An Integrated Framework for the Spatio–Temporal–Spectral Fusion of Remote Sensing Images

Huanfeng Shen, *Senior Member, IEEE*, Xiangchao Meng, and Liangpei Zhang, *Senior Member, IEEE*

Abstract—Remote sensing satellite sensors feature a tradeoff between the spatial, temporal, and spectral resolutions. In this paper, we propose an integrated framework for the spatio–temporal–spectral fusion of remote sensing images. There are two main advantages of the proposed integrated fusion framework: it can accomplish different kinds of fusion tasks, such as multiview spatial fusion, spatio–spectral fusion, and spatio–temporal fusion, based on a single unified model, and it can achieve the integrated fusion of multisource observations to obtain high spatio–temporal–spectral resolution images, without limitations on the number of remote sensing sensors. The proposed integrated fusion framework was comprehensively tested and verified in a variety of image fusion experiments. In the experiments, a number of different remote sensing satellites were utilized, including IKONOS, the Enhanced Thematic Mapper Plus (ETM+), the Moderate Resolution Imaging Spectroradiometer (MODIS), the Hyperspectral Digital Imagery Collection Experiment (HYDICE), and Système Pour l' Observation de la Terre-5 (SPOT-5). The experimental results confirm the effectiveness of the proposed method.

Index Terms—Image fusion, integrated framework, remote sensing, spatial resolution, spectral resolution, temporal resolution.

I. Introduction

WITH the rapid development of remote sensing sensor networks, massive volumes of remote sensing images can be now obtained every day. However, due to the technical limitations of the sensors and other factors, the existing optical remote sensing sensors have to make a fundamental tradeoff between the spatial, temporal, and spectral resolutions, which greatly limits the potential applications of remote sensing images [1]–[3]. Fortunately, remote sensing image fusion is an effective way to obtain superior high spatio–temporal–spectral resolution images by merging the complementary information.

It is defined as the synergic combination of two or more image data sets, which is aimed at producing a knowledge of phenomena under investigation better than the knowledge achievable from individual data sets [4].

To date, a variety of remote sensing image fusion methods [5]–[10] have been proposed. In general, the existing image fusion methods are performed at three different processing levels, i.e., data level, feature level, and decision level [6], [11]. In this paper, we focus on the data-level fusion of remote sensing images. According to the different aims, there are several categories of remote sensing image fusion, such as multiview spatial fusion, spatio–spectral fusion, and spatio–temporal fusion. However, on one hand, most of the existing methods have been developed independently, and few studies have been dedicated to studying the relationships between them, and thus, the fusion frameworks lack versatility. On the other hand, most of the existing fusion methods can be only applied to integrate two of the spatial, temporal, and spectral resolution indexes, and thus, they cannot obtain the highest spatio–temporal–spectral resolution fused images. In addition, they are designed to merge the complementary information from one or two sensors, and thus, they cannot take full advantage of the useful complementary information from more sensor observations.

Therefore, in this paper, we propose an integrated framework for the spatio–temporal–spectral fusion of remote sensing images. In the proposed fusion framework, the maximum *a posteriori* (MAP) theory is utilized to describe the inverse fusion problem. The spatial, temporal, and spectral relationships between the desired image and the multisource remote sensing observations are then thoroughly analyzed to construct an integrated relationship model. Finally, the fused image is solved by the classical conjugate gradient optimization algorithm. The proposed integrated fusion framework can simultaneously accomplish multiview spatial fusion, spatio–spectral fusion, spatio–temporal fusion, etc., based on a single unified model. In addition, it can obtain high spatio–temporal–spectral resolution fused images by taking full advantage of the useful complementary information from multisource observations. Furthermore, it does not have any limitation on the number of sensors.

The rest of this paper is organized as follows. Section II presents the related work about the existing remote sensing image fusion methods. Section III presents the proposed integrated fusion framework. The optimization solution and the parameter determination are presented in Section IV. In Section V, the experimental results and analysis are provided. Finally, the conclusion is drawn in Section VI.

II. RELATED WORK

A. Multiview Spatial Fusion

Multiview spatial fusion refers to the process of combining a sequence of multiview low-spatial-resolution remote sensing images, generally acquired by one imaging system, to produce a higher spatial resolution image. It is also known as multiframe superresolution reconstruction [12]–[15]. Multiview spatial fusion was first proposed by Tsai and Huang [16] in the frequency domain, and it has since been successfully applied to improve the hardware performance of remote sensing sensors. For example, it has been successfully applied to the Système Pour l' Observation de la Terre-5 (SPOT-5) satellite system to produce a 2.5-m "THR" (a French acronym that stands for "very high resolution") panchromatic (PAN) image with two 5-m half-pixel shift images delivered by the double charge-coupled device linear array [17]. To date, multiview spatial fusion has been focused on the fusion of multitemporal and multiangle image sequences [18], [19]. To the best of our knowledge, Shen *et al.* [20] and Merino and Nunez [21] earlier proposed multiview spatial fusion methods for multitemporal images. Li *et al.* [22] subsequently presented a MAP-based method with a universal hidden Markov tree model for use with Landsat-7 PAN images of different dates. Dually, Zhang *et al.* [19] and Ma *et al.* [23] proposed fusion methods for multiangle remote sensing data sets. It is noteworthy that most of the existing multiview spatial fusion methods are based on an observation model considering the imaging process of the remote sensing sensors. In addition, a notable feature is that they take advantage of the complementary information of multiframe degraded observations to improve the spatial resolution. However, due to the lack of auxiliary higher spatial resolution images, the effect of the spatial resolution improvement is not significant.

B. Spatio–Spectral Fusion

Spatio–spectral fusion [7]–[9] is an important approach in remote sensing image fusion. It is aimed at obtaining a fused image with both high spatial and spectral resolutions. The classical spatio–spectral fusion methods include PAN/multispectral (MS) fusion [7], [8], PAN/hyperspectral (HS) fusion [9], [24], and MS/HS fusion [9], [25].

PAN/MS fusion, which is popularly referred to as "pansharpening," is a particular case of the spatio–spectral fusion methods. It is used to fuse the high-spatial-resolution PAN image (a single band) and several bands of the low-spatial-resolution MS image simultaneously acquired over the same area [8]. PAN/MS fusion originated in the 1980s [26]–[28]; since 1986, the Système Pour l' Observation de la Terre-1 (SPOT-1) system has provided two MS images together with one PAN image; thus, PAN/MS fusion methods have got rapid development over a period of nearly 30 years. In general, most of the existing PAN/MS fusion methods can be grouped into four categories: 1) component substitution (CS)-based approaches; 2) multiresolution analysis (MRA)-based approaches; 3) model-based optimization (MBO) approaches; and 4) sparse reconstruction (SR)-based approaches. Among them, the CS-based [29]–[31] and MRA-based [32]–[34] fusion methods are the most popular due to their simplicity and high efficiency, and they have been formulated as general image fusion frameworks [29], [35]. The MBO fusion approach [36] is another important category. These methods regard the solution of the fused image as an ill-posed problem, and they mainly rely on constructing a relationship model between the desired fused image and the observed PAN/MS images. In general, the MBO fusion method has higher precision than the CS-based and MRA-based fusion methods; however, it is a relatively complex and time-consuming approach. The SR-based PAN/MS fusion approaches [37]–[40] have attracted increasing attention in recent years, and they are considered to be a promising new generation of fusion methods [41]. Their basic idea is that the reconstructed signal is sparse and it can be represented as a linear combination of a few base elements. This category of fusion methods mainly relies on dictionary construction. More detailed comparisons and discussions of the pansharpening methods can be found in [4], [7], [8], and [42]–[44].

In addition to PAN/MS fusion, PAN/HS fusion and MS/HS fusion are another two essential spatio–spectral fusion categories. To date, many different PAN/MS fusion methods have been applied and extended to PAN/HS fusion [9], [24] and MS/HS fusion [9], [45], [46]. In addition, a large number of model-based or unmixing-based PAN/HS [47], [48] and MS/HS fusion methods [25], [49]–[51] have been developed. On the whole, PAN/HS fusion and MS/HS fusion are more challenging than PAN/MS fusion because of the increased number of spectral bands and the more extensive spectral wavelength [52].

C. Spatio–Temporal Fusion

Spatio–temporal fusion is another fundamental approach to remote sensing image fusion. It can make use of the complementary spatial and temporal information of multisource remote sensing observations to obtain a fused image with both a high spatial resolution and a frequent temporal coverage [53]. Most of the spatio–temporal fusion methods can be classified into three categories: 1) filter-based methods; 2) unmixing-based methods; and 3) learning-based methods. Among them, the filter-based methods [3], [53]–[56] are the most popular. This category of methods utilizes the weighted sum of the neighboring similar pixels of the input observations to reconstruct the fused image [10]. Gao *et al.* [53] proposed the spatial and temporal adaptive reflectance fusion model (STARFM) to fuse Landsat and Moderate Resolution Imaging Spectroradiometer (MODIS) observations. Subsequently, a number of popular spatio–temporal fusion methods have been further developed, including the enhanced STARFM [56] and the extended spatio–temporal fusion method considering sensor observation differences [54]. In addition, the unmixing-based methods [57], [58] and the learning-based methods [59], [60] have been attracting ever-increasing attention.

D. Integrated Fusion Framework

Although many different kinds of image fusion methods have been proposed, they have all been developed independently, and there have been few studies exploring the relationships between them. In addition, the aforementioned fusion methods

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

SHEN *et al.*: FRAMEWORK FOR THE SPATIO–TEMPORAL–SPECTRAL FUSION OF REMOTE SENSING IMAGES 3
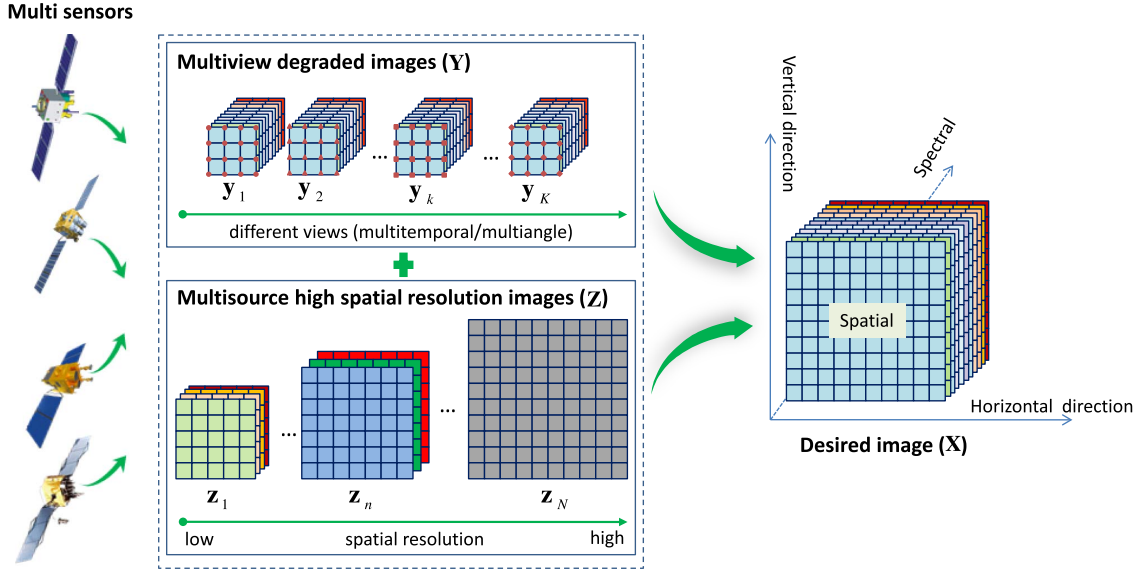

Fig. 1. Schematic of the proposed integrated fusion framework.

are dedicated to fusing two of the spatial, temporal, and spectral resolution indexes, and they are also only focused on one or two sensors. Therefore, in this paper, we propose an integrated fusion framework.

