# A Remote Sensing Image Fusion Method Based on the Analysis Sparse Model

Chang Han, Hongyan Zhang, *Member, IEEE*, Changxin Gao, Cheng Jiang, Nong Sang, and Liangpei Zhang, *Senior Member, IEEE*

***Abstract*—This paper addresses the remote sensing image fusion problem from the perspective of the analysis sparse model. As an alternative to the synthesis sparse representation approach, the analysis sparse model can yield richer feature representations and better results for image restoration. We, therefore, propose an image fusion method for remote sensing images based on the analysis sparse model. In this method, the analysis operators for the high-resolution multispectral (HR MS) image are trained band by band, directly from the source images, which can greatly improve the adaptability. During the analysis operator learning stage, the geometric analysis operator learning (GOAL) algorithm is utilized with the upsampled low-resolution MS (LR MS) image and the HR panchromatic (HR PAN) image, which does not require an external HR MS image data set. Moreover, the imagery system modulation transfer function (MTF) is considered during the LR MS imaging modeling process, which greatly extends the practical application potential of the proposed method. The simulated and real-data experimental results on IKONOS and QuickBird data sets show that the proposed method can effectively preserve the spectral information and the spatial detail of the image. The fused HR MS images produced by the proposed method are comparable and even superior to the images fused by the other state-of-the-art methods.**

***Index Terms*—Analysis operator, analysis sparse model, image fusion, inverse problem, remote sensing.**

## I. Introduction

HIGH-spatial-resolution (HR) multispectral (MS) remote sensing images are widely used in many application areas [1], such as land-use classification, change detection, map updating, disaster monitoring, and so on. However, due to the fact that sensors always have a physical tradeoff between the spatial and the spectral resolutions, the remote sensing image data which are provided by earth observation satellites, such as IKONOS, QuickBird, GeoEye-1, WorldView-1, and WorldView-2, are often composed of an HR panchromatic (HR PAN) channel and several MS channels with a lower spatial resolution (LR). Therefore, the remote sensing image fusion technique is designed to effectively integrate the spatial detail of the PAN image and the spectral information of the MS image to acquire the desired HR MS image [2]. The fusion of the PAN channel and spectral channels is also called "image pansharpening," and it aims at fully utilizing the complementary information of the LR MS image and the HR PAN image to obtain the HR MS image [3]. It is a typical data fusion method, and we refer to it as "image fusion" in this paper.

In the past few years, various types of image fusion methods have been proposed to obtain the HR MS remote sensing images. Many of them, such as intensity-hue-saturation (IHS) [4], principal component analysis (PCA) [5], and the Gram–Schmidt (GS) [6] algorithm, are based on a transformation and substitution strategy. These methods usually consist of the following steps: 1) upsample the LR MS image to the PAN scale, and forward transform the resampled LR MS image into a certain new image space; 2) match the histogram of the PAN image to that of the main component of the new image space, then replace this component with the newly matched PAN image; and 3) finally, obtain the fused result by inverse transformation over the substituted image. These methods can often obtain an HR spatial quality, but are subject to serious spectral degradation in most cases. The Brovey transform [7], which is based on the assumption that the LR PAN image can be considered as a linear combination of the original MS image, is another popular image fusion method that has been widely applied. However, due to the significant difference in the digital number (DN) values between PAN and MS images, because of their different wavelength ranges, the Brovey method often suffers from spectral distortion.

A number of well-known image fusion methods based on multiresolution analysis (MRA) [8], [9] have been proposed, such as wavelet transform [10], Laplacian pyramids [11], and contourlet transform [12]. Although these kinds of methods have limitations in keeping the spatial details of the image, they can effectively preserve the spectral information. Another MRA method named the coupled multiresolution decomposition model (CMD), which takes the imagery system modulation transfer function (MTF) into consideration, was recently investigated in [13]. This new fusion scheme allows the reconstruction of an HR MS image given its approximation, which is obtained by MTF-tailored downsampling and wavelet decomposition.

Manuscript received January 14, 2015; revised November 06, 2015; accepted December 03, 2015. Date of publication January 05, 2016; date of current version January 28, 2016. This work was supported in part by the 863 Program under Grant 2013AA12A301, in part by the National Natural Science Foundation of China under Grant 41571362, Grant 61201342, and Grant 61433007, and in part by the Key Laboratory of Agri-informatics, Ministry of Agriculture, P.R. China, Beijing, 100081, China. *(Corresponding author: Hongyan Zhang.)*

C. Han, C. Gao, and N. Sang are with the School of Automation, Huazhong University of Science and Technology, Wuhan 430074, China.

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

C. Jiang is with the School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China.

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/JSTARS.2015.2507859

In recent years, Li *et al.* [14] addressed the image fusion problem with inverse problem-based methods. However, the solution of this ill-posed inverse problem is often not unique. The sparse regularization method, which reflects the inherent property of the natural signals, is particularly suitable to resolve this ill-posed inverse problem [15], [16]. In [14], it was suggested that the MS image patches could be sparsely represented by atoms of a dictionary that was randomly sampled from HR MS images acquired by "comparable" sensors. Although this method can acquire competitive results when compared with the aforementioned methods, it requires training images from an HR MS image set which may not be available, which limits its practical application. To overcome this problem, Jiang *et al.* [17] proposed a joint dictionary learning algorithm based on compressed sensing theory. Furthermore, the joint dictionary was constructed with the available upsampled LR MS and HR PAN images to make the image fusion more practical. Zhu and Blamer [3] proposed a new method named sparse fusion of images (SparseFI), which involves constructing LR/HR dictionary pairs from the HR PAN image and its downsampled version. This method explores the same sparse coefficient vector of the corresponding HR/LR MS image patches over the coupled HR/LR dictionaries. As external image data sets are not required for the dictionary learning, this method is considered to be a promising approach with a broader application potential. Recently, Jiang *et al.* [18] proposed a two-step sparse coding approach with patch normalization (PN-TSSC) for image fusion. This method exploits the local similarity between the MS and PAN images, and overcomes the instability of choosing dictionary atoms compared with one-step sparse coding. Li *et al.* [19] also recently developed a restoration-based remote sensing image fusion method with sparse regularization, in which the dictionary is adaptively learned with the available source image. Guo *et al.* [20] proposed an online coupled dictionary learning approach, in which the coupled dictionaries are iteratively updated with the fused HR MS image. In this method, the atoms of the constructed dictionary are more relevant to the MS image patches, and lead to a better fusion result.

As mentioned above, the sparse regularization method which has been utilized in these image fusion techniques is a generative model known as "synthesis sparsity." This model has performed very well in many signal processing applications [21]–[26]. However, when an overcomplete dictionary with high mutual coherence is used in the synthesis model [27], [28], it becomes difficult to determine where a signal may lie under the synthesis sparsity [29]. Recently, an alternative to the synthesis model named the analysis sparse model was proposed in [29]–[32]. In this model, all the atoms of the analysis operator take an equal part in describing the signal, thus minimizing the dependence on each individual atom, and stabilizing the recovery process [33]–[36].

In this paper, we propose a novel remote sensing image fusion method under the analysis sparse model. Compared with other methods presented recently, the contributions of our method are twofold. 1) To the best of our knowledge, the analysis sparse model has not yet been exploited for remote sensing image fusion, and this is a new attempt in the remote sensing image fusion domain. 2) In our method, the analysis operators for the HR MS image are trained band by band, directly from the available source images, which can greatly improve the adaptability. During the analysis operator learning stage, the geometric analysis operator learning (GOAL) algorithm is utilized with the LR MS image and the HR PAN image, which does not require an HR MS image data set. Furthermore, in the observation model, the MTF of the imagery system is taken into account, which greatly extends the practical application potential of the proposed method. Two groups of simulated experiments were carried out with two kinds of satellite sensor data, i.e., IKONOS and QuickBird. A real-data experiment with an IKONOS data set was also implemented. The image fusion results were then compared with the results of the other state-of-the-art fusion methods.

