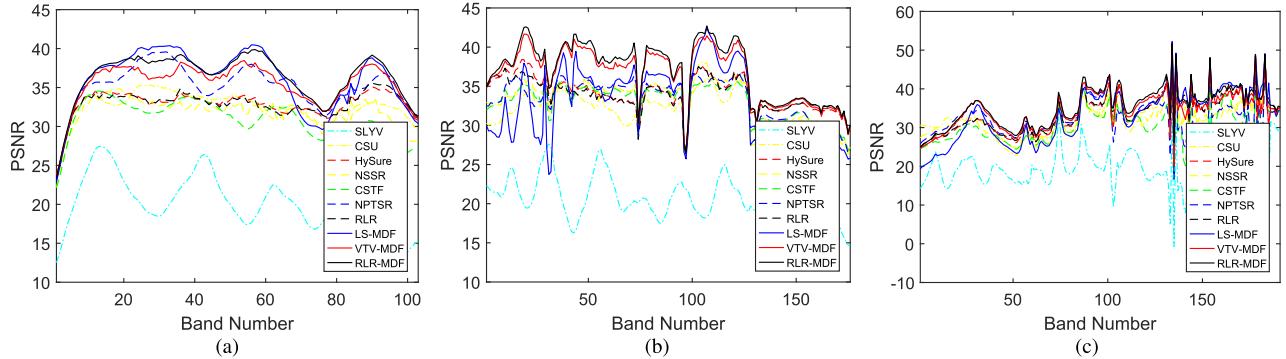


Fig. 9. PSNR as a function of spectral band. (a) University of Pavia data set. (b) Moffett field data set. (c) Washington DC Mall data set.

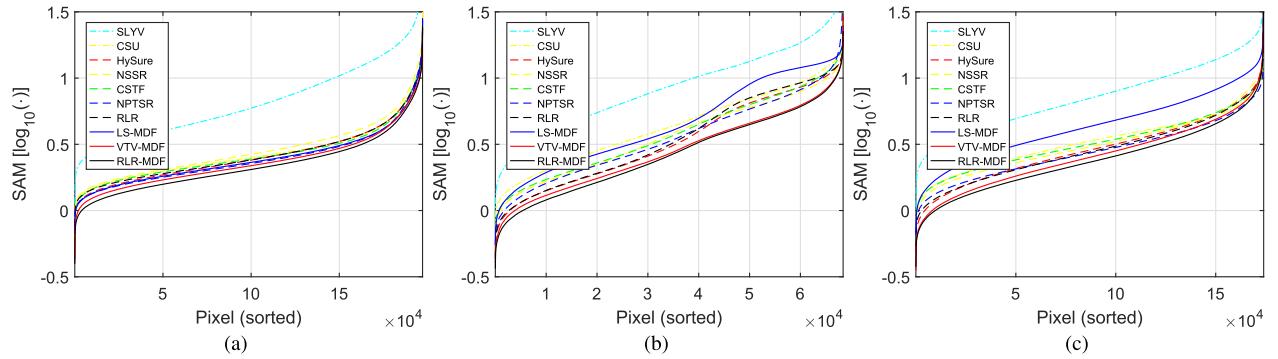


Fig. 10. SAM as a function of sorted pixel. (a) University of Pavia data set. (b) Moffett field data set. (c) Washington DC Mall data set.

(c) illustrate the PSNR and SAM as functions of spectral band and sorted pixel. It can be seen that RLR-MDF performs best in terms of band-level PSNR and pixel-level SAM.

E. Influence of Noise

This set of experiments is used to show how the noise level affects the performance of the compared methods mentioned

