Pansharpening for Cloud-Contaminated Very High-Resolution Remote Sensing Images

Xiangchao Meng, *Member, IEEE*, Huanfeng Shen, *Senior Member, IEEE*, Qiangqiang Yuan, *Member, IEEE*, Huifang Li, *Member, IEEE*, Liangpei Zhang, *Senior Member, IEEE*, and Weiwei Sun, *Member, IEEE*

***Abstract*— The optical remote sensing images not only have to make a fundamental tradeoff between the spatial and spectral resolutions, but also are inevitable to be polluted by the clouds; however, the existing pansharpening methods mainly focus on the resolution enhancement of the optical remote sensing images without cloud contamination. How to fuse the cloud-contaminated images to achieve the joint resolution enhancement and cloud removal is a promising and challenging work. In this paper, a pansharpening method for the challenging cloud-contaminated very high-resolution remote sensing images is proposed. Furthermore, the cloud-contaminated conditions for the practical observations with all the thick clouds, the thin clouds, the haze, and the cloud shadows are comprehensively considered. In the proposed methods, a two-step fusion framework based on multisource and multitemporal observations is presented: 1) the thin clouds, the haze, and the light cloud shadows are proposed to be first jointly removed and 2) a variational-based integrated fusion model is then proposed to achieve the joint resolution enhancement and missing information reconstruction for the thick clouds and dark cloud shadows. Through the proposed fusion method, a promising cloud-free fused image with both high spatial and high spectral resolutions can be obtained. To comprehensively test and verify the proposed method, the experiments were implemented based on both the cloud-free and cloud-contaminated images, and a number of different remote sensing satellites including the IKONOS, the QuickBird, the Jilin (JL)-1, and the Deimos-2 images were utilized. The experimental results confirm the effectiveness of the proposed method.**

***Index Terms*— Pansharpening, cloud contamination, integrated model, image fusion, remote sensing.**

I. Introduction

DUE to the physical constraints of satellite sensors and other factors [1], [2], such as the incoming energy to the sensors, the on-board storage, and the data transmission, the remote sensing images have to feature a fundamental tradeoff between the spatial and spectral resolutions. Accordingly, more than 70% of the optical earth observation satellites provide the bundled high spatial resolution (HR) panchromatic (PAN) image with a single-band and the low spatial resolution (LR) multispectral (MS) image with relatively higher spectral resolution [3].

PAN/MS fusion, typically referred to as pansharpening, was proposed to integrate the complementary spatial and spectral resolutions of the HR PAN and LR MS images to obtain an HR MS image [2]. The pansharpening methods originated in the 1980s [1], [4], [5]. Since 1986, the Système Pour l' Observation de la Terre-1 system has provided two LR MS images together with one HR PAN image, the pansharpening methods have got a rapid development over a period of 30 years. To date, there have proposed a large number of pansharpening methods. In general, most of them can be classified into three major categories [6]: 1) the component substitution (CS)-based methods; 2) the multiresolution analysis (MRA)-based methods; and 3) the variational optimization (VO)-based methods. In addition, the deep learning (DL)-based pansharpening methods have been proposed, and they are attracting increasing attention.

Among different categories of pansharpening methods, the CS-based and MRA-based methods are the most popular due to their simplicity and high efficiency.

1) *CS-Based Pansharpening Methods*: They are based on the substitution of an LR component of the MS image by the HR PAN image to obtain the fused HR MS image, and the LR component is generally obtained by spectral transformation of the MS image. On the whole, the development process of CS-based methods includes two stages: 1) the traditional understanding [7], [8] characterized by the process of the "forward transformation—CS—inverse transformation" and 2) the general understanding with a unified framework [9]–[11], which makes a major contribution to the improvement and

Manuscript received July 1, 2018; revised September 28, 2018; accepted October 11, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 41801252 and Grant 61671334, and in part by the K. C. Wong Magna Fund in Ningbo University. *(Corresponding author: Weiwei Sun.)*

X. Meng is with the Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China (e-mail: mengxiangchao@nbu.edu.cn).

H. Shen and H. Li are with the School of Resources and Environmental Sciences, Wuhan University, Wuhan 430079, China (e-mail: shenhf@whu.edu.cn; huifangli@whu.edu.cn).

Q. Yuan is with the School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China (e-mail: yqiang86@gmail.com).

L. Zhang is with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan 430079, China, and also with the Collaborative Innovation Center for Geospatial Technology, Wuhan University, Wuhan 430079, China (e-mail: zlp62@whu.edu.cn).

W. Sun is with the Department of Geography and Spatial Information Techniques, Ningbo University, Ningbo 315211, China (e-mail: sunweiwei@nbu.edu.cn).

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

Digital Object Identifier 10.1109/TGRS.2018.2878007

development of this kind of methods. The CS-based methods generally produce fused image with sharp spatial structures, and they are also robust to the MS-to-PAN misalignment. However, the traditional CS-based methods, such as intensity–hue–saturation (IHS) method [7] and principal component analysis method [7], generally produce serious spectral distortions [12], [13]. But, it is noteworthy that this is not true for all the CS-based methods, especially for some improved CS-based methods [6], [14], [15].

2) *MRA-Based Pansharpening Methods*: This category of pansharpening methods is based on the injection of the high frequency, extracted by the spatial decomposition on the HR PAN image, into the LR MS image to obtain the fused HR MS image. Similar to the CS-based methods, the MRA-based pansharpening methods have experienced a development process from the traditional understanding to the general understanding. The traditional understanding is strictly based on MRA algorithms, such as the wavelet transform [16] and the Laplacian pyramid [17]. This is characterized by a complex process with "multiresolution decomposition—fusion—image reconstruction." Then, the MRA-based methods are simplified and extended to a general fusion framework [10], [11], [18], which has promoted the emerge in large numbers of improved methods [13], [19], [20]. In contrast to the CS-based methods, the MRA-based pansharpening methods are relatively sensitive to the MS-to-PAN misalignment, whereas they are more robust to the temporal and instrumental misalignment between the PAN and MS observations [21].
3) *VO-Based Pansharpening Methods*: The VO-based methods are based on the variational theory, which features the optimal solution of an energy functional. The representative methods include the popular model-based methods [22], [23] and the sparse-based methods [24], [25]. For the model-based pansharpening methods, they regard the fusion process as an ill-posed inverse problem, based on the observation model by considering the imaging process of the observations [26], [27]. For the sparse-based pansharpening methods, the first work was proposed by Li and Yang [24]. They are characterized by the dictionary learning, which includes the early offline learning methods based on the external database [24], [28] and the mainstream online learning methods based on the observations to be fused [25], [29].
4) *DL-Based Pansharpening Methods*: In addition to the above three categories of pansharpening methods, the DL-based pansharpening methods [30] have been attracting ever-increasing attention. They are based on the DL theory [31]–[33]. Such as Masi *et al.* [34] and Zhong *et al.* [35] proposed pansharpening methods based on convolutional neural networks, respectively. Yuan *et al.* [36] proposed the multiscale and multidepth convolutional neural network for pansharpening of remote sensing images. Though the existing DL-based pansharpening works have shown good performance, whether the DL could better solve the problems for pansharpening should be further fully tested and verified.

Although a large number of pansharpening methods have been proposed, most of them do not consider the possible cloud contamination [37], [38] in the optical remote sensing observations. However, it is noteworthy that at any one time, approximately 35% of the global land surface is covered by clouds [38], [39]. In addition, for the actual optical remote sensing observations, especially for the very high-resolution (VHR) satellite images, multiple cloud-contaminated factors, including the thick clouds, the thin clouds, the haze, and the cloud shadows, are generally simultaneously existed. Therefore, for the cloud-contaminated VHR satellite PAN and MS images, how to effectively achieve the joint resolution enhancement and cloud removal to obtain the cloud-free HR MS image is a promising and challenging work. To the best of our knowledge, few papers [40] have addressed this problem, especially for the challenging VHR remote sensing scenes with multiple cloud-contaminated factors simultaneously, including the thick clouds, the thin clouds, and the cloud shadows.

Based on the above problems, we proposed a pansharpening method for the cloud-contaminated remote sensing observations. The proposed method is based on multisource and multitemporal PAN and MS observations, and multiple cloud-contaminated factors including the thick clouds, the thin clouds, the haze, and the cloud shadows are comprehensively considered. Through the proposed method, the cloud-free image with both high spatial and high spectral resolutions will be obtained. Two contributions are made in this paper.

1) This is the first work to propose the pansharpening method for the images with challenging complex cloud-contaminated conditions, including all the thick clouds, the thin clouds, and the cloud shadows.
2) A variational-based integrated fusion model was proposed to achieve the joint resolution enhancement and information reconstruction.

The rest of this paper is organized as follows. Section II presents the description of the proposed fusion framework. Section III presents the variational-based integrated fusion model in detail. The experimental results and analysis are presented in Section IV, and Section V shows the conclusion.

II. Proposed Fusion Framework

In this paper, a pansharpening method for the cloud-contaminated VHR remote sensing images is proposed. The proposed fusion framework is based on multisource and multitemporal LR MS and HR PAN observations, as shown in Fig. 1. The target LR MS and HR PAN images are both contaminated by clouds, and the cloud-free auxiliary LR MS and HR PAN images, which should be performed geometrical registration and radiometric normalization, are applied to provide the complementary information. It should be noted that the auxiliary images with approaching phase should be chosen as far as possible to ensure the consistent ground features to the target images. The aim of the proposed method is to obtain the desired cloud-free HR MS image at the target time. To achieve

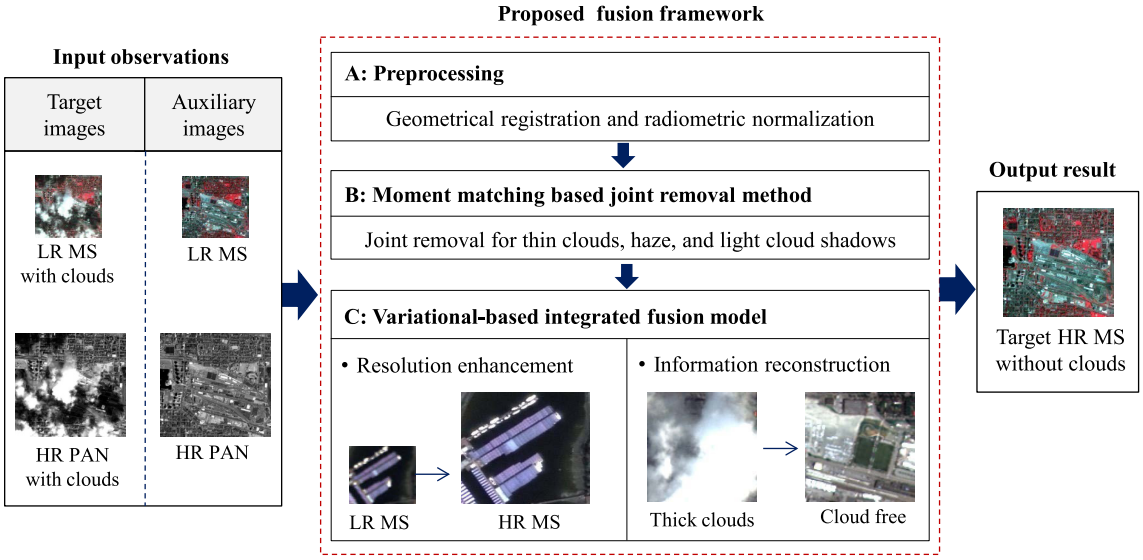


Fig. 1. Schematic of the proposed fusion framework.

this aim, two challenging problems should be solved: 1) how to construct an effective fusion model to achieve the joint resolution enhancement and cloud removal to avoid the error accumulation of the step-by-step fusion and 2) how to process the optical remote sensing scenes with the great challenging complex cloud degradation conditions, such as the thick clouds, the thin clouds, the haze, and the cloud shadows existed simultaneously. To solve the challenges, a two-step integrated fusion strategy is proposed, by comprehensively considering the relations and difference among the multiple degradation factors. On the one hand, in consideration of the common semitransparent visibility of ground features for the thin clouds, the haze, and the cloud shadows, an integrated removal strategy based on moment matching is applied to perform the joint information restoration for them. On the other hand, a variational-based integrated fusion model is further proposed, and the resolution enhancement and the missing information reconstruction for the thick clouds are jointly achieved based on the proposed integrated variational fusion model. Through the proposed fusion method, the cloud-free image with both high spatial and high spectral resolutions can be obtained.

