

# DDLPS: Detail-Based Deep Laplacian Pansharpening for Hyperspectral Imagery

Kaiyan Li<sup>1</sup>, Student Member, IEEE, Weiying Xie<sup>1</sup>, Member, IEEE, Qian Du<sup>2</sup>, Fellow, IEEE, and Yunsong Li

**Abstract**—In this paper, we propose a new pansharpening method called detail-based deep Laplacian pansharpening (DDLPS) to improve the spatial resolution of hyperspectral imagery. This method includes three main components: upsampling, detail injection, and optimization. In particular, a deep Laplacian pyramid super-resolution network (LapSRN) improves the resolution of each band. Then, a guided image filter and a gain matrix are used to combine the spatial and spectral details with an optimization problem, which is formed to adaptively select an injection coefficient. The DDLPS method is compared with 11 state-of-the-art or traditional pansharpening approaches. The experimental results demonstrate the superiority of the DDLPS method in terms of both quantitative indices and visual appearance. In addition, the training of LapSRN is based on the data sets of traditional RGB images, which overcomes the practical difficulty of insufficient training samples for pansharpening.

**Index Terms**—Guided image filter, hyperspectral (HS) imaging, Laplacian pyramid super-resolution network (LapSRN), pansharpening, super-resolution, Sylvester equation.

## I. INTRODUCTION

**H**YPERSPECTRAL (HS) imaging can simultaneously acquire images of the same scene across hundreds of wavelengths [1]. As a result, the HS imagery has a very high spectral resolution and has been widely used in many domains, such as military surveillance [2], environmental monitoring [3], mineral exploration [4], [5], and agriculture [6], [7]. Commercial products, such as Google Earth, use pansharpened images, making pansharpening an important preliminary step for several image analysis tasks, e.g., change detection [8]. In the design of optical remote sensing systems, because of the limited amount of incident

energy, there are critical tradeoffs among the spatial resolution, spectral resolution, and signal-to-noise ratio (SNR). Thus, optical systems can provide data with a high spatial resolution (HR) but a small number of spectral bands or with a high spectral resolution but a low spatial resolution (LR) [9]. Therefore, to acquire an HR-HS image, an HR panchromatic (PAN) image is often fused with an LR-HS image of the same scene. Compared with the recent approach of fusing an LR-HS image with a high spatial multispectral (LR-MS) image, the reconstruction of an HR-HS image via a PAN image is more difficult because of the low spectral resolution in a single-band PAN image.

Various data fusion techniques have been developed in the last decade to enhance the spatial resolution of HS imagery [10]. The representative traditional techniques can be roughly categorized as follows: 1) component substitution (CS) [11]–[13]; 2) multiresolution analysis (MRA) [14], [15]; 3) matrix factorization [16]–[18]; and 4) Bayesian methods [19]–[21]. CS methods include the following algorithms: intensity–hue–saturation (IHS) [11], [22], [23], Brovey transform (BT) [24], principal component analysis (PCA) [25], Gram–Schmidt (GS) [12], adaptive GS (GSA) [13], and partial replacement adaptive CS (PRACS) [26]. These CS-based methods employ the PAN image to replace the spatial component by employing the inverse transformation, which works well from a spatial aspect and is notably robust to coregistration errors but may result in spectral distortion. The MRA approach is based on the injection of spatial details that are obtained via a multiscale decomposition of the PAN image into the LR-HS image. The typical MRA algorithms include the decimated wavelet transform (DWT) [27], Laplacian pyramid [28], modulation transfer function (MTF) generalized Laplacian pyramid (MTF-GLP) [15], and MTF-GLP with high-pass modulation (MTF-GLP-HPM) [29]. The MRA approach mainly suffers from spatial distortions (e.g., ringing artifacts) but well preserves the spectral information but well preserves the spectral information by injecting details into the HS images that are obtained from the PAN image. The Bayesian approach relies on the use of a posterior distribution of the desired HR-HS image for the given LR-HS and PAN images [9]. The Bayesian sparsity promoted Gaussian prior (Bayesian sparse) [19], Bayesian naive Gaussian prior (Bayesian naive) proposed by Wei *et al.* [20], and Bayesian HySure proposed by Simões *et al.* [21] belong to this approach. Pansharpening based on matrix factorization is effective. For example,

Manuscript received March 7, 2019; accepted May 10, 2019. Date of publication June 12, 2019; date of current version September 25, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 61801359, Grant 61571345, Grant 91538101, Grant 61501346, Grant 61502367, and Grant 61701360, in part by the 111 Project under Grant B08038, in part by the Fundamental Research Funds for the Central Universities under Grant JB180104, in part by the Natural Science Basic Research Plan in Shaanxi Province of China under Grant 2016JQ6023 and Grant 2016JQ6018, in part by the General Financial Grant from the China Postdoctoral Science Foundation under Grant 2017M620440, and in part by the Yangtze Rive Scholar Bonus Schemes under Grant CJT160102. (Corresponding authors: Weiying Xie; Yunsong Li.)

K. Li, W. Xie, and Y. Li are with the State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China (e-mail: ky-li@outlook.com; wyxie@xidian.edu.cn; ysl@mail.xidian.edu.cn).

Q. Du is with the Department of Electronic and Computer Engineering, Mississippi State University, Starkville, MS 39759 USA (e-mail: du@ece.msstate.edu).

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

Digital Object Identifier 10.1109/TGRS.2019.2917759

the coupled nonnegative matrix factorization (CNMF) method proposed by Yokoya *et al.* [18], which alternately unmixes both sources of data to obtain the endmember spectra and HR abundance maps, can produce high quality in both spatial and spectral aspects. These methods separate the connection between the spatial and the spectral information and may cause spatial or spectral distortion. Recently, convolutional neural network (CNN) has been successfully applied to computer vision tasks [30]–[32], and CNN-based approaches have become the new trend in pansharpening with excellent fusion ability [33]–[35]. The recent CNN-based pansharpening methods learn the mapping between HR/LR PAN images (because these methods assume that the relationship of HR/LR HS images is identical to that between HR/LR PAN images [35]).

The first method using the CNN was proposed by Huang *et al.* [33]. Then, Masi *et al.* [34] transferred the architecture proposed for super-resolution [30] to pansharpening problem, which obtains the state-of-the-art performance and shows the feasibility of using the super-resolution method to solve pansharpening problems. Although these methods achieve a good performance, they treat pansharpening as a black-box deep learning problem and miss two aspects of a pansharpening method, spectral and spatial preservations, and they miss the final goal of combining these two aspects. Then, Yang *et al.* [35] presented a deep network for pansharpening called PanNet, which automatically learns the mapping purely from data, incorporates problem-specific knowledge into the deep learning framework, and focuses on both spatial and spectral preservations. To the best of our knowledge, PanNet achieves the best performance among all CNN-based methods. Recent CNN-based pansharpening methods apply the bicubic interpolation to the LR-HS image to make its scale equal to the PAN image. However, the lost spectral information during the procedure may limit the performance of the state-of-the-art methods. Our approach to this problem is motivated by recent success in CNN-based super-resolution methods [30]–[32], which show great spectral preservation ability and are used in the proposed method. To make them computationally feasible, recent CNN-based pansharpening methods usually estimate their performance on data sets with a few spectral bands. For example, Huang *et al.* [33] evaluated their method on the QuickBird (4 bands) and IKONOS data sets (4 bands), Masi *et al.* [34] used IKONOS, GeoEye1 (4 bands), and WorldView2 (8 bands) data sets, and Yang *et al.* [35] tested PanNet on the WorldView2, WorldView3 (16 bands), and IKONOS data sets.

In this paper, we propose a detail-based deep Laplacian pansharpening (DDLPS) method for the fusion of a simulated PAN image and an LR-HS image with the following features.

- 1) Considering insufficient training samples of HS image, the current CNN-based methods usually employ multispectral satellite images, which may cause that the methods degrade when applied to hundreds of bands. The proposed method uses the RGB image to train the network and achieves a satisfactory performance

when applied to both multispectral images and HS images.

- 2) The traditional CNN-based methods treat pansharpening problem as a black-box problem, while the proposed method focuses on the two aspects of pansharpening: spatial and spectral preservations.
- 3) The proposed DDLPS uses a guided filter [36] to transfer the spatial details of the PAN image to the LR-HS image because of its efficiency and strong ability to achieve an appropriate balance between spectral and spatial preservations.
- 4) We propose a novel adaptive parameter selection method instead of manually tuning the parameter and turning the optimization into solving a Sylvester equation. Thus, the detail extraction can better adapt to the fusion of complex HS images.
- 5) To demonstrate the better performance on images with more bands, we estimate the DDLPS on data sets with hundreds of bands, e.g., Moffett Field (176 bands), Chikusei (128 bands), and Salinas (204 bands) data sets. The experimental results reflect its universality in HS image pansharpening
- 6) Unlike the traditional CNN-based methods, which use bicubic interpolation to upsample the LR-HS image, we use the deep Laplacian pyramid super-resolution network (LapSRN) proposed by Lai *et al.* [37] to conduct super-resolution on the LR-HS image. As shown in Fig. 1, the LapSRN can better preserve the spatial and spectral information than the bicubic interpolation.

The remainder of this paper is organized as follows. Section II formulates the problem of HS image pansharpening. The proposed DDLPS approach is discussed in Section III. Section IV presents the experimental results of three databases and the corresponding discussions. Finally, the conclusions are drawn in Section V.

## II. PROPOSED DDLPS APPROACH

The proposed DDLPS method considers both spatial content of the PAN image and spectral content of the LR-HS image. The DDLPS method mainly includes the following steps: conduct super-resolution via LapSRN, inject details, and turn the optimization problem into solving a Sylvester equation. The detail injection includes two steps: guided filtering and detail extraction. The diagram of the proposed method is shown in Fig. 2.

### A. Problem Formulation

Let the desired HR-HS image be denoted as  $\mathbf{X} \in \mathbb{R}^{S \times N}$ , where  $S$  is the number of spectral bands,  $N$  is the number of pixels, and  $\mathbf{X}_l$  is the  $l$ th band of the image. Let the LR-HS image be denoted as  $\mathbf{Y} \in \mathbb{R}^{S \times n}$ , where  $n = N/d^2$  is the number of pixels with  $d$  being the scale factor. Let the PAN image be denoted as  $\mathbf{P} \in \mathbb{R}^{1 \times N}$ .  $\mathbf{X}$  and  $\mathbf{Y}$  have the same numbers of bands, but  $\mathbf{X}$  is spatially downsampled;  $\mathbf{X}$  and  $\mathbf{P}$  have the same numbers of pixels, but  $\mathbf{P}$  is spectrally downsampled.

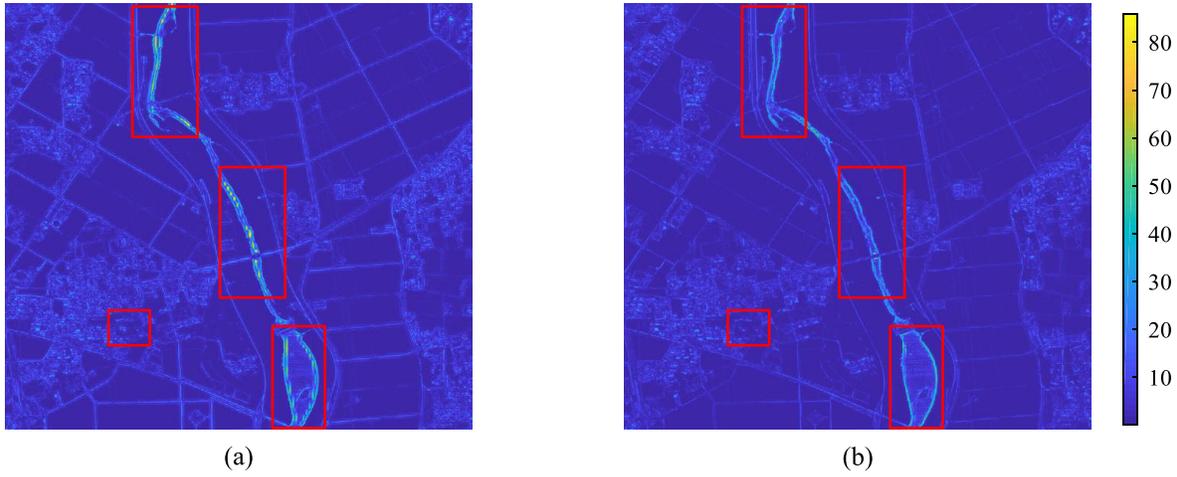


Fig. 1. Spectral angle mapping of the Chikusei data set using (a) bicubic interpolation and (b) LapSRN.