In this paper, we define the integrated fusion framework as taking full advantage of the multicomplementary spatial, temporal, and spectral information to obtain superior fused images, such as images with high spatial, high temporal, and high spectral resolutions. In general, it has two meanings. On one hand, it is dedicated to explore the relationship among different types of image fusion methods to formulate the unified fusion framework. On the other hand, it aims at fusing more complementary information from multisensor observations to obtain better fused results than one-sensor or two-sensor fusion. To the best of our knowledge, Shen [61] was the first to attempt to formulate an integrated fusion framework for multiple spatio–temporal–spectral remote sensing images; however, the method was only tested on simulated data sets. Huang *et al.* [62] proposed an extended version of the model in [61] by exploring the relationship between the spatio–spectral and spatio–temporal fusion methods. This method was tested in real-data experiments; however, it was only performed on two sensors. Wu *et al.* [3] and Meng *et al.* [63] proposed integrated fusion frameworks for multiple sensors, but they did not consider the simultaneous fusion of multiple spatial, temporal, and spectral features. Therefore, as an extension of [64], this paper presents a unified framework for the integrated fusion of complementary spatial, temporal, and spectral information in multisource observations.

III. PROPOSED FUSION FRAMEWORK

A. Framework Description

In this paper, we propose an integrated framework for the spatio–temporal–spectral fusion of remote sensing images. It aims to obtain fused images with the highest spatial, temporal, and spectral resolutions among the input multisource remote

TABLE I
PROPOSED FUSION FRAMEWORK AS A UNIFIED MODEL
FOR DIFFERENT KINDS OF FUSION TASKS

Fusion tasks	Input observations
Multiview spatial fusion	Only multiview degraded images $\mathbf{Y}$
Spatio-spectral fusion	One image in $\mathbf{Y}$ and the corresponding image in $\mathbf{Z}$. It may be MS/HS fusion, PAN/HS fusion, or PAN/MS fusion, etc.
Spatio-temporal fusion	One image in $\mathbf{Y}$ at the target time, and one or more pairs of images in $\mathbf{Y}$ and $\mathbf{Z}$ at other times
Spatio-temporal -spectral fusion	All the useful observations of $\mathbf{Y}$ and $\mathbf{Z}$ with multiple spatial, temporal, and spectral features

sensing observations. In ideal conditions, it is expected to produce time series of high spatio–spectral resolution images simultaneously. However, difficulties in model establishment and computation load make this infeasible. Therefore, the proposed fusion framework is designed to obtain a fused image with high spatial and spectral resolutions for a specific time k, as shown in Fig. 1. The prerequisite for the fusion process is that there is an observation $\mathbf{y}_k$ at the given time k. The multiview (multitemporal or multiangle) images $\mathbf{Y} = \{\mathbf{y}_1, \ldots, \mathbf{y}_k, \ldots, \mathbf{y}_K\}$ of $\mathbf{y}_k$ may be also used in the fusion process, with k being the total number of images. In general, the images $\mathbf{Y}$ have higher spectral and temporal resolutions, but with a lower spatial resolution. Therefore, we regard $\mathbf{Y}$ as the spatially degraded images. As the enhanced counterpart of $\mathbf{y}_k$ by the fusion process, $\mathbf{x}$ is the desired image with the higher spatial resolution. For a better solution of $\mathbf{x}$, auxiliary multisource observations $\mathbf{Z} = \{\mathbf{z}_1, \ldots, \mathbf{z}_n, \ldots, \mathbf{z}_N\}$ are often necessary. Here, $\mathbf{z}_n$ represents the nth image, and it is generally a PAN or an MS image with a higher spatial resolution, but with a lower spectral or temporal resolution. N is the total number of auxiliary images.

As aforementioned, the proposed integrated fusion framework can accomplish different kinds of fusion tasks, including multiview spatial fusion, spatio–spectral fusion, spatio–temporal fusion, and spatio–temporal–spectral fusion, based on a single unified model. Table I shows the different input requirements of the different fusion tasks.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

4

IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING

B. Relationship Model

The integrated relationship model is constructed to build the relationship between the desired image $\mathbf{x}$ and the multisource remote sensing observations $\mathbf{Y}$, $\mathbf{Z}$. It includes the spatial degradation model and the spatio–temporal–spectral relationship model.

1) Spatial Degradation Model: The construction of the relationship between the desired image $\mathbf{x}$ and the observed degraded observations $\mathbf{Y}$ requires a comprehensive analysis of the integrated fusion problem to formulate a suitable spatial degradation model. It is assumed that the observed degraded images can be obtained by warping, blurring, downsampling, and noise operators performed on the desired image [12], [13], [15], [65], and it is denoted by the following expression:

$$\mathbf{y}_{k,b} = \mathbf{D}\mathbf{S}_{k,b}\mathbf{M}_k\mathbf{x}_b + \mathbf{v}_{k,b} \qquad (1 \le k \le K) \qquad (1)$$

where $\mathbf{y}_{k,b}$ denotes the bth band of the kth image in the collection of degraded observed images $\mathbf{Y}$. $\mathbf{x}_b$ represents the bth band of the desired fused image $\mathbf{x}$, $\mathbf{M}_k$ denotes the motion matrix, $\mathbf{S}_{k,b}$ represents the point spread function blurring [25], [48], $\mathbf{D}$ is the downsampling matrix, and $\mathbf{v}_{k,b}$ is the zero-mean Gaussian noise [9], which is caused by sensors and the external environment. For convenience, (1) can be expressed as

$$\mathbf{y}_{k,b} = \mathbf{A}_{y,k,b}\mathbf{x}_b + \mathbf{v}_{k,b} \qquad (1 \le k \le K) \qquad (2)$$

where $\mathbf{A}_{y,k,b} = \mathbf{D}\mathbf{S}_{k,b}\mathbf{M}_k$.

2) Spatio–Temporal–Spectral Relationship Model: The spatio–temporal–spectral relationship model relates the desired image $\mathbf{x}$ to the multisource high-spatial-resolution observations $\mathbf{Z}$. The multisource images in the collection $\mathbf{Z}$ may have different spatial scales, different spectral resolutions and spectral ranges, and different temporal resolutions and imaging times. Therefore, the relationship between $\mathbf{x}$ and $\mathbf{Z}$ should be comprehensively analyzed to construct the spatio–temporal–spectral relationship model. For convenience, the model in terms of $\mathbf{z}_{n,q}$ is represented as follows:

$$\mathbf{z}_{n,q} = \mathbf{\Psi}_{n,q}\mathbf{C}_{n,q}\mathbf{A}_{z,n,q}\mathbf{x} + \boldsymbol{\tau}_{n,q}$$
$$+ \boldsymbol{v}_{n,q} \,(1 \le q \le B_{z,n}, \quad 0 \le n \le N) \quad (3)$$

where $\mathbf{z}_{n,q}$ denotes the qth band of the nth image $\mathbf{z}_n$ in the collection of multisource high-spatial-resolution observations $\mathbf{Z}$, the total number of spectral bands of $\mathbf{z}_n$ is $B_{z,n}$, and $\mathbf{A}_{z,n,q}$ denotes the spatial degradation matrix between the desired image and the multiscale high-spatial-resolution images; it is similar to $\mathbf{A}_{y,k,b}$ in (2). $\mathbf{C}_{n,q}$ denotes the spectral correlation matrix, $\mathbf{\Psi}_{n,q}$ is the temporal correlation matrix considering the temporal features between the desired image and $\mathbf{z}_{n,q}$, $\boldsymbol{\tau}_{n,q}$ is the offset, and $\boldsymbol{v}_{n,q}$ denotes the noise.

C. Integrated Fusion Model

1) MAP Formulation: The MAP theory has been widely applied in the description of ill-posed inverse problems such as denoising [66], destriping and inpainting [67], and image fusion [68]. In this paper, the MAP theory is used to describe the proposed integrated fusion problem, and the aim is to estimate the desired image $\mathbf{x}$ based on the multisource observed images $\mathbf{Y}$ and $\mathbf{Z}$. We can write this as

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} \ p(\mathbf{x}|\mathbf{Y}, \mathbf{Z}). \qquad (4)$$

Applying Bayes' rules, (4) becomes

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} \frac{p(\mathbf{Y}, \mathbf{Z}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{Y}, \mathbf{Z})}. \qquad (5)$$

Since $p(\mathbf{Y}, \mathbf{Z})$ is independent of $\mathbf{x}$, it can be thus considered as a constant to be removed in (5), and hence, we can write (5) as

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} p(\mathbf{Y}, \mathbf{Z}|\mathbf{x})p(\mathbf{x})$$
$$= \arg\max_{\mathbf{x}} p(\mathbf{Y}|\mathbf{x})p(\mathbf{Z}|\mathbf{x}, \mathbf{Y})p(\mathbf{x})$$
$$= \arg\max_{\mathbf{x}} p(\mathbf{Y}|\mathbf{x})p(\mathbf{Z}|\mathbf{x})p(\mathbf{x}). \qquad (6)$$

Since $\mathbf{Y}$ and $\mathbf{Z}$ are known quantities, it is therefore tenable that $p(\mathbf{Z}|\mathbf{x}, \mathbf{Y}) = p(\mathbf{Z}|\mathbf{x})$ in (6).

2) Fusion Model: It can be seen that there are three probability density functions involved in (6). The first function $p(\mathbf{Y}|\mathbf{x})$ provides a measure of the coherence between the desired image $\mathbf{x}$ and the degraded observed images $\mathbf{Y}$, as given by the spatial degradation model (2), and it is determined by the probability density function of the noise in (2). The noise is assumed to be Gaussian, independent and identically distributed [69], [70]; hence, $p(\mathbf{Y}|\mathbf{x})$ can be written as

$$p(\mathbf{Y}|\mathbf{x}) = \prod_{k=1}^{K}\prod_{b=1}^{B_x} p(\mathbf{y}_{k,b}|\mathbf{x}_b) \qquad (7)$$

$$p(\mathbf{y}_{k,b}|\mathbf{x}_b) = \frac{1}{(2\pi a_{y,k,b})^{\Phi_1\Phi_2/2}}$$
$$\exp\left\{-\|\mathbf{y}_{k,b} - \mathbf{A}_{y,k,b}\mathbf{x}_b\|_2^2/2a_{y,k,b}\right\} \qquad (8)$$

where $a_{y,k,b}$ is the variance of the noise $\mathbf{v}_{k,b}$, B_x is the number of spectral bands, $\Phi_1\Phi_2$ denotes the spatial dimension of $\mathbf{y}_{k,b}$, and $\|\cdot\|_2$ denotes the ℓ_2-norm.