As aforementioned, the proposed fusion method can achieve the joint resolution enhancement and cloud removal to obtain the cloud-free HR MS image. Furthermore, it can be applied to the fusion of the challenging cloud-contaminated optical remote sensing observations with multiple complex cloud degradation conditions, including the thick clouds, the thin clouds, the haze, and the cloud shadows.

A. Preprocessing

The proposed method involves multisource and multitemporal images, and the preprocessing including the geometrical registration and radiometric normalization should be performed first. For the geometrical registration, the processing by professional software of the environment for visualizing images can meet the requirements for fusion. For the radiometric normalization, a two-step radiometric normalization method for the auxiliary images is proposed, as shown in Fig. 2.

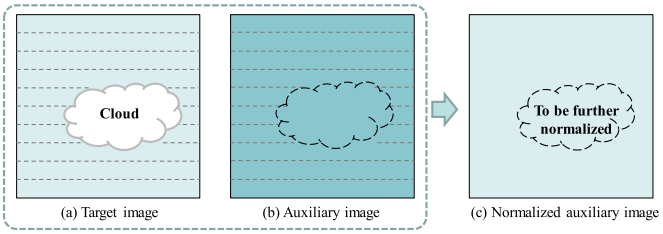


Fig. 2. Schematic of the radiometric normalization.

The radiometric normalization includes two steps.

1) The auxiliary MS and PAN images are normalized to the target MS and PAN images, respectively, based on the relationships calculated by the common cloud-free regions as shown in the dashed areas of Fig. 2.
2) In consideration of the possible slight radiometric difference between the target and auxiliary images after the first normalization, and also considering that the auxiliary images mainly provide useful information for the cloud regions of the target images.

We propose to further normalize the local regions of the auxiliary images corresponding to the cloud regions of the target image, as shown in Fig. 2, based on the Poisson equation. Following the proposed two-step radiometric normalization processing, the radiometric difference between the target and auxiliary images is minimized.

1) *The First-Step Radiometric Normalization:* It is assumed that the relationship between multitemporal scenes can be expressed by a linear model [6], [38], [41]. For convenience, the target image is denoted by $\mathbf{O}_0$, and the auxiliary image is denoted by $\mathbf{O}_1$. It is represented as

$$\mathbf{O}_0^{\Omega} = c_1 \mathbf{O}_1^{\Omega} + c_2 \tag{1}$$

where Ω denotes the common cloud-free regions of the target and auxiliary images, and c_1, c_2 are the relationship coefficients, which are solved based on the robust regression analysis using iteratively reweighted least squares [42]. Then, the normalized auxiliary image is obtained by

$$\tilde{\mathbf{O}}_1 = c_1 \mathbf{O}_1 + c_2 \tag{2}$$

where $\tilde{\mathbf{O}}_1$ denotes the normalized auxiliary image.

2) *The Second-Step Radiometric Normalization:* The Poisson editing [43] has been successfully applied to a variety of applications, such as the seamless interpolation, the local illumination change, and color transfer. In this paper, the Poisson editing is applied for the local radiometric normalization. It is represented as

$$\hat{\mathbf{O}}_1^{\Omega-} = \tilde{\mathbf{O}}_1^{\Omega-} + \mathbf{g} \tag{3}$$

where $\Omega-$ represents the regions, corresponding to the cloud regions of the target image, to be further normalized, and $\hat{\mathbf{O}}_1^{\Omega-}$ is the final normalized result. $\mathbf{g}$ represents the possible slight radiation difference between the target image and the first-step normalized auxiliary image. It is solved by

$$\Delta \mathbf{g} = 0 \text{ over } \Omega-, \quad \mathbf{g}|_{\partial\Omega-} = (\mathbf{O}_0 - \tilde{\mathbf{O}}_1)|_{\partial\Omega-} \tag{4}$$

where $\partial\Omega-$ denotes the boundary of the local regions.

B. Moment Matching-Based Joint Removal

How to process the optical remote sensing observations with multiple cloud degradation conditions, such as the thick clouds, thin clouds, the haze, and the cloud shadows existed simultaneously, is a challenging work. On the one hand, to obtain the cloud-free image with both high spatial and high spectral resolutions, all the information restoration for the thin clouds, the haze, the light cloud shadows, the missing information reconstruction for the thick clouds, and the resolution enhancement should be achieved. On the other hand, error accumulation during the above multiple complex processing tasks in pansharpening of the cloud-contaminated images should be minimized.

For the removal of the thin cloud, the haze, and the light cloud shadows, the sliding window-based local moment matching method [44] is introduced by comprehensively considering the possible ground feature changes between the target image and the auxiliary image. This can preserve the spatial structures of the target images as much as possible, by only adjusting the local mean value and the variance. The schematic of the sliding window-based local moment matching method is shown in Fig. 3.

As shown in Fig. 3, the normalized cloud-free auxiliary image is regarded as the reference, and the processing is performed pixel by pixel by making use of the contextual information. It is represented as follows:

$$\tilde{\mathbf{O}}_0 = \frac{\left(\hat{\mathbf{O}}_1^{\rho}\right)_{\text{std}}}{\left(\mathbf{O}_0^{\rho}\right)_{\text{std}}} \times \left(\mathbf{O}_0^{\rho} - \left(\mathbf{O}_0^{\rho}\right)_{\text{mean}}\right) + \left(\hat{\mathbf{O}}_1^{\rho}\right)_{\text{mean}} \tag{5}$$

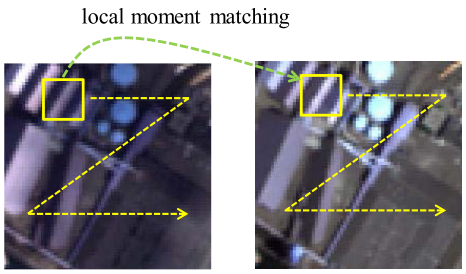


Fig. 3. Schematic of sliding window-based local moment matching for haze and light cloud shadow removal. (a) Target image with cloud shadows. (b) Auxiliary image.

where $\tilde{\mathbf{O}}_0$ denotes the result, and the superscript ρ denotes the patch extraction. The subscript std denotes the standard deviation, and subscript mean represents the mean value. Through the sliding window-based local moment matching method, the thin cloud, the haze, and the light cloud shadows of the target images can be jointly removed.

C. Variational-Based Integrated Fusion Model

After the thin clouds, the haze, and the light cloud shadows have been removed, the target MS and PAN images are still polluted by the thick clouds and the dark cloud shadows, and the information in these regions can be assumed to be completely missing. Furthermore, for the VHR satellite images, the cloud-contaminated regions are generally large. Therefore, how to achieve the joint missing information reconstruction and the resolution enhancement to obtain the desired cloud-free image with both high spatial and high spectral resolutions is challenging. In this paper, a variational-based integrated fusion model is proposed. The proposed model is based on the multisource and multitemporal images, and it can achieve the joint resolution enhancement and the missing information reconstruction to obtain the desired cloud-free HR MS image.

III. Variational-Based Integrated Fusion Model

Let $\mathbf{Y} = \{\mathbf{y}_p, \mathbf{y}_q\}$ denotes the group of the target and auxiliary LR MS images, where $\mathbf{y}_p$ denotes the target image with thin cloud and light cloud shadow being removed, and $\mathbf{y}_q$ denotes the normalized auxiliary image. $\mathbf{Z} = \{\mathbf{z}_p, \mathbf{z}_q\}$ represents the group of the target and auxiliary HR PAN images, where $\mathbf{z}_p$ represents the target HR PAN image with thin cloud and light cloud shadow being removed, and $\mathbf{z}_q$ represents the normalized auxiliary HR PAN image. The desired fused image is denoted by $\mathbf{x}$.

A. Spectral Fidelity Term for the Integrated Model

In this paper, a 3-D spectral fidelity term is proposed. The spectral fidelity term relates the desired fused image $\mathbf{x}$ and the LR MS images $\mathbf{Y}$. It is based on the observation model of the LR MS image [1], [23], [45]. For convenience, let $\mathbf{y}$ denotes a general LR MS image. The observation model is represented as

$$\mathbf{y} = \mathbf{D}\,\mathbf{S}\mathbf{x} + \mathbf{v} \tag{6}$$

where $\mathbf{S}$ represents the blurring matrix, $\mathbf{D}$ denotes the downsampling matrix, and $\mathbf{v}$ denotes the additive noise. Hence, the traditional spectral fidelity term [23], [46] is generally represented as

$$E_{\text{spe}} = \sum_{b=1}^{B} \|\mathbf{y}_b - \mathbf{D}\mathbf{S}_b\mathbf{x}_b\|_2^2 \tag{7}$$

where b denotes the band number, and B denotes the total number of spectral bands.