This paper is organized as follows. Section II briefly describes the inverse problem and the sparse regularization method. Thereafter, the scheme of the proposed image fusion algorithm is reported in Section III. In Section IV, the experiments with IKONOS and QuickBird data are described, demonstrating the effectiveness of the proposed method with respect to the visual, spatial, and spectral quality. Finally, conclusion is drawn in Section V.

## II. Related Works

### A. Inverse Problem

Many image processing problems can be modeled as linear inverse problems, including image denoising, deblurring, inpainting, super-resolution, and so on. In these applications, the purpose is to reconstruct an unknown image $\mathbf{s} \in \mathbb{R}^d$ as accurately as possible from the potentially contaminated measurements $\mathbf{y} \in \mathbb{R}^{\mathrm{n}}$, with $n < d$. For the generic inverse problem, an incomplete set of linear observations $\mathbf{y} \in \mathbb{R}^{\mathrm{n}}$ is available as [37]

$$\mathbf{y} = \mathbf{\Phi s} + \boldsymbol{\epsilon} \tag{1}$$

where $\boldsymbol{\epsilon} \in \mathbb{R}^n$ is the additive bounded noise and sampling model errors, and $\mathbf{\Phi} \in \mathbb{R}^{n \times d}$ is the observation matrix modeling the sampling process. In many cases, reconstructing the original signal $\mathbf{s}$ by inverting (1) is ill-posed because it admits an infinite number of solutions (since $n < d$). However, a stable approximate solution of (1) can be achieved if we have some prior information about the underlying object $\mathbf{s}$. Regularization is a popular way to impose such prior information. The general regularization model can be written as

$$\min_{\mathbf{s} \in \mathbb{R}^d} \frac{1}{2} \|\mathbf{y} - \mathbf{\Phi s}\|_2^2 + \lambda g(\mathbf{s}) \tag{2}$$

where the first term is the data fidelity term, and $g(\bullet)$ is an appropriate regularization (prior) term.

To date, a number of different regularization methods have been proposed. In [38], Tikhonov regularization with an $\ell_2$ norm penalty was found to be able to acquire a meaningful approximate solution and make the inversion less sensitive to perturbations. However, the $\ell_2$-based norm penalty has a spatial smoothing effect on the solution, which can lead to the loss of meaningful features [39]. The basis pursuit framework proposed by Donoho [40] and the $\ell_1$-regularization proposed by

Tibshirani [41] provide alternative sparse methods for signal recovery. Compared with the $\ell_2$-based regularization methods, the sparse regularization method with $\ell_1$-regularization can more effectively preserve sharp edges, which is vitally important in image processing applications [42]. According to the terminology of [37], sparse $\ell_1$-regularization can be divided into an $\ell_1$-synthesis prior and an $\ell_1$-analysis prior. In the following part, we briefly review these two models.

### *B. Synthesis and Analysis Sparse Models*

During the last decade, the synthesis sparse representation model has established itself and has led to state-of-the-art results in many image processing areas. In the synthesis-based model, signal $\mathbf{s} \in \mathbb{R}^d$ is sparsely represented over a given dictionary $\mathbf{D} \in \mathbb{R}^{d\times k}$, with $d \ll k$ when it is composed of a linear combination of only a few of the atoms from dictionary $\mathbf{D}$. Thus, the image signal can be represented as $\mathbf{s} = \mathbf{D}\boldsymbol{\alpha}$. The sparse representation task is formulated as the following minimization problem:

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_0 \quad s.t. \quad \|\mathbf{y} - \boldsymbol{\Phi}\mathbf{D}\boldsymbol{\alpha}\|_2^2 \le \varepsilon \tag{3}$$

where $\boldsymbol{\alpha}$ is the vector of the sparse coefficients, with most of the coefficients being close to or equal to zero, and it matches the measurements up to a specified tolerance of $\varepsilon$. Since problem (3) is NP-hard [43], it is often typically relaxed as

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \frac{1}{2}\|\mathbf{y} - \boldsymbol{\Phi}\mathbf{D}\boldsymbol{\alpha}\|_2^2 + \lambda\|\boldsymbol{\alpha}\|_1. \tag{4}$$

It then becomes an $\ell_1$-norm constrained convex optimization problem, where $\lambda$ is a suitable regularization parameter. Many different kinds of $\ell_1$-norm optimization methods have been proposed in the last few years [44].

Recently, as a new sparse representation approach, the analysis sparse model has been proposed and is gradually attracting researchers' attention [30], [31]. Suppose that the suitable $\boldsymbol{\Omega} \in \mathbb{R}^{p\times d}$ for noise-free training of $\mathbf{s}$, whose rows constitute analysis atoms, is a linear analysis operator. The sparse coefficient vector $\boldsymbol{\alpha} \in \mathbb{R}^p$ can then be approximated as $\boldsymbol{\alpha} = \boldsymbol{\Omega}\mathbf{s}$, which with*l*zero elements is as sparse as possible [30]. It can be formally expressed as the following optimization problem:

$$\boldsymbol{\Omega} \in \arg\min \; g(\boldsymbol{\Omega}\mathbf{s}) \tag{5}$$

where $\boldsymbol{\Omega}$ is subject to certain constraints, and $\mathrm{g}(\bullet)$ is the function that measures the sparsity of the matrix $\boldsymbol{\Omega}\mathbf{s}$.

In the analysis sparse model, the signal information is encoded in the zero entries of $\boldsymbol{\alpha}$. The number of zero elements in the coefficient vector, i.e., $l = p - \|\alpha\|_0$, is called the co-support of the signal [31]. This implies that there is a subset $\Lambda$ of the rows of $\boldsymbol{\Omega}$, the co-support $\boldsymbol{\Omega}_\Lambda$, such that [45]

$$\boldsymbol{\Omega}_\Lambda \mathbf{s} = 0. \tag{6}$$

The subspace is spanned by the rows of $\boldsymbol{\Omega}_\Lambda$ to which $\mathbf{s}$ is orthogonal. Geometrically, $\mathbf{s}$ lies in the intersection of all the hyperplanes whose normal vectors are in the set of the co-support. The issue of recovering $\mathbf{s}$ from the corrupted measurements is conducted by solving the following generic minimization problem [31]:

$$\hat{\mathbf{s}} = \arg\min_{s} g(\boldsymbol{\Omega}\mathbf{s}) \quad \text{s.t} \quad \|\mathbf{y} - \boldsymbol{\Phi}\mathbf{s}\|_2 \le \varepsilon \tag{7}$$

where $\mathrm{g}(\bullet)$ is the sparsity-promoting function, and $\boldsymbol{\epsilon} \in \mathbb{R}^+$ is an estimated upper bound on the noise power.