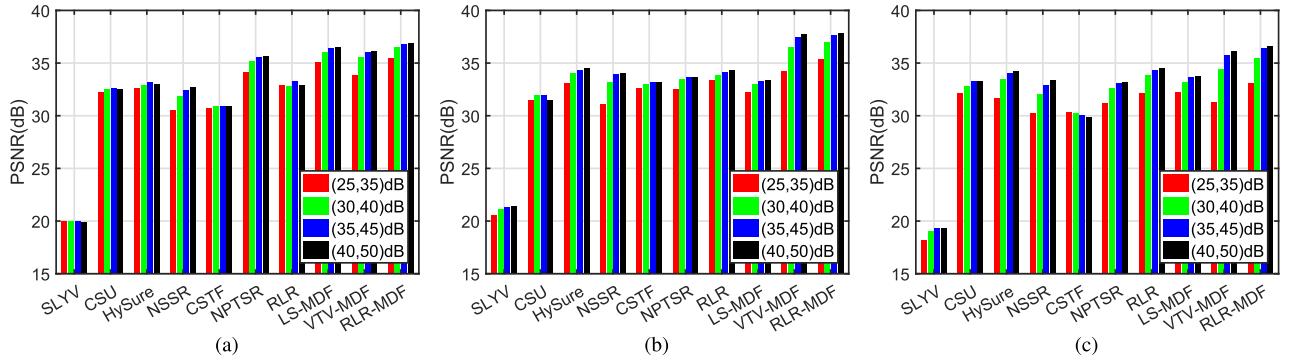


Fig. 11. PSNRs of the compared methods under noise of different powers (in the parentheses). (a) University of Pavia data set. (b) Moffett field data set. (c) Washington DC Mall data set.

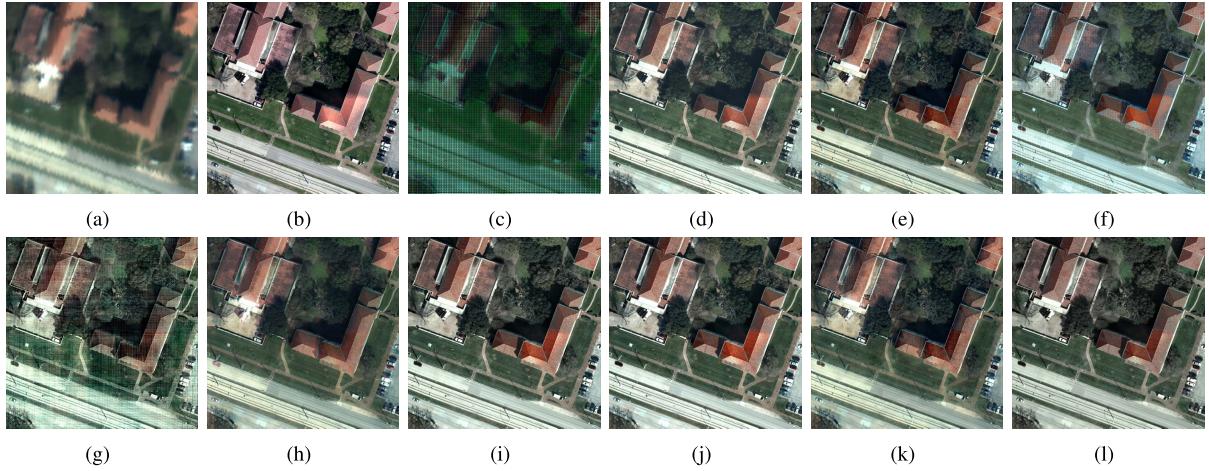


Fig. 12. RGB images of HSI super-resolution results when applied to real data set. (a) LR-HSI. (b) HR-RGB image. (c) SYLV. (d) CSU. (e) HySure. (f) NSSR. (g) CSTF. (h) NPTSR. (i) RLR. (j) LS-MDF. (k) VTV-MDF. (l) RLR-MDF.

in Section V-D. As with the previous synthetic data sets, Gaussian noise is added to the generated LR-HSI and HR-MSI. Four different levels of noise are investigated. Specifically, for LR-HSI the noise power is varied to be 25 dB, 30 dB, 35 dB and 40 dB, and for HR-MSI the noise power is varied to be 35 dB, 40 dB, 45 dB and 50 dB. Fig. 11 shows the PSNR results of the compared methods under noise of different powers. It is obvious that, in most cases, the performance of the compared methods increases as the noise level decreases. The experiment results of different noisy cases further confirm the reports in Section V-D.