It can be seen that the traditional spectral fidelity term only builds the top–down relationship for each band of the LR MS image and the desired fused image. However, it does not comprehensively consider the interband spectral correlation among the MS bands. For the pansharpening of the cloud-contaminated images, the spectral fidelity is more challenging. Therefore, to further improve the ability of the spectral fidelity, a 3-D spectral fidelity term is proposed, represented as

$$E_{\text{3Dspe}} = \sum_{b=1}^{B}\sum_{k=0}^{B} \|(\mathbf{y}_b - \mathbf{y}_k) - \mathbf{D}\,(\mathbf{S}_b\mathbf{x}_b - \mathbf{S}_k\mathbf{x}_k)\|_2^2 \tag{8}$$

where k denotes the kth band, and it should be noted that $\mathbf{y}_k, \mathbf{x}_k, \mathbf{S}_k$ will be zero matrix with $k = 0$. Hence, the proposed spectral fidelity term for the integrated fusion model based on target and auxiliary images is represented as follows:

$$E'_{\text{3Dspe}} = \sum_{b=1}^{B}\sum_{k=0}^{B} \left\|\left(\mathbf{Y}'_b - \mathbf{Y}'_k\right) - \mathbf{MD}\,(\mathbf{S}_b\mathbf{x}_b - \mathbf{S}_k\mathbf{x}_k)\right\|_2^2 \tag{9}$$

where $\mathbf{M} = [\mathbf{M}_p, \mathbf{M}_q]^T$ denotes the image masks; here, $\mathbf{M}_p$ is the mask for the target image, and the pixel value of $\mathbf{M}_p$ is 0 for the missing information regions, and it is 1 for the cloud-free regions. $\mathbf{M}_q$, in which zero value for the cloud-free regions and one value for the missing information regions, is the mask for the auxiliary image. $\mathbf{Y}'_b = [\mathbf{M}_p\mathbf{y}_{p,b}, \mathbf{M}_q\mathbf{y}_{q,b}]^T$ and $\mathbf{Y}'_k = [\mathbf{M}_p\mathbf{y}_{p,k}, \mathbf{M}_q\mathbf{y}_{q,k}]^T$ denote the group of the corresponding target and auxiliary LR MS images.

B. Spatial Enhancement Term for the Integrated Model

The spatial enhancement term relates the desired fused image $\mathbf{x}$ and the HR PAN image $\mathbf{Z}$. It is based on spatial structures other than the gray value [22], [26], and this will weaken the negative influence of low frequency of the HR PAN images, especially for the auxiliary HR PAN image. It is assumed that the desired HR MS image has similar spatial structures to the HR PAN image [22], [47], represented as $\nabla\mathbf{z} \approx \nabla\mathbf{x}_b$, where $\mathbf{z}$ denotes an HR PAN image, $\nabla = [\nabla_{\text{H}}^T, \nabla_{\text{V}}^T]^T$ denotes the gradient operators, and $\nabla_{\text{H}} = [-1,\ 1]$, $\nabla_{\text{V}} = [-1,\ 1]^T$ are the horizontal and vertical gradient operators, respectively. In this paper, moment matching of the gradient features is introduced to further enhance the robustness. The spatial enhancement term for each spectral band is represented as

$$E_{\text{spa}}^b = \|\nabla\mathbf{Z}' - f(\hat{\mathbf{M}}\nabla\mathbf{x}_b)\|_2^2 \tag{10}$$

where $f(\cdot)$ denotes the moment matching of the gradient features. $\hat{\mathbf{M}} = [\hat{\mathbf{M}}_p, \hat{\mathbf{M}}_q]^T$ represents the image masks, corresponding to $\mathbf{M}$ in (9), at the spatial scale of the HR PAN images, and $\mathbf{Z}' = [\hat{\mathbf{M}}_p\mathbf{z}_p, \hat{\mathbf{M}}_q\mathbf{z}_q]^T$ denotes the group of the HR PAN images.

C. Total Energy Functional

Based on the proposed spectral fidelity term and the spatial enhancement term, the total energy functional is finally obtained. It is represented as follows:

$$\begin{aligned} E(\mathbf{x}) = {} & \lambda_1 \sum_{b=1}^{B}\sum_{k=0}^{B} \|(\mathbf{Y}'_b - \mathbf{Y}'_k) - \mathbf{MD}\,(\mathbf{S}_b\mathbf{x}_b - \mathbf{S}_k\mathbf{x}_k)\|_2^2 \\ & + \sum_{b=1}^{B} w_b\|\nabla\mathbf{Z}' - f(\hat{\mathbf{M}}\nabla\mathbf{x}_b)\|_2^2 + \lambda_2 \sum_{b=1}^{B} \|\mathbf{W}_b\Delta\mathbf{x}_b\|_2^2 \end{aligned} \tag{11}$$

where the first term is the spectral fidelity term, the second term is the spatial enhancement term, and the third term is the prior term. In this paper, the spatial weighted Laplacian prior is developed based on [1] and [48]. Δ denotes the Laplacian operator, and $\mathbf{W}_b$ denotes the adaptive weight for each pixel location (i, j) calculated as $1/(1+\text{abs}(\nabla\mathbf{x}_b(i, j)))$. λ_1 and λ_2 denote the model parameters to control the relative contribution of the three terms. Furthermore, the weight w_b in the spatial enhancement term is developed by considering the correlation among the spectral bands to adaptively adjust the spatial enhancement for different spectral bands. It is represented as

$$w_b = \frac{\text{cov}(\mathbf{M}_p\mathbf{y}_{p,b}, \mathbf{M}_p\mathbf{I})/\text{var}(\mathbf{M}_p\mathbf{I})}{\sum_{b=1}^{B}\text{cov}(\mathbf{M}_p\mathbf{y}_{p,b}, \mathbf{M}_p\mathbf{I})/\text{var}(\mathbf{M}_p\mathbf{I})} \tag{12}$$

where $\mathbf{I}$ denotes the intensity calculated by the generalized IHS transformation [10] of the spectral bands of MS image.

D. Optimization Procedure

The fused image is solved by the classical gradient descent method, and the fused image is obtained by the iterative optimization solution. First, differentiating the energy function (11) with respect to $\mathbf{x}_b$, it is represented as

$$\begin{aligned} & \nabla E(\mathbf{x}_b) \\ & = -\lambda_1 \sum_{k=0}^{B}\sum_{t=p,q} \mathbf{S}_b^T\mathbf{D}^T\mathbf{M}_t^T[\mathbf{M}_t(\mathbf{y}_{t,b} - \mathbf{y}_{t,k}) - \mathbf{M}_t\mathbf{D}(\mathbf{S}_b\mathbf{x}_b - \mathbf{S}_k\mathbf{x}_k)] \\ & \quad - w_b \sum_{t=p,q} \hat{\mathbf{M}}_t^T[\hat{\mathbf{M}}_t\Delta\mathbf{z}_t - f(\hat{\mathbf{M}}_t\Delta\mathbf{x}_b)] + \lambda_2\mathbf{W}_b^T\mathbf{W}_b\Delta^T\Delta\mathbf{x}_b. \end{aligned} \tag{13}$$

The desired fused image can then be solved by successive approximation iterations, and it is represented as

$$\mathbf{x}_{b,n+1} = \mathbf{x}_{b,n} - \tau_{b,n}\nabla E(\mathbf{x}_{b,n}) \tag{14}$$

where n denotes the iteration number, $\mathbf{x}_{b,n}$ and $\mathbf{x}_{b,n+1}$ denote the results of the nth iteration and $(n+1)$th iteration, respectively, and $\tau_{b,n}$ denotes the step size of the nth iteration,

TABLE I
DATA SETS IN THE EXPERIMENTS

Experimental settings		Satellites	Spatial and spectral dimensions	Region
Cloud-free image fusion experiments	Simulated experiments	IKONOS	PAN: spatial size 256×256 MS: spatial size 64×64 with 4bands	Wuhan city, Hubei province, China
	Real experiments	QuickBird	PAN: spatial size 400×400 MS: spatial size 100×100 with 4 bands	Wuhan university in Hubei province, China.
Cloud-contaminated image fusion experiments	Simulated experiments	Jinlin(JL)-1	PAN: spatial size 600×600 MS: spatial size 150×100 with 3 bands	Kansas city, Missouri, America
	Real experiments	Deimos-2	PAN: spatial size 1600×1600 MS: spatial size 400×400 with 4 bands	Vancouver, Canada

represented as

$$\tau_{b,n} = \frac{[\nabla E(\mathbf{x}_{b,n})]^T \nabla E(\mathbf{x}_{b,n})}{[\nabla E(\mathbf{x}_{b,n})]^T \left[\lambda_1 \sum_{k=0}^{B} \sum_{t=p,q} \mathbf{S}_b^T \mathbf{D}^T \mathbf{M}_t^T \mathbf{M}_t \mathbf{D} \mathbf{S}_b\right) + w_b \sum_{t=p,q} \hat{\mathbf{M}}_t^T f(\hat{\mathbf{M}}_t \Delta) + \lambda_2 \mathbf{W}_b^T \mathbf{W}_b \Delta^T \Delta] \nabla E(\mathbf{x}_{b,n}). \quad (15)$$

The fused image is updated in each iteration, and it is terminated when

$$\|\mathbf{x}_{n+1} - \mathbf{x}_n\|^2 / \|\mathbf{x}_n\|^2 \leq d \quad (16)$$

where d is the threshold value.

IV. EXPERIMENTS

A. Data Sets and Experimental Settings

To comprehensively test and verify the performance of the proposed method, the experiments were performed based on both the cloud-free and cloud-contaminated data. Furthermore, the performance of the proposed method was evaluated on both the simulated and real experiments from the qualitative and quantitative aspects. For cloud-free image fusion experiments, the simulated experimental data were obtained by the spatial degradation of the original PAN and MS images based on Wald's protocol [49], and the original MS image was used as the reference image for evaluation. For cloud-contaminated image fusion experiments, they are implemented based on the target and auxiliary images. The simulated data were acquired by low-pass filtering of the original cloud-free target and auxiliary images, using Gaussian modulation transfer function blurring and then downsampling by a factor of the spatial resolution ratio between the HR PAN image and the LR MS image. Then, the missing information of the thick cloud corruption was simulated for the target MS and PAN images, and the original cloud-free target HR MS image was used as the reference for the evaluation. In addition, in the fusion experiments, a number of different remote sensing satellites including the IKONOS, QuickBird, Jilin(JL)-1, and Deimos-2 images were employed, as shown in Table I.

In addition, a number of state-of-the-art pansharpening methods were utilized for comparison. On the one hand, for the cloud-free image fusion experiments, the popular Gram–Schimidt (GS) pansharpening method [8], the partial replacement adaptive component substitution (PRACS) [14] pansharpening method, the additive wavelet luminance proportional (AWLP) method [13], the modulation transfer function (MTF)-generalized Laplacian pyramid (GLP) method [19], and the adjustable model-based data fusion (AMDF) pansharpening methods proposed in [23] were compared. On the other hand, for the cloud-contaminated image fusion experiments, in consideration that the existing pansharpening methods cannot obtain the desired cloud-free fused HR MS image, the proposed integrated fusion model was compared with the step-by-step processing way, i.e., the thick clouds of the target MS and PAN images were first removed by using the state-of-the-art missing information reconstruction methods; then, the cloud-free LR MS and HR PAN were fused based on the proposed variational fusion model. In addition, to ensure the objectivity and sufficiency, two popular missing information reconstruction methods of weighted linear regression (WLR) method [50], which can be downloaded from the website (http://sendimage.whu.edu.cn/send-resource-download/), and the spatio-temporal Markov random fields (STMRF) method [51] are both applied. In the experiments, unless otherwise specified, the model parameter λ_1 was set to 20, and λ_2 was set to 0.1.

B. Experimental Results and Analysis

The experiments were performed based on both the cloud-free and cloud-contaminated images, including: 1) pansharpening experiments for cloud-free images and 2) pansharpening experiments for cloud-contaminated images.