The second probability density function $p(\mathbf{Z}|\mathbf{x})$ provides a measure of the accordance of $\mathbf{Z}$ and $\mathbf{x}$ to the spatio–temporal–spectral relationship model (3). It is determined by the probability distribution of the noise in (3), and we can argue the expression in

$$p(\mathbf{Z}|\mathbf{x}) = \prod_{n}^{N}\prod_{q=1}^{B_{z,n}} p(\mathbf{z}_{n, q}|\mathbf{x}) \qquad (9)$$

$$p(\mathbf{z}_{n, q}|\mathbf{x}) = \frac{1}{(2\pi\alpha_{z,n,q})^{H_{n,1}H_{n,2}/2}}$$
$$\exp\{-\|\mathbf{z}_{n,q} - \mathbf{\Psi}_{n,q}\mathbf{C}_{n,q}\mathbf{A}_{z,n,q}\mathbf{x} - \boldsymbol{\tau}_{n,q}\|_2^2/2\alpha_{z,n,q}\} \qquad (10)$$

where $a_{z,n,q}$ denotes the variance of the noise, and $H_{n,1}H_{n,2}$ denotes the spatial dimension of $\mathbf{z}_{n,q}$

The third probability density function $p(\mathbf{x})$ is the image prior term. Inspired by [71], an adaptive weighted 3-D spatio–spectral Laplacian prior is introduced, which is represented as

$$p(\mathbf{x}) = \prod_{b=1}^{B_x} \frac{1}{(2\pi a_{x,b})^{L_1L_2/2}} \exp\left\{-\|\mathbf{Q}\mathbf{x}_b\|_2^2/2a_{x,b}\right\} \qquad (11)$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

SHEN *et al.*: FRAMEWORK FOR THE SPATIO–TEMPORAL–SPECTRAL FUSION OF REMOTE SENSING IMAGES

5

$$\theta_d = \frac{\sum\limits_{b=1}^{B_x} [\nabla F(\mathbf{x}_b)_d]^T [\nabla F(\mathbf{x}_b)_d]}{\sum\limits_{b=1}^{B_x} [\mathbf{e}_{b,d}]^T \left[\sum\limits_{k=1}^{K} \mathbf{A}_{y,k,b}^T \mathbf{A}_{y,k,b} + \lambda_1 \sum\limits_{n}^{N} \sum\limits_{q=1}^{B_{z,n}} w_{n,q} \mathbf{A}_{z,n,q}^T \mathbf{C}_{n,q,b}^T \mathbf{\Psi}_{n,q}^T \mathbf{\Psi}_{n,q} \mathbf{C}_{n,q} \mathbf{A}_{z,n,q} + \lambda_2 \mathbf{Q}^T \mathbf{Q} \right] [\mathbf{e}_{b,d}]} \tag{18}$$

where $\alpha_{x,b}$ is the variance of the noise, $L_1 L_2$ denotes the spatial dimension of $\mathbf{x}$, and $\mathbf{Q}$ indicates the adaptive weighted 3-D Laplacian matrix. For convenience, we write this as the expression

$$\begin{aligned}
Q\mathbf{x}_b(i,j) &= Q_{\text{spa}}\mathbf{x}_b(i,j) + \beta Q_{\text{spe}}\mathbf{x}_b(i,j) \\
&= \mathbf{x}_b(i+1,j) + \mathbf{x}_b(i-1,j) \\
&\quad + \mathbf{x}_b(i,j+1) + \mathbf{x}_b(i,j-1) - 4\mathbf{x}_b(i,j) \\
&\quad + \beta \left(\frac{\|\tilde{\boldsymbol{x}}_b\|_2}{\|\tilde{\boldsymbol{x}}_{b+1}\|_2} \mathbf{x}_{b+1}(i,j) \right. \\
&\quad\quad \left. + \frac{\|\tilde{\boldsymbol{x}}_b\|_2}{\|\tilde{\boldsymbol{x}}_{b-1}\|_2} \mathbf{x}_{b-1}(i,j) - 2\mathbf{x}_b(i,j) \right)
\end{aligned} \tag{12}$$

where $\tilde{\mathbf{x}}$ is the initial estimate of the fused image obtained by resampling of the corresponding observation, and β is the parameter, which is adaptively determined as

$$\beta = \begin{cases} \exp\left(-\frac{1}{B_x} \sum\limits_{b=1}^{B_x} \|\nabla\tilde{\boldsymbol{x}}_b\|_2 / L_1 L_2 B_x \right), & B_x > u \\ 0, & B_x \le u \end{cases} \tag{13}$$

where u is the threshold parameter. It is assumed that, when the desired image has only a few spectral bands, their spectral curves are discontinuous; therefore, it is tenable and more robust to set $\beta = 0$. Conversely, if the desired image is a hyperspectral image, the spectral curve can be assumed to be approximately continuous, and thus, the adaptive weighted prior term can effectively preserve the shape of the spectral curve to decrease the spectral distortion. It is assumed that the smaller the difference between spectral bands, the stronger the spectral constraints. $\nabla\tilde{\mathbf{x}}_b$ denotes the gradient in the spectral dimension of the bth band. Here, the parameter u is set to 8, with the spectral band number of WorldView-2 multispectral images as a reference.

Substituting (7)–(11) into (6), through the implementation of the monotonic logarithm function and the simplification, many parameters can be safely dropped, and the final energy function can be expressed as the regularized minimization problem in

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}}[F(\mathbf{x})] \tag{14}$$

$$\begin{aligned}
F(\mathbf{x}) &= \frac{1}{2} \sum_{k=1}^{K} \sum_{b=1}^{B_x} \|\mathbf{y}_{k,b} - \mathbf{A}_{y,k,b}\mathbf{x}_b\|_2^2 \\
&\quad + \frac{\lambda_1}{2} \sum_{n}^{N} \sum_{q=1}^{B_{z,n}} w_{n,q} \|\mathbf{z}_{n,q} - \mathbf{\Psi}_{n,q}\mathbf{C}_{n,q}\mathbf{A}_{z,n,q}\mathbf{x} - \boldsymbol{\tau}_{n,q}\|_2^2 \\
&\quad + \frac{\lambda_2}{2} \sum_{b=1}^{B_x} \|\mathbf{Q}\mathbf{x}_b\|_2^2
\end{aligned} \tag{15}$$

where the first term denotes the coherence between $\mathbf{x}$ and $\mathbf{Y}$, the second term represents the relationship between $\mathbf{x}$ and $\mathbf{Z}$, and the third term denotes the image prior. $w_{n,q}$ denotes the relative contribution of $\mathbf{z}_{n,q}$ to the desired image $\mathbf{x}$, and it is adaptively calculated as $w_{n,q} = \lambda'_{n,q} U_n$. U_n is calculated by the correlation [72] between the degraded version of $\mathbf{z}_n$ and the observed image corresponding to $\mathbf{x}$, and it is assumed that the greater the correlation, the greater the contribution of the observation $\mathbf{z}_n$. The auxiliary parameter $\lambda'_{n,q}$ is used to adaptively adjust the balance of the spatial details to each band of the fused image, and it is expressed as $\lambda'_{n,q} = f(\mathbf{z}_{n,q}, \mathbf{x})/\min[f(\mathbf{z}_{1,1}, \mathbf{x}), \ldots, f(\mathbf{z}_{1,q}, \mathbf{x}), \ldots, f(\mathbf{z}_{n,q}, \mathbf{x}), \ldots]$, where $f(\mathbf{z}_{n,q}, \mathbf{x})$ denotes the number of spectral bands to be fused by $\mathbf{z}_{n,q}$. λ_1 and λ_2, which are related to the variance of the noise, are the parameters used to control the relative contribution of the three terms. Because there are parameters in the second and third terms, there is no need to add extra parameters in the first term.

IV. OPTIMIZATION SOLUTION AND PARAMETER DETERMINATION

A. Optimization Procedure

The fused image is solved by the conjugate gradient algorithm [73], differentiating the energy function (15) with respect to $\mathbf{x}_b$, as

$$\begin{aligned}
\nabla F(\mathbf{x}_b) &= -\sum_{k=1}^{K} \mathbf{A}_{y,k,b}^T (\mathbf{y}_{k,b} - \mathbf{A}_{y,k,b}\mathbf{x}_b) \\
&\quad - \lambda_1 \sum_{n}^{N} \sum_{q=1}^{B_{z,n}} w_{n,q} \mathbf{A}_{z,n,q}^T \mathbf{C}_{n,q,b}^T \mathbf{\Psi}_{n,q}^T (\mathbf{z}_{n,q} \\
&\quad - \mathbf{\Psi}_{n,q}\mathbf{C}_{n,q}\mathbf{A}_{z,n,q}\mathbf{x} - \boldsymbol{\tau}_{n,q}) + \lambda_2 \mathbf{Q}^T \mathbf{Q}\mathbf{x}_b.
\end{aligned} \tag{16}$$

The desired fused image can be then solved by successive approximation iterations, i.e.,

$$\mathbf{x}_{b,d+1} = \mathbf{x}_{b,d} + \theta_d \mathbf{e}_{b,d} \tag{17}$$

where $\mathbf{e}_{b,d}$ denotes the search direction of the dth iteration, and the initial value is negative of the gradient $\nabla F(\mathbf{x}_b)$, i.e., $\mathbf{e}_{b,1} = -\nabla F(\mathbf{x}_b)_1$. θ_d is the step size of the dth iteration, as in (18), shown at the top of the page.

The next search direction is then relevant to the current iteration and the previous search direction and is represented as

$$\mathbf{e}_{b,d+1} = -\nabla F(\mathbf{x}_b)_{d+1} + \gamma_d \mathbf{e}_{b,d} \tag{19}$$

with (20) and (21), shown at the bottom of the next page.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

6   IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING

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

$$\|\hat{\boldsymbol{x}}_{d+1} - \hat{\boldsymbol{x}}_d\|_2^2 / \|\hat{\boldsymbol{x}}_d\|_2^2 \leq \varsigma \tag{22}$$

where ς is a predetermined threshold value.

B. Determination of the Spectral and Temporal Correlations

The spatial, temporal, and spectral relationships between the desired image and the multisource remote sensing observations are comprehensively considered in the integrated relationship model. In the model, two issues need to be solved, i.e., the determination of the spectral correlation matrix and the temporal correlation matrix in (3).