According to the theoretical basis in [30], the following optimization problem is a special case of (7):

$$\hat{\mathbf{s}} = \arg\min_{s} \|\boldsymbol{\Omega}\mathbf{s}\|_0 \quad \text{s.t.} \quad \|\mathbf{y} - \boldsymbol{\Phi}\mathbf{s}\|_2 \le \varepsilon. \tag{8}$$

In addition, the solving of problem (8) is an NP-complete problem [46], and thus an approximate solving method is required. Many approximate solving methods have been proposed in recent years. One option for approximating (8) is to use a greedy strategy, such as greedy analysis pursuit (GAP) [30], analysis iterative hard thresholding (AIHT), analysis hard thresholding pursuit (AHTP), analysis compressive sampling matching pursuit (ACoSaMP), analysis subspace pursuit (ASP) [47], [48], and so on. Another way is to use $\ell_1$-relaxation to approximate (8). The analysis $\ell_1$-minimization, which replaces the $\ell_0$ with $\ell_1$ in (8), is well known and has been widely applied in practice.

In [32], Yaghoobi *et al.* utilized a convex $\ell_1$-norm, i.e., $\mathrm{g}(\boldsymbol{\Omega}\mathbf{s}) = \|\boldsymbol{\Omega}\mathbf{s}\|_1$, as the sparsity-promoting function, and the uniformly normalized tight frame (UNTF) constraint for the analysis operator was exploited to solve problem (5). The trained operator by this algorithm has been applied to natural face image denoising. In [35], Have *et al.* proposed an analysis operator learning algorithm called GOAL, which exploits the constraint of full-rank matrices with normalized columns. The conjugate gradient method on manifolds is utilized to solve this optimization problem. The experimental results for the classic image restoration problems have shown the competitive performance of GOAL when compared with the other state-of-the-art techniques. Thereafter, Worman *et al.* [49] applied this operator learning model to image reconstruction based on blind compressive sensing, and also achieved a desirable result. Chen *et al.* [50] proposed a bi-level training approach for analysis operator learning, which is effectively solved with the limited-memory Broyden-Fletcher-Goldfarb-Shanno quasi-Newton method (L-BFGS). This method excludes the trivial solution of the analysis operator without imposing any additional constraints, and has also obtained desirable results.

## III. Proposed Image Fusion Scheme

For image fusion, the goal is to recover an HR MS image from a single HR PAN image and its corresponding LR MS image. This means that we are increasing the spatial resolution of the LR MS image while keeping its spectral information with the help of the HR PAN image.

### *A. Observation Model*

First of all, we need to build the observation model which establishes the mathematical relationship between the desired

HR MS image and the observed LR MS image and HR PAN image. According to [14], the observed LR MS image can be modeled as a decimated, blurred, and noisy version of the corresponding HR MS image. Considering the MTF of the satellite imagery system, the linear degradation model between the HR MS image and the LR MS image is formulated as

$$\mathbf{Y}_{MS} = \mathbf{LFS} + \mathbf{v}_1 \tag{9}$$

where $\mathbf{Y}_{MS} \in \mathbb{R}^{BHW}$ and $\mathbf{S} \in \mathbb{R}^{\rho^2 BHW}$ are column vectors representing the lexicographically reordered LR MS image and HR MS image, respectively, with $B$, $H$, and $W$ representing the number of bands, the height, and the width of the LR MS image, and $\rho$ being the spatial resolution ratio between the low-resolution MS image and the PAN image. $\mathbf{F} \in \mathbb{R}^{\rho^2 BHW \times \rho^2 BHW}$ represents the MTF-tailored low-pass filter for the HR MS image, $\mathbf{L} \in \mathbb{R}^{BHW \times \rho^2 BHW}$ denotes the decimation operator, and $\mathbf{v}_1 \in \mathbb{R}^{BHW}$ is the noise vector. Usually, the MTF is bell-shaped, and its magnitude value at the cutoff Nyquist frequency needs to be much lower than 0.5 to prevent aliasing [11]. We denote $\mathbf{M_1} = \mathbf{LF} \in \mathbb{R}^{BHW \times \rho^2 BHW}$, and then (9) can be rewritten as

$$\mathbf{Y}_{MS} = \mathbf{M}_1\mathbf{S} + \mathbf{v}_1. \tag{10}$$

As we know, the range of the wavelength spectrum of the PAN modality is usually overlapped or partly overlapped with the MS image. This overlapping characteristic suggests that the HR PAN image can be approximated as a linear combination of the desired HR MS image [14]. The linear model can be modeled as

$$\mathbf{Y}_{PAN} = \sum_{b}^{B} \theta^b \mathbf{S}^b + \mathbf{v}_2^b = \mathbf{M}_2\mathbf{S} + \mathbf{v}_2 \tag{11}$$

where $\mathbf{Y}_{PAN} \in \mathbb{R}^{\rho^2 HW}$ and $\mathbf{S}^b \in \mathbb{R}^{\rho^2 HW}$ represent the lexicographically reordered version of the PAN image and the $b$th band of the HR MS image, respectively, and $\theta^b$ is the weight of the $b$th band. $\mathbf{M}_2 \in \mathbb{R}^{\rho^2 HW \times \rho^2 BHW}$ is the linear combination matrix which denotes the spectral response of the different remote sensors. $\mathbf{v}_2 \in \mathbb{R}^{\rho^2 HW}$ is the additive zero-mean Gaussian noise. The parameters $\theta^b$ and $\mathbf{M}_2$ are determined by the specific satellite sensors. Recently, Xavier [9] proposed an $\theta^b$ parameter estimation technique that takes into account the physical electromagnetic spectrum responses of the sensors during the procedure of image fusion. In this paper, we use Xavier's method to estimate parameter $\theta^b$.

Equations (10) and (11) can then be incorporated by the following formulation:

$$\mathbf{Y} = \mathbf{MS} + \mathbf{v} \tag{12}$$

where $\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_{MS} \\ \beta * \mathbf{Y}_{PAN} \end{bmatrix} \in \mathbb{R}^{(B+\rho^2)HW}$ and $\mathbf{M} = \begin{bmatrix} \mathbf{M}_1 \\ \beta * \mathbf{M}_2 \end{bmatrix} \in \mathbb{R}^{(B+\rho^2)HW \times \rho^2 BHW}$ represent the measurement vector and the observation matrix, respectively. In addition, $\mathbf{v} \in \mathbb{R}^{(B+\rho^2)HW}$ is the model noise. Clearly, the factor $\beta$ is used as a tradeoff parameter to balance the relative contribution of the LR MS image and the HR PAN image to the final result of the image fusion [15], [17]. In the experimental part, we investigate the influence of parameter $\beta$ on the results.

It can now be observed that (1) and (12) are surprisingly similar. Therefore, the remote sensing image fusion problem is a typical ill-posed inverse problem. As shown in Section II, the goal is to recover $\mathbf{S}$ from $\mathbf{Y}$ with the regularization method, and the analysis sparse model is applied to solve this problem here.

### *B. Image Fusion Scheme Based on the Analysis Sparse Model*

According to sparse regularization theory, the original signal can be accurately reconstructed from a set of incomplete measurements. Thus, the remote sensing image fusion problem of (12) can be formulated as an analysis sparse minimization problem

$$\hat{\mathbf{S}} = \arg\min_{\mathbf{S}} g(\mathbf{\Theta S}) \quad + \lambda \left\| \mathbf{Y} - \mathbf{MS} \right\|_2^2 \tag{13}$$

where $\mathbf{\Theta} \in \mathbb{R}^{P \times \rho^2 BHW}$ is the analysis operator for the entire image $\mathbf{S}$, and the analyzed vector can be written as $\boldsymbol{\gamma} = \mathbf{\Theta S} \in \mathbb{R}^P$. $\mathrm{g}(\bullet)$ is the analysis sparsity-promoting function, which promotes the product of the vector $\mathbf{S}$ and analysis operator $\mathbf{\Theta}$ to be the sparsest.