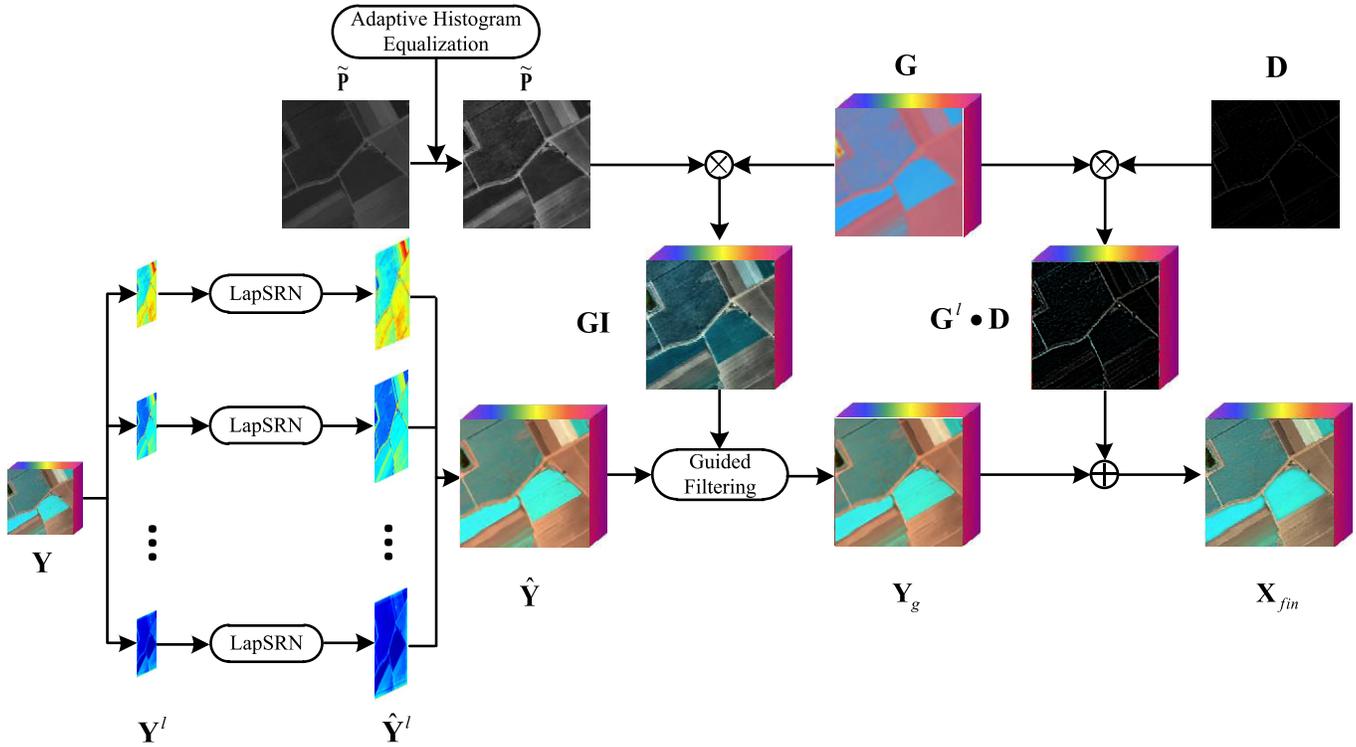


Fig. 2. Diagram of the proposed DDLPS method.  $Y$ ,  $\hat{Y}$ ,  $\tilde{P}$ ,  $\hat{P}$ ,  $D$ ,  $G$ ,  $GI$ ,  $Y_g$ , and  $X_{fin}$  represent the LR-HS image, image after upsampling, PAN images before and after the adaptive histogram equalization, extracted details [see (8)], gain matrix [see (3) and (4)], guidance image, image after the guided image filtering, and output HR-HS image, respectively.

### B. Upsampling via LapSRN

The LR-HS image is first upsampled to the scale of the PAN image. Recent pansharpening methods use bicubic interpolation to perform upsampling. However, the bicubic interpolation may not have the ideal spectral information preservation ability according to Fig. 1, whereas LapSRN [37] satisfies our demands. This method demonstrates excellent performance in preserving spatial and spectral information.

Super-resolution is successively conducted on each band of the origin LR-HS image successively

$$\hat{Y}^l = \text{LapSRN}(Y^l) \quad (1)$$

where  $\hat{Y}$  is the image after upsampling and  $\hat{Y}^l$  is its  $l$ th band. The architecture and theory of LapSRN are demonstrated in Fig. 3.

### C. Guided Filtering

The guided image filter is proposed by He *et al.* [36] and demonstrates its efficiency and strong ability to transfer the structures in a guidance image to the filtering output. The guided filter principal component analysis (GFPCA) [38] method first uses the guided image filter to solve pansharpening problems because it can transfer the spatial structures

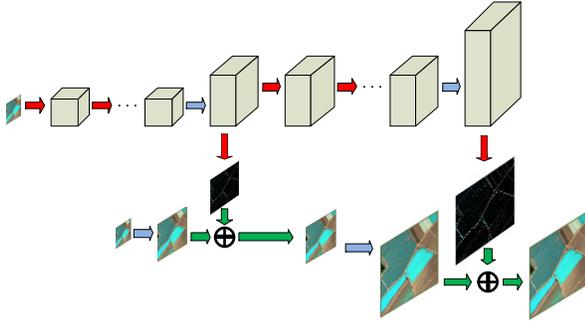


Fig. 3. Network architecture of LapSRN. The red, blue, and green arrows indicate the convolution layers, transposed convolution (upsampling), and elementwise operator, respectively.

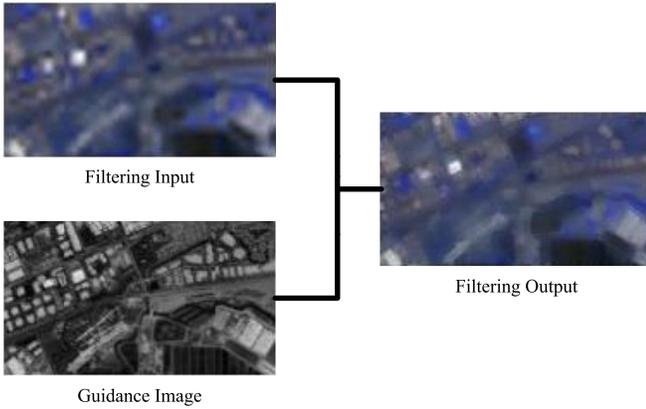


Fig. 4. Example of guided image filtering.

in the PAN image to enhance the HS image. Fig. 4 shows an example of guided image filtering.

To acquire the guidance image, enhancing the spatial information of the PAN image and constructing a gain matrix to preserve the spectral information are essential.

With respect to  $\mathbf{P} \in \mathbb{R}^{1 \times N}$ , we simply obtain matrix  $\tilde{\mathbf{P}}$  that contains the spatial details of the PAN image using the following formula:

$$\tilde{\mathbf{P}} = \text{AHE}(\mathbf{P}) \quad (2)$$

where  $\text{AHE}(\cdot)$  is the adaptive histogram equalization function [39] and  $\tilde{\mathbf{P}} \in \mathbb{R}^{1 \times N}$ . We use  $\tilde{\mathbf{P}}$  to denote the PAN image in the remaining of this paper.

Constructing a gain matrix can control the spatial and spectral distortions. To reduce the spectral distortion, the relationship between the spectral bands should be unchanged when mapping to an HR-HS image. As mentioned in [40], the gain matrix has the following relationship with  $\hat{\mathbf{Y}}$ :

$$\mathbf{G}^s \propto \frac{\hat{\mathbf{Y}}^s}{\frac{1}{S} \cdot \sum_{s=1}^S \hat{\mathbf{Y}}^s} \quad (3)$$

where  $\mathbf{G} \in \mathbb{R}^{S \times N}$  is the gain matrix,  $\mathbf{G}^s$  is the  $s$ th band of the gain matrix,  $\hat{\mathbf{Y}}$  is the LR-HS image after the super-resolution, and  $\hat{\mathbf{Y}}^s$  is its  $s$ th band. To tune the method to a balance between spectral and spatial preservations, we introduce a

tradeoff parameter  $\delta$  to control the injected details

$$\mathbf{G}^s = \delta \cdot \frac{\hat{\mathbf{Y}}^s}{\frac{1}{S} \cdot \sum_{s=1}^S \hat{\mathbf{Y}}^s}. \quad (4)$$

Then, the spatial and spectral details are injected into the guidance image and transferred to the LR-HS image

$$\mathbf{GI}^s = \hat{\mathbf{P}} \cdot \mathbf{G}^s \quad (5)$$

where  $\mathbf{GI}$  is the guidance image,  $\mathbf{GI}^s$  is the  $l$ th band of the guidance image, and  $\cdot$  is the elementwise multiplication.

Given the guidance image  $\mathbf{GI}$  and input image  $\hat{\mathbf{Y}}$ , the output image can be obtained as

$$\mathbf{Y}_g^s = \text{GF}(\mathbf{GI}^s, \hat{\mathbf{Y}}^s) \quad (6)$$

where  $\text{GF}(\cdot)$  is the guided filter function,  $\mathbf{Y}_g$  is the guided filter output, and  $\mathbf{Y}_g^s$  is the  $l$ th band of output image  $\mathbf{Y}_g$ . The kernel of the guided image filter is [25]

$$W_{i,j}(I) = \frac{1}{|\omega|^2} \sum_{k(i,j) \in \omega_k} \left( 1 + \frac{(I_i - \mu_k)(I_j - \mu_k)}{\sigma_k^2 + \varepsilon} \right) \quad (7)$$

where  $I$  is the guidance image and  $\mu_k$  and  $\sigma_k^2$  are the mean and variance of  $I$  in area  $\omega_k$ , respectively.

#### D. Detail Extraction

Although the guided image filter successfully enhances the details in the edge of the image, some spatial and spectral information in the flat area (their pixel values are similar to those around them) is not enhanced through guided filtering. Thus, injecting details to them can greatly improve their appearance.

First, a Gaussian filter is performed on the PAN image to remove the details (high-frequency component), and we obtain the low-frequency component. Then, details will be obtained when the low-frequency component from the original PAN image is subtracted as

$$\mathbf{D} = \hat{\mathbf{P}} - \hat{\mathbf{P}} * g \quad (8)$$

where  $\mathbf{D}$  denotes the details of the image (high-frequency component),  $g$  is the Gaussian filter, and  $*$  is the convolution operation.

We can use the gain matrix to inject the details into the spectral information. Then, the output HR-HS image  $\mathbf{X}_{\text{fin}}$  can be obtained as

$$\begin{cases} \mathbf{X}_{\text{fin}} = \mathbf{Y}_g + \mathbf{G} \cdot \tilde{\mathbf{D}} \\ \tilde{\mathbf{D}}^s = \mathbf{D} \end{cases} \quad (9)$$

where  $\mathbf{X}_{\text{fin}}, \tilde{\mathbf{D}} \in \mathbb{R}^{S \times N}$ ,  $\tilde{\mathbf{D}}^s$  is the  $s$ th band of  $\tilde{\mathbf{D}}$ , and  $\cdot$  is the elementwise multiplication.