F. Experiment Results on Real Data

In this section, we evaluate the performance the proposed approach in practical application by using a real data set. This data set consists of a LR-HSI collected by ITRES CASI-1500 sensor and a high-resolution RGB (HR-RGB) image collected by DiMAC ULTRALIGHT+ sensor. Both images are taken from the multi-modal optical remote sensing data sets released by the 2018 IEEE GRSS Data Fusion Contest [72]. The original data is acquired by the National Center for Airborne Laser Mapping (NCALM) at the University of Houston (UH), covering the UH campus and its surrounding urban areas. For the original LR-HSI, it contains 4172×1202 pixels at a 1-m ground sampling distance and 48 spectral bands

covering the spectral range $0.38\text{--}1.05 \mu\text{m}$, and we select an area of $116 \times 116 \times 48$ as our test data. For the original HR-RGB image, it contains 83440×24040 pixels, and we select and downsample the corresponding area to be a $580 \times 580 \times 3$ -size image. So the downsampling ratio of the real data set is 5. As before, the spatial blur \mathbf{B} and the spectral responses of the sensor \mathbf{R} are estimated by the method proposed in [27], and for CSTF the spatial blur kernel is further factorized by rank-1 MD. Fig. 12 illustrates the RGB images of the real data set and the fusion results of the compared methods mentioned in Section V-D. It can be seen that the proposed LS-MDF and RLR-MDF methods perform well for real data super-resolution.

VI. CONCLUSION

This paper has proposed an MDF approach for HSI super-resolution. By exploiting three specific properties of MD, an MDF framework is proposed. Different from NMF and dictionary learning, the MDF framework does not require the nonnegativity of the result matrixes and includes only one iteration. Specifically, the MDF framework consists of four steps. The first step is to initialize the spectral matrix, which is done by performing VCA on LR-HSI. The second step is to capture the spatial information of HR-HSI from the degraded observations. Two models are proposed in this step. One is

a simple and fast case that estimates the spatial matrix from HR-MSI by solving a LS-based problem, and the other is a well-designed case that captures the superpixel-level low-rank characteristics of HR-HSI by introducing a RLR regularizer. The first model is convex and smooth, and thus has an analytic solution. The second model is solved by ADMM, where an optimized stopping condition is utilized to terminate the algorithm ahead of time. The third step is to update the spectral matrix, which is done on LR-HSI by solving a LS-based problem. The last step is to construct the desired HR-HSI by combining the estimated spectral and spatial matrixes. According to the two models of the second step, two methods are generated, i.e., LS-MDF and RLR-MDF. The proposed approach has been experimentally tested using three synthetic data sets and one real data set, and the experiment results demonstrate the effectiveness of the proposed MDF framework and its two instantiations, i.e., LS-MDF and RLR-MDF.

ACKNOWLEDGMENT

The authors would like to thank the authors of [25], [27], [33], [40], [45], [71] for providing their codes. They would like to thank NCALM and the Hyperspectral Image Analysis Laboratory at UH for providing the UH data sets, and the Image Analysis and Data Fusion Technical Committee of the IEEE GRSS for supporting the annual Data Fusion Contest. They would like to thank the anonymous reviewers for their constructive comments on this paper.