To reduce the computational burden, a patch-wise processing strategy [51] is adopted for the proposed method in this paper. That is to say, the algorithm is carried out patch by patch, and the final desired HR MS image can be synthesized from all the fused HR MS image patches. We denote the height and width of the LR MS image patch as $h$ and $w$, respectively. To avoid any confusion, the patches of the desired HR MS image and the observations are represented by the lower case $\mathbf{s} \in \mathbb{R}^{\rho^2 Bhw}$ and $\mathbf{y} \in \mathbb{R}^{(\rho^2+B)hw}$, and the analysis operator $\mathbf{\Theta}$ is substituted by $\mathbf{\Omega} \in \mathbb{R}^{p \times \rho^2 Bhw}$, which is the analysis operator for the image patches. For the purpose of continuity, the definition and the construction of the analysis operator $\mathbf{\Omega}$ is explained later in Section III-D. We can obtain the unconstrained optimization model for image fusion with respect to the image patches as follows:

$$\hat{\mathbf{s}} = \arg\min_{s} g(\mathbf{\Omega s}) \quad + \lambda \left\| \mathbf{y} - \mathbf{Ms} \right\|_2^2. \tag{14}$$

Herein, $\mathbf{M} \in \mathbb{R}^{(\rho^2+B)hw \times \rho^2 Bhw}$, consisting of $\mathbf{M_1} \in \mathbb{R}^{Bhw \times \rho^2 Bhw}$ and $\mathbf{M}_2 \in \mathbb{R}^{\rho^2 hw \times \rho^2 Bhw}$, represents the patch-level observation matrix, and $\lambda \in \mathbb{R}^+$ balances the sparsity of the solution's analysis coefficients and the solution's fidelity to the measurements. To measure the sparsity of the analyzed patches, a differentiable sparsity-promoting function is adopted

$$g\left(\mathbf{\Omega s}\right) = \sum_{j=1}^{p} \log(1 + \mu(\mathbf{\Omega s})_j^{\,2}) \tag{15}$$

where $\mu$ is a positive weighting factor, which serves as an appropriate sparsity measure, and $p$ is the number of elements in the analyzed vector $\mathbf{\Omega s}$.

The proposed reconstruction algorithm can be solved based on the conjugate gradient method. The solution is iteratively updated by

$$\mathbf{s}^{(i+1)} = \mathbf{s}^{(i)} + \sigma \hbar^{(i)} \tag{16}$$

where the scalar $\sigma$ is the step size, and $\mathbf{s}^{(i)}$and $\hbar^{(i)}$ are the local image patch vector and the descent direction at the $i$th iteration, respectively. We define

$$f(\mathbf{s}) = \sum_{j=1}^{p} \log(1 + \mu(\mathbf{\Omega s})_j^2) \quad + \lambda \left\| \mathbf{y} - \mathbf{M s} \right\|_2^2. \tag{17}$$

Let $g^{(i)} := \nabla f(s^{(i)})$ be the gradient of the cost function at the $i$th iteration

$$g^{(i)} := \sum_{j=1}^{p} \frac{2\mu \mathbf{\Omega}^T \left( \mathbf{\Omega s}^{(i)} \right)_j}{1 + \mu \left( \mathbf{\Omega s}^{(i)} \right)_j^2} + 2\lambda \mathbf{M} \left( \mathbf{M}^T \mathbf{s}^{(i)} - \mathbf{y} \right). \tag{18}$$

The descent direction can be initiated with $\hbar^{(0)} = -g^{(0)}$. The Hestenes–Stiefel formula shown in [52] is given by

$$HS^{(i)} = \frac{(g^{(i)})^T (\hbar^{(i)})}{(d^{(i)})^T \hbar^{(i)}} \tag{19}$$

where $\hbar^{(i)} = g^{(i+1)} - g^{(i)}$ and $d^{(i)} = \begin{cases} -g^{(i)} & for\ i = 1 \\ -g^{(i)} + HS^{(i)} d^{(i-1)} & for\ i \geq 2 \end{cases}$. The new search direction is computed by

$$\hbar^{(i+1)} = -g^{(i+1)} + HS^{(i)} d^{(i)}. \tag{20}$$

The step size $\sigma$ is computed by backtracking the line search. In the subsequent iterations, the backtracking line search is initialized by the previous step size divided by the line search parameter, i.e., $\sigma_m = \frac{\sigma_{m-1}}{\eta^2}$, where $\eta$ is a certain constant. In our implementation, the line search parameter is set as $\eta = 0.9$. The relative variation of the estimated fusion images in two consecutive iterations is set as $\delta = \| \frac{\mathbf{s}^{(i)} - \mathbf{s}^{(i-1)}}{\mathbf{s}^{(i-1)}} \|$. The threshold $\delta \leq 0.01$ is set as the stopping criterion.

### C. Patch-Based Processing Strategy

Because of the huge computational load for a whole image [51], we adopt a patch-based processing strategy in the proposed method. The remote sensing images utilized in this study were IKONOS and QuickBird data sets, whose spatial resolution ratio $\rho$ between the LR MS image and the PAN image is 4. We set the size of the image patch in the HR MS images as $8 \times 8 \times 4$, which corresponds to an $8 \times 8$image patch in the PAN image and a $2 \times 2 \times 4$ image patch in the LR MS image. Therefore, the decimation operator matrix $\mathbf{L} \in \mathbb{R}^{16 \times 256}$ in (9) can be constructed as [14], [17]

$$\mathbf{L} = ((1/16) \times \mathbf{I}_{8\times 8} \otimes (\mathbf{1}^T \otimes (\mathbf{I}_{2\times 2} \otimes \mathbf{1}^T))) \tag{21}$$

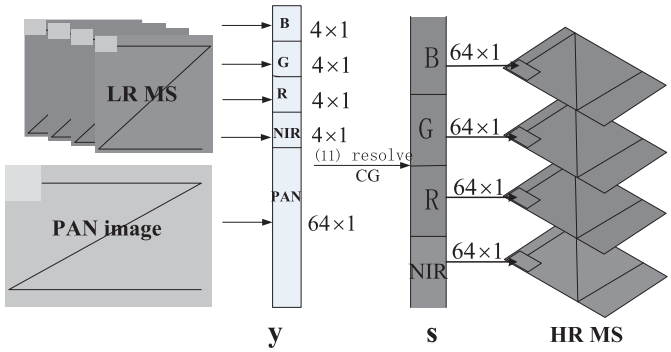



Fig. 1. Correspondence of the image patches.

where $\mathbf{1}$ is a $4 \times 1$ vector with all entries equal to one, and $\mathbf{I}_{8\times 8}$and $\mathbf{I}_{2\times 2}$ are diagonal matrices whose elements on the primary diagonal are ones with the size of $8 \times 8$ and $2 \times 2$, respectively. $\otimes$ represents the Kronecker product operator, details of which can be found in [14] and [17]. The MTF-tailored low-pass filter $\mathbf{F} \in \mathbb{R}^{256\times 256}$ can be formulated as

$$\mathbf{F} = \begin{pmatrix} \mathbf{F}_{Blue} & & & \\ & \mathbf{F}_{Green} & & \\ & & \mathbf{F}_{Red} & \\ & & & \mathbf{F}_{NIR} \end{pmatrix} \tag{22}$$

where $\mathbf{F}_{Blue}$, $\mathbf{F}_{Green}$, $\mathbf{F}_{Red}$, and $\mathbf{F}_{NIR} \in \mathbb{R}^{64\times 64}$ represent the MTF-tailored low-pass filter matrix for the blue, green, red, and near-infrared bands with different cutoff frequencies, respectively, where the optical point spread function (PSF) is assumed to be a Gaussian kernel with the support size of $5 \times 5$.