#### E. Optimization

To generate the final HR-HS  $\mathbf{X}_{\text{fin}}$ , we must find the corresponding optimal  $\delta$ . Instead of tuning the value of  $\delta$ , we find the corresponding  $\mathbf{G}$  by solving an optimization problem.

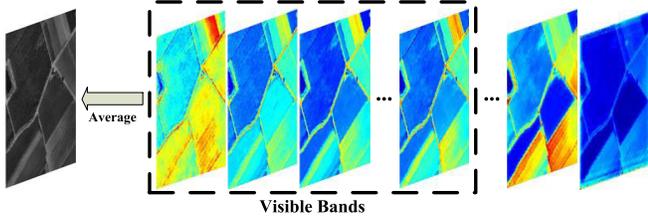


Fig. 5. Creation of the PAN image.

**Algorithm 1** Closed-Form Solution of Sylvester Equation With Respect to  $\mathbf{G}$

**Input:**  $\mathbf{R}, \mathbf{B}, \mathbf{S}, \mathbf{Y}, \mathbf{Y}_g, \tilde{\mathbf{P}}$ .

- 1: Eigendecomposition of  $\mathbf{B} : \mathbf{B} = \mathbf{FDF}^H$
- 2:  $\mathbf{D} = D(\tilde{\mathbf{I}}_d \otimes \mathbf{I}_m)$
- 3: Eigendecomposition of  $\mathbf{C}_1 : \mathbf{C}_1 = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}$
- 4:  $\mathbf{C} = \tilde{\mathbf{Q}}^{-1}\mathbf{C}\mathbf{F}$
- 5: **for**  $s=1 \rightarrow S$  **do**
- 6:  $\check{\mathbf{g}}_l = \lambda_s^{-1}\check{\mathbf{c}}_s - \lambda_s^{-1}\check{\mathbf{c}}_s\check{\mathbf{D}}(\lambda_s d\mathbf{I}_m + \sum_{i=1}^d \mathbf{D}_i^2)^{-1}\check{\mathbf{D}}^H$  ▷  
 Compute auxiliary matrix  $\mathbf{G}$  band by band:
- 7: **end for**
- 8: **Set**  $\mathbf{G} = \mathbf{Q}\check{\mathbf{G}}\mathbf{F}^H$

**Output:**  $\mathbf{G}$

Since the objective is to estimate the HR-HS image  $\mathbf{X}_{\text{fin}}$  from  $\mathbf{Y}$  and  $\mathbf{P}$ , the fusion problem can be written as

$$\mathcal{L} = \min_{\mathbf{G}} [f_1(\mathbf{G}, \tilde{\mathbf{P}}) + f_2(\mathbf{G}, \mathbf{Y})] \quad (10)$$

where functions  $f_1(\cdot)$  and  $f_2(\cdot)$  focus on the spatial consistency and spectral consistency, respectively. We motivate this formulation by considering our approach in terms of the two goals of pansharpening. The spatial information of the HR-HS image mainly comes from the PAN image, and an image generated by optimal pansharpening after spectral sampling should be identical to the PAN image. Thus, in this paper, the spatial consistency is evaluated as

$$f_1(\mathbf{G}, \tilde{\mathbf{P}}) = \left\| \frac{1}{S} \sum_{s=1}^S (\mathbf{Y}_g^s + \mathbf{G}^s \cdot \tilde{\mathbf{D}}^s) - \tilde{\mathbf{P}} \right\|_F^2 \quad (11)$$

where  $\|\cdot\|_F$  is the Frobenius norm. The spectral information of the HR-HS image is provided by the LR-HS image, and an optimal HR-HS image after spectral sampling should be identical to the PAN image. Thus, the spectral consistency is

$$f_2(\mathbf{G}, \mathbf{Y}) = \|\downarrow(\mathbf{Y}_g + \mathbf{G} \cdot \tilde{\mathbf{D}}) - \mathbf{Y}\|_F^2 \quad (12)$$

where  $\downarrow$  denotes the bicubic interpolation.

To simplify the computational work, we denote the operation  $(1/S) \sum_{s=1}^S (\cdot)$  as  $\mathbf{R} \in \mathbb{R}^{1 \times S}$ . In this paper, we consider blurring and downsampling [26], which are denoted as  $\mathbf{B} \in \mathbb{R}^{N \times N}$  and  $\mathbf{S} \in \mathbb{R}^{N \times n}$ , respectively. The bicubic interpolation can be decomposed into blurring and downsampling. Some state-of-the-art methods [1], [19], [21] assume that matrix  $\mathbf{B}$  is a block circulant matrix with circulant blocks and can be decomposed as  $\mathbf{B} = \mathbf{FDF}^H$ , where  $\mathbf{F}$  is the DFT matrix

TABLE I  
QUANTITATIVE RESULTS OF TEN METHODS ESTIMATED ON THE MOFFETT DATA SET

Quantitative Indices	CC	SAM	RMSE	ERGAS
GS	0.8559	13.8738	0.0584	10.0608
GSA	0.9526	7.2831	0.0347	5.9500
SFIM	<u>0.9676</u>	5.4691	<u>0.0291</u>	<u>4.7024</u>
MTF-GLP	0.9663	<u>5.3889</u>	0.0302	4.9279
MTF-GLP-HPM	0.9652	5.4420	0.0308	4.8970
Bayesian Naive	0.9477	5.8143	0.0329	5.9741
Bayesian Sparse	0.9576	7.5509	0.0318	5.4736
HySure	0.9656	5.8689	0.0298	4.9555
GFPCA	0.9258	8.4540	0.0416	7.2285
CNMF	0.9465	6.5848	0.0355	6.1638
Proposed	<b>0.9761</b>	<b>4.8476</b>	<b>0.0241</b>	<b>4.0521</b>
Optimal Value	1	0	0	0

TABLE II  
SUM OF ABSOLUTE DIFFERENCE VALUES OF EACH PIXEL ESTIMATED ON THE MOFFETT DATA SET

Methods	(a)	(b)	(c)	(d)
GS	5.4295	8.4713	14.2378	4.9433
GSA	10.3782	9.7134	8.9213	5.0495
SFIM	7.6947	2.4763	6.7826	4.0723
MTF-GLP	<u>5.0309</u>	3.8654	7.0338	3.5762
MTF-GLP-HPM	5.1012	2.1136	9.2138	3.5945
Bayesian Naive	13.3000	2.9019	8.2074	3.4123
Bayesian Sparse	8.7181	2.0396	3.3136	5.0731
HySure	5.6417	4.9055	4.7616	<u>3.3972</u>
GFPCA	11.2694	7.0202	3.7811	6.7495
CNMF	9.9379	<u>1.4864</u>	<u>1.9216</u>	6.4955
Proposed	<b>3.7584</b>	<b>0.9824</b>	<b>1.0166</b>	<b>2.9898</b>
Optimal Value	0			

( $\mathbf{F}\mathbf{F}^H = \mathbf{I}_n$ ) and  $\mathbf{D}$  is a diagonal matrix that contains the eigenvalues of  $\mathbf{B}$ .

Thus,  $f_1(\cdot)$  and  $f_2(\cdot)$  can be simplified as

$$\begin{cases} f_1(\mathbf{G}, \tilde{\mathbf{P}}) = \|\mathbf{R}(\mathbf{Y}_g + \mathbf{G} \cdot \tilde{\mathbf{D}}) - \tilde{\mathbf{P}}\|_F^2 \\ f_2(\mathbf{G}, \mathbf{Y}) = \|(\mathbf{Y}_g + \mathbf{G} \cdot \tilde{\mathbf{D}})\mathbf{B}\mathbf{S} - \mathbf{Y}\|_F^2. \end{cases} \quad (13)$$

Based on the above-mentioned model, we estimate the HR-HS image by solving the following optimization problem:

$$\mathcal{L} = \min_{\mathbf{G}} \underbrace{\|\mathbf{R}(\mathbf{Y}_g + \mathbf{G} \cdot \tilde{\mathbf{D}}) - \tilde{\mathbf{P}}\|_F^2}_{\text{Spatial Presentation}} + \underbrace{\|(\mathbf{Y}_g + \mathbf{G} \cdot \tilde{\mathbf{D}})\mathbf{B}\mathbf{S} - \mathbf{Y}\|_F^2}_{\text{Spectral Preservation}}. \quad (14)$$

To minimize (14), we force the derivative of (14) for  $\mathbf{G}$  to be zero, and it becomes a Sylvester equation

$$\mathbf{C}_1\mathbf{G} + \mathbf{G}\mathbf{C}_2 = \mathbf{C}_3 \quad (15)$$

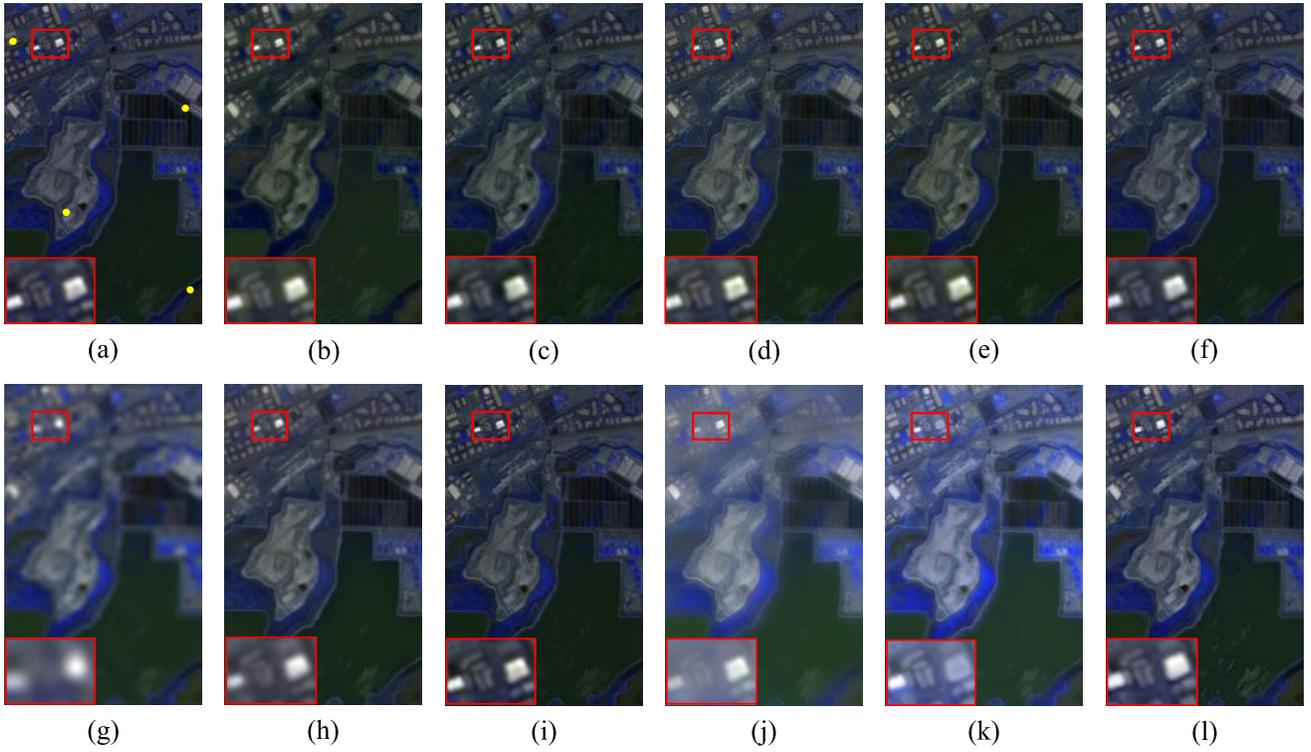


Fig. 6. Ground truth and the pansharpening images generated by different methods based on the Moffett Field data set. (a) Ground truth. (b) GS. (c) GSA. (d) SFIM. (e) MTF-GLP. (f) MTF-GLP-HPM. (g) Bayesian Naive. (h) Bayesian Sparse. (i) HySure. (j) GFPCA. (k) CNMF. (l) DDLPS.