1) Pansharpening for Cloud-Free Images:

a) Simulated experiment: The simulated experiment was performed based on IKONOS images acquired over the Wuhan, Hubei, China. The IKONOS is the first commercial operational satellite of the world; it provides PAN image

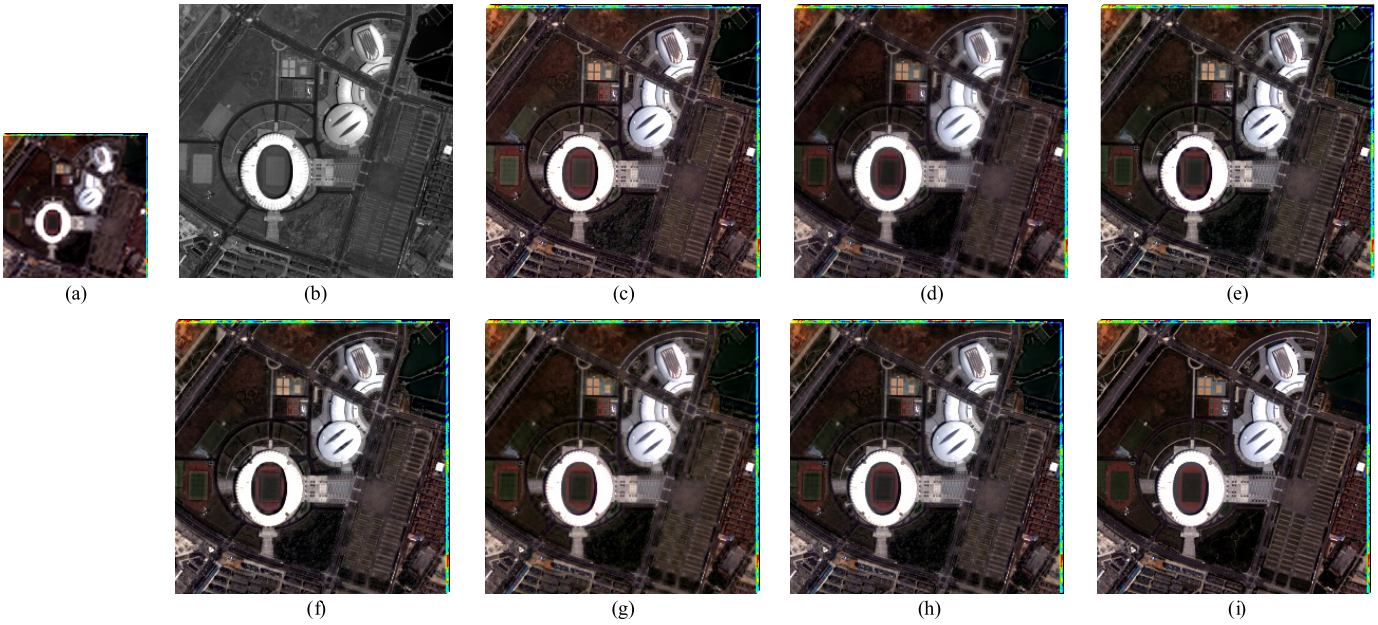


Fig. 4. Fusion results of the simulated experiment for cloud-free images. (a) LR MS image. (b) HR PAN image. (c) GS fusion result. (d) PRACS fusion result. (e) AWLP fusion result. (f) MTF-GLP fusion result. (g) AMDF fusion result. (h) Proposed fusion result. (i) Original MS image.

with a spatial resolution of 1 m and MS image [blue, green, red, and near infrared response (NIR) spectral bands] with a spatial resolution of 4 m. In the experiment, the simulated LR MS, shown in Fig. 4(a), and HR PAN, shown in Fig. 4(b), are 4 and 16 m, respectively, and the original MS image with a spatial resolution of 4 m, as shown in Fig. 4(i), was regarded as the reference for the quantitative evaluation. In the experiments, five widely used quality indices in pansharpening were applied for the quantitative evaluation. They are the correlation coefficient [23], the peak signal-to-noise ratio [52], the dimensionless global error in synthesis [49], the spectral angle mapper (SAM) [53], and the $Q2^n$-index [54].

Fig. 4 shows the experimental results. The proposed fused result is shown in Fig. 4(h), and the fused results of other state-of-the-art pansharpening methods are shown in Fig. 4(c)–(g), respectively. It can be seen that the proposed method and other pansharpening methods can all obtain good fused results for cloud-free images. On the one hand, visually, all the fused results show sharp spatial details, and they also show similar colors with the reference image. On the other hand, to more objectively evaluate the performance, the quantitative evaluation is performed and shown in Table II. It is shown that the proposed method shows the best results in most of the quantitative evaluation indices.

b) Real experiment: The real experiment for pansharpening of the cloud-free PAN and MS images was implemented based on QuickBird data. The QuickBird satellite provides HR PAN and LR MS images (blue, green, red, and NIR spectral bands) with a spatial resolution of 0.61 and 2.44 m, respectively, and the images in the experiment were acquired over the Wuhan University, Wuhan.

Fig. 5 shows the fusion results. The LR MS and HR PAN images are shown in Fig. 5(a) and (b), respectively. The proposed fusion result is shown in Fig. 5(h), and other fusion results of the state-of-the-art pansharpening methods are shown in Fig. 5(c)–(g), respectively. For comparison, the fusion results of the GS, AWLP, and MTF-GLP methods have good spatial details; however, they show slight a little spectral distortions visually. In addition, the quantitative evaluation results are shown in Table III for objective evaluation. It should be noted that due to the lack of the referenced images in the real experiment. The quantitative evaluation was comprehensively performed from two aspects. First, the fused image was spatially degraded to the spatial dimension of the LR MS image, and the LR MS image was applied as the reference image for the quantitative evaluation. Second, the fused results were evaluated based on nonreference quality evaluation indices, i.e., the popular quality with no reference [55]. Table III shows that the proposed method has the best performance in most of the quantitative evaluation indices.

TABLE II

QUANTITATIVE EVALUATION RESULTS IN THE SIMULATED EXPERIMENT FOR CLOUD-FREE IMAGES

	GS	PRACS	AWLP	MTF-GLP	AMDF	Proposed
CC	0.965	0.973	0.972	0.971	0.977	**0.978**
PSNR	36.269	38.952	37.992	37.207	38.976	**39.069**
ERGAS	2.048	1.730	1.747	1.798	1.580	**1.549**
SAM	2.912	2.689	2.649	2.666	2.518	**2.467**
$Q2^n$-Index	0.841	0.876	0.879	0.873	**0.894**	0.892

2) Pansharpening for Cloud-Contaminated Images:

a) Simulated experiment: The simulated experiment was implemented based on the multisource and multitemporal JL-1 MS and PAN images. The JL-1 is the first commercial operational satellite of China. It provides the HR PAN image with a spatial resolution of 0.72 m and the LR MS image (blue, green, and red spectral bands) with a spatial resolution of 2.88 m. the data sets were acquired on February 22, 2019

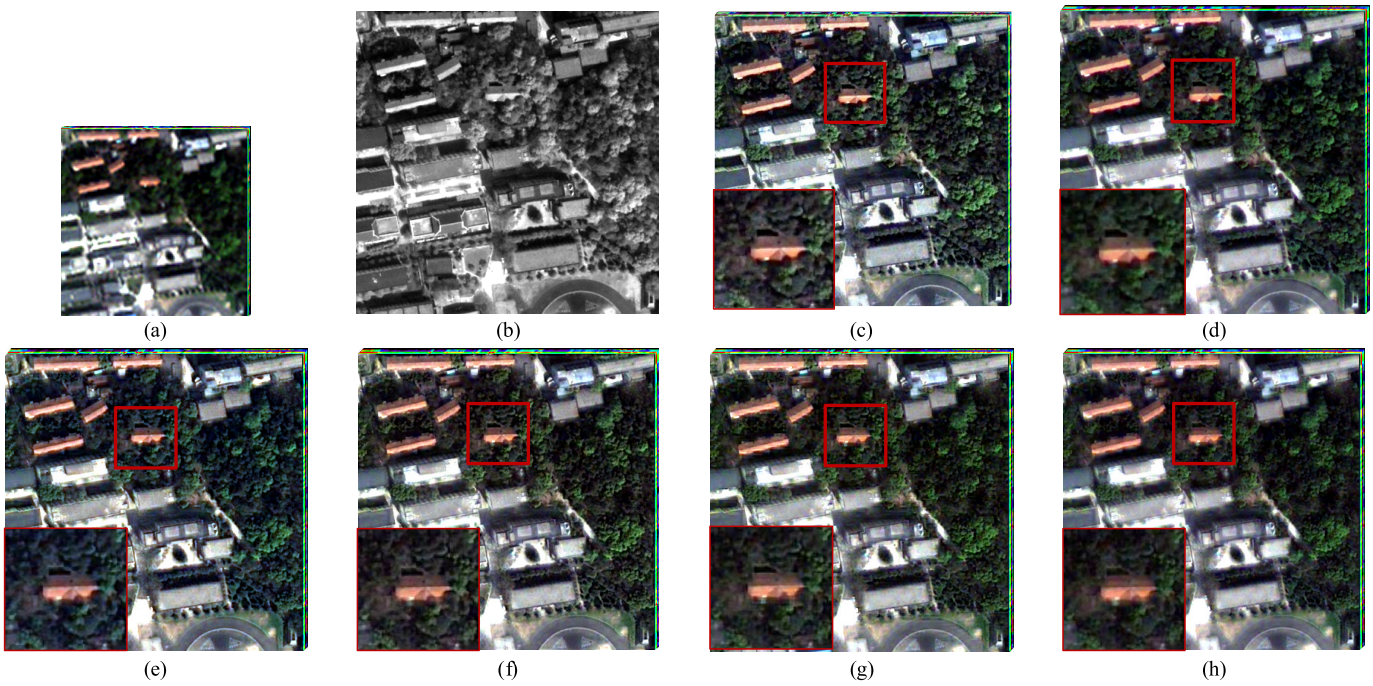


Fig. 5. Fusion results of the real experiment for cloud-free images. (a) LR MS image. (b) HR PAN image. (c) GS fusion result. (d) PRACS fusion result. (e) AWLP fusion result. (f) MTF-GLP fusion result. (g) AMDF fusion result. (h) Proposed fusion result.

TABLE III
QUANTITATIVE EVALUATION RESULTS IN THE REAL EXPERIMENT FOR CLOUD-FREE IMAGES

	GS	PRACS	AWLP	MTF-GLP	AMDF	Proposed
CC	0.942	0.974	0.984	0.990	0.995	**0.998**
PSNR	47.938	53.234	48.758	55.412	57.324	**59.618**
ERGAS	1.590	1.128	0.741	0.740	0.558	**0.268**
SAM	0.898	0.739	0.270	0.553	0.408	**0.190**
$Q2^n$-Index	0.861	0.940	0.938	0.966	0.976	**0.992**
QNR	0.836	0.892	0.811	0.878	0.943	**0.949**
D_λ	0.063	0.052	0.111	0.067	0.033	**0.025**
D_S	0.107	0.059	0.088	0.060	**0.025**	0.026

and April 13, 2016, over Kansas, MO, USA. The simulated target LR MS and HR PAN images with thick clouds on February 22, 2016 are shown in Fig. 6(a) and (b), respectively, and the information in the thick cloud regions can be assumed to be completely missing. The auxiliary LR MS and HR PAN images on April 13, 2016 are shown in Fig. 6(c) and (d), respectively. The main task is to obtain cloud-free HR MS images. The original cloud-free target MS image on February 22, 2016 shown in Fig. 6(h) is employed to evaluation.