REFERENCES

- [1] D. W. J. Stein, S. G. Beaven, L. E. Hoff, E. M. Winter, A. P. Schaum, and A. D. Stocker, "Anomaly detection from hyperspectral imagery," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 58–69, Jan. 2002.
- [2] H. Van Nguyen, A. Banerjee, and R. Chellappa, "Tracking via object reflectance using a hyperspectral video camera," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 44–51.
- [3] Z. Pan, G. Healey, M. Prasad, and B. Tromberg, "Face recognition in hyperspectral images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 12, pp. 1552–1560, Dec. 2003.
- [4] M. Uzair, A. Mahmood, and A. Mian, "Hyperspectral face recognition with spatiotemporal information fusion and PLS regression," *IEEE Trans. Image Process.*, vol. 24, no. 3, pp. 1127–1137, Mar. 2015.
- [5] G. Camps-Valls, D. Tuia, L. Bruzzone, and J. A. Benediktsson, "Advances in hyperspectral image classification: Earth monitoring with statistical learning methods," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 45–54, Jan. 2014.
- [6] J. Liu, Z. Wu, J. Li, A. Plaza, and Y. Yuan, "Probabilistic-kernel collaborative representation for spatial-spectral hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 2371–2384, Apr. 2016.
- [7] J. Xie, N. He, L. Fang, and A. Plaza, "Scale-free convolutional neural network for remote sensing scene classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6916–6928, Sep. 2019.
- [8] J. Liu, Z. Wu, L. Xiao, and H. Yan, "Learning multiple parameters for kernel collaborative representation classification," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jan. 17, 2020, doi: 10.1109/TNNLS.2019.2962878.
- [9] L. Loncan *et al.*, "Hyperspectral pansharpening: A review," *IEEE Trans. Geosci. Remote Sens.*, vol. 3, no. 3, pp. 27–46, Sep. 2015.
- [10] M. Selva, B. Aiazzi, F. Butera, L. Chiarantini, and S. Baronti, "Hyper-sharpening: A first approach on SIM-GA data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 3008–3024, Jun. 2015.
- [11] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. M. Bruce, "Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3012–3021, Oct. 2007.
- [12] G. Vivone *et al.*, "A critical comparison among pansharpening algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2565–2586, Oct. 2015.
- [13] T.-M. Tu, S.-C. Su, H.-C. Shyu, and P. S. Huang, "A new look at IHS-like image fusion methods," *Inf. Fusion*, vol. 2, no. 3, pp. 177–186, Sep. 2001.
- [14] F. Nencini, A. Garzelli, S. Baronti, and L. Alparone, "Remote sensing image fusion using the curvelet transform," *Inf. Fusion*, vol. 8, no. 2, pp. 143–156, Apr. 2007.
- [15] C. Ballester, V. Caselles, L. Igual, J. Verdera, and B. Roug  , "A variational model for P+XS image fusion," *Int. J. Comput. Vis.*, vol. 69, no. 1, pp. 43–58, Aug. 2006.
- [16] S. Li, H. Yin, and L. Fang, "Remote sensing image fusion via sparse representations over learned dictionaries," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 9, pp. 4779–4789, Sep. 2013.
- [17] X. X. Zhu and R. Bamler, "A sparse image fusion algorithm with application to pan-sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2827–2836, May 2013.
- [18] X. He, L. Condat, J. M. Bioucas-Dias, J. Chanussot, and J. Xia, "A new pansharpening method based on spatial and spectral sparsity priors," *IEEE Trans. Image Process.*, vol. 23, no. 9, pp. 4160–4174, Sep. 2014.
- [19] X. X. Zhu *et al.*, "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 8–36, Dec. 2017.
- [20] J. Yang, X. Fu, Y. Hu, Y. Huang, X. Ding, and J. Paisley, "PanNet: A deep network architecture for pan-sharpening," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 5449–5457.
- [21] Q. Yuan, Y. Wei, X. Meng, H. Shen, and L. Zhang, "A multiscale and multidepth convolutional neural network for remote sensing imagery pan-sharpening," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 978–989, Mar. 2018.
- [22] W. Xie, J. Lei, Y. Cui, Y. Li, and Q. Du, "Hyperspectral pansharpening with deep priors," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 5, pp. 1529–1543, May 2020.
- [23] K. Zhang, M. Wang, and S. Yang, "Multispectral and hyperspectral image fusion based on group spectral embedding and low-rank factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 3, pp. 1363–1371, Mar. 2017.
- [24] 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.
- [25] 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.
- [26] 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.
- [27] 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.
- [28] Q. Wei, J. Bioucas-Dias, N. Dobigeon, and J.-Y. Tourneret, "Hyperspectral and multispectral image fusion based on a sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3658–3668, Jul. 2015.
- [29] 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.
- [30] R. Kawakami, Y. Matsushita, J. Wright, M. Ben-Ezra, Y.-W. Tai, and K. Ikeuchi, "High-resolution hyperspectral imaging via matrix factorization," in *Proc. CVPR*, Jun. 2011, pp. 2329–2336.
- [31] R. Dian, S. Li, and L. Fang, "Non-local sparse representation for hyperspectral image super-resolution," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 63–78.
- [32] H. Song, B. Huang, K. Zhang, and H. Zhang, "Spatio-spectral fusion of satellite images based on dictionary-pair learning," *Inf. Fusion*, vol. 18, pp. 148–160, Jul. 2014.
- [33] W. Dong *et al.*, "Hyperspectral image super-resolution via non-negative structured sparse representation," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 2337–2352, May 2016.
- [34] M. A. Veganzones, M. Simoes, G. Licciardi, N. Yokoya, J. M. Bioucas-Dias, and J. Chanussot, "Hyperspectral super-resolution of locally low rank images from complementary multisource data," *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 274–288, Jan. 2016.