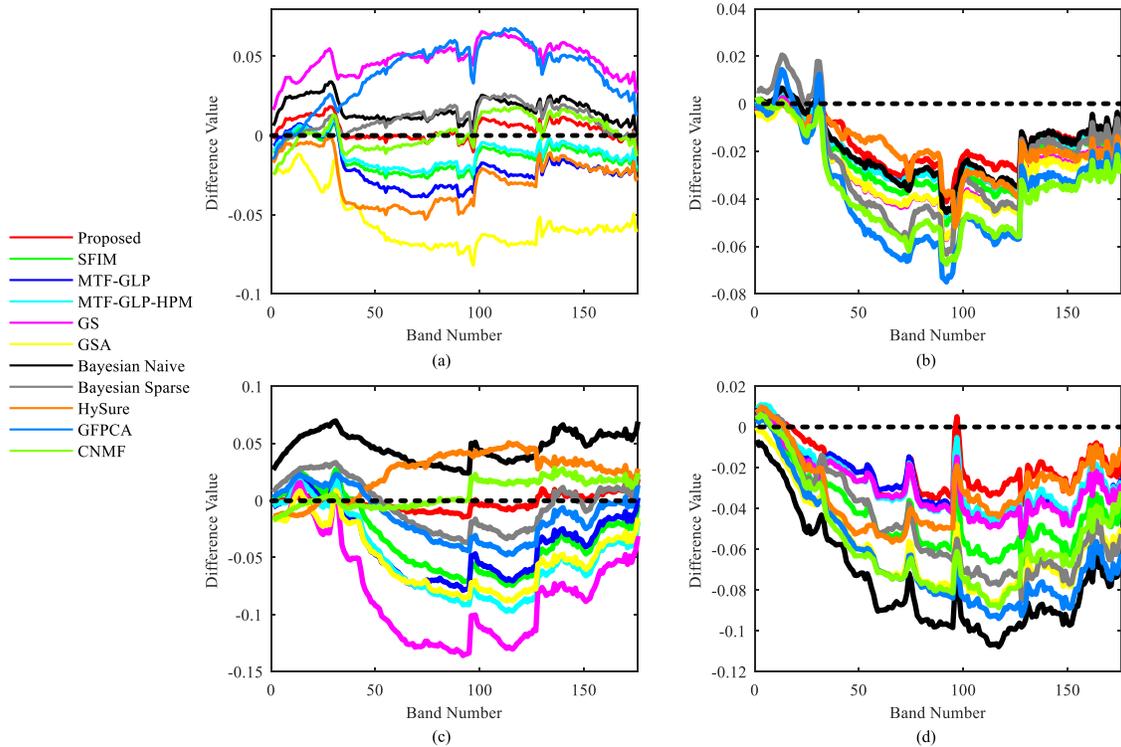


Fig. 7. Comparison of the spectral reflectance difference values on four pixels in Fig. 6(a). (a)–(d) Four pixels marked in yellow in Fig. 6.

where

$$\begin{cases} \mathbf{C}_1 = \mathbf{R}^T \mathbf{R} \\ \mathbf{C}_2 = (\mathbf{B}\mathbf{S})(\mathbf{B}\mathbf{S})^T \\ \mathbf{C}_3 = \mathbf{R}^T (\tilde{\mathbf{P}} - \mathbf{R}\mathbf{Y}_g) + (\mathbf{Y} - \mathbf{Y}_g \mathbf{B}\mathbf{S})(\mathbf{B}\mathbf{S})^T. \end{cases} \quad (16)$$

Equation (15) is known to have a unique solution if and only if the arbitrary sum of the eigenvalues of  $\mathbf{C}_1$  and  $\mathbf{C}_2$  is not equal to zero [41].  $\mathbf{R}$  is the mean operation, so  $\mathbf{C}_1$  is a positive matrix; thus, the eigenvalues of  $\mathbf{C}_1$  are positive.  $\mathbf{C}_2$  is a semipositive number, and its eigenvalues are semipositive

TABLE III  
QUANTITATIVE RESULTS OF TEN METHODS ESTIMATED  
ON THE CHIKUSEI DATA SET

Quantitative Indices	CC	SAM	RMSE	ERGAS
GS	0.8632	4.6403	0.0514	7.7978
GSA	0.9478	3.0232	0.0355	<b>3.6037</b>
SFIM	0.9384	2.9753	0.0388	3.9547
MTF-GLP	0.9196	3.5539	0.0462	4.1036
MTF-GLP-HPM	0.9359	2.9212	0.0405	3.6596
Bayesian Naive	0.9132	3.3739	<u>0.0335</u>	7.0120
Bayesian Sparse	0.8463	4.8240	<u>0.0582</u>	7.4081
HySure	0.9286	3.0897	0.0411	4.2758
GFPCA	0.8703	4.3642	0.0472	7.7978
CNMF	<u>0.9476</u>	<u>2.9125</u>	0.0337	<u>3.6517</u>
Proposed	<b>0.9552</b>	<b>2.5588</b>	<b>0.0298</b>	4.1393
Optimal Value	1	0	0	0

TABLE IV  
SUM OF ABSOLUTE DIFFERENCE VALUES OF EACH PIXEL  
ESTIMATED ON THE CHIKUSEI DATA SET

Methods	(a)	(b)	(c)	(d)
GS	2.9458	2.5497	3.5546	1.6725
GSA	1.2503	<u>0.7498</u>	2.1160	1.6498
SFIM	2.2597	<b>0.7229</b>	2.4229	1.4554
MTF-GLP	1.7440	0.8981	2.9153	1.4650
MTF-GLP-HPM	1.5354	0.8493	<u>1.0302</u>	1.2501
Bayesian Naive	<u>1.0935</u>	0.8067	2.0556	<u>1.1014</u>
Bayesian Sparse	1.3130	1.8672	2.2845	3.4775
HySure	4.8030	2.7957	1.8834	4.3394
GFPCA	1.4307	1.7191	3.5502	1.6378
CNMF	4.3062	1.2900	4.5306	1.3511
Proposed	<b>0.7009</b>	0.8391	<b>0.5397</b>	<b>0.9064</b>
Optimal Value		0		

values. Therefore, the solution of Sylvester equation is unique. The approach to solve (15) is summarized in Algorithm 1, and more details can be found in [20].

Based on (4) and Algorithm 1, the selection of  $\delta$  can be written as

$$\begin{cases} \delta = \mu \left( \mathbf{G}^s \cdot \frac{\sum_{s=1}^S \hat{\mathbf{Y}}^s}{S \hat{\mathbf{Y}}^s} \right) \\ \mathbf{G} = \mathbf{Q} \hat{\mathbf{G}} \mathbf{F}^H \end{cases} \quad (17)$$

where  $\mu(\cdot)$  is the mean function and the multiplication and division in the first equation of (17) are elementwise.

### III. EXPERIMENT

#### A. Experimental Setup

In this section, we validate the performance of the proposed method on three public data sets: Moffett Field, Chikusei, and Salinas. Discarding the water absorption bands and low-SNR bands, the Moffett Field data set includes 176 bands of 400–2500 nm. The Chikusei data set comprises 128 bands in the spectral range of 363–1018 nm. The Salinas data set,

TABLE V  
QUANTITATIVE RESULTS OF TEN METHODS ESTIMATED  
ON THE SALINAS DATA SET

Quantitative Indices	CC	SAM	RMSE	ERGAS
GS	0.8162	4.7600	0.0507	4.6606
GSA	0.9613	2.1875	0.0148	2.6563
SFIM	0.9657	1.9182	0.0182	2.3973
MTF-GLP	<u>0.9661</u>	2.0553	0.0191	<b>2.3039</b>
MTF-GLP-HPM	0.9605	1.9699	0.0203	5.0392
Bayesian Naive	0.9581	<u>1.7133</u>	<u>0.0142</u>	2.7844
Bayesian Sparse	0.9438	2.7773	0.0195	2.3737
HySure	0.9493	2.0210	0.0211	2.9067
CNMF	0.9466	2.1314	0.0194	3.1621
Proposed	<b>0.9688</b>	<b>1.3608</b>	<b>0.0131</b>	<u>2.3348</u>
Optimal Value	1	0	0	0

TABLE VI  
SUM OF ABSOLUTE DIFFERENCE VALUES OF EACH PIXEL  
ESTIMATED ON THE SALINAS DATA SET

Methods	(a)	(b)	(c)	(d)
GS	3.5559	2.5775	<b>3.0454</b>	2.5539
GSA	4.0225	4.0426	6.0594	<b>1.2984</b>
SFIM	<u>1.1500</u>	2.0404	5.9648	2.9961
MTF-GLP	1.2224	1.8549	3.4664	3.6708
MTF-GLP-HPM	12.7681	6.4517	4.7796	2.6186
Bayesian Naive	1.3765	1.4843	3.5805	1.934
Bayesian Sparse	2.8992	2.2688	3.3256	5.5838
HySure	1.6054	2.5207	4.2105	3.3529
GFPCA	3.4605	<b>0.9023</b>	3.2609	3.5530
CNMF	1.3056	1.6978	1.6579	1.6579
Proposed	<b>0.4416</b>	<u>1.1475</u>	<u>3.2286</u>	<u>1.3480</u>
Optimal Value		0		

which was taken over Salinas valley in Southern California, comprises 204 bands (244 bands with 20 bands removed) in the spectral of 400–2500 nm [42]. Because the human visible wavelength is approximately 400–760 nm, the PAN image can be simulated by averaging the visible bands, as shown in Fig. 5.

For the quantitative evaluation, the HS image is used as the ground truth, and the LR-HS image is obtained by being filtered by a  $9 \times 9$  Gaussian smoothing kernel and downsampled by a factor of 4. Thus, the spatial sizes of LR-HS images in the Moffett Field, Chikusei, and Salinas data sets are  $65 \times 40$ ,  $150 \times 150$ , and  $50 \times 50$ , respectively, and the spatial resolutions of the PAN images are  $260 \times 160$ ,  $600 \times 600$ , and  $200 \times 200$ , respectively.

We use 91 images from [43] and 200 images from the training set of the Berkeley Segmentation data set [44] to train the LapSRN. We randomly sample 64 patches with the size of  $128 \times 128$  in each training batch.

The experiment on the three data sets is conducted using MATLAB R2014b and performed on a server with an RHEL 6.5 operating system, an Intel Xeon E5-2650 V4 CPU at 2.20 GHz, an NVIDIA K80 GPU and 128-GB memory.