Fig. 6 shows the experimental results. The proposed fused result is shown in Fig. 6(g), and the fused results of the GS method and the PRACS method are shown in Fig. 6(e) and (f) to demonstrate the cloud removal effect of the proposed fusion model. It can be seen that the proposed variational-based integrated model can obtain the cloud-free HR MS image, and it is similar to the original target image. For the traditional pansharpening methods, though they can realize the resolution enhancement, the missing information has still existed. In addition, it should be noted that due to the existing pansharpening methods cannot realize the missing information reconstruction as well as to improve the resolutions for the cloud-contaminated images and to comprehensively test and verify the proposed integrated fusion method, the proposed integrated variational model is further compared with the step-by-step fusion. The fused results and the corresponding local zoomed-in areas are shown in Fig. 7.

For the step-by-step fusion, the missing formation in the thick cloud regions of the target LR MS image and the HR PAN image was first reconstructed by the state-of-the-art missing information reconstruction methods of WLR and STMRF methods, based on the auxiliary images, and the reconstructed results were then fused to obtain the final cloud-free HR MS image. Fig. 7(c) and (g) shows the fused result with the proposed integrated fusion model, Fig. 7(a) and (e) and Fig. 7(b) and (f) show the step-by-step fused results with the state-of-the-art WLR and STMRF methods, respectively. The original target MS images are presented in Fig. 7(d) and (h). It can be seen that the proposed integrated fusion model and the step-by-step fusion can both obtain the cloud-free HR MS image. However, the zoomed-in fusion results show that the step-by-step fusion generally introduces the spatial errors, and the proposed integrated fusion result is more consistent to the original target MS image. It stems from the fact that the error accumulation was generally introduced for the step-by-step fusion.

To evaluate the proposed methods objectively, the quantitative evaluation results are shown in Table IV. It is shown that the proposed integrated fusion model achieves the best

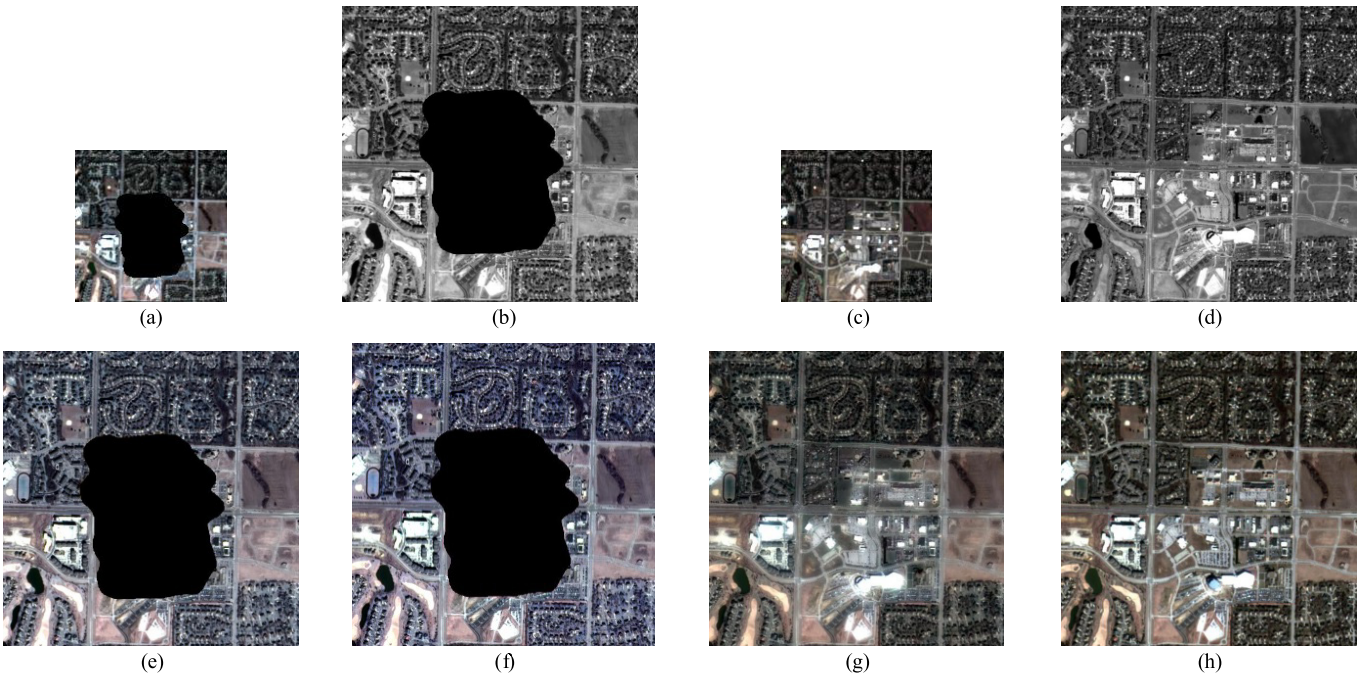


Fig. 6. Fusion results of the simulated experiment for cloud-contaminated images. (a) Target LR MS image with simulated cloud contamination. (b) Target HR PAN image with simulated cloud contamination. (c) Auxiliary LR MS image. (d) Auxiliary HR PAN image. (e) GS fusion result. (f) PRACS fusion result. (g) Proposed fusion result. (h) Original target MS image.

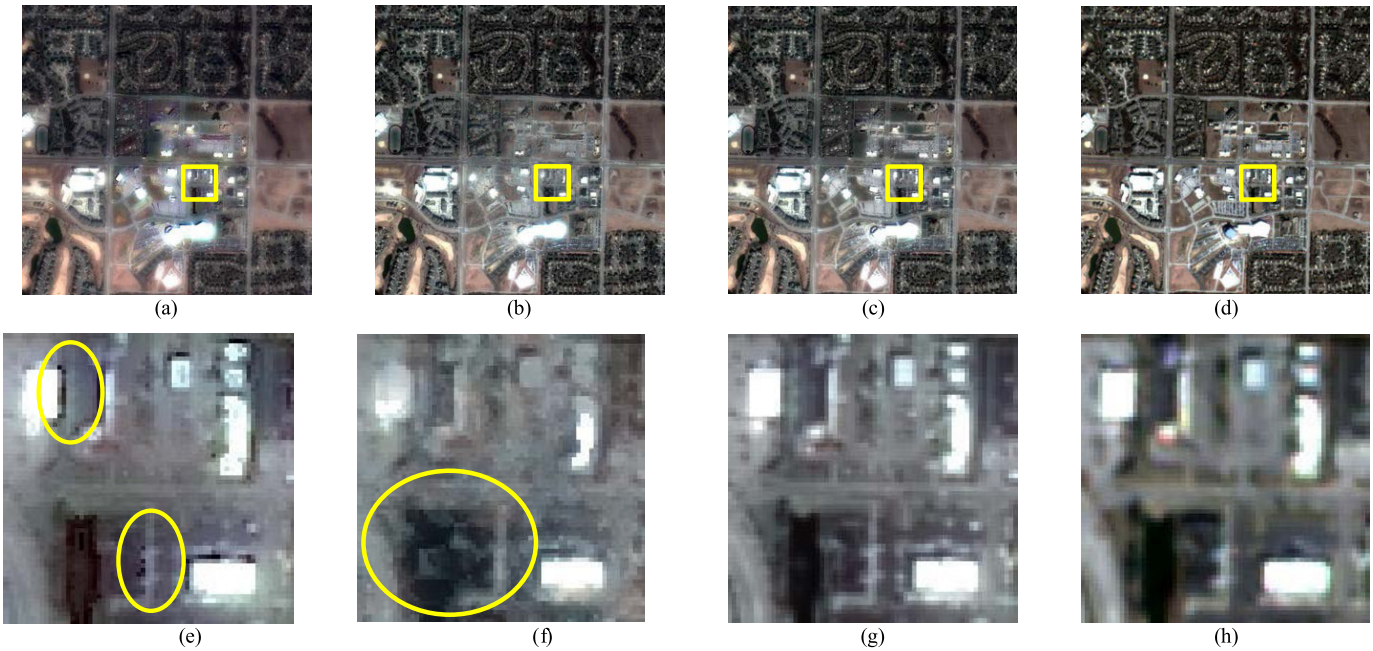


Fig. 7. Comparison between the proposed integrated fusion model and the step-by-step fusion. (a) Step-by-step fusion result with WLR method. (b) Step-by-step fusion result with STMRF method. (c) Integrated fusion result. (d) Original target MS image. (e)–(h) Local zoomed-in areas corresponding to (a)–(d), respectively.

quantitative evaluation result in all the quantitative evaluation indices.

b) Real experiment: The real experiments were implemented based on Deimos-2 satellite images at level 1 C. The Deimos-2 satellite provides HR PAN and LR MS images (blue, green, red, and NIR spectral bands) with a spatial resolution of 1 and 4 m, respectively. The data sets were acquired over Vancouver, BC, Canada (N 49° 15′, W 123° 6′). The target LR MS and HR PAN images were acquired on March 31, 2015, as shown in Fig. 8(a) and (b), respectively, and the auxiliary LR MS and HR PAN images were acquired on May 30, 2015, as shown in Fig. 8(c) and (d), respectively. It can be seen that the target LR MS and HR PAN observations were both contaminated by the thick clouds, the thin clouds, and the cloud shadows. The aim of the experiment was to obtain the desired cloud-free image on March 31, 2015, with both high