1) Spectral Correlation $\mathbf{C}$*:* The spectral correlation matrix plays an important role in the integrated relationship model. In general, the images in the collection of multisource high-spatial-resolution observations $\mathbf{Z}$ have a relatively wider bandwidth than the desired image $\mathbf{x}$. Therefore, a popular approach is to relate $\mathbf{x}$ and the images in $\mathbf{Z}$ by linear band combinations [36]–[38] , [61]–[63], [68]. For convenience, we take PAN/MS fusion [68] as an example, which is written as

$$\mathbf{z}(i,j) = \sum_{b=1}^{B_x} c_b \mathbf{x}_b(i,j) + \boldsymbol{\tau} + \boldsymbol{v}(i,j) \tag{23}$$

where $\mathbf{z}(i,j)$ is the brightness value of the wideband PAN image, with (i,j) being the pixel position, and $\{c_b\}$ is the linear combination coefficients.

To date, a variety of fusion methods have been involved in searching for the solution of the linear combination coefficients $\{c_b\}$, and most of them employ the spectral response functions of the sensors to calculate the linear coefficients [68], [74]. However, in fact, the spectral response functions of the sensors may differ from the responses of the observed data sets due to the atmospheric effects, illumination conditions, etc. [51], [75]. Furthermore, if we cannot easily obtain the spectral response functions, then the image fusion methods will not work. In this paper, it is assumed that the linear coefficients $\{c_b\}$ are not only about the fitting between $\mathbf{z}$ and $\mathbf{x}$; furthermore, they also play an important role in assigning the relative weight of the spatial structure information to each band of the fused image. Inspired by this idea, a novel and simple solution for $\{c_b\}$ is proposed by
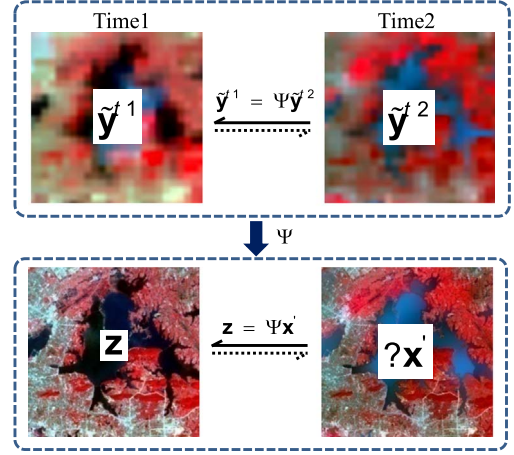

Fig. 2. Schematic for determination of the temporal correlation.

considering the internal correlation between the spectral bands. This is represented as

$$c_b = \frac{\mathrm{cov}(\tilde{\boldsymbol{x}}_b, \mathbf{I}_{\mathrm{syn}})}{\mathrm{var}(\mathbf{I}_{\mathrm{syn}})} \tag{24}$$

where $\mathrm{cov}(\cdot)$ denotes the covariance, and $\mathrm{var}(\cdot)$ denotes the variance. $\tilde{\mathbf{x}}_b$ is the resampled bth band of the observed narrowband image, and $\mathbf{I}_{\mathrm{syn}}$ denotes the intensity by the generalized intensity–hue–saturation [29], [76] transformation of the corresponding spectral bands. Finally, the linear combination coefficients are obtained by normalization as $c_b/\mathrm{sum}(\{c_b\})$.

2) Temporal Correlation $\boldsymbol{\Psi}$*:* It is assumed that the relationship between the multitemporal scenes can be expressed by a linear model [62], [77]–[79]. To improve the robustness, a space-varying multitemporal linear relationship model is adopted by comprehensively analyzing the integrated fusion problem. The schematic for the determination of the temporal relationship is shown in Fig. 2. It is shown that the temporal relationship $\boldsymbol{\Psi}$ can be approximately obtained by the corresponding multitemporal low-spatial-resolution remote sensing images $\tilde{\mathbf{y}}^{t1}$, $\tilde{\mathbf{y}}^{t2}$ [62], [79]. To ensure the effectiveness of the relationship mapping, the spectral range of $\tilde{\mathbf{y}}^{t1}$, $\tilde{\mathbf{y}}^{t2}$ should be consistent with $\mathbf{z}$; for example, in the case of spatio–temporal–spectral fusion, $\tilde{\mathbf{y}}^{t1}$, $\tilde{\mathbf{y}}^{t2}$ should be first obtained by band combination of the corresponding observations. Here, $\mathbf{x}'$ has the same spatial scale and spectral range as $\mathbf{z}$, and it can be approximately represented as $\mathbf{x}' = \mathbf{CAx}$, where $\mathbf{A}$ denotes the spatial degraded matrix, and $\mathbf{C}$ represents the spectral correlation matrix, as shown in (3), and when it is a case of spatio–temporal fusion, $\mathbf{A}$ and $\mathbf{C}$ are generally the identity matrices.

$$\nabla F(\mathbf{x}_b)_{d+1} = \nabla F(\mathbf{x}_b)_d + \theta_d \left[\sum_{k=1}^{K} \mathbf{A}_{y,k,b}^T \mathbf{A}_{y,k,b} + \lambda_1 \sum_{n}^{N} \sum_{q=1}^{B_{z,n}} w_{n,q} \mathbf{A}_{z,n,q}^T \mathbf{C}_{n,q,b}^T \boldsymbol{\Psi}_{n,q}^T \boldsymbol{\Psi}_{n,q} \mathbf{C}_{n,q} \mathbf{A}_{z,n,q} + \lambda_2 \mathbf{Q}^T \mathbf{Q} \right] \mathbf{e}_{b,d} \tag{20}$$

$$\gamma_d = \frac{\sum_{b=1}^{B_x} [\nabla F(\mathbf{x}_b)_{d+1}]^T [\nabla F(\mathbf{x}_b)_{d+1}]}{\sum_{b=1}^{B_x} [\nabla F(\mathbf{x}_b)_d]^T [\nabla F(\mathbf{x}_b)_d]} \tag{21}$$

TABLE II
QUANTITATIVE EVALUATION RESULTS OF THE MULTIVIEW
SPATIAL FUSION EXPERIMENTS

	Nearest neighbor	Bilinear	Proposed method
CC	0.9342	0.9428	0.9763
SSIM	0.5544	0.5570	0.7930
RMSE	19.9851	18.8745	12.1370
PSNR	22.1166	22.6132	26.4485

The space-varying temporal relationship is solved with the aid of neighboring similar pixels. The initial similar pixel locations are determined in the high-spatial-resolution image $\mathbf{z}$, and they are selected based on a nonlocal strategy [80]. The selection rule is

$$\exp\left(\frac{-\|\mathbf{G}(\mathbf{O}_z(i) - \mathbf{O}_z(j))\|_2^2}{h^2}\right) \geq \eta, \qquad j \in W_i \quad (25)$$

where $\mathbf{O}_z(i)$ denotes the 3-D patch at the central pixel position i, and $\mathbf{O}_z(j)$ denotes the 3-D neighboring patch at pixel position j within a window W_i. $\mathbf{G}$ is the kernel function with inverse distance weighting, h is the standard deviation of the central pixel patch, and η is the threshold value.

After the initial similar pixel locations are determined, an additional filtering process is applied to the candidates to improve the accuracy. The basic rule is that the more consistent the temporal change between the similar pixels and the central pixel, the greater the probability of the similar pixels' selection. This is expressed as

$$\mathbf{E} = \{\tilde{\mathbf{y}}_j^{t1} - \tilde{\mathbf{y}}_j^{t2} | j \in \Omega_i\}, \Rightarrow \text{abs}(\mathbf{E}_j - \mathbf{E}_i) \leq \delta_E \; j \in \Omega_i \quad (26)$$

where $\mathbf{E}$ denotes the temporal change of the similar pixels, Ω_i denotes the collection of initial similar pixel locations for the central pixel position i, abs$(\cdot)$ represents the absolute value function, and σ_E is the standard variation of $\mathbf{E}$.

After the similar pixels are finally selected in $\tilde{\mathbf{y}}^{t1}$, $\tilde{\mathbf{y}}^{t2}$, robust regression analysis using iteratively reweighted least squares with a Huber weighting function [81] is utilized to solve the multitemporal correlation.

V. EXPERIMENTS

A. Data Sets and Experimental Settings

To comprehensively test the performance of the proposed integrated fusion framework, a variety of image fusion experiments were implemented with both simulated and real data. The experiments consisted of the following: 1) multiview spatial fusion experiments; 2) spatio–spectral fusion experiments; 3) spatio–temporal fusion experiments; and 4) spatio–temporal–spectral fusion experiments. Furthermore, the proposed method was comprehensively tested and verified using data from multiple different remote sensing satellites, including IKONOS, Landsat Enhanced Thematic Mapper Plus (ETM+), MODIS, Hyperspectral Digital Imagery Collection Experiment (HYDICE), and SPOT-5.

There are several parameters in the experiments. However, it is noteworthy that most of the parameters are not sensitive to the fusion model for different fusion tasks, and hence, most of them were empirically set to be fixed values. The parameter λ_1 was the sole tunable parameter in the experiments. Unless
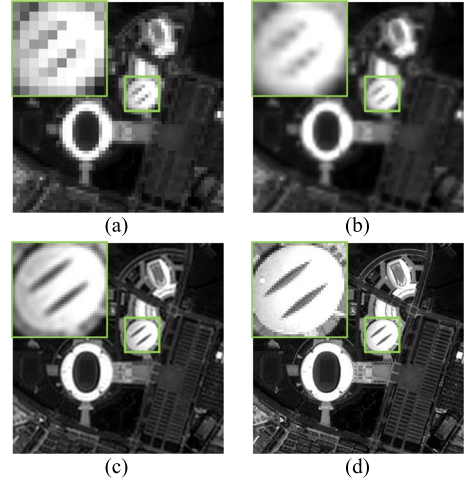

Fig. 3. Experimental results in multiview spatial fusion. (a) Nearest neighbor interpolation. (b) Bilinear interpolation. (c) Result of the proposed method. (d) Original image.

otherwise specified, the parameter λ_2 was empirically set to 0.001, and the iterative threshold value ς was empirically set to 1e-7. In the spatio–temporal and spatio–temporal–spectral fusion experiments, the patch size in (25) was set to 7×7, the window size of W was set to 23×23, and the threshold value η was set to 1e-3 empirically. In addition, the motion and the blurring of (3) have little influence to the experiments, considering the efficiency and effectiveness; they are hence set to be identity matrix.

B. Experimental Results and Analysis

1) Multiview Spatial Fusion Experiments: The multiview spatial fusion experiments were implemented with simulated data, and a 256×256 pixel subarray of the blue band of the original IKONOS MS image [see Fig. 3(d)], which was acquired in Hubei Province, China, on September 4, 2009, was used to generate four simulated low-spatial-resolution images with subpixel displacements. The subpixel shifts were (0, 0), (0.5, 0), (0, 0.5), and (0.5, 0.5), respectively; and the spatial dimension size of each degraded image was 64×64. The popular quantitative indexes of the correlation coefficient (CC) [20], [68], the structural similarity index (SSIM) [82], [83], the peak signal-to-noise ratio (PSNR) [82], [84], and the root-mean-square error (RMSE) [68] were utilized to evaluate the fused results.