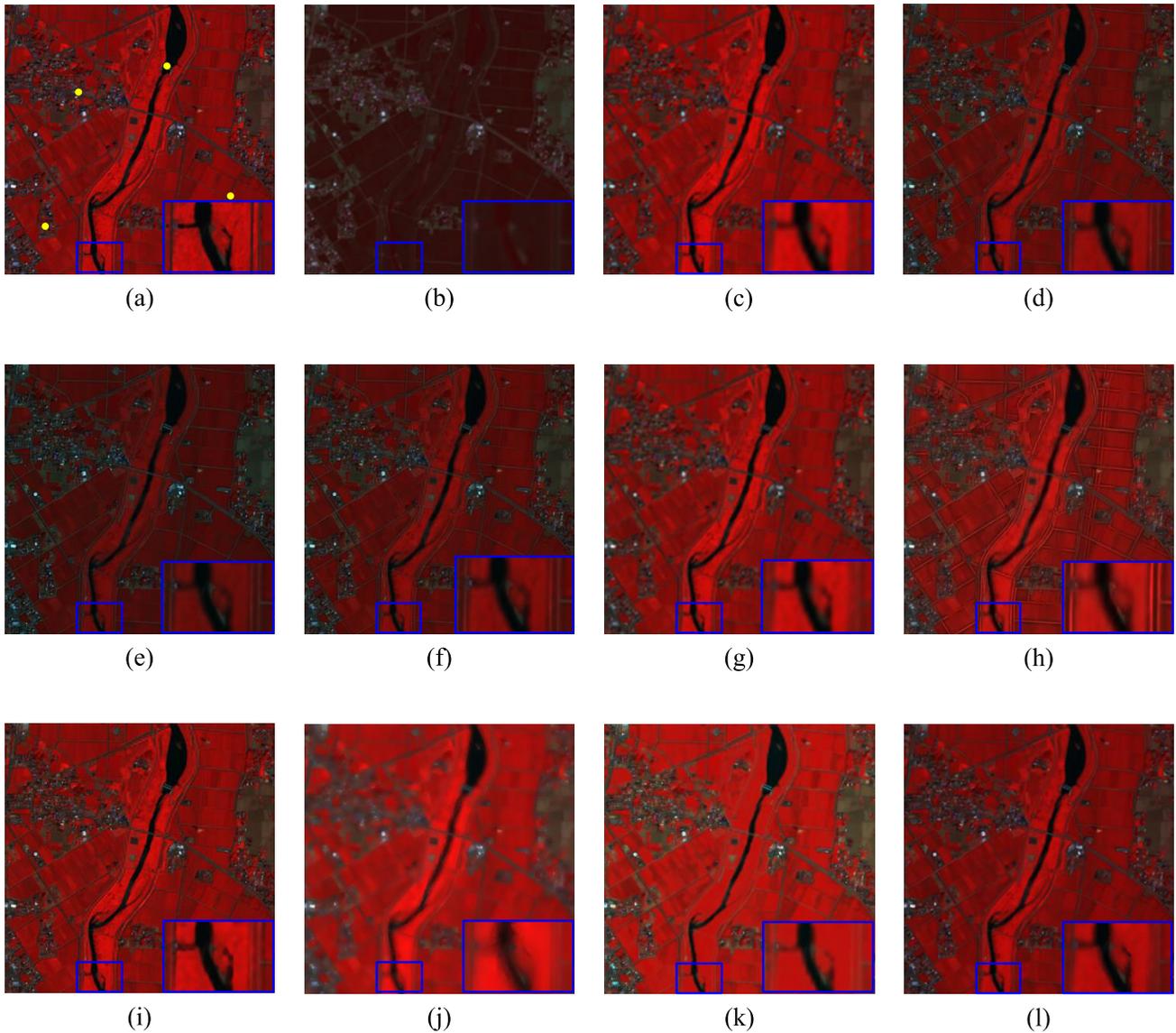


Fig. 8. Ground truth and the pansharpening images generated by different methods based on the Chikusei data set. (a) Ground truth. (b) GS. (c) GSA. (d) SFIM. (e) MTF-GLP. (f) MTF-GLP-HPM. (g) Bayesian naïve. (h) Bayesian Sparse. (i) HySure. (j) GFPCA. (k) CNMF. (l) DDLPS.

### B. Comparative Methods

The proposed DDLPS method is compared with 11 recent state-of-the-art or representative methods: the CS approach (GS and GSA), MRA approach [smoothing filter-based intensity modulation (SFIM) [45], MTF-GLP [15], and MTF-GLP-HPM [29]], Bayesian approach (Bayesian Naive [20], Bayesian Sparse [19], and HySure [21]), hybrid approach (GFPCA [38]), and matrix factorization approach (CNMF [18]). With respect to CNN-based approaches, we compare the proposed method with PanNet [35], which is the best CNN-based method. Because all 176 bands of the Moffett data set cannot be performed on an NVIDIA K80 GPU via PanNet (the required memory is far more than the GPU memory), we spectrally downsample the Moffett data set and create a new data set with 8 bands, which is called Moffett-8. Thus, the proposed method and PanNet are compared on the Moffett-8 data set.

### C. Quantitative Indices

For quantitative evaluation, we use four widely used indices: the correlation coefficient (CC), spectral angle mapper (SAM) [46], root-mean-squared error (RMSE), and erreur relative globale adimensionnelle de synthèse (ERGAS) [47]. The CC focuses on the spatial quality with the best value of 1. The SAM reflects the degree of spectral similarity. The RMSE and ERGAS measure the overall quality of the HR-HS image. The optimal values of the SAM, RMSE, and ERGAS are 0.

### D. Experimental Results

1) *Moffett Field Data Set*: In this section, we compare our DDLPS method with the aforementioned methods on the Moffett Field data set. As shown in Table I, the proposed method demonstrates the best performance with respect to the four quality indices: CC, SAM, RMSE, and ERGAS.

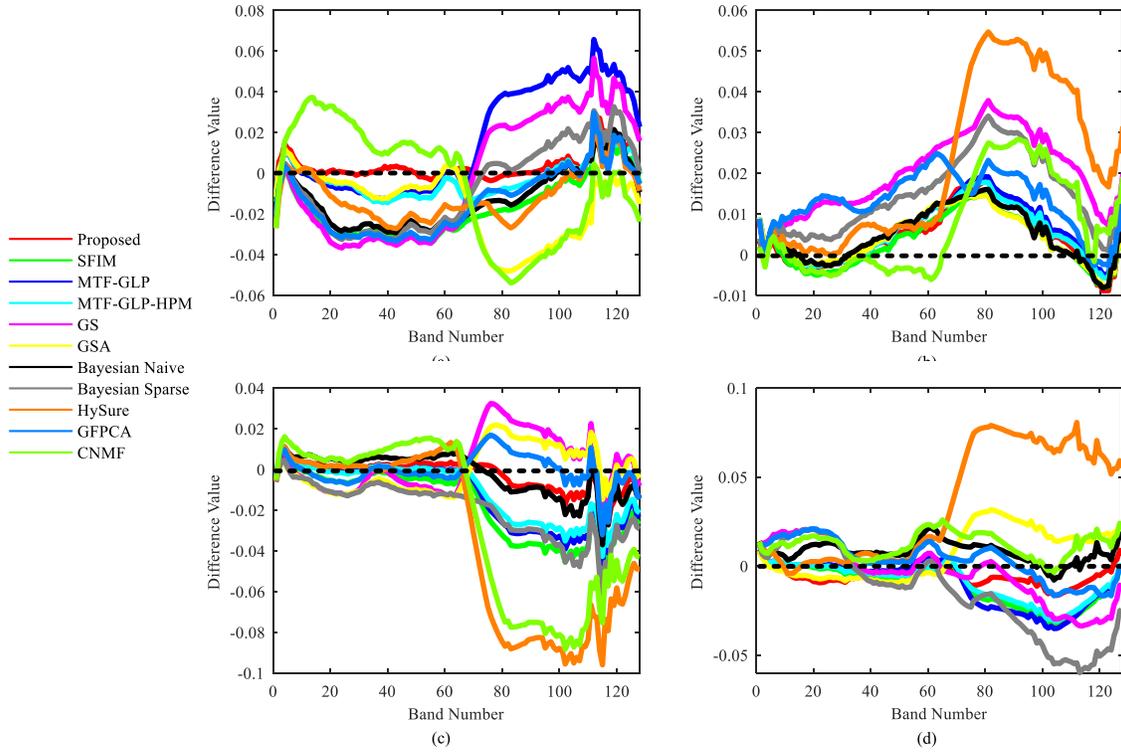


Fig. 9. Comparison of the spectral reflectance difference values on four pixels in Fig. 8(a). (a)–(d) Four pixels marked in yellow in Fig. 8.

TABLE VII  
QUANTITATIVE RESULTS OF TEN METHODS ESTIMATED ON THE MOFFETT-8 DATA SET

Quantitative Indices	CC	SAM	RMSE	ERGAS
PanNet	0.9649	4.9907	0.0297	4.5270
Proposed	<b>0.9780</b>	<b>3.7634</b>	<b>0.0239</b>	<b>3.6686</b>
Optimal Value	1	0	0	0

TABLE VIII  
SUM OF ABSOLUTE DIFFERENCE VALUES OF EACH PIXEL ESTIMATED ON THE MOFFETT-8 DATA SET

Quantitative Indices	(a)	(b)	(c)	(d)
GS	0.4201	0.3131	0.1767	0.1658
Proposed	<b>0.1379</b>	<b>0.0633</b>	<b>0.0925</b>	<b>0.1547</b>
Optimal Value	0			

The visual appearance of the DDLPS and comparing methods is shown in Fig. 6. Although MTF-GLP-HPM, HySure, and DDLPS have approximately identical ability to preserve spatial details’ information according to visual appearance, their overall image spatial quality is not as good as that of the DDLPS. To show the spectral distortion, the spectral reflectance difference values between the ground truth and each fused image are compared [40]. Four pixels of Fig. 7 are marked in yellow in Fig. 6(a). With respect to Fig. 7, being closer to the dotted line indicates better preserved spectral

information. Obviously, the proposed DDLPS has excellent spectral preservation ability.

Since small spectral difference values cannot be perceived from the figures, we compute the sums of the absolute difference values of each pixel, which are shown in Table II.

2) *Chikusei Data Set*: For the Chikusei data set, Table III shows that the proposed DDLPS generally outperforms all other methods (all quantitative indices rank first except ERGAS). As shown in the fused images, the proposed method holds more spatial information. With respect to lines (e.g., roads and rivers), the proposed method shows the outstanding performance in preserving spatial details, as shown in Fig. 8. The GS method suffers from an obvious spectral distortion, and the Bayesian sparse and GFPCA methods cannot preserve details well. The proposed DDLPS can prevent the shapes of the details from suffering distortion during the fusion procedure. Fig. 9 shows the spectral difference values of the four pixels marked in yellow in Fig. 8(a). The four pixels are at the edge of the image or in areas with details, so Fig. 9 can reflect the pansharpening performance on a single pixel, which is more representative than the global quantitative indices. As shown in Fig. 9, the difference values of the proposed DDLPS are closest to the dotted line, which implies that the proposed method offers the best spectral preservation. The sum of the absolute difference values of each pixel is reported in Table IV.

As shown in Fig. 9 and Table IV, the proposed DDLPS has excellent spectral preservation ability.