The matrix $\mathbf{M}_2 \in \mathbb{R}^{64\times 256}$ in (11) can be written as

$$\mathbf{M}_2 = (\theta_1 \cdot \mathbf{I}_{64\times 64}\ \theta_2 \cdot \mathbf{I}_{64\times 64}\ \theta_3 \cdot \mathbf{I}_{64\times 64}\ \theta_4 \cdot \mathbf{I}_{64\times 64}) \tag{23}$$

where $\mathbf{I} \in \mathbb{R}^{64\times 64}$ is a $64 \times 64$ diagonal matrix with elements equal to one. The correspondence relationship between the LR MS image patch and the HR PAN image patch is illustrated in Fig. 1.

As can be seen from Fig. 1, all of the image patches (PAN and MS image patches) are processed in raster-scan order, from left-top to right-bottom. In the PAN image, the scan step size is four pixels, which corresponds to one pixel in the LR MS image. Therefore, the overlapped size of the HR MS image patch is $4 \times 4$. As shown in Fig. 1, the observed image patch $\mathbf{y}$, which is an $80 \times 1$ column vector, is composed of four $4 \times 1$ column bands of LR MS image patches and one $64 \times 1$column PAN image patch. The overlapped pixels of the image patches are counted when the HR MS image patches are reconstructed. Finally, the HR MS image is synthesized by averaging the overlapping image patches.

### D. Analysis Operator Learning

The ability of sparse representation for a certain class of signals is the key to its application [35], [53]. For the analysis sparse model, this ability depends on the subspace precision of the analysis operator. In this study, the analysis operator for image fusion must have the capability of simultaneously

carrying spatial and spectral information. Therefore, the analysis operator $\mathbf{\Omega}_b$ for the $b$th band ($\mathbf{\Omega}_b$ can also be represented with $\mathbf{\Omega}_{Red}$, $\mathbf{\Omega}_{Green}$, $\mathbf{\Omega}_{Blue}$, and $\mathbf{\Omega}_{NIR}$ for the red, green, blue, and near-infrared bands, respectively) is trained separately with training samples that are composed of the PAN image patches and the upsampled MS image patches.

We take the operator $\mathbf{\Omega}_{Red}$ for the red band as an example to illustrate the construction of the analysis operator for the MS image. The PAN image and the upsampled red band image are cropped into image patches with the size of $8 \times 8$. We then choose 5000 image patches randomly from the PAN image and the upsampled red band image, respectively. These 10 000 image patches are then collected as the training samples for the analysis operator learning of the red band image.

Assuming we have a set of $K$ training samples $\left\{\mathbf{s}_k \in \mathbb{R}^{64}\right\}_{k=1}^{K}$ (where $K = 10\,000$), the goal of analysis operator learning is to find the suitable matrix $\mathbf{\Omega}_{Red} \in \mathbb{R}^{p\times 64}$ with $p \geq 64$, which leads to the analyzed vector $\mathbf{\Omega}_{Red}\mathbf{s}_k$ being as sparse as possible for the training sample $\mathbf{s}_k$. Therefore, the analysis operator for a given set of image patches $\left\{\mathbf{s}_k \in \mathbb{R}^{64}\right\}_{k=1}^{K}$ is formally denoted by

$$\mathbf{\Omega}_{Red} \in \arg\min \sum \|\mathbf{\Omega}_{Red} s_k\|_0 \tag{24}$$

Obviously, this is nonconvex and discontinuous, and it is difficult to obtain the optimal solution. One tractable approach is to replace it with a smooth approximation function. Therefore, the $\ell_0-$pseudo-norm in (24) can be replaced by

$$\mathbf{\Omega}_{Red} \in \arg\min \sum_{k}^{K} \sum_{j=1}^{p} \log(1 + \mu(\mathbf{\Omega}_{Red}\mathbf{s})_{k,j}^2) \tag{25}$$

which is a smooth log-square function [54], where $k$ represents the $k$th training sample. Similar to the log-sum sparsity measure proposed by Candes *et al.* [55], when $\mu$ reaches an infinite value, the log-square function in (25) tends to a constant which is a good approximation of $\ell_0$-sparsity.

For problem (24), without additional prior assumptions on $\mathbf{\Omega}_{Red}$, $\mathbf{\Omega}_{Red} = 0$ is the global minimum solution. To avoid this kind of trivial solution, we impose certain constraints on $\mathbf{\Omega}_{Red}$. One direct constraint is that the rows of $\mathbf{\Omega}_{Red}$ have unit Euclidean norm. This means that we can restrict the transpose of the analysis operator to a manifold structure, which is known as an oblique manifold [56]

$$\mathbf{OB}(64, p) := \left\{\chi \in \mathbb{R}^{64\times p} | rk(\chi) = 64, ddiag\left(\chi^T\chi\right) = \mathrm{I}_p\right\}. \tag{26}$$

Here, we assume that $\boldsymbol{\gamma} = \chi^T\chi$, $ddiag(\boldsymbol{\gamma})$ is a diagonal matrix whose entries on the diagonal are those of $\boldsymbol{\gamma}$. $rk(\chi)$ represents the rank of the matrix $\chi$. Since we require the rows of $\mathbf{\Omega}_{Red}$ to have unit Euclidean norm, we restrict $\mathbf{\Omega}_{Red}^T$ to be an element of $\mathbf{OB}(64, p)$. We then enforce the full-rank constraint with (27) on the analysis operator $\mathbf{\Omega}_{Red}$ [35]

$$h(\mathbf{\Omega}_{Red}) = -\frac{1}{64\log(64)} \log\det\left(\frac{1}{p}\mathbf{\Omega}_{Red}^T\mathbf{\Omega}_{Red}\right). \tag{27}$$

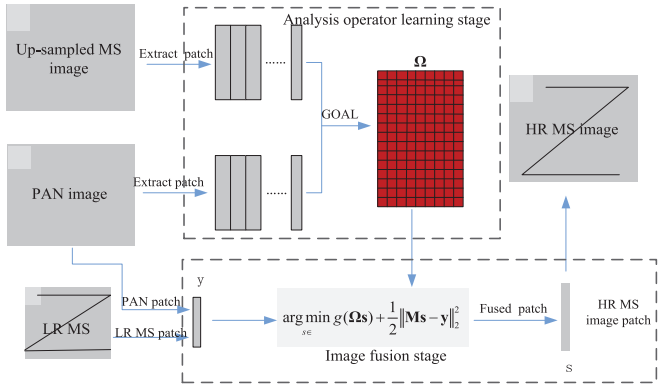



Fig. 2. Proposed image fusion scheme.

(a) (b)

Fig. 3. IKONOS images. (a) Degraded MS image at a 16-m spatial resolution. (b) Degraded PAN image at a 4-m spatial resolution.