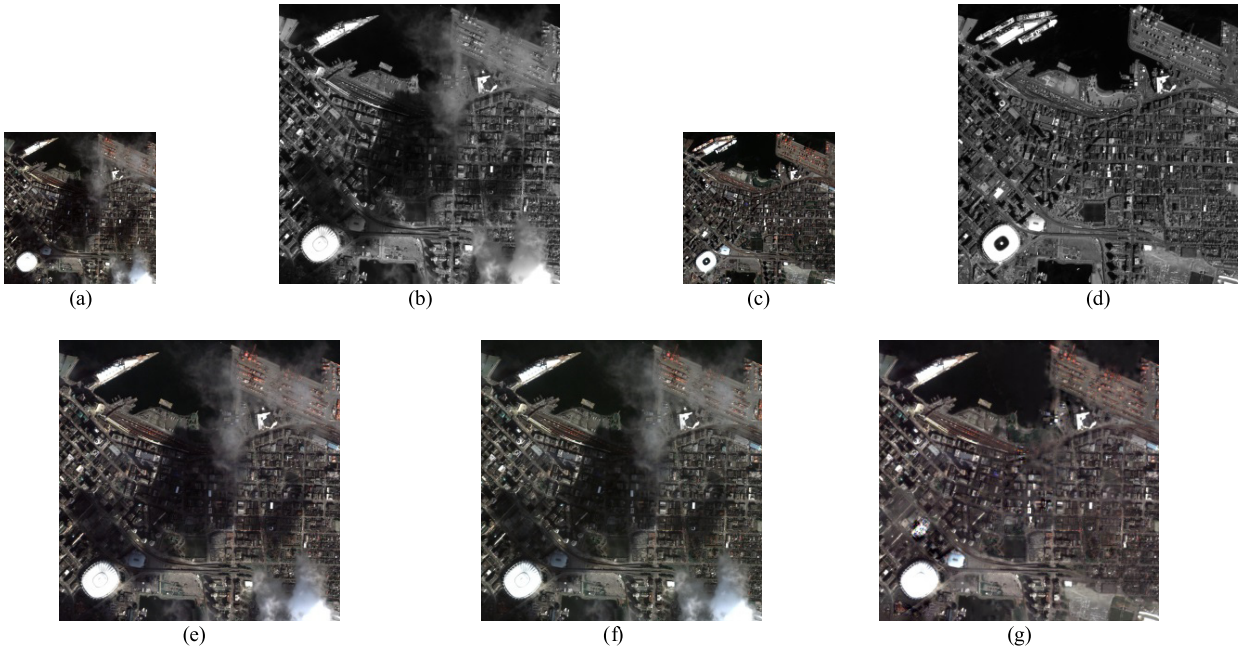


Fig. 8. Fusion results of the real experiment for cloud-contaminated images. (a) Target LR MS image. (b) Target HR PAN image. (c) Auxiliary LR MS image. (d) Auxiliary HR PAN image. (e) GS fusion result. (f) PRACS fusion result. (g) Proposed fusion result.

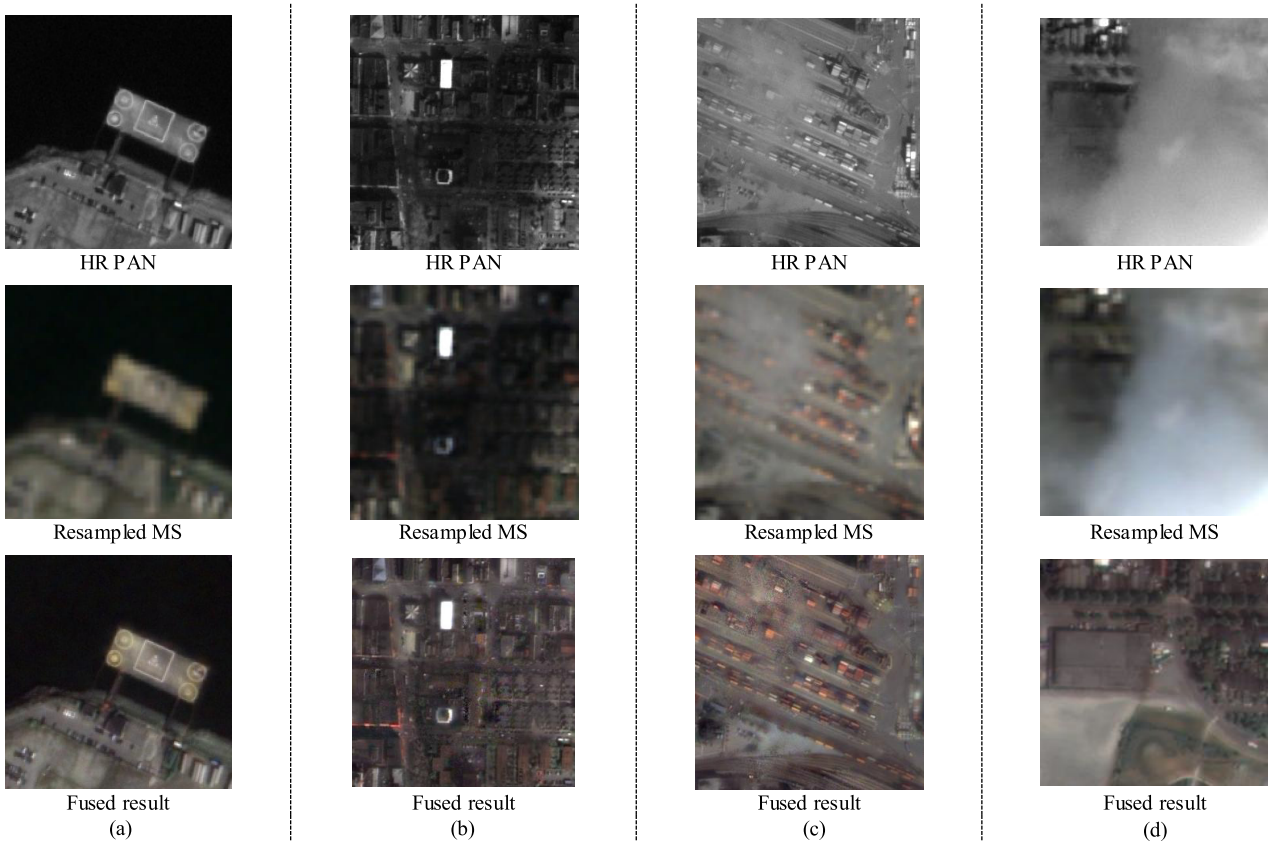


Fig. 9. Four typical zoomed-in areas of the proposed fusion result. (a) Cloud-free region. (b) Cloud shadow region. (c) Thin cloud region. (d) Thick cloud region.

spatial and high spectral resolution. The window size in (5) was set to be 31 by experience. The experimental results are presented in Fig. 8.

The proposed fused result is shown in Fig. 8(g), and the fused result of the GS and PRACS pansharpening methods is shown in Fig. 8(e) and (f), respectively. It can be seen

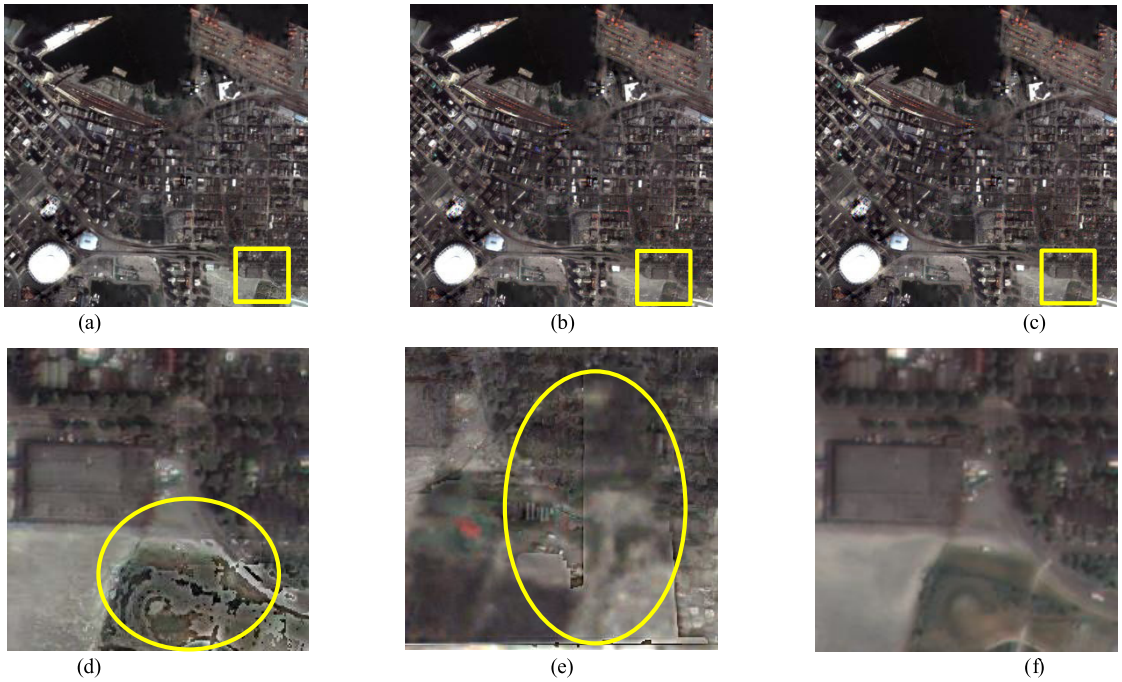


Fig. 10. Comparison between the integrated fusion model and the step-by-step fusion. (a) Step-by-step fusion result with WLR method. (b) Step-by-step fusion result with STMRF method. (c) Fusion result of the variational-based integrated model. (d)–(f) Zoomed-in areas corresponding to (a)–(c), respectively.

TABLE IV
COMPARISON OF THE QUANTITATIVE EVALUATION RESULTS BETWEEN THE INTEGRATED FUSION AND THE STEP-BY-STEP FUSION

	Step-by-step fusion with WLR	Step-by-step fusion with STMRF	Integrated fusion
CC	0.867	0.929	**0.950**
PSNR	41.025	41.238	**41.646**
ERGAS	6.727	6.615	**6.258**
SAM	1.519	1.223	**1.212**
$Q2^n$-Index	0.621	0.686	**0.710**

that the traditional pansharpening method can only achieve the resolution enhancement; however, the clouds and cloud shadows have still existed. On the contrary, the proposed pansharpening method can effectively achieve the joint resolution enhancement and cloud removal, and the cloud-free HR MS image was obtained. For details, a cloud-free region and three typical cloud polluted local regions are shown in Fig. 9. It can be obviously seen that the proposed method can achieve the resolution enhancement for the cloud-free regions, and for the cloud and cloud shadow regions, not only the resolution can be enhanced, but also the cloud and cloud shadows have been effectively removed.

To more comprehensively verify the performance of the proposed variational-based integrated fusion methods, it is further compared with the step-by-step fusion results. The experimental results are presented in Fig. 10. It can be seen that the error accumulation was introduced for the step-by-step fusion results based on both the WLR method and the STMRF method. On the contrary, no extra spatial distortions and errors were introduced for the proposed integrated fused result. Since the ground truth cannot be obtained, the quantitative evaluation was not given.

C. Parameter Analysis of the Variational-Based Integrated Fusion Model

The parameters of the proposed variational-based integrated fusion model were analyzed and discussed.

1) Convergence Analysis and Determination of the Iterative Threshold: The convergence of the proposed fusion model and the determination of the iterative threshold corresponding to simulated cloud-contaminated image fusion experiments were presented, as shown in Fig. 11. For the convergence analysis of the fusion model, the relative error versus iterations was plotted, and the relative error was computed as (16). It can be seen that the proposed fusion model is convergent effectively. The determination of the iterative threshold was performed based on the statistical quantitative evaluation from both the spatial and spectral fidelity of the fused results with different iterative thresholds. Among them, the spectral distortion, represented by SAM, versus the iterative threshold line, and the spatial distortion, denoted by D_S, versus the iterative threshold line were plotted. It can be seen that both the SAM and D_S are getting better and better with the threshold value decreasing, and they achieve the best fusion results when the iterative threshold is 1E-7. Then, the fusion result becomes stable with the increasing of the iterative threshold. Therefore, iterative threshold value of 1E-7 was selected in the experiments.