3) *Salinas Data Set*: As shown in Table V, the proposed method generally outperforms all comparing methods (all

TABLE IX  
EFFECTIVENESS ESTIMATION ON THREE DATA SETS USING ONE STEP

Data Set	Moffett Field				Chikusei				Salinas			
Quantitative Indices	CC	SAM	RMSE	ERGAS	CC	SAM	RMSE	ERGAS	CC	SAM	RMSE	ERGAS
$X_{sr}$	0.9597	<u>4.9675</u>	0.0288	5.2344	<u>0.9405</u>	<u>2.6501</u>	<b>0.0279</b>	<u>5.7523</u>	0.9643	<b>1.3476</b>	<b>0.0109</b>	2.6845
$X_{gf}$	0.9553	5.6463	0.0313	5.5476	0.9267	3.2413	0.0341	6.1310	0.9653	1.5914	0.0135	2.4565
$X_{de}$	<u>0.9643</u>	5.6181	<u>0.0281</u>	<u>4.9475</u>	0.9291	3.2437	0.0343	5.8476	<u>0.9676</u>	1.6107	0.0140	<b>2.3105</b>
$X_{fin}$	<b>0.9761</b>	<b>4.8476</b>	<b>0.0241</b>	<b>4.0521</b>	<b>0.9552</b>	<b>2.5588</b>	<u>0.0298</u>	<b>4.1393</b>	<b>0.9688</b>	<u>1.3608</u>	<u>0.0131</u>	<u>2.3348</u>
Optimal Value	1	0	0	0	1	0	0	0	1	0	0	0

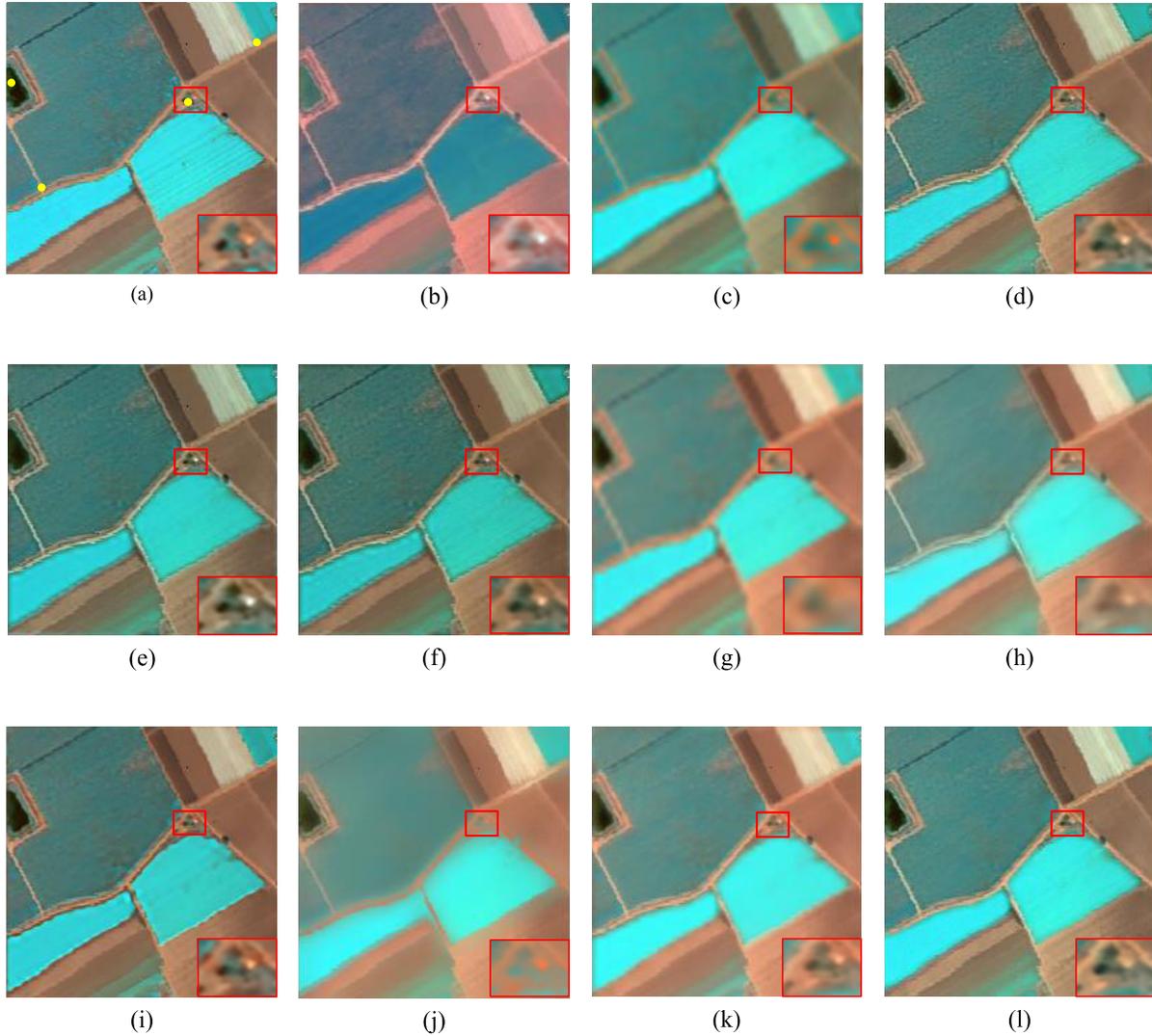


Fig. 10. Ground truth and the pansharpening images generated by different methods based on Salinas data set. (a) Ground truth. (b) GS. (c) GSA. (d) SFIM. (e) MTF-GLP. (f) MTF-GLP-HPM. (g) Bayesian Naive. (h) Bayesian Sparse. (i) HySure. (j) GFPCA. (k) CNMF. (l) DDLPS.

quantitative indices rank first except ERGAS; the MTF-GLP ranks first in ERGAS). Fig. 10 also shows that the DDLPS holds more spatial details than other methods and retains better spectral information. With respect to a single pixel, Fig. 11 shows the spectral preservation performance of each method, and they are quantitatively reported in Table VI. The proposed DDLPS demonstrates the best spectral preservation performance on a single pixel (the DDLPS generally has the

closest spectral difference values to the dotted line, and its sum of absolute difference values is closest to zero). Although the CNMF method demonstrates great visual appearance, the produced image does not contain more details than the DDLPS. The MTF-GLP-HPM produces similar images, but the quantitative indices reflect that it suffers from distortion.

4) *Moffett-8 Data Set*: We compare PanNet and the proposed DDLPS on a Moffett data set with fewer bands because

TABLE X  
EFFECTIVENESS ESTIMATION ON THREE DATA SETS USING TWO STEPS

Data Set	Moffett Field				Chikusei				Salinas			
Quantitative Indices	CC	SAM	RMSE	ERGAS	CC	SAM	RMSE	ERGAS	CC	SAM	RMSE	ERGAS
$X_{srgf}$	0.9697	<u>4.8554</u>	0.0263	4.5897	<u>0.9519</u>	<b>2.5448</b>	<u>0.0285</u>	<u>4.7510</u>	0.9698	1.3478	<u>0.0117</u>	2.4166
$X_{srde}$	<u>0.9694</u>	4.9960	<u>0.0259</u>	<u>4.5963</u>	0.9461	2.6636	<b>0.0289</b>	5.1972	0.9616	<u>1.3537</u>	<b>0.0106</b>	6.5178
$X_{gfde}$	0.9667	5.6727	0.0281	4.8417	0.9334	3.2482	0.0353	5.4345	<u>0.9683</u>	1.6490	0.0147	<b>2.2759</b>
$X_{fin}$	<b>0.9770</b>	<b>4.7711</b>	<b>0.0237</b>	<b>4.0088</b>	<b>0.9567</b>	<u>2.5499</u>	0.0284	<b>4.3877</b>	<b>0.9692</b>	<b>1.2803</b>	0.0115	<u>5.4832</u>
Optimal Value	1	0	0	0	1	0	0	0	1	0	0	0

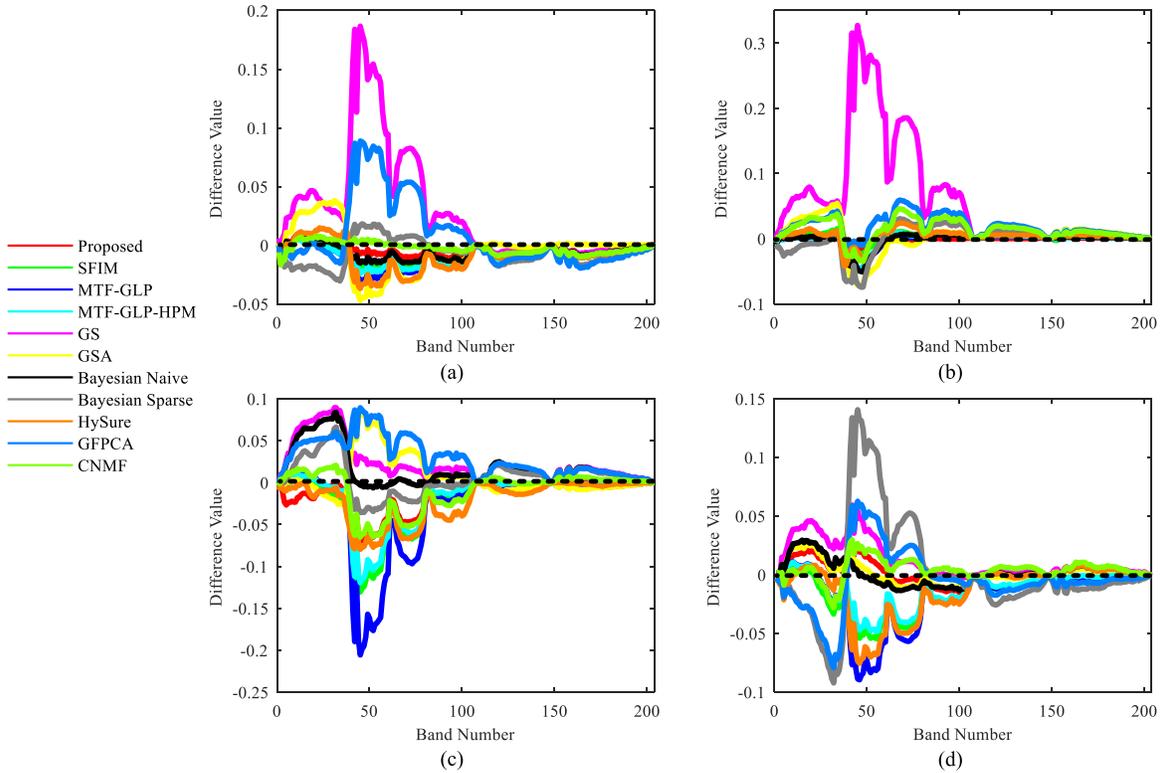


Fig. 11. Comparison of the spectral reflectance difference values on four pixels in Fig. 10(a). (a)–(d) Four pixels marked in yellow in Fig. 10.

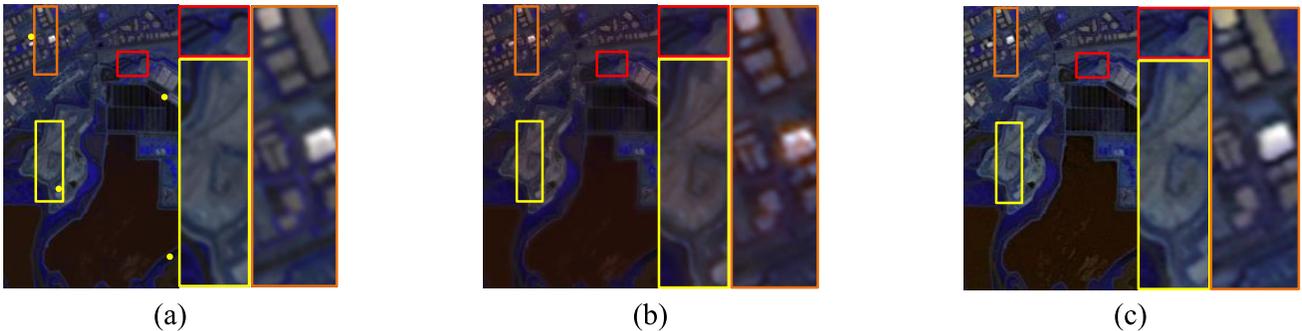


Fig. 12. Ground truth and the pansharpening images generated by PanNet and the proposed DDLPS. (a) Ground truth. (b) PanNet. (c) DDLPS.