Fig. 3(c) shows the fused result obtained by the proposed integrated fusion framework with $\lambda_1 = 0$. Fig. 3(a) and (b) shows the results of the nearest neighbor and bilinear interpolation algorithms, respectively, for comparison. It can be observed that the result of the proposed fusion framework has more detailed information and a better visual quality than the traditional interpolation methods. A more detailed visual comparison can be made from the subregions of the images. The quantitative evaluation results in Table II also show that the proposed fusion framework produces results that are superior to the results of the traditional methods. Therefore, it is demonstrated that the proposed integrated fusion framework can make full use of the redundancy and complementary information to improve the spatial resolution of the degraded images.
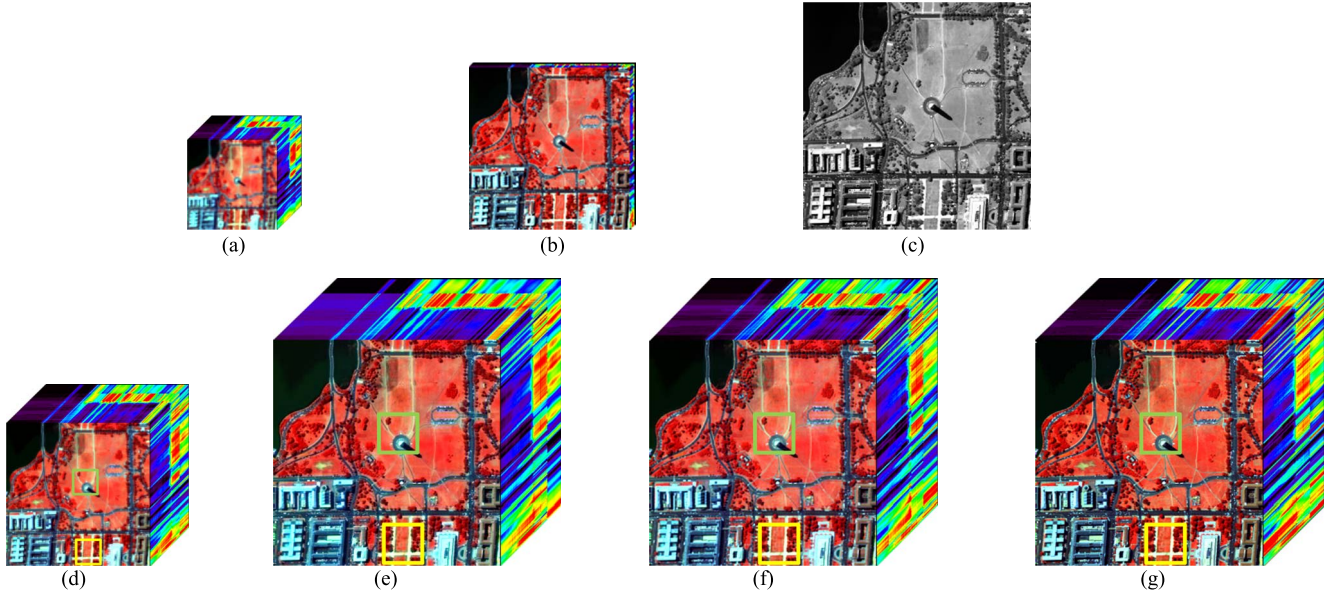
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

8   IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING

Fig. 4.   Experimental results in spatio–spectral fusion (displayed as false color: near infrared, green, and blue). (a) HS image. (b) MS image. (c) PAN image. (d) MS/HS fusion result. (e) PAN/HS fusion result. (f) PAN/MS/HS fusion result. (g) Original HS image.

TABLE III
QUANTITATIVE EVALUATION RESULTS OF
THE SPATIO–SPECTRAL EXPERIMENTS

	MS/HS	PAN/HS	PAN/MS/HS
CC	0.9355	0.9479	0.9731
SSIM	0.6084	0.6499	0.7557
PSNR	16.7513	17.0602	17.8347
ERGAS	9.8455	9.6495	8.8213
SAM	7.2499	10.7833	8.4942
Q$_{-avg}$	0.7300	0.7490	0.8101

2) Spatio–Spectral Fusion Experiments: In the spatio–spectral fusion experiments, three fusion cases were implemented by the proposed integrated fusion framework: 1) PAN/HS fusion; 2) MS/HS fusion; and 3) PAN/MS/HS fusion.

The experiments were performed on the simulated data sets based on an original HYDICE HS image [85] acquired over the Washington DC Mall. The simulated HS image was obtained by low-pass filtering using Gaussian modulation transfer function blurring and downsampling by a factor of 4 in the spatial domain. The spatial resolution of the simulated degraded HS image was 4 m, and it had 79 bands with the spectral range of 450–1750 nm. The MS image was obtained by degradation in both the spatial and spectral domains. In the spatial dimension, it was downsampled by a factor of 2, and its spatial resolution was 2 m; in the spectral dimension, it was produced according to the spectral characteristics of Landsat-7 ETM+ bands 1–5, which cover the 450- to 515-nm, 525- to 605-nm, 630- to 690-nm, 750- to 900-nm, and 1550- to 1750-nm regions, respectively. The PAN image with a 1-m spatial resolution was degraded in the spectral dimension, and it was produced according to the spectral characteristics of a Landsat-7 PAN image covering 520–900 nm. In the experiments, several widely used quality indexes were applied for the quantitative evaluation: CC, SSIM, PSNR, the dimensionless global error in synthesis (ERGAS) [39], [68], the spectral angle mapper (SAM) [39], [68], and Q$_{-avg}$ [8], [72].
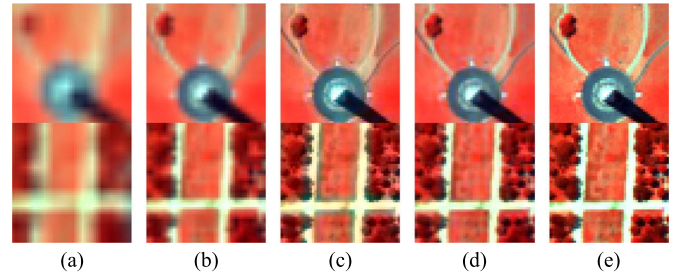

Fig. 5.   Local zoomed-in areas of the fusion results. (a) Resampled HS image with bilinear interpolation to the spatial scale of the original HS image. (b) Resampled MS/HS fusion result with bilinear interpolation to the spatial scale of the original HS image. (c) PAN/HS fusion result. (d) PAN/MS/HS fusion result. (e) Original HS image.

Fig. 4 shows the experimental results. The PAN/MS/HS fusion result is shown in Fig. 4(f), and the MS/HS and PAN/HS fusion results are shown in Fig. 4(d) and (e), respectively. For a more detailed comparison, local zoomed-in areas are shown in Fig. 5. It is noteworthy that we set the same parameter $\lambda_1 = 1$ for all three fusion cases, for a fair comparison. It can be seen that all three fusion cases can improve the spatial resolution of the HS image. However, there is still a distinct difference between the three fusion results. On one hand, in the aspect of spatial resolution enhancement, the MS/HS fusion can only improve the spatial resolution of the HS image from 4 to 2 m, whereas the PAN/HS and PAN/MS/HS fusion cases can improve the spatial resolution of the HS image from 4 to 1 m. Relatively speaking, the PAN/MS/HS fusion result has slightly better spatial details than the PAN/HS fusion result because of the incorporation of the MS images. On the other hand, in the aspect of spectral fidelity, the PAN/MS/HS fusion and the MS/HS fusion can effectively enhance all the spectral bands of the HS image; however, the PAN/HS fusion with a single PAN image cannot enhance all the spectral bands due to the incomplete spectral coverage between the PAN (520–900 nm) and HS (450–1320 nm) images. In addition, the PAN/MS/HS
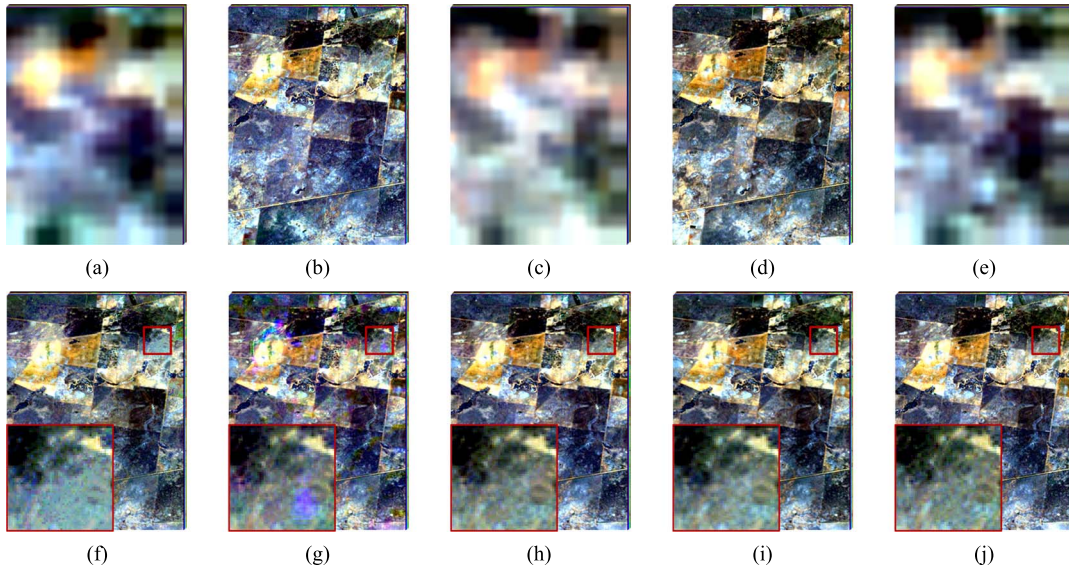
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

SHEN *et al.*: FRAMEWORK FOR THE SPATIO–TEMPORAL–SPECTRAL FUSION OF REMOTE SENSING IMAGES

9

Fig. 6. Experimental results in spatio–temporal fusion (displayed as band combination: red, green, blue). (a) MODIS image on October 8, 2001. (b) ETM+ image on October 8, 2001. (c) MODIS image on November 2, 2001. (d) ETM+ image on November 2, 2001. (e) MODIS image on October 17, 2001. (f) Predicted result on October 17, 2001 by the STARFM method. (g) Predicted result on October 17, 2001 by the SPSTFM method. (h) Predicted result on October 17, 2001 by the EBSPTF method. (i) Predicted result on October 17, 2001 by the proposed fusion framework. (j) Observed ETM+ image on October 17, 2001.

fusion result shows better spectral information preservation than the other two fusion cases. Table III shows the quantitative evaluation results. It should be noted that the MS/HS fusion was evaluated by bilinear resampling to the spatial scale of the original HS image, for a fair comparison. It can be seen that most of the evaluation indexes of the PAN/MS/HS fusion result are the best. On the whole, the PAN/MS/HS fusion result is the best in terms of both the spatial enhancement and the spectral fidelity. This is because the PAN/MS/HS fusion can make full use of the complementary spatial and spectral information of the PAN and MS groups.