- [35] L. Fang, H. Zhuo, and S. Li, "Super-resolution of hyperspectral image via superpixel-based sparse representation," *Neurocomputing*, vol. 273, pp. 171–177, Jan. 2018.
- [36] X.-H. Han, B. Shi, and Y. Zheng, "Self-similarity constrained sparse representation for hyperspectral image super-resolution," *IEEE Trans. Image Process.*, vol. 27, no. 11, pp. 5625–5637, Nov. 2018.
- [37] C. Yi, Y.-Q. Zhao, and J. C.-W. Chan, "Hyperspectral image super-resolution based on spatial and spectral correlation fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 7, pp. 4165–4177, Jul. 2018.
- [38] 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.
- [39] 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.
- [40] C. Lanaras, E. Baltsavias, and K. Schindler, "Hyperspectral super-resolution by coupled spectral unmixing," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 3586–3594.
- [41] Q. Wei, J. Bioucas-Dias, N. Dobigeon, J.-Y. Tourneret, M. Chen, and S. Godsill, "Multiband image fusion based on spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7236–7249, Dec. 2016.
- [42] Y. Zhou, L. Feng, C. Hou, and S.-Y. Kung, "Hyperspectral and multispectral image fusion based on local low rank and coupled spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 10, pp. 5997–6009, Oct. 2017.
- [43] C.-H. Lin, F. Ma, C.-Y. Chi, and C.-H. Hsieh, "A convex optimization-based coupled nonnegative matrix factorization algorithm for hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1652–1667, Mar. 2018.
- [44] 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.
- [45] 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.
- [46] 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.
- [47] Y. Li, J. Hu, X. Zhao, W. Xie, and J. Li, "Hyperspectral image super-resolution using deep convolutional neural network," *Neurocomputing*, vol. 266, pp. 29–41, Nov. 2017.
- [48] F. Palsson, J. R. Sveinsson, and M. O. Ulfarsson, "Multispectral and hyperspectral image fusion using a 3-D-Convolutional neural network," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 5, pp. 639–643, May 2017.
- [49] 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.
- [50] J. Yang, Y.-Q. Zhao, and J. Chan, "Hyperspectral and multispectral image fusion via deep two-branches convolutional neural network," *Remote Sens.*, vol. 10, no. 5, p. 800, May 2018.
- [51] Q. Xie, M. Zhou, Q. Zhao, D. Meng, W. Zuo, and Z. Xu, "Multispectral and hyperspectral image fusion by MS/HS fusion net," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 1585–1594.
- [52] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [53] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2016.
- [54] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440.
- [55] N. He, L. Fang, S. Li, J. Plaza, and A. Plaza, "Skip-connected covariance network for remote sensing scene classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 5, pp. 1461–1474, May 2020.
- [56] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, p. 788, 1999.
- [57] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, Aug. 2009.
- [58] J. Liu, Z. Xiao, and L. Xiao, "Superpixel-guided multiscale kernel collaborative representation for hyperspectral image classification," *Remote Sens. Lett.*, vol. 7, no. 10, pp. 975–984, Oct. 2016.
- [59] J. Liu, Z. Wu, Z. Xiao, and J. Yang, "Region-based relaxed multiple kernel collaborative representation for hyperspectral image classification," *IEEE Access*, vol. 5, pp. 20921–20933, 2017.
- [60] M.-Y. Liu, O. Tuzel, S. Ramalingam, and R. Chellappa, "Entropy-rate clustering: Cluster analysis via maximizing a submodular function subject to a matroid constraint," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 1, pp. 99–112, Jan. 2014.
- [61] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.
- [62] J. M. Bioucas-Dias *et al.*, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.
- [63] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.
- [64] J. Liu, Z. Wu, Z. Wei, L. Xiao, and L. Sun, "A novel sparsity constrained nonnegative matrix factorization for hyperspectral unmixing," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2012, pp. 1389–1392.
- [65] Y. Qian, S. Jia, J. Zhou, and A. Robles-Kelly, "Hyperspectral unmixing via $L_{1/2}$ sparsity-constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4282–4297, Nov. 2011.
- [66] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, Mar. 2011.
- [67] N. Zhao, Q. Wei, A. Basarab, N. Dobigeon, D. Kouame, and J. Tourneret, "Fast single image super-resolution using a new analytical solution for ℓ_2 - ℓ_2 problems," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3683–3697, May 2016.
- [68] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, Jan. 2010.
- [69] N. Yokoya, C. Grohnfeldt, and J. Chanussot, "Hyperspectral and multispectral data fusion: A comparative review of the recent literature," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 2, pp. 29–56, Jun. 2017.
- [70] T. Ranchin and L. Wald, "Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation," *Photogramm. Eng. Remote Sens.*, vol. 66, no. 1, pp. 49–61, Jan. 2000.
- [71] 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.
- [72] Y. Xu *et al.*, "Advanced multi-sensor optical remote sensing for urban land use and land cover classification: Outcome of the 2018 IEEE GRSS data fusion contest," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 6, pp. 1709–1724, Jun. 2019.