176 bands are too large for PanNet. As shown in Fig. 12, the results of PanNet suffer from spectral distortion. With respect to spatial preservation, PanNet has deficient improvement of spatial quality in lines (e.g., roads and rivers) and small objects (e.g., roofs). The results of the four quality

indices are reported in Table VII, which shows that the proposed DDLPS demonstrates a better pansharpening performance than PanNet. The comparison of spectral preservation between the two methods is conducted via spectral difference values as shown in Fig. 13, where four pixels are marked

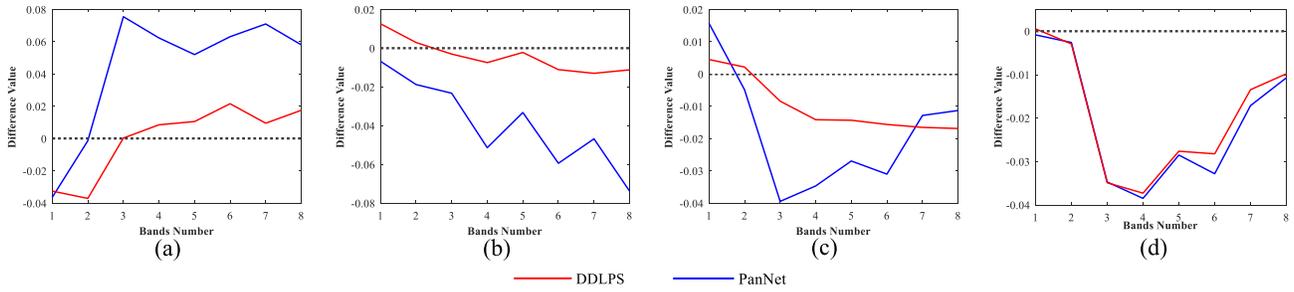


Fig. 13. Comparison of the spectral reflectance difference values on four pixels in Fig. 12(a). (a)–(d) Four pixels marked in yellow in Fig. 12.

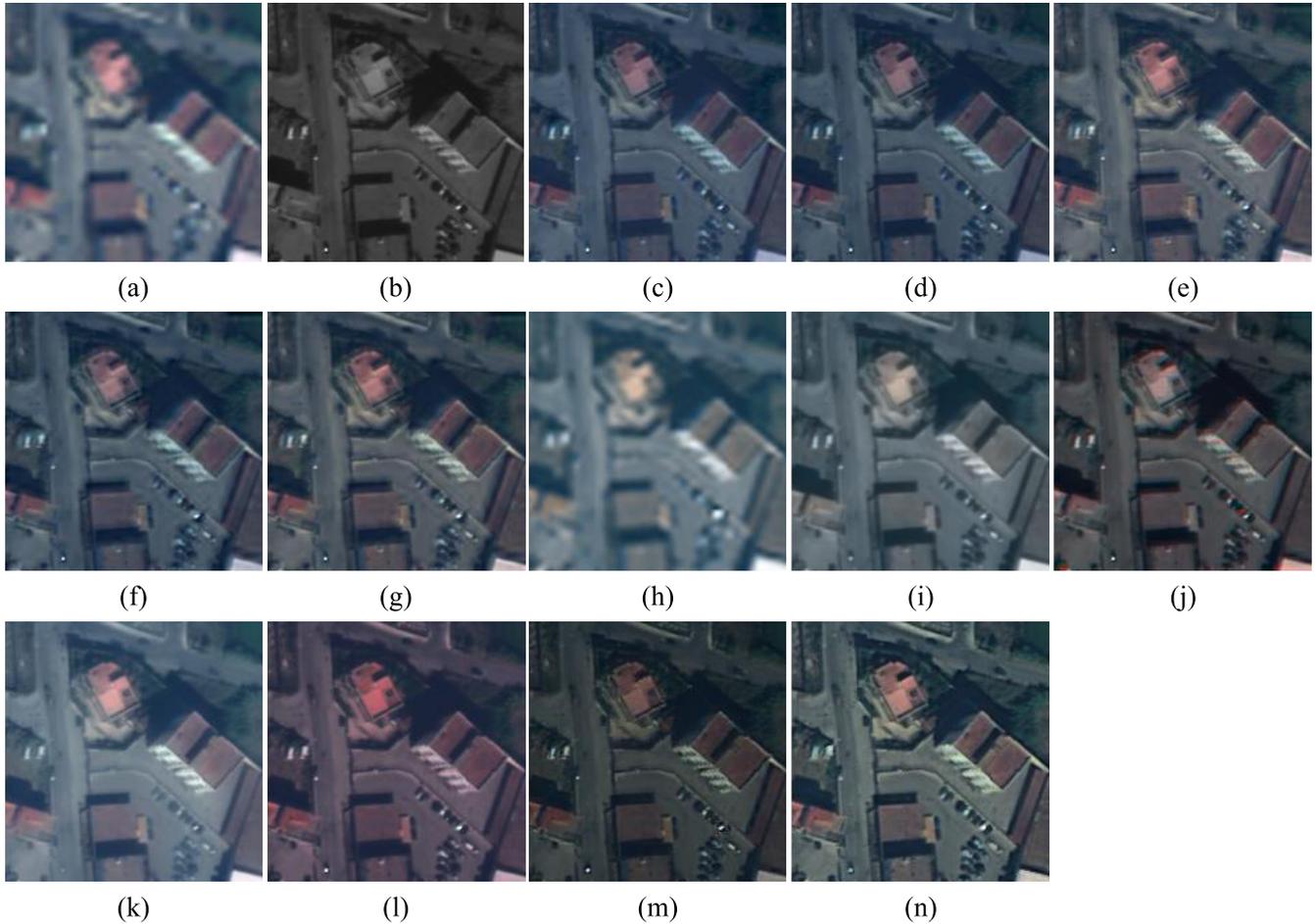


Fig. 14. Pansharpening images generated by different methods. (a) Interpolated. (b) PAN. (c) GS. (d) GSA. (e) SFIM. (f) MTF-GLP. (g) MTF-GLP-HPM. (h) Bayesian Naive. (i) Bayesian Sparse. (j) HySure. (k) GFPCA. (l) CNMF. (m) PanNet. (n) DDLPS.

in yellow in Fig. 12(a). The spectral difference values of the DDLPS are the closest to the benchmark line, which implies that the proposed method can preserve spectral details better than the PanNet. To quantitatively assess the spectral difference values of the two methods, the sums of the absolute difference values are shown in Table VIII.

#### E. Experiments on Real Data

In Fig. 14, we show the actual full-resolution images obtained through pansharpening. During this time, no ground truth exists. However, we employ the upsampled HS images

with bicubic interpolation as a reference. As shown in Fig. 14, most generated images exhibit a very good quality with great improvement with respect to the upsampled image. In particular, the proposed method appears as one of the more effective methods. Generally, most pansharpened images show a satisfactory quality compared with the interpolated image. However, there still exist some problems. HySure, CNMF, and PanNet suffer from spectral distortion, while SFIM and Bayesian Naive introduce spatial blurring, though less than in the interpolated case. According to Fig. 14, the proposed method also suffers from slight spectral distortion.

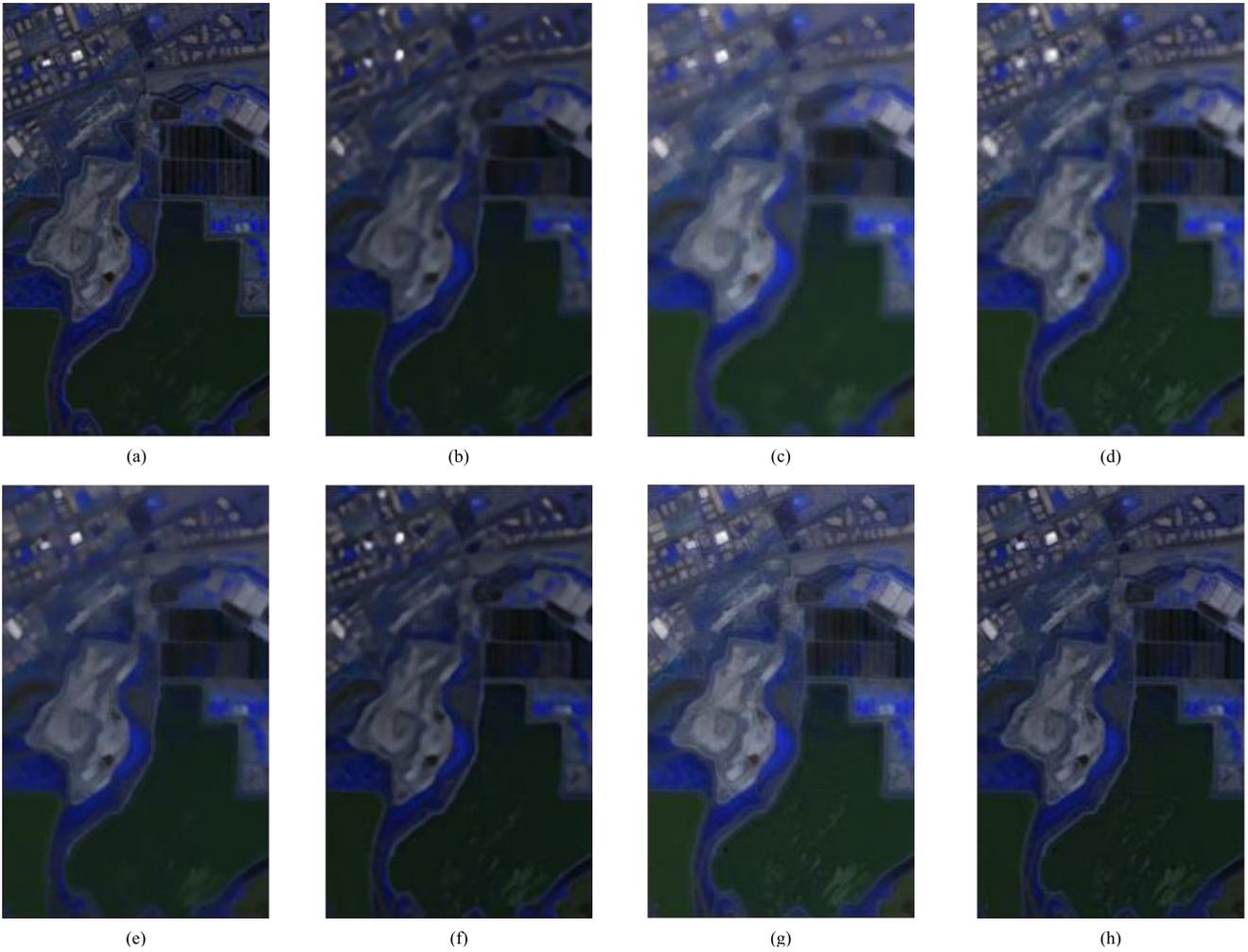


Fig. 15. Ground truth and pansharpening images generated by different combinations of different steps. (a) Ground truth. (b)  $\mathbf{X}_{\text{sr}}$ . (c)  $\mathbf{X}_{\text{gf}}$ . (d)  $\mathbf{X}_{\text{de}}$ . (e)  $\mathbf{X}_{\text{srgf}}$ . (f)  $\mathbf{X}_{\text{srde}}$ . (g)  $\mathbf{X}_{\text{gfde}}$ . (h)  $\mathbf{X}_{\text{fin}}$ .

#### F. Component Analysis

1) *Using One Step*: As mentioned in Section III, three operations are performed on the LR-HS image: super-resolution via LapSRN, guided image filtering, and detail extraction. In this section, we illustrate the effectiveness of the three operations from the perspective of objective experimental results. Let  $\mathbf{X}_{\text{sr}}$ ,  $\mathbf{X}_{\text{gf}}$ , and  $\mathbf{X}_{\text{de}}$  denote the HR-HS images generated by three operations in the DDLPS method: upsampling, guided image filtering, and detail extraction, respectively.  $\mathbf{X}_{\text{gf}}$  and  $\mathbf{X}_{\text{de}}$  are upsampled to the scale of the PAN image via bicubic interpolation. As shown in Table IX,  $\mathbf{X}_{\text{sr}}$  demonstrates the spectral preservation ability with its SAM values ranking second (only worse than  $\mathbf{X}_{\text{fin}}$ ) in all data sets. Thus, LapSRN can better preserve the spectral information than bicubic interpolation, and  $\mathbf{X}_{\text{fin}}$  has the best overall performance on the three data sets.