3) Spatio–Temporal Fusion Experiments: The spatio–temporal fusion experiments were implemented on Landsat-7 ETM+ and MODIS terra MOD09GA image pairs [86], with spatial resolutions of 25 and 500 m, respectively. The images were from southern New South Wales (NSW, Australia; 34.0034° E, 145.0675° S). Fig. 6(a)–(d) shows two pairs of ETM+/MODIS observations on October 8, 2001 and November 2, 2001, respectively. Fig. 6(e) shows the MODIS image on October 17, 2001. The task of the fusion was to predict a Landsat-like image on October 17, 2001 by making use of Fig. 6(a)–(e). The observed ETM+ image shown in Fig. 6(j) is for reference. Fig. 6(i) shows the result of the proposed fusion framework with $\lambda_1 = 3$, and Fig. 6(f)–(h) shows the fusion results of the popular STARFM method [53], the SParse-representation-based SpatioTemporal reflectance Fusion Model (SPSTFM) [59], and the error-bound-regularized sparse coding (EBSPTM) [87] method, respectively, for comparison. Here, the default parameters consistent to the original published papers are used. For a more detailed comparison, local zoomed-in areas are shown. It can be observed that the proposed fusion framework can effectively predict the ETM+ image on October 17, 2001, and it is competitive compared with other methods. The quantitative results are shown in Table IV. It also demonstrates the good performance of the proposed fusion framework.

TABLE IV
QUANTITATIVE EVALUATION RESULTS OF THE
SPATIO–TEMPORAL FUSION EXPERIMENTS

	STARFM	SPSTFM	EBSPFM	Proposed
CC	0.8912	0.8833	0.9052	0.9074
SSIM	0.9625	0.9583	0.9650	0.9683
PSNR	41.1070	39.6698	40.9760	42.3954
ERGAS	2.2305	2.4208	2.2222	1.9979
SAM	2.1420	2.1860	1.9119	1.8270
Q-avg	1.4925	1.4922	1.4936	1.4935

4) Spatio–Temporal–Spectral Fusion Experiments: In the spatio–temporal–spectral fusion experiments, a SPOT-5 PAN image (5 m), a SPOT-5 MS image (10 m, band numbers: 1–4), an ETM+ PAN image (15 m), an ETM+ MS image (30 m, band numbers: 1–5, 7), and MODIS images (MOD02 1 km, band numbers: 1–12, 17–19, 26) were used. The spectral bands of the MODIS images were arranged according to the spectra of the SPOT-5 and ETM+ images. Fig. 7(a)–(g) shows the experimental data sets with multiple spatial, temporal, and spectral features. The main task of these spatio–temporal–spectral fusion experiments was to obtain high spatial–spectral resolution images on October 6, 2011. To comprehensively test the performance of the proposed integrated fusion framework, three groups of fusion experiments were implemented: 1) two-sensor fusion; 2) three-sensor fusion; and 3) five-sensor fusion. Table V shows the input data sets of the three experimental groups. The parameter λ_1 was set to 1 for all three cases.

The experimental results of the two-sensor fusion, the three-sensor fusion, and the five-sensor fusion are shown in Fig. 7(h)–(j), respectively. For a detailed comparison, local zoomed-in areas are shown. It can be clearly seen that all three fusion cases can effectively predict the result on October 6, 2011. In comparison, the five-sensor fusion result has the highest spatial resolution of 5 m, the three-sensor fusion result has a spatial resolution of 15 m, and the two-sensor fusion result has a spatial resolution of 30 m. Furthermore, although
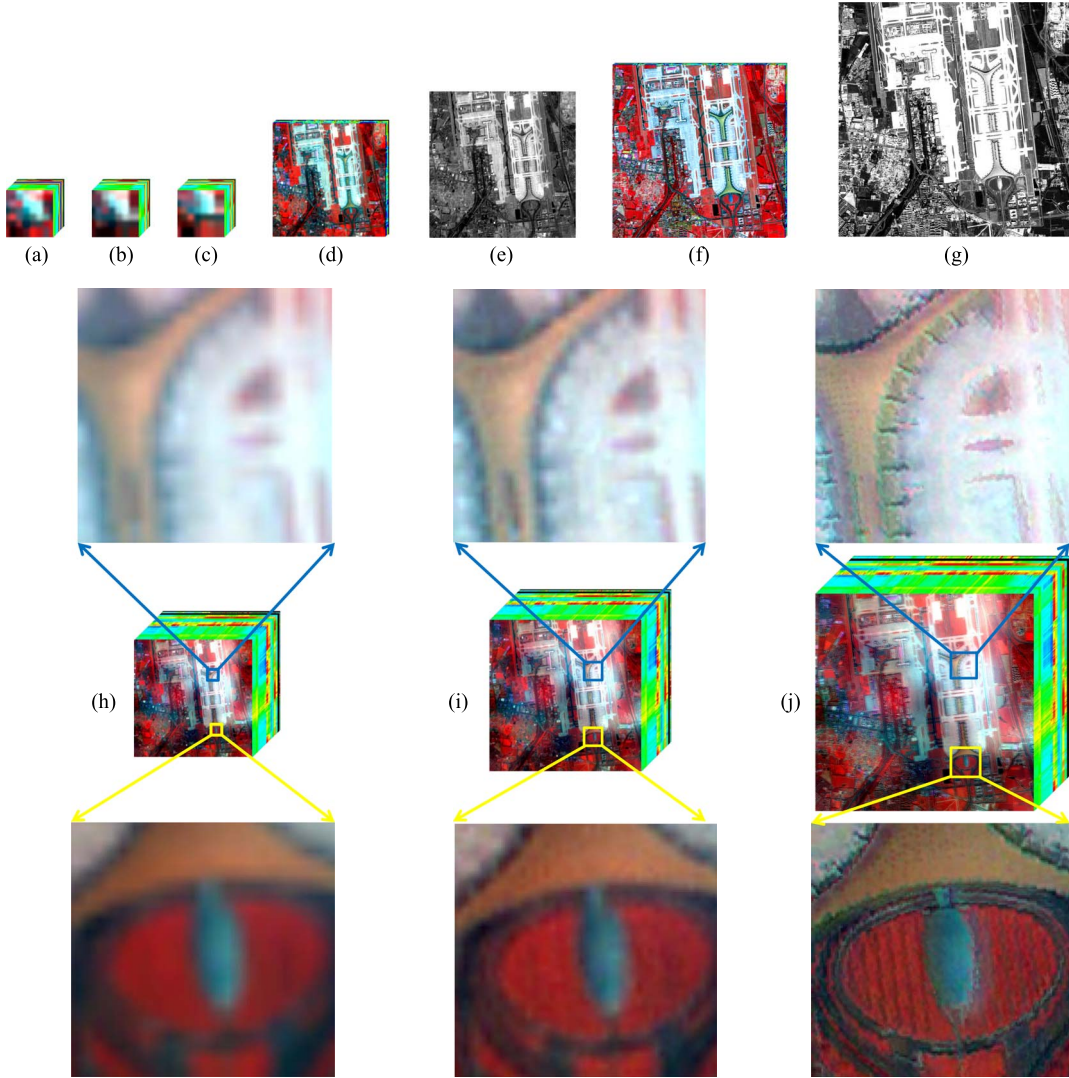
Fig. 7. Experiments in spatio–temporal–spectral fusion (displayed as false color). (a) MODIS image on September 4, 2011. (b) MODIS image on October 6, 2011. (c) MODIS image on September 22, 2011. (d) ETM+ MS image on September 4, 2011. (e) ETM+ PAN image on September 4, 2011. (f) SPOT-5 MS image on September 22, 2011. (g) SPOT-5 PAN image on September 22, 2011. (h) Fusion result on October 6, 2011 by two sensors, where the inputs are observations (a), (b), and (d). (i) Fusion result on October 6, 2011 by three sensors, where the inputs are observations (a), (b), (d), and (e). (j) Fusion result on October 6, 2011 by five sensors, where the inputs are all the observations (a)–(g).

TABLE V
THREE GROUPS OF SPATIO–TEMPORAL–SPECTRAL
FUSION EXPERIMENTS

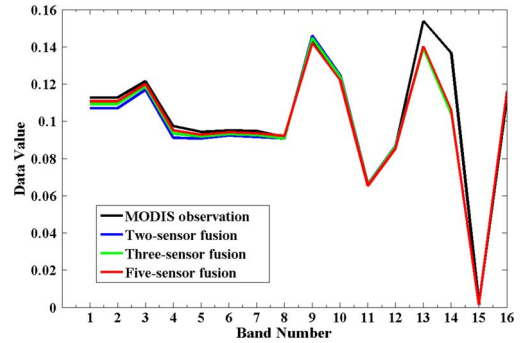| Experimental groups	Input observations	Sensors
Two-sensor fusion	Fig. 7 (a), Fig. 7 (b), and Fig. 7 (d)	ETM+ MS and MODIS
Three-sensor fusion	Fig. 7 (a), Fig. 7 (b), Fig. 7 (d), and Fig. 7 (e)	ETM+ PAN, ETM+ MS, and MODIS
Five-sensor fusion	All observations of Fig. 7 (a)–Fig. 7 (g)	SPOT-5 PAN, SPOT-5 MS, ETM+ PAN, ETM+ MS, and MODIS

Fig. 8. Spectral profiles of the original MODIS observation on October 6, 2011 and the fusion results of the three experimental groups, i.e., two-sensor fusion, three-sensor fusion, and five-sensor fusion.