Jianjun Liu (Member, IEEE) received the B.S. degree in applied mathematics and the Ph.D. degree in pattern recognition and intelligence system from the Nanjing University of Science and Technology, China, in 2009 and 2014, respectively.

He is currently an Associate Professor with the School of Artificial Intelligence and Computer Science, Jiangnan University, China. Since 2018, he has been a Postdoctoral Researcher with the Department of Electrical Engineering, City University of Hong Kong, China. His research interests are in the areas of hyperspectral image classification, super-resolution, spectral unmixing, sparse representation, computer vision, and pattern recognition.



Zebin Wu (Senior Member, IEEE) received the B.S. and Ph.D. degrees in computer science from the Nanjing University of Science and Technology, China, in 2003 and 2008, respectively.

He is currently a Professor with the School of Computer Science, Nanjing University of Science and Technology. His research interests include hyperspectral image processing, high-performance computing, and computer simulation.



Liang Xiao (Member, IEEE) received the B.S. degree in applied mathematics and the Ph.D. degree in computer science from the Nanjing University of Science and Technology, China, in 1999 and 2004, respectively.

From 2009 to 2010, he was a Postdoctoral Fellow with the Rensselaer Polytechnic Institute, Troy, NY, USA. He is currently a Professor with the School of Computer Science, Nanjing University of Science and Technology. His main research areas include inverse problems in image processing, scientific computing, data mining, and pattern recognition.



Jun Sun (Member, IEEE) received the Ph.D. degree in control theory and engineering and the M.Sc. degree in computer science and technology from Jiangnan University, China, in 2009 and 2003, respectively.

He is currently a Full Professor with the School of Artificial Intelligence and Computer Science, Jiangnan University, China. He is also the Vice Director of the Jiangsu Provincial Engineering Laboratory of Pattern Recognition and Computational Intelligence, Jiangsu. His major research areas are related to computational intelligence, machine learning, bioinformatics, and among others. He published more than 150 papers in journals, conference proceedings, and several books in the above areas.



Hong Yan (Fellow, IEEE) received the B.S. degree from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 1982, the M.S. degree from the University of Michigan, Ann Arbor, MI, USA, in 1984, and the Ph.D. degree from Yale University, New Haven, CT, USA, in 1989, all in electrical engineering.

From 1986 to 1989, he was a Research Scientist with General Network Corporation, New Haven, where he researched on design and optimization of computer and telecommunications networks.

He joined The University of Sydney, Sydney, NSW, Australia, in 1989, and became a Professor of imaging science in 1997. He is currently a Professor of electrical engineering with the City University of Hong Kong, Hong Kong. His research interests include image processing, pattern recognition, and bioinformatics. He has more than 600 publications in these areas. He was elected an IAPR Fellow for contributions to document image analysis and the IEEE Fellow for contributions to image recognition techniques and applications. He received the 2016 Norbert Wiener Award from the IEEE SMC Society for contributions to image and biomolecular pattern recognition techniques.