2) Parameter Analysis of λ_1 and λ_2: The selection of the model parameters λ_1 and λ_2 in (11) has been discussed and is represented in Fig. 12. They are analyzed based on the statistical quantitative evaluation from both the spatial and spectral aspects of the fused results with a variety of different λ_1 and λ_2. The statistical results based on spectral quality evaluation by the SAM and spatial quality evaluation

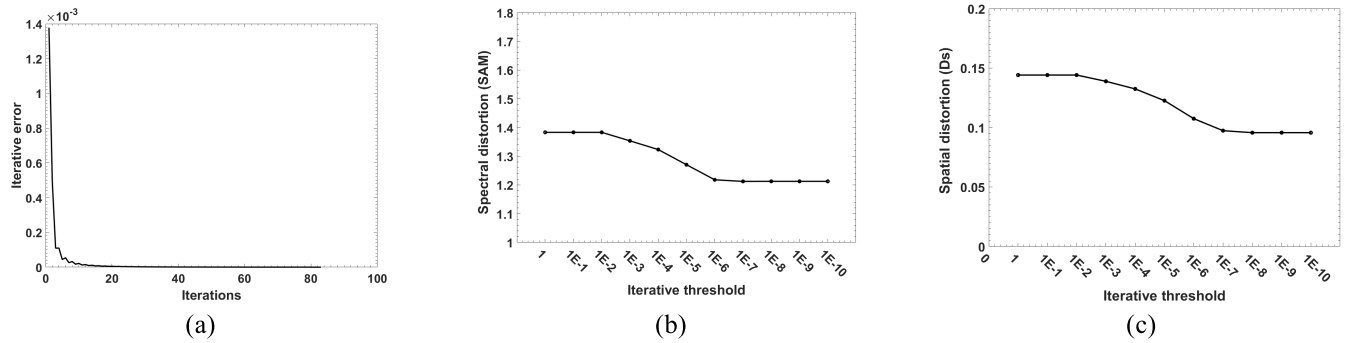


Fig. 11. Convergence analysis and determination of the iterative threshold. (a) Convergence curve of the relative error versus iterations. (b) Spectral quality evaluation (SAM) versus iterative threshold. (c) Spatial quality evaluation (D_S) versus iterative threshold.

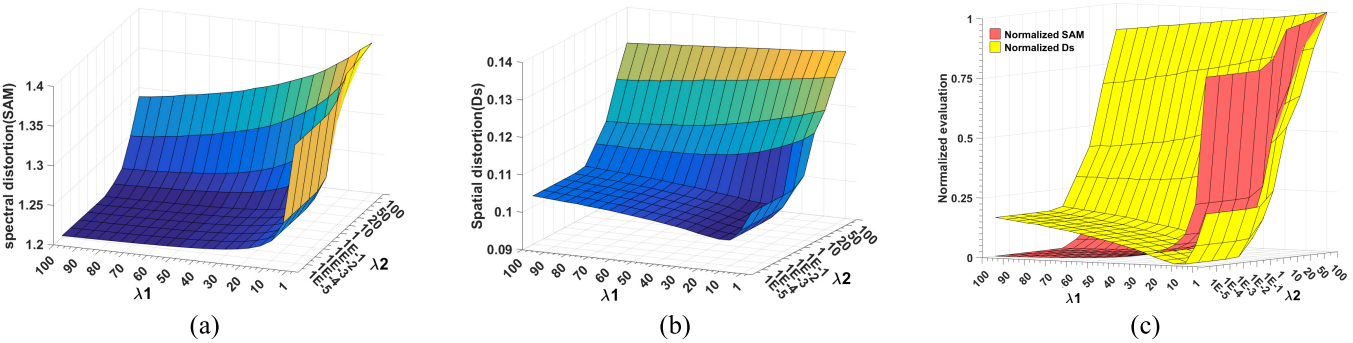


Fig. 12. Parameter analysis of the model parameters λ_1 and λ_2. (a) Statistical spectral quality evaluation (SAM) with different λ_1 and λ_2. (b) Statistical spatial quality evaluation (D_S) with different λ_1 and λ_2. (c) Normalized SAM and D_S with different λ_1 and λ_2.

by D_S are plotted in Fig. 12(a) and (b), respectively. In addition, the quantitative evaluation results of the SAM and D_S are both normalized to 0–1 and shown in the same coordinate space. It is shown that both the spectral evaluation index of SAM and spatial evaluation index of D_S are relatively stable with a different λ_1, as shown in Fig. 12(a) and (b), respectively. Specifically, when λ_1 is set to 10–30, the fusion model shows relatively better performance in both SAM and D_S, as shown in Fig. 12(c). Therefore, λ_1 is set to 20 in the experiments. For the model parameter λ_2, the quantitative evaluations show good performance when $\lambda_2 \leq$ 1E-1, and it is noteworthy that with the increasing of this parameter within this range values, the fusion model generally show robust performance for different experimental data sets. When $\lambda_2 >$ 1E-1, with the increasing of λ_2, the SAM and D_S show worse and worse performance. Therefore, the model parameter λ_2 was set to 0.1 in the experiments.

V. Conclusion

This paper has presented a pansharpening method for the cloud-contaminated VHR remote sensing images. Furthermore, the complex cloud contaminations, i.e., the simultaneous cloud contamination with the thick clouds, the thin clouds, the haze, and the cloud shadows are comprehensively considered. In the proposed method, considering the common semitransparent characteristics of the thin clouds, the haze, and the light cloud shadows, they are first jointly removed based on the sliding window-based local moment matching method. Then, an integrated variational fusion model is proposed to achieve the joint resolution enhancement and missing information reconstruction. The proposed method was finally tested and verified by a number of different remote sensing satellite images, including the IKONOS, QuickBird, JL-1, and deimos-2, based on both the simulated and real experiments. Furthermore, the pansharpening experiments for both the cloud-free and cloud-contaminated images are implemented to comprehensively test and verify the proposed methods. The experimental results show that the proposed method can effectively obtain the cloud-free image with both high spatial and spectral resolutions.

However, it should be noted that there is still room for the proposed fusion method to be further improved. The first limitation is the determination of the cloud masks. In this paper, the cloud masks are determined by manual selection. This is because the thin clouds and the thick clouds should be distinguished for the proposed framework; in addition, the light cloud shadows and dark cloud shadows should also be classified. To the best of our knowledge, this may be challenging as well as a complex problem for most of the cloud detection methods. Therefore, they are selected manually, and this can be further improved by developing adaptive determination method. The second limitation is that this paper assumed that the auxiliary images with approaching phase should be chosen as far as possible to ensure the consistent ground features to the target images. However, if we cannot obtain such auxiliary images, how to realize the

integrated fusion? This may be further improved by making use of the useful complementary and redundancy information from both multitemporal auxiliary images and the synthetic aperture radar observations. Finally, it is the efficiency of the integrated fusion model; a faster optimization algorithm and acceleration strategies such as parallel computing could be incorporated into the integrated fusion framework.

ACKNOWLEDGMENT

The authors would like to thank Deimos Imaging for acquiring and providing the data used in this paper and the IEEE Geoscience and Remote Sensing Society Image Analysis and Data Fusion Technical Committee. They would also like to thank the anonymous reviewers for their careful reading and valuable comments.

REFERENCES

[1] H. Shen, X. Meng, and L. Zhang, "An integrated framework for the spatio–temporal–spectral fusion of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7135–7148, Dec. 2016.

[2] Y. Zhang, "Understanding image fusion," *Photogramm. Eng. Remote Sens.*, vol. 70, no. 6, pp. 657–661, 2004.

[3] Y. Zhang and R. K. Mishra, "From UNB PanSharp to fuze go—The success behind the pan-sharpening algorithm," *Int. J. Image Data Fusion*, vol. 5, no. 1, pp. 39–53, 2014.

[4] R. Haydn *et al.*, "Application of the IHS color transform to the processing of multisensor data and image enhancement," in *Proc. Int. Symp. Remote Sens.*, Cairo, Egypt, 1982, pp. 599–616.

[5] R. A. Schowengerdt, "Reconstruction of multispatial, multispectral image data using spatial frequency content," *Photogramm. Eng. Remote Sens.*, vol. 46, no. 10, pp. 1325–1334, 1980.

[6] X. Meng, H. Shen, H. Li, L. Zhang, and R. Fu, "Review of the pansharpening methods for remote sensing images based on the idea of meta-analysis: Practical discussion and challenges," *Inf. Fusion*, vol. 46, pp. 101–113, Mar. 2019.

[7] P. S. Chavez, Jr., 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.

[8] C. A. Laben and B. V. Brower, "Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening," U.S. Patent 6 011 875, Jan. 4, 2000.

[9] W. Dou, Y. Chen, X. Li, and D. Z. Sui, "A general framework for component substitution image fusion: An implementation using the fast image fusion method," *Comput. Geosci.*, vol. 33, no. 2, pp. 219–228, 2007.

[10] 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.

[11] Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li, "A comparative analysis of image fusion methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1391–1402, Jun. 2005.

[12] A. G. Mahyari and M. Yazdi, "Panchromatic and multispectral image fusion based on maximization of both spectral and spatial similarities," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 1976–1985, Jun. 2011.

[13] X. Otazu, M. González-Audícana, O. Fors, and J. Núñez, "Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 10, pp. 2376–2385, Oct. 2005.

[14] 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.

[15] A. Garzelli, F. Nencini, and L. Capobianco, "Optimal MMSE pan sharpening of very high resolution multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 1, pp. 228–236, Jan. 2008.

[16] H. Li, B. S. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," *Graph. Models Image Process.*, vol. 57, no. 3, pp. 235–245, 1995.

[17] 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.

[18] 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.

[19] 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.

[20] X. Meng, J. Li, H. Shen, L. Zhang, and H. Zhang, "Pansharpening with a guided filter based on three-layer decomposition," *Sensors*, vol. 16, no. 7, p. 1068, 2016.

[21] B. Aiazzi, L. Alparone, S. Baronti, R. Carlà, A. Garzelli, and L. Santurri, "Sensitivity of pansharpening methods to temporal and instrumental changes between multispectral and panchromatic data sets," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 1, pp. 308–319, Jan. 2017.

[22] C. Ballester, V. Caselles, L. Igual, J. Verdera, and B. Rougé, "A variational model for P+XS image fusion," *Int. J. Comput. Vis.*, vol. 69, no. 1, pp. 43–58, 2006.

[23] L. Zhang, H. Shen, W. Gong, and H. Zhang, "Adjustable model-based fusion method for multispectral and panchromatic images," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 6, pp. 1693–1704, Dec. 2012.

[24] S. Li and B. Yang, "A new pan-sharpening method using a compressed sensing technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 738–746, Feb. 2011.

[25] 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.

[26] Z. Li and H. Leung, "Fusion of multispectral and panchromatic images using a restoration-based method," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1482–1491, May 2009.

[27] X. Meng *et al.*, "Improving the spatial resolution of hyperspectral image using panchromatic and multispectral images: An integrated method," in *Proc. 7th Workshop Hyperspectral Image Signal Process., Evol. Remote Sens. (WHISPERS)*, Tokyo, Japan, 2015, pp. 1–4.