the spatial resolution of the five-sensor fusion result has been greatly improved, the spectral information is preserved very well, which can be clearly seen in the local zoomed-in areas in Fig. 7. In addition, this can be also seen in the spectral profiles in Fig. 8. On the whole, all the spectral profiles of the three fusion cases are consistent with the MODIS observations. However, there is a slight difference in the spectral profiles of the 13th and 14th bands between the fusion results and the MODIS observations. This is because the wavelength of the 13th spectral band is far from the relevant spectral bands of the SPOT-5 and ETM+ images. Thus, it causes slight spectral distortion of the 13th and 14th bands.
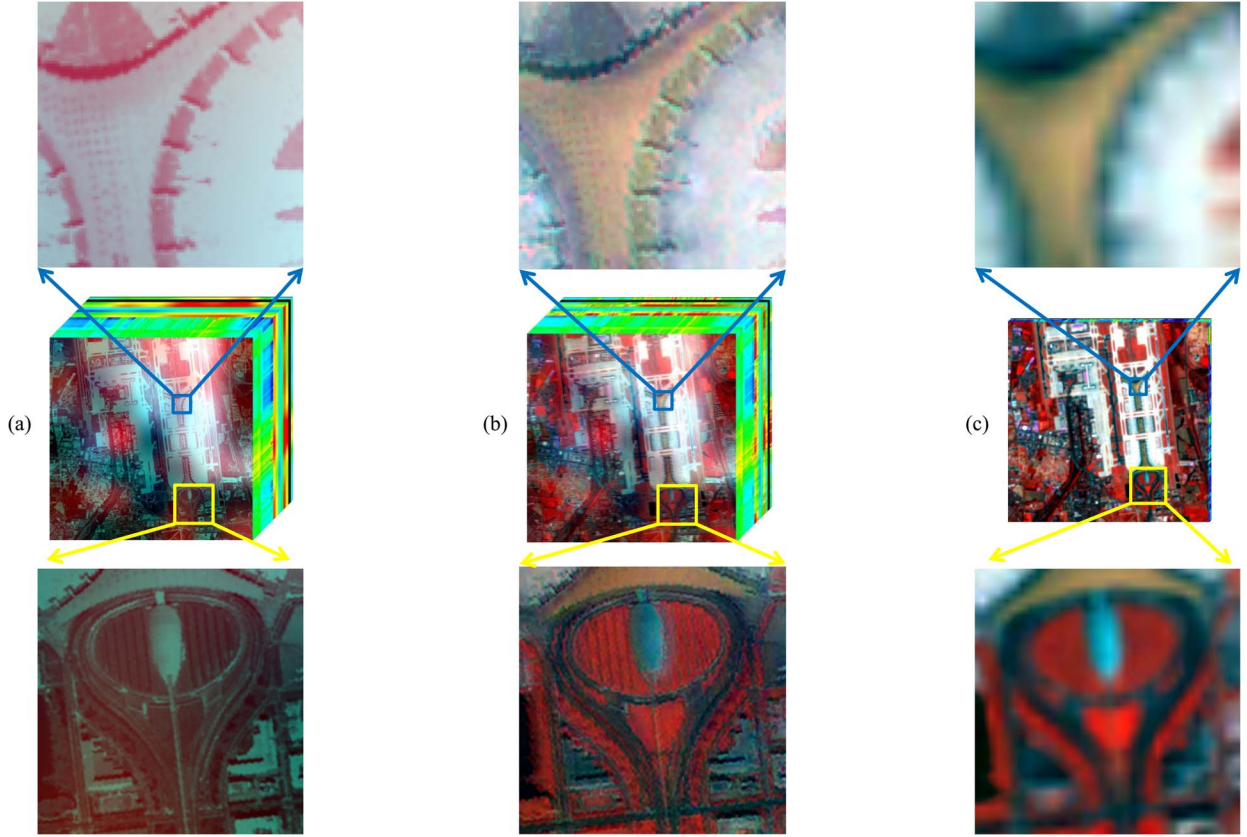
Fig. 9. Performance verification of multisensor fusion. (a) Fusion result on October 6, 2011 by two sensors with only the highest and lowest spatial resolution observations, i.e., Fig. 7(b), (c), and (g). (b) Fusion result on October 6, 2011 by five sensors using Fig. 7(a)–(g). (c) Bilinear interpolated ETM+ MS on October 6, 2011 with the consistent spectral display for the spectral feature validation.

To further verify the effectiveness of the multisensor fusion, a two-sensor fusion experiment with only the highest and lowest spatial resolution observations was implemented, i.e., the observations in Fig. 7(b), (c), and (g). A comparison between the two-sensor and five-sensor fusion results [i.e., Fig. 7(j)] is shown in Fig. 9. The bilinear interpolated ETM+ MS image with consistent spectra on the same day, i.e., October 6, 2011, is shown in Fig. 9(c) for the identification of the spectral features. Visually, it can be clearly seen that both the two-sensor and five-sensor fusion results have promoted the spatial structure; however, some regions of the two-sensor fusion result are too homogenized. In addition, the two-sensor fusion result shows a serious loss of spectral information. This can be clearly seen by a comparison with Fig. 9(c), and it can be also seen in the spectral profiles in Fig. 10. This is because the difference in spatial resolution between the SPOT-5 PAN (5 m) and MODIS observations (1000 m) is too great, and the complementary information is not enough for such a challenging fusion process. Conversely, the five-sensor fusion can make full use of the complementary information of the multisource observations, which fully demonstrates the advantage of the proposed integrated fusion framework. However, this does not mean that more sensors will always lead to better fusion results. Several factors are involved, including the redundancy, the complementarity, and the image quality.
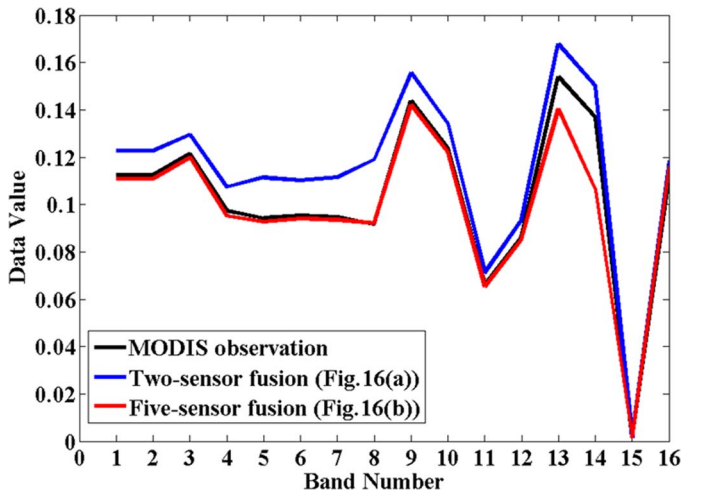

Fig. 10. Spectral profiles of the two-sensor fusion result with only the highest and lowest spatial resolution observations, the five-sensor fusion result, and the original MODIS observation on October 6, 2011.

VI. Conclusion

This paper has presented an integrated framework for the spatio–temporal–spectral fusion of remote sensing images. The proposed integrated fusion framework was comprehensively tested and verified by a variety of image fusion experiments involving multiple remote sensing satellites, including IKONOS,

ETM+, MODIS, HYDICE, and SPOT-5. The experimental results confirm the good performance of the proposed fusion framework.

In conclusion, there are two main advantages of the proposed integrated fusion framework. First, the proposed integrated fusion framework can accomplish different kinds of image fusion tasks based on a single unified model, including multiview spatial fusion, spatio–spectral fusion (PAN/MS fusion, PAN/HS fusion, and MS/HS fusion), and spatio–temporal fusion. Second, the proposed integrated fusion framework can obtain promising high spatio–temporal–spectral resolution fused images by taking full advantage of the complementary information from multisource observations. Furthermore, it does not have limitations on the number of remote sensing sensors.

However, it is noteworthy that there is still room for the proposed fusion framework to be further improved. The first limitation is the efficiency of the algorithm, particularly when there are quite a few sensor images to be fused. Therefore, a faster optimization algorithm and acceleration strategies such as parallel computing could be incorporated into the integrated fusion framework. Second, the proposed integrated fusion framework can be only used for homogenous optical remote sensing images. Fusion of multisource heterogeneous data (optical, thermal infrared, radar, etc.) is a more challenging task. Furthermore, new computation frameworks, including sparse representation, low-rank approximation, and tensor analysis, could be considered to further improve the robustness of the proposed fusion framework.

REFERENCES

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

[2] H. Song, B. Huang, Q. Liu, and K. Zhang, "Improving the spatial resolution of Landsat TM/ETM+ through fusion with SPOT5 images via learning-based super-resolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1195–1204, Mar. 2015.

[3] P. Wu, H. Shen, L. Zhang, and F.-M. Göttsche, "Integrated fusion of multi-scale polar-orbiting and geostationary satellite observations for the mapping of high spatial and temporal resolution land surface temperature," *Remote Sens. Environ.*, vol. 156, pp. 169–181, 2015.

[4] L. Alparone, B. Aiazzi, S. Baronti, and A. Garzelli, *Remote Sensing Image Fusion*. Boca Raton, FL, USA: CRC Press, 2015.

[5] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. M. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.

[6] C. Pohl and J. L. Van Genderenc, "Multisensor image fusion in remote, sensing: Concepts, methods, application," *Int. J. Remote Sens.*, vol. 19, no. 5, pp. 823–854, 1998.

[7] C. Thomas, T. Ranchin, L. Wald, and J. Chanussot, "Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1301–1312, May 2008.

[8] L. Alparone *et al.*, "A critical comparison among pansharpening algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2565–2586, May 2015.

[9] L. Loncan *et al.*, "Hyperspectral pansharpening: A review," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 3, pp. 27–46, Mar. 2015.

[10] B. Chen, B. Huang, and B. Xu, "Comparison of spatiotemporal fusion models: A review," *Remote Sens.*, vol. 7, no. 2, pp. 1798–1835, 2015.

[11] J. Zhang, "Multi-source remote sensing data fusion: Status and trends," *Int. J. Image Data Fusion*, vol. 1, no. 1, pp. 5–24, 2010.

[12] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, May 2003.

[13] Q. Yuan, L. Zhang, and H. Shen, "Multiframe super-resolution employing a spatially weighted total variation model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 3, pp. 379–392, Mar. 2012.

[14] L. Yue, H. Shen, J. Li, Q. Yuan, H. Zhang, and L. Zhang, "Image super-resolution: The techniques, applications, and future," *Signal Process.*, vol. 128, pp. 389–408, 2016.

[15] H. Shen, L. Peng, L. Yue, Q. Yuan, and L. Zhang, "Adaptive norm selection for regularized image restoration and super-resolution," *IEEE Trans. Cybern.*, vol. 46, no. 6, pp. 1388–1399, Jun. 2016.

[16] R. Tsai and T. S. Huang, "Multiframe image restoration and registration," *Adv. Comput. Vis. Image Process.*, vol. 1, no. 2, pp. 317–339, 1984.

[17] C. Latry and B. Rougé, "Super resolution: Quincunx sampling and fusion processing," in *Proc. IEEE IGARSS*, Toulouse, France, 2003, pp. 315–317.

[18] J. Ma, J. C.-W. Chan, and F. Canters, "An operational superresolution approach for multi-temporal and multi-angle remotely sensed imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 1, pp. 110–124, Feb. 2012.

[19] H. Zhang, Z. Yang, L. Zhang, and H. Shen, "Super-resolution reconstruction for multi-angle remote sensing images considering resolution differences," *Remote Sens.*, vol. 6, no. 1, pp. 637–657, 2014.

[20] H. Shen, M. K. Ng, P. Li, and L. Zhang, "Super-resolution reconstruction algorithm to MODIS remote sensing images," *Comput. J.*, vol. 52, no. 1, pp. 90–100, 2009, (online 2007).

[21] M. T. Merino and J. Nunez, "Super-resolution of remotely sensed images with variable-pixel linear reconstruction," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1446–1457, May 2007.

[22] F. Li, X. Jia, D. Fraser, and A. Lambert, "Super resolution for remote sensing images based on a universal hidden Markov tree model," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1270–1278, Mar. 2010.

[23] J. Ma, J. C.-W. Chan, and F. Canters, "Robust locally weighted regression for superresolution enhancement of multi-angle remote sensing imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1357–1371, Apr. 2014.

[24] M. Cetin and N. Musaoglu, "Merging hyperspectral and panchromatic image data: Qualitative and quantitative analysis," *Int. J. Remote Sens.*, vol. 30, no. 7, pp. 1779–1804, 2009.