2) *Using Two Steps*: To comprehensively investigate the impact of each step on the final performance, we illustrate the effectiveness of the combinations using two steps in terms of experimental results. Let  $\mathbf{X}_{\text{srgf}}$ ,  $\mathbf{X}_{\text{srde}}$ , and  $\mathbf{X}_{\text{gfde}}$  denote the HR-HS images generated by removing the detail extraction, guided image filtering, and upsampling, respectively.  $\mathbf{X}_{\text{gfde}}$  is upsampled to the scale of the PAN image by bicubic

interpolation. As shown in Table X,  $\mathbf{X}_{\text{srgf}}$  and  $\mathbf{X}_{\text{srde}}$  show the best ability in both spatial and spectral preservations with respect to the quantitative indices.  $\mathbf{X}_{\text{fin}}$  has the best overall performance on the three data sets.

3) *Summary*: The visual appearance of pansharpening images generated by different combinations of different steps is shown in Fig. 15. LapSRN can simultaneously improve the spatial resolution and preserve the spectral information. The guided image filtering can transfer the structure of the PAN image to the HR-HS image [36] and make the fusion performance better on both spectral and spatial preservations [38]. The detail extraction can improve the fusion performance mainly in terms of visual appearance and quantitative indices. Although LapSRN makes the most contributions to the improvement of quantitative indices, its contributions to visual appearance are quite limited. Detail extraction and guided image filtering contribute less to the quantitative indices, but they are quite important in improving the visual appearance.

#### IV. CONCLUSION

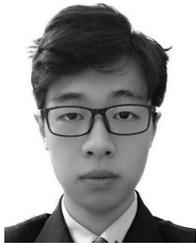
In this paper, we have proposed the DDLPS method, which outperforms the state-of-the-art and traditional methods in HS image pansharpening. Instead of using bicubic

interpolation to conduct the upsampling operation, DDLPS uses the LapSRN to conduct super-resolution on the LR-HS image. To preserve both spatial information and spectral information, we use a gain matrix to transfer the spectral details to spatial information and simultaneously consider the two aspects. Then, we turn the optimization problem into solving a Sylvester equation. Compared with traditional methods (CS, MRA, Bayesian, hybrid, and matrix factorization), the proposed method demonstrates the ability to preserve spatial and spectral information in both quantitative indices and visual appearance. The proposed DDLPS also shows universality on HS images with many spectral bands in comparison with the recent CNN-based methods.

## REFERENCES

- [1] 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.
- [2] A. Ertürk, M.-D. Iordache, and A. Plaza, "Sparse unmixing with dictionary pruning for hyperspectral change detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 1, pp. 321–330, Jan. 2017.
- [3] R. J. Ellis and P. W. Scott, "Evaluation of hyperspectral remote sensing as a means of environmental monitoring in the St. Austell China clay (kaolin) region, Cornwall, UK," *Remote Sens. Environ.*, vol. 93, no. 1, pp. 118–130, 2004.
- [4] T. A. Carrino, A. P. Crósta, C. L. B. Toledo, and A. M. Silva, "Hyperspectral remote sensing applied to mineral exploration in southern Peru: A multiple data integration approach in the Chapi Chiara gold prospect," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 64, pp. 287–300, Feb. 2018.
- [5] C. A. Bishop, J. G. Liu, and P. J. Mason, "Hyperspectral remote sensing for mineral exploration in Pulang, Yunnan Province, China," *Int. J. Remote Sens.*, vol. 32, no. 9, pp. 2409–2426, 2011.
- [6] S. Mahesh, D. S. Jayas, J. Paliwal, and N. White, "Hyperspectral imaging to classify and monitor quality of agricultural materials," *J. Stored Products Res.*, vol. 61, pp. 17–26, Mar. 2015.
- [7] C. Chion, J. A. Landry, and L. D. Costa, "A genetic-programming-based method for hyperspectral data information extraction: Agricultural applications," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2446–2457, Aug. 2008.
- [8] C. Souza, Jr., L. Firestone, L. M. Silva, and D. Roberts, "Mapping forest degradation in the Eastern Amazon from SPOT 4 through spectral mixture models," *Remote Sens. Environ.*, vol. 87, no. 4, pp. 494–506, Nov. 2003.
- [9] L. Loncan *et al.*, "Hyperspectral pansharpening: A review," *IEEE Trans. Geosci. Remote Sens.*, vol. 3, no. 3, pp. 27–46, Sep. 2015.
- [10] 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.
- [11] W. J. Carper, T. M. Lillesand, and R. W. Kiefer, "The use of intensity-hue-saturation transformations for merging SPOT panchromatic and multispectral image data," *Photogramm. Eng. Remote Sens.*, vol. 56, no. 4, pp. 459–467, Apr. 1990.
- [12] C. A. Laben, B. V. Brower, C. A. Laben, and B. V. Brower, "Process for enhancing the spatial resolution of multispectral imagery using pansharpening," U.S. Patent 6011 875 A, Jan. 4, 2000.
- [13] B. Aiazzi, S. Baronti, and M. Selva, "Improving component substitution pansharpening through multivariate regression of MS+pan data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3230–3239, Oct. 2007.
- [14] B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli, "Context-driven fusion of high spatial and spectral resolution images based on over-sampled multiresolution analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 10, pp. 2300–2312, Oct. 2002.
- [15] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "MTF-tailored multiscale fusion of high-resolution MS and PAN imagery," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 5, pp. 591–596, May 2006.
- [16] P. O. Hoyer, "Non-negative sparse coding," in *Proc. 12th IEEE Workshop Neural Netw. Signal Process.*, Sep. 2002, pp. 557–565.
- [17] R. Kawakami, Y. Matsushita, J. Wright, M. Ben-Ezra, Y.-W. Tai, and K. Ikeuchi, "High-resolution hyperspectral imaging via matrix factorization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 2329–2336.
- [18] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 528–537, Feb. 2012.
- [19] 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.
- [20] Q. Wei, N. Dobigeon, and J. 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.
- [21] 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.
- [22] 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.
- [23] P. Chavez, S. C. Sides, and J. A. Anderson, "Comparison of three different methods to merge multiresolution and multispectral data-Landsat TM and SPOT panchromatic," *Photogramm. Eng. Remote Sens.*, vol. 57, no. 3, pp. 295–303, 1991.
- [24] A. R. Gillespie, A. B. Kahle, and R. E. Walker, "Color enhancement of highly correlated images. II. Channel ratio and 'chromaticity' transformation techniques," *Remote Sens. Environ.*, vol. 22, no. 3, pp. 343–365, 1987.
- [25] P. S. Chavez, Jr., and A. Y. Kwarteng, "Extracting spectral contrast in Landsat thematic mapper image data using selective principal component analysis," *Photogramm. Eng. Remote Sens.*, vol. 55, no. 3, pp. 339–348, 1989.
- [26] J. Choi, K. Yu, and Y. Kim, "A new adaptive component-substitution-based satellite image fusion by using partial replacement," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 295–309, Jan. 2011.
- [27] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.
- [28] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. COM-31, no. 4, pp. 532–540, Apr. 1983.
- [29] G. Vivone, R. Restaino, M. Dalla Mura, G. Licciardi, and J. Chanussot, "Contrast and error-based fusion schemes for multispectral image pansharpening," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 5, pp. 930–934, May 2014.
- [30] 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. 2015.
- [31] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian Denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.
- [32] M. Haris, G. Shakhnarovich, and N. Ukita, "Deep back-projection networks for super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2018, pp. 1664–1673.
- [33] W. Huang, L. Xiao, Z. Wei, H. Liu, and S. Tang, "A new pan-sharpening method with deep neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 5, pp. 1037–1041, May 2015.
- [34] G. Masi, D. Cozzolino, L. Verdoliva, and G. Scarpa, "Pansharpening by convolutional neural networks," *Remote Sens.*, vol. 8, no. 7, p. 594, Jul. 2016.
- [35] 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. 1753–1761.
- [36] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1397–1409, Jun. 2013.
- [37] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Deep Laplacian pyramid networks for fast and accurate super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 2, no. 3, Jul. 2017, pp. 5835–5843.
- [38] L. Wenzhi *et al.*, "Processing of multiresolution thermal hyperspectral and digital color data: Outcome of the 2014 IEEE GRSS data fusion contest," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2984–2996, Jun. 2015.

- [39] S. M. Pizer *et al.*, "Adaptive histogram equalization and its variations," *Comput. Vis., Graph., Image Process.*, vol. 39, no. 3, pp. 355–368, 1987.
- [40] J. Qu, J. Lei, Y. Li, W. Dong, Z. Zeng, and D. Chen, "Structure tensor-based algorithm for hyperspectral and panchromatic images fusion," *Remote Sens.*, vol. 10, no. 3, p. 373, 2018.
- [41] R. H. Bartels and G. W. Stewart, "Solution of the matrix equation  $AX + XB = C$  [F4]," *Commun. ACM*, vol. 15, no. 9, pp. 820–826, 1972.
- [42] Y. Li, W. Xie, and H. Li, "Hyperspectral image reconstruction by deep convolutional neural network for classification," *Pattern Recognit.*, vol. 63, pp. 371–383, Mar. 2017.
- [43] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.
- [44] P. Arbeláez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, May 2011.
- [45] J. G. Liu, "Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3461–3472, Nov. 2000.
- [46] 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.
- [47] L. Zhang, L. Zhang, D. Tao, and X. Huang, "On combining multiple features for hyperspectral remote sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 879–893, Mar. 2012.



**Kaiyan Li** (S'18) is currently pursuing the B.E. degree with the School of Telecommunications Engineering, Xidian University, Xi'an, China.

Since 2017, he has been a Research Assistant with the State Key Laboratory of Integrated Service Network, Image Coding and Processing Center, Xidian University. His research interests include hyperspectral image processing, biomedical signal/image analysis, and deep learning.



**Weiying Xie** (M'18) received the B.S. degree in electronic information science and technology from the University of Jinan, Jinan, China, in 2011, the M.S. degree in communication and information systems from Lanzhou University, Lanzhou, China, in 2014, and the Ph.D. degree in communication and information systems from Xidian University, Xi'an, China, in 2017.

She is currently a Post-Doctoral Researcher and a Teacher with Xidian University. Her research interests include neural networks, deep learning, hyperspectral image processing, and high-performance computing.



**Qian Du** (S'98–M'00–SM'05–F'18) received the Ph.D. degree in electrical engineering from the University of Maryland at Baltimore County, Baltimore, MD, USA, in 2000.

She is currently a Bobby Shackouls Professor with the Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS, USA. Her research interests include hyperspectral remote sensing image analysis and applications, pattern classification, data compression, and neural networks.

Dr. Du is a fellow of the SPIE-International Society for Optics and Photonics. She was a recipient of the 2010 Best Reviewer Award from the IEEE Geoscience and Remote Sensing Society (GRSS). She was the Co-Chair of the Data Fusion Technical Committee of the IEEE GRSS from 2009 to 2013. She was the Chair of the Remote Sensing and Mapping Technical Committee of the International Association for Pattern Recognition from 2010 to 2014. She was the General Chair of the fourth IEEE GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing held at Shanghai, China, in 2012. She served as an Associate Editor for the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (JSTARS), the *Journal of Applied Remote Sensing*, and the IEEE SIGNAL PROCESSING LETTERS. Since 2016, she has been the Editor-in-Chief of the IEEE JSTARS.



**Yunsong Li** received the M.S. degree in telecommunication and information systems and the Ph.D. degree in signal and information processing from Xidian University, Xi'an, China, in 1999 and 2002, respectively.

In 1999, he joined the School of Telecommunications Engineering, Xidian University, where he is currently a Professor. He is also the Director of the State Key Laboratory of Integrated Service Networks, Image Coding and Processing Center, Xidian University. His research interests include image and video processing and high-performance computing.