[28] C. Jiang, H. Zhang, H. Shen, and L. Zhang, "A practical compressed sensing-based pan-sharpening method," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 4, pp. 629–633, Jul. 2012.

[29] M. Guo, H. Zhang, J. Li, L. Zhang, and H. Shen, "An online coupled dictionary learning approach for remote sensing image fusion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1284–1294, Apr. 2014.

[30] 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.

[31] X. X. Zhu *et al.* (2017). "Deep learning in remote sensing: A review." [Online]. Available: https://arxiv.org/abs/1710.03959

[32] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, Jun. 2016.

[33] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 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. Zhong, B. Yang, G. Huang, F. Zhong, and Z. Chen, "Remote sensing image fusion with convolutional neural network," *Sens. Imag.*, vol. 17, no. 1, p. 10, 2016.

[36] 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.

[37] G. Hu, X. Li, and D. Liang, "Thin cloud removal from remote sensing images using multidirectional dual tree complex wavelet transform and transfer least square support vector regression," *J. Appl. Remote Sens.*, vol. 9, no. 1, pp. 095053:1–095053:19, 2015.

[38] H. Shen *et al.*, "Missing information reconstruction of remote sensing data: A technical review," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 3, pp. 61–85, Sep. 2015.

[39] X. Meng, H. Shen, Q. Yuan, H. Li, and L. Zhang, "An integrated fusion framework for joint information reconstruction and resolution enhancement," in *Proc. Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, Wuhan, China, 2017, pp. 831–835.

[40] C. Shi, F. Liu, L.-L. Li, and H.-X. Hao, "Pan-sharpening algorithm to remove thin cloud via mask dodging and nonsampled shift-invariant shearlet transform," *J. Appl. Remote Sens.*, vol. 8, no. 1, p. 083658, 2014.

[41] D. Fasbender, V. Obsomer, J. Radoux, P. Bogaert, and P. Defourny, "Bayesian data fusion: Spatial and temporal applications," in *Proc. Int. Workshop Anal. Multi-Temporal Remote Sens. Images*, Leuven, Belgium, 2007, pp. 1–6.

[42] P. J. Huber, "Robust estimation of a location parameter," *Ann. Math. Statist.*, vol. 35, no. 1, pp. 73–101, 1964.

[43] P. Pérez, M. Gangnet, and A. Blake, "Poisson image editing," *ACM Trans. Graph.*, vol. 22, no. 3, pp. 313–318, 2003.

[44] J. Storey, P. Scaramuzza, and G. Schmidt, "Landsat 7 scan line corrector-off gap-filled product development," in *Proc. Global Priorities Land Remote Sens.*, Sioux Falls, South Dakota, 2005, pp. 23–27.

[45] M. Joshi and A. Jalobeanu, "MAP estimation for multiresolution fusion in remotely sensed images using an IGMRF prior model," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1245–1255, Mar. 2010.

[46] P. Liu, L. Xiao, J. Zhang, and B. Naz, "Spatial-Hessian-feature-guided variational model for pan-sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 2235–2253, Apr. 2016.

[47] J. Wen, Y. Li, and H. Gong, "Remote sensing image fusion on gradient field," in *Proc. 18th Int. Conf. Pattern Recognit. (ICPR)*, Hong Kong, 2006, pp. 643–646.

[48] M. K. Ng, W. C. Kwan, and R. H. Chan, "A fast algorithm for high-resolution color image reconstruction with multisensors," in *Proc. Int. Conf. Numer. Anal. Appl.*, 2000, pp. 615–627.

[49] L. Wald, "Quality of high resolution synthesised images: Is there a simple criterion?" in *Proc. 3rd Conf. Fusion Earth Data, Merging Point Meas., Raster Maps Remotely Sensed Images*, Sophia Antipolis, France, 2000, pp. 99–103.

[50] C. Zeng, H. Shen, and L. Zhang, "Recovering missing pixels for Landsat ETM + SLC-off imagery using multi-temporal regression analysis and a regularization method," *Remote Sens. Environ.*, vol. 131, pp. 182–194, Apr. 2013.

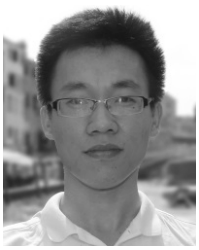
[51] Q. Cheng, H. Shen, L. Zhang, Q. Yuan, and C. Zeng, "Cloud removal for remotely sensed images by similar pixel replacement guided with a spatio-temporal MRF model," *ISPRS J. Photogramm. Remote Sens.*, vol. 92, pp. 54–68, Jun. 2014.

[52] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? A new look at signal fidelity measures," *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98–117, Jan. 2009.

[53] 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.

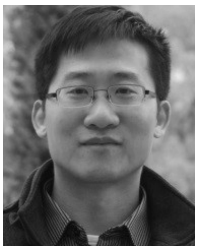
[54] A. Garzelli and F. Nencini, "Hypercomplex quality assessment of multi/hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 662–665, Oct. 2009.

[55] L. Alparone, B. Aiazzi, S. Baronti, A. Garzelli, F. Nencini, and M. Selva, "Multispectral and panchromatic data fusion assessment without reference," *Photogramm. Eng. Remote Sens.*, vol. 74, no. 2, pp. 193–200, Feb. 2008.



Xiangchao Meng (M'18) received the B.S. degree in geographic information system from the Shandong University of Science and Technology, Qingdao, China, in 2012, and the Ph.D. degree in cartography and geography information system from Wuhan University, Wuhan, China, in 2017.

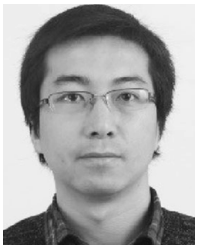
He is currently a Lecturer with the Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo, China. His research interests include variational methods and remote sensing image fusion.



Huanfeng Shen (M'10–SM'13) received the B.S. degree in surveying and mapping engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2002 and 2007, respectively.

In 2007, he joined the School of Resource and Environmental Sciences, Wuhan University, where he is currently a Luojia Distinguished Professor. He has authored more than 100 research papers. His research interests include image quality improvement, remote sensing mapping and application, data fusion and assimilation, and regional and global environmental change.

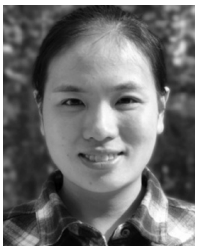
Dr. Shen is currently a member of the Editorial Board of the *Journal of Applied Remote Sensing*. He has been supported by several talent programs, such as The Youth Talent Support Program of China in 2015, the China National Science Fund for Excellent Young Scholars in 2014, and the New Century Excellent Talents by the Ministry of Education of China in 2011.



Qiangqiang Yuan (M'13) received the B.S. degree in surveying and mapping engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2006 and 2012, respectively.

In 2012, he joined the School of Geodesy and Geomatics, Wuhan University, where he is currently an Associate Professor. He has authored more than 50 research papers, including more than 30 peer-reviewed articles in international journals, such as the IEEE TRANSACTIONS ON IMAGE PROCESSING and the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. His research interests include image reconstruction, remote sensing image processing and application, and data fusion.

Dr. Yuan was a recipient of the Top-Ten Academic Star of Wuhan University, in 2011, and the Hong Kong Scholar Award from the Society of Hong Kong Scholars and the China National Postdoctoral Council, in 2014. He has frequently served as a referee for more than 20 international journals for remote sensing and image processing.



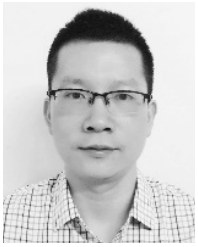
Huifang Li (M'14) received the B.S. degree in geographical information science from the China University of Mining and Technology, Xuzhou, China, in 2008, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2013.

She is currently a Lecturer with the School of Resource and Environmental Sciences, Wuhan University. Her research interests include variational methods, dehazing, and deshadowing of remote sensing images.

Liangpei Zhang (M'06–SM'08) received the B.S. degree in physics from Hunan Normal University, Changsha, China, in 1982, the M.S. degree in optics from the Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, China, in 1988, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1998.

He was a Principal Scientist for the China State Key Basic Research Project from 2011 to 2016 appointed by the Ministry of National Science and Technology of China to lead the Remote Sensing Program in China. He is currently the Head of the Remote Sensing Division, State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University. He is also a "Chang-Jiang Scholar" Chair Professor appointed by the Ministry of Education of China. He has authored more than 450 research papers and five books. He holds 15 patents. His research interests include hyperspectral remote sensing, high-resolution remote sensing, image processing, and artificial intelligence.

Dr. Zhang is a fellow of the Institution of Engineering and Technology and an Executive Member (Board of Governor) of the China National Committee of International Geosphere–Biosphere Program and the China Society of Image and Graphics. He was a recipient of the 2010 Best Paper Boeing Award, the 2013 Best Paper ERDAS Award from the American Society of Photogrammetry and Remote Sensing, and the best reviewer awards from IEEE Geoscience and Remote Sensing Society (GRSS) for his service to the IEEE JOURNAL OF SELECTED TOPICS IN EARTH OBSERVATIONS AND APPLIED REMOTE SENSING (JSTARS) in 2012 and IEEE GEOSCIENCE AND REMOTE SENSING LETTERS in 2014. His research teams won the top three prizes of the IEEE GRSS 2014 Data Fusion Contest, and his students have been selected as the winners or finalists of the IEEE International Geoscience and Remote Sensing Symposium Student Paper Contest in recent years. He is the Founding Chair of the IEEE GRSS Wuhan Chapter. He was the General Chair for the 4th IEEE GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing and the Guest Editor of JSTARS. He regularly serves as a Co-Chair of the series SPIE conferences on multispectral image processing and pattern recognition, conference on Asia remote sensing, and many other conferences. He edits several conference proceedings, issues, and geoinformatics symposiums. He also serves as an Associate Editor for the *International Journal of Ambient Computing and Intelligence*, the *International Journal of Image and Graphics*, the*International Journal of Digital Multimedia Broadcasting*, the *Journal of Geo-spatial Information Science*, and the *Journal of Remote Sensing*, and the Guest Editor of the *Journal of Applied Remote Sensing* and the *Journal of Sensors*. He is currently serving as an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.



Weiwei Sun (M'15) received the B.S. degree in surveying and mapping and the Ph.D. degree in cartography and geographic information engineering from Tongji University, Shanghai, China, in 2007 and 2013, respectively.

From 2011 to 2012, he was a Visiting Scholar with the Department of Applied Mathematics, University of Maryland at College Park, College Park, MD, USA, with Prof. J. Benedetto to study on the dimensionality reduction of hyperspectral image. From 2014 to 2016, he was a Post-Doctoral Researcher with the State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, to study intelligent processing in hyperspectral imagery. From 2017 to 2018, he was a Visiting Scholar with the Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS, USA, with Prof. Q. Du. He is currently an Associate Professor with Ningbo University, Ningbo, China. He has authored more than 60 journal papers. His research interests include remote sensing and machine learning.