The mutual coherence of the analysis operator can be controlled via (28) [35]

$$r\left(\mathbf{\Omega}_{Red}\right) = -\sum_{1\leq i\leq j\leq p} \log\left(1 - \left(\boldsymbol{\omega}_{i,:}^{\mathbf{T}}\boldsymbol{\omega}_{j,:}\right)^2\right) \tag{28}$$

where $\boldsymbol{\omega}_i$ and $\boldsymbol{\omega}_j$ represent the transpose of the $i$th row and $j$th row of $\mathbf{\Omega}_{Red}$. This constraint can ensure that the analysis operator does not have linear dependent rows. Combining (15), (25), (27), and (28), the cost function of the analysis operator learning can be expressed as

$$\begin{aligned}\mathbf{\Omega}_{Red} : = \arg\min_{\mathbf{\Omega}_{Red}^T \in \mathbf{OB}(64,p)} & \sum_{k}^{K} \mathrm{g}\left(\mathbf{\Omega}_{Red}\mathbf{s}_k\right) + k_1 \cdot h\left(\mathbf{\Omega}_{Red}\right) \\ & + \mu_1 \cdot r\left(\mathbf{\Omega}_{Red}\right).\end{aligned} \tag{29}$$

The conjugate gradient on the oblique manifold is employed to solve the optimization problem (29) [35]. The parameters $k_1$ and $\mu_1$ influence the condition number and the mutual coherence for the analysis operator. For an in-depth introduction to the GOAL method, we refer the interested reader to [35].

In this study, by setting $p = 128$, the initial matrix $\mathbf{\Omega}_{Red_0}$ is generated randomly with the size of $128 \times 64$, and the entries of the matrix are random numbers between 0 and 1. After solving (29), the optimal analysis operator $\mathbf{\Omega}_{Red}$ for the red band is obtained. In this way, the analysis operators $\mathbf{\Omega}_{Green}$, $\mathbf{\Omega}_{Blue}$, and $\mathbf{\Omega}_{NIR}$, which correspond to the green, blue, and near-infrared bands, respectively, can be constructed.

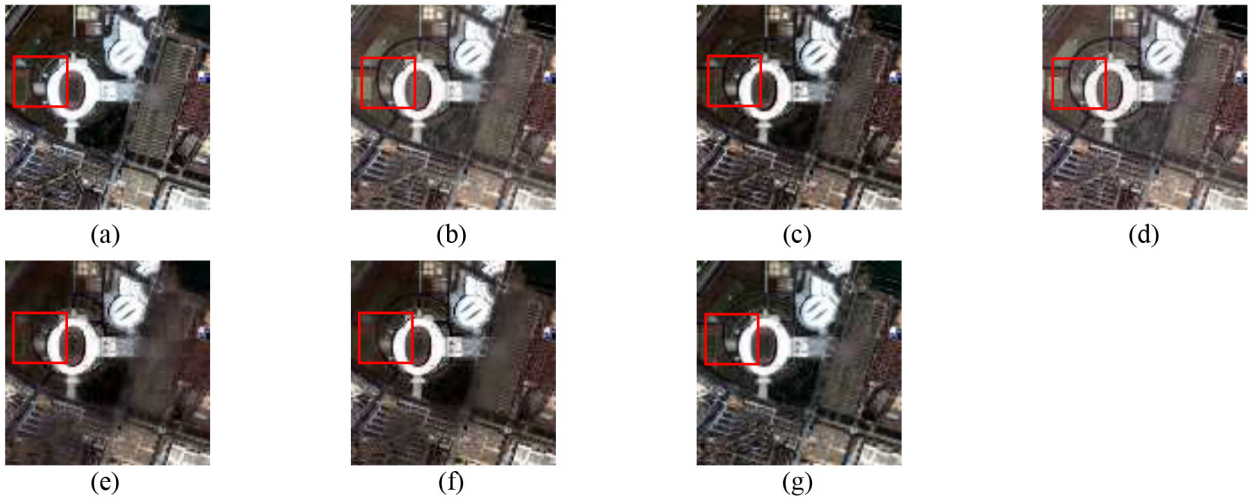



Fig. 4. Results of the image fusion with the degraded IKONOS image. (a) Original. (b) GIHS. (c) AWLP. (d) GS. (e) SparseFI. (f) PN-TSSC. (g) Proposed method.

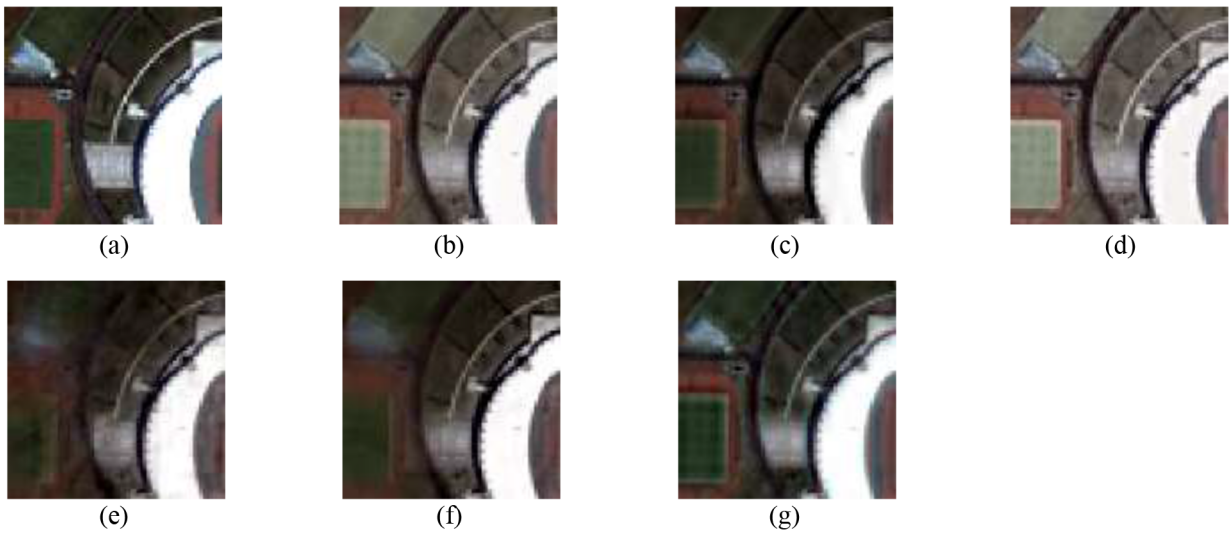



Fig. 5. Close-ups from Fig. 4. (a) Original. (b) GIHS. (c) AWLP. (d) GS. (e) SparseFI. (f) PN-TSSC. (g) Proposed method.

In this linear transformation system, the matrix $\mathbf{\Omega}$ can be formulated as the direct sum of these four operator matrices

$$\mathbf{\Omega} = \bigoplus_{b=1}^{4} \mathbf{\Omega}_b = \mathbf{\Omega}_{Blue} \oplus \mathbf{\Omega}_{Green} \oplus \mathbf{\Omega}_{Red} \oplus \mathbf{\Omega}_{NIR} = \begin{bmatrix} \mathbf{\Omega}_{Blue} & & & \\ & \mathbf{\Omega}_{Green} & & \\ & & \mathbf{\Omega}_{Red} & \\ & & & \mathbf{\Omega}_{NIR} \end{bmatrix}. \quad (30)$$

With the operator $\mathbf{\Omega}$, the analyzed vector of the fused MS image patch can be represented as

$$\alpha = \begin{bmatrix} \mathbf{\Omega}_{Blue} & & & \\ & \mathbf{\Omega}_{Green} & & \\ & & \mathbf{\Omega}_{Red} & \\ & & & \mathbf{\Omega}_{NIR} \end{bmatrix} \begin{bmatrix} \mathbf{s}_{BP} \\ \mathbf{s}_{GP} \\ \mathbf{s}_{RP} \\ \mathbf{s}_{NP} \end{bmatrix} = \mathbf{\Omega}\mathbf{s} \quad (31)$$

where $\boldsymbol{\alpha}$ is the sparse coefficient of the MS image patch, and $\mathbf{s}_{BP}$, $\mathbf{s}_{GP}$, $\mathbf{s}_{RP}$, and $\mathbf{s}_{NP}$ are column vectors representing the lexicographically reordered MS image patches for the blue, green, red, and near-infrared bands, respectively. This also implies that $\boldsymbol{\alpha}$ has the sparsest representation of the remote sensing image patch at a given overcomplete analysis operator $\mathbf{\Omega} \in \mathbb{R}^{512\times256}$.