[25] 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.

[26] 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.

[27] G. Cliche, F. Bonn, and P. Teillet, "Integration of the SPOT panchromatic channel into its multispectral mode for image sharpness enhancement," *Photogramm. Eng. Remote Sens.*, vol. 51, no. 3, pp. 311–316, 1985.

[28] W. A. Hallada and S. Cox, "Image sharpening for mixed spatial and spectral resolution satellite systems," presented at the International Symposium Remote Sensing Environment, Ann Arbor, MI, USA, May 9–13, 1983.

[29] 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, 2001.

[30] 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.

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

[32] 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.

[33] B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli, "Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 10, pp. 2300–2312, Oct. 2002.

[34] 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, 2016, Art. no. E1068.

[35] 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.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

SHEN *et al.*: FRAMEWORK FOR THE SPATIO–TEMPORAL–SPECTRAL FUSION OF REMOTE SENSING IMAGES
13

[36] 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.

[37] 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.

[38] 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.

[39] C. Jiang, H. Zhang, H. Shen, and L. Zhang, "Two-step sparse coding for the pan-sharpening of remote sensing images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 5, pp. 1792–1805, May 2014.

[40] S. Li, H. Yin, and L. Fang, "Remote sensing image fusion via sparse representations over learned dictionaries," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 9, pp. 4779–4789, Sep. 2013.

[41] R. Carla, L. Santurri, B. Aiazzi, and S. Baronti, "Full-scale assessment of pansharpening through polynomial fitting of multiscale measurements," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6344–6355, Dec. 2015.

[42] 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.

[43] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "Twenty-five years of pansharpening: A critical review and new developments," in *Signal and Image Processing for Remote Sensing*, 2nd ed., C.-H. Chen, Ed. Boca Raton, FL, USA: CRC Press, 2012, pp. 533–548.

[44] G. Vivone, "Multispectral and hyperspectral pansharpening: A critical examination and new developments," Ph.D. dissertation, Dept. Inf. Eng., Electr. Eng. Appl., Univ. Salerno, Fisciano, Italy, 2014.

[45] M. Moeller, T. Wittman, and A. L. Bertozzi, "A variational approach to hyperspectral image fusion," in *Proc. SPIE Defense, Security, Sens.*, 2009, pp. 1–10.

[46] Z. Chen, H. Pu, B. Wang, and G.-M. Jiang, "Fusion of hyperspectral and multispectral images: A novel framework based on generalization of pan-sharpening methods," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 8, pp. 1418–1422, Aug. 2014.

[47] M. E. Winter and E. M. Winter, "Physics-based resolution enhancement of hyperspectral data," in *Proc. SPIE Algorithms Technol. Multispectr., Hyperspectr., Ultraspectr. Imagery VIII*, 2002, pp. 580–587.

[48] R. C. Hardie, M. T. Eismann, and G. L. Wilson, "MAP estimation for hyperspectral image resolution enhancement using an auxiliary sensor," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1174–1184, Sep. 2004.

[49] M. T. Eismann and R. C. Hardie, "Hyperspectral resolution enhancement using high-resolution multispectral imagery with arbitrary response functions," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 455–465, Mar. 2005.

[50] 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.

[51] M. Simões, 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.

[52] A. Plaza *et al.*, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, pp. S110–S122, 2009.

[53] F. Gao, J. Masek, M. Schwaller, and F. Hall, "On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 8, pp. 2207–2218, Aug. 2006.

[54] H. Shen, P. Wu, Y. Liu, T. Ai, Y. Wang, and X. Liu, "A spatial and temporal reflectance fusion model considering sensor observation differences," *Int. J. Remote Sens.*, vol. 34, no. 12, pp. 4367–4383, 2013.

[55] H. Shen, L. Huang, L. Zhang, P. Wu, and C. Zeng, "Long-term and fine-scale satellite monitoring of the urban heat island effect by the fusion of multi-temporal and multi-sensor remote sensed data: A 26-year case study of the city of Wuhan in China," *Remote Sens. Environ.*, vol. 172, pp. 109–125, 2016.

[56] X. Zhu, J. Chen, F. Gao, X. Chen, and J. G. Masek, "An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions," *Remote Sens. Environ.*, vol. 114, no. 11, pp. 2610–2623, 2010.

[57] C. M. Gevaert and F. J. García-Haro, "A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion," *Remote Sens. Environ.*, vol. 156, pp. 34–44, 2015.

[58] J. Amorós-López *et al.*, "Multitemporal fusion of Landsat/TM and ENVISAT/MERIS for crop monitoring," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 23, pp. 132–141, 2013.

[59] B. Huang and H. Song, "Spatiotemporal reflectance fusion via sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 3707–3716, Oct. 2012.

[60] H. Song and B. Huang, "Spatiotemporal satellite image fusion through one-pair image learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 1883–1896, Apr. 2013.

[61] H. Shen, "Integrated fusion method for multiple temporal–spatial–spectral images," in *Proc. 22nd Congr. ISPRS*, Melbourne, Vic., Australia, 2012, pp. 407–410.

[62] B. Huang, H. Zhang, H. Song, J. Wang, and C. Song, "Unified fusion of remote-sensing imagery: Generating simultaneously high-resolution synthetic spatial–temporal–spectral earth observations," *Remote Sens. Lett.*, vol. 4, no. 6, pp. 561–569, 2013.

[63] X. Meng, H. Shen, H. Li, Q. Yuan, H. Zhang, and L. Zhang, "Improving the spatial resolution of hyperspectral image using panchromatic and multispectral images: An integrated method," presented at the Proceedings 7th WHISPERS, Tokyo, Japan, 2015.

[64] X. Meng, H. Shen, L. Zhang, Q. Yuan, and H. Li, "A unified framework for spatio–temporal–spectral fusion of remote sensing images," in *Proc. IEEE IGARSS*, Milan, Italy, 2015, pp. 2584–2587.

[65] H. Shen, L. Zhang, B. Huang, and P. Li, "A MAP approach for joint motion estimation, segmentation, and super resolution," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 479–490, Feb. 2007.

[66] Q. Yuan, L. Zhang, and H. Shen, "Hyperspectral image denoising employing a spectral–spatial adaptive total variation model," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 3660–3677, Oct. 2012.

[67] H. Shen and L. Zhang, "A MAP-based algorithm for destriping and inpainting of remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1492–1502, May 2009.

[68] 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.

[69] 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.

[70] M. V. Joshi, L. Bruzzone, and S. Chaudhuri, "A model-based approach to multiresolution fusion in remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2549–2562, Sep. 2006.

[71] M. K. Ng, W. C. Kwan, and R. H. Chan, "A fast algorithm for high-resolution color image reconstruction with multisensors," in *Numerical Analysis and Its Applications*. Berlin, Germany: Springer-Verlag, 2001, pp. 615–627.

[72] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.

[73] R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," *Comput. J.*, vol. 7, no. 2, pp. 149–154, 1964.

[74] 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.

[75] N. Yokoya, N. Mayumi, and A. Iwasaki, "Cross-calibration for data fusion of EO-1/Hyperion and Terra/ASTER," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 419–426, Apr. 2013.

[76] 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.

[77] 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, 2013.

[78] 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.

[79] 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.

[80] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2005, pp. 60–65.

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

[82] S. Li, Z. Li, and J. Gong, "Multivariate statistical analysis of measures for assessing the quality of image fusion," *Int. J. Image Data Fusion*, vol. 1, no. 1, pp. 47–66, 2010.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

14   IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING

[83] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[84] 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.

[85] R. W. Basedow, D. C. Carmer, and M. E. Anderson, "HYDICE system: Implementation and performance," in *Proc. SPIE Symp. OE/Aerosp. Sens. Dual Use Photon.*, 1995, pp. 258–267.

[86] I. V. Emelyanova, T. R. McVicar, T. G. Van Niel, L. T. Li, and A. I. van Dijk, "Assessing the accuracy of blending Landsat-MODIS surface reflectances in two landscapes with contrasting spatial and temporal dynamics: A framework for algorithm selection," *Remote Sens. Environ.*, vol. 133, pp. 193–209, 2013.

[87] B. Wu, B. Huang, and L. Zhang, "An error-bound-regularized sparse coding for spatiotemporal reflectance fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6791–6803, Dec. 2015.
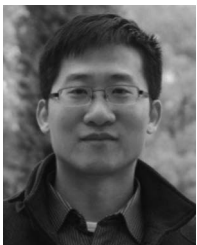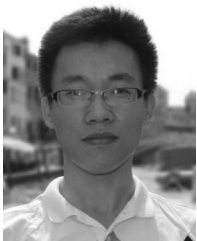
Huanfeng Shen (M'10–SM'13) received the B.S. degree in surveying and mapping engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2002 and 2007, respectively.

In July 2007, he joined the School of Resources and Environmental Sciences, Wuhan University, where he is currently a Luojia Distinguished Professor. He has authored or coauthored over 100 research papers. He has been supported by several talent programs, such as the New Century Excellent Talents by the Ministry of Education of China (2011), the China National Science Fund for Excellent Young Scholars (2014), and the Youth Talent Support Program of China (2015). 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*.

Xiangchao Meng received the B.S. degree in geographic information system from Shandong University of Science and Technology, Qingdao, China, in 2012. He is currently working toward the Ph.D. degree in the School of Resources and Environmental Sciences, Wuhan University, Wuhan, China.

His research interests include image fusion and remote sensing image processing.

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 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 is currently the Head of the Remote Sensing Division, State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing (LIESMARS), Wuhan University. He is also a Chang Jiang Scholar Chair Professor appointed by the Ministry of Education of China. He is currently a Principal Scientist for the China State Key Basic Research Project (2011–2016) appointed by the Ministry of National Science and Technology of China to lead the remote sensing program in China. He has over 450 research papers and five books. He is the holder of 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, an Executive Member (Board of Governors) of the Chinese National Committee for the International Geosphere-Biosphere Program, an Executive Member of the China Society of Image and Graphics, and others. He is the Founding Chair of the IEEE Geoscience and Remote Sensing Society (GRSS) Wuhan Chapter. He was the General Chair for the fourth IEEE GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS). He regularly serves as a Cochair of the series SPIE conferences on multispectral image processing and pattern recognition, conference on Asia remote sensing, and many other conferences. He was a Guest Editor of the IEEE JOURNAL OF SELECTED TOPICS IN EARTH OBSERVATIONS AND APPLIED REMOTE SENSING (JSTARS). 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 *Geo-spatial Information Science*, and the *Journal of Remote Sensing* and a Guest Editor for the *Journal of Applied Remote Sensing* and the *Journal of Sensors*. He is also currently serving as an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He was a recipient of the best reviewer awards from the IEEE GRSS for his service to the IEEE JSTARS in 2012 and the 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 (IGARSS) student paper contest in recent years. He was also a recipient of the 2010 best paper Boeing award and the 2013 best paper ERDAS award from the American Society of Photogrammetry and Remote Sensing.