### *E. Flowchart*

The scheme of the proposed image fusion method is illustrated in Fig. 2.

The whole procedure of the proposed remote sensing image fusion scheme can be divided into two stages: 1) the analysis operator learning stage; and 2) the image fusion stage. First, the resampled LR MS images and the PAN image are randomly abstracted into image patches. These image patches are then exploited for training the analysis operators by solving (29). Second, during the image fusion stage, the HR PAN and the LR MS image patches are jointly stacked as the $\mathbf{y}$ measurements. By solving (14), the HR MS image patches are obtained. Finally, all the obtained HR MS image patches are synthesized as the desired HR MS image.

TABLE I
QUANTITATIVE ASSESSMENT RESULTS OF THE SIMULATED EXPERIMENT SHOWN IN FIG. 4

| | | GS | GIHS | AWLP | SparseFI | PN-TSSC | **Proposed** |
|---|---|---|---|---|---|---|---|
| CC | B | 0.9128 | 0.9257 | 0.9440 | 0.9372 | 0.9481 | **0.9490** |
| | G | 0.9136 | 0.9314 | 0.9507 | 0.9448 | 0.9544 | **0.9652** |
| | R | 0.9000 | 0.9235 | 0.9430 | 0.9375 | 0.9456 | **0.9597** |
| | NIR | 0.9095 | 0.8968 | 0.8720 | 0.9058 | 0.9179 | **0.9275** |
| | Avg | 0.9090 | 0.9193 | 0.9274 | 0.9313 | 0.9415 | **0.9503** |
| RMSE | B | 25.0405 | 27.8421 | 20.5372 | 21.3120 | 19.1888 | **18.9910** |
| | G | 24.9956 | 27.2382 | 19.3278 | 20.0953 | 18.1256 | **15.9294** |
| | R | 27.1059 | 28.4827 | 21.1962 | 21.5845 | 20.0468 | **17.4291** |
| | NIR | 20.2932 | 27.0878 | 29.6334 | 21.0680 | 19.5106 | **19.1536** |
| | Avg | 24.3588 | 27.6627 | 22.6736 | 21.0150 | 19.2180 | **17.8758** |
| SSIM | B | 0.9153 | 0.9098 | **0.9433** | 0.9256 | 0.9398 | 0.9411 |
| | G | 0.9206 | 0.9138 | 0.9505 | 0.9343 | 0.9469 | **0.9583** |
| | R | 0.8977 | 0.8941 | 0.9391 | 0.9243 | 0.9348 | **0.9490** |
| | NIR | 0.9381 | 0.9406 | 0.8975 | 0.9302 | 0.9408 | **0.9422** |
| | Avg | 0.9179 | 0.9146 | 0.9326 | 0.9286 | 0.9406 | **0.9476** |
| SAM | | 10.1552 | 8.1511 | 7.5478 | 8.1723 | 8.1345 | **7.3476** |
| ERGAS | | 6.3619 | 7.0508 | 5.05911 | 5.3395 | 4.8799 | **4.4850** |
| $Q_4$ | | 0.6987 | 0.7294 | 0.7360 | 0.5802 | 0.6861 | **0.7701** |

## IV. EXPERIMENTS AND DISCUSSIONS

### A. Experimental Setting

To evaluate the effectiveness of the proposed method, we adopted Wald's protocol, which states that a synthetic image should be similar to the image that the corresponding sensor would observe at the highest spatial resolution [57]. We carried out two groups of simulated experiments with two kinds of satellite sensor data, i.e., QuickBird and IKONOS, and a real-data experiment with an IKONOS data set. To quantitatively assess the quality of the simulated experiment results, we need a ground-truth image as the reference. A common method for this is to degrade the original PAN images and the MS images into an inferior resolution level. The experiments are then performed on the degraded images, and the original MS images are selected as the real HR images to compare with the fused images. During this stage, the MTF variation of the imagery system should be taken into account in the simulated experiments. Referring to [19], we obtained the degraded MS image by filtering the original MS image with MTF-tailored low-pass filters and then downsampling by a factor of four. The approximated Gaussian filters with different Nyquist cutoff frequencies simulate the MTF of satellites. The Nyquist cutoff frequencies of QuickBird and IKONOS for different spectral bands are provided in [19] and [58]. The cutoff frequencies for the blue, green, red, and near-infrared bands are around 0.34, 0.32, 0.30, and 0.24 for QuickBird and 0.27, 0.28, 0.29, and 0.28 for IKONOS, respectively. The degraded PAN image was also generated by an MTF-tailored low-pass filter with a cutoff value of 0.17 for IKONOS and 0.15 for QuickBird, respectively.

The common evaluation methods for image fusion results consist of two aspects: 1) visual inspection and 2) quantitative evaluation. For the quantitative evaluation, the following typical evaluation indexes were utilized: the correlation coefficient (CC) [59], the structural similarity metric (SSIM) [60], and the root-mean-square error (RMSE). These indexes were calculated for each band between the fused MS images and the original reference MS image. A small value of spectral angle mapper (SAM) [59] and the *erreur relative globale adimensionnelle de synthèse* (ERGAS) [61] indicates good fusion results. The $Q_4$ quality index is a generalization for four-band images of the $Q$ index [60], [62], which is a specific index known as the universal image quality index (UIQI) that was proposed in [59]. The $Q_4$ index gives values in the range of [0, 1], with 1 being the ideal value.

(a) (b)

Fig. 6. Input QuickBird images. (a) Degraded MS image at an 11.2-m spatial resolution. (b) Degraded PAN image at a 2.8-m spatial resolution.

Another quality measurement protocol which does not require an HR reference MS image is called the quality not requiring a reference (QNR) index, which was proposed by Alparone *et al.* [63]. The results of the real experiments were confirmed quantitatively in terms of QNR, $D_\lambda$, and $D_s$. The spectral distortion $D_\lambda$ is calculated as follows [63]:

$$D_\lambda = \frac{2}{B(B-1)} \sum_{b=1}^{B} \sum_{k=1}^{b-1} \times \left| Q(MS_{low,b}, MS_{low,k}) - Q(MS_b, MS_k) \right| \quad (32)$$

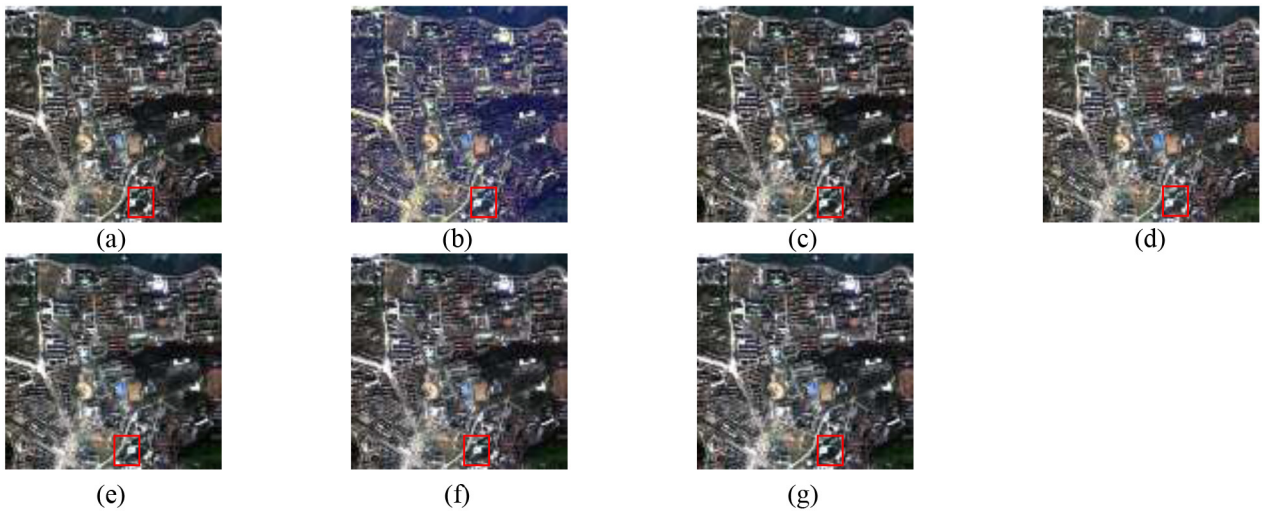



Fig. 7. Image fusion results with the degraded QuickBird image. (a) Original. (b) GIHS. (c) AWLP. (d) GS. (e) SparseFI. (f) PN-TSSC. (g) Proposed method.

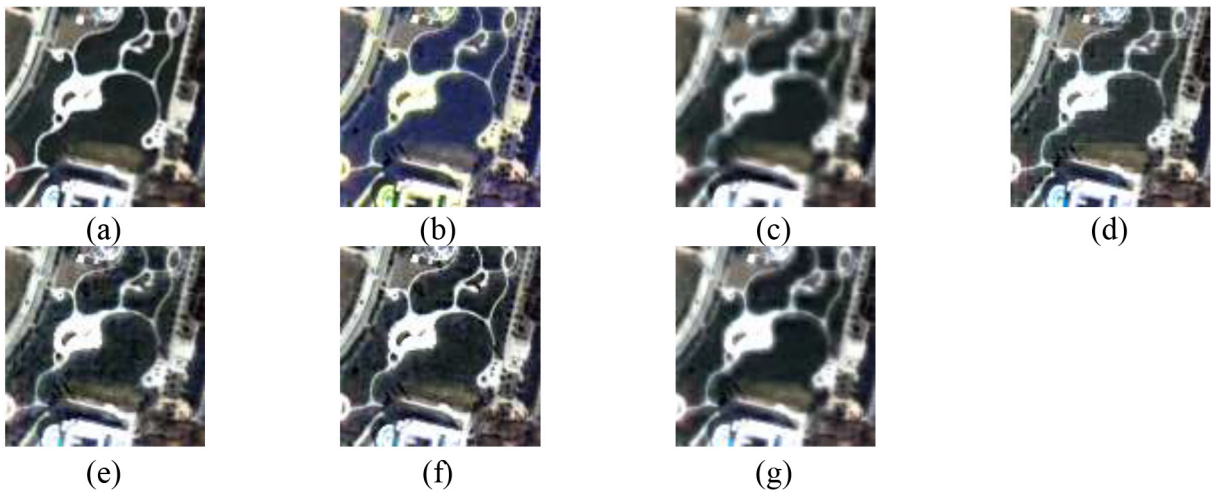



Fig. 8. Close-ups from Fig. 7. (a) Original. (b) GIHS. (c) AWLP. (d) GS. (e) SparseFI. (f) PN-TSSC. (g) Proposed method.

where $B$ is the number of bands of the MS image, and $MS_{low,b}$ and $MS_b(b \in \{1, \ldots, B\})$ are the $b$th low-resolution and sharpened MS bands, respectively. Q is the UIQI index.

The spatial distortion $D_s$ is calculated as follows:

$$D_s = \frac{1}{B}\sum_{b=1}^{B} |Q(MS_{low,\,b}, PAN_{low}) - Q(MS_b, PAN)| \tag{33}$$

where PAN and $PAN_{low}$ are the full-resolution and downsampled PAN images, respectively. Then, taking into account the spatial and spectral distortion, the QNR is written as [63]

$$QNR = (1 - D_\lambda)(1 - D_s) \tag{34}$$

We compared our image fusion results with several typical fusion methods: the traditional GS method [6]; the generalized IHS (GIHS) method [4]; and the additive wavelet luminance proportional (AWLP) approach [9], which was implemented via a two-level pyramidal decomposition. Two state-of-the-art sparse regularization based methods were also used: SparseFI [3] and the two-step sparse coding with patch normalization (PN-TSSC) method [18].

To obtain the analysis operator for the MS images, the concrete values of $k_1$ and $\mu_1$ in (29) were empirically set as $k_1 = 40$ and $\mu_1 = 1000$, which led to the optimal analysis operator in this experiment. In addition, the parameter $\mu$ in (25) was set as $\mu = 10^4$ for all the experiments in this paper. The parameter $\lambda$ in (14) was set as $\lambda = 2.8 \times 10^6$ for IKONOS and $\lambda = 2.9 \times 10^4$ for QuickBird, respectively. For $\beta$ in (12), we empirically set it as $\beta = 0.075$ for IKONOS and $\beta = 0.8$ for QuickBird. The weights in (11) was set as follows [19]: 1) $\theta^1 = 0.1071$, $\theta^2 = 0.2646$, $\theta^3 = 0.2696$, and $\theta^4 = 0.3587$ for IKONOS; and 2) $\theta^1 = 0.1139$, $\theta^2 = 0.2315$, $\theta^3 = 0.2308$, and $\theta^4 = 0.4239$ for QuickBird.

The optimal size of image patch for the LR MS image was $9 \times 9$ for the SparseFI method, and we chose an image patch size of $7 \times 7$ for the PN-TSSC method. The overlapping area size was set as $9 \times 4$ pixels for the SparseFI method and $7 \times 4$ pixels for the TN-TSSC method, respectively. The size of image patch and the parameters for the other methods were set as recommended. For the SparseFI and PN-TSSC methods, the

TABLE II
QUANTITATIVE ASSESSMENT RESULTS OF THE SIMULATED EXPERIMENT SHOWN IN FIG. 8

| | | GS | GIHS | AWLP | SparseFI | PN-TSSC | **Proposed** |
|---|---|---|---|---|---|---|---|
| CC | B | 0.9429 | 0.7923 | 0.9458 | 0.9464 | 0.9455 | **0.9605** |
| | G | 0.9550 | 0.9613 | 0.9485 | 0.9595 | 0.9612 | **0.9683** |
| | R | 0.9537 | 0.9633 | 0.9508 | 0.9597 | 0.9611 | **0.9693** |
| | NIR | 0.9542 | 0.9594 | 0.9544 | 0.9614 | 0.9655 | **0.9687** |
| | Avg | 0.9515 | 0.9191 | 0.9499 | 0.9568 | 0.9583 | **0.9667** |
| RMSE | B | 11.2778 | 20.6586 | 10.3649 | 10.3966 | 10.3997 | **8.8923** |
| | G | 21.2271 | 21.4900 | 20.8534 | 18.7767 | 18.1458 | **16.4458** |
| | R | 22.3163 | 22.0130 | 21.4588 | 19.7404 | 19.0814 | **17.0241** |
| | NIR | 29.3940 | 29.3254 | 29.1024 | 27.3347 | 25.3993 | **24.2778** |
| | Avg | 21.0538 | 23.3717 | 20.4449 | 19.0261 | 18.2566 | **16.6600** |
| SSIM | B | 0.9813 | 0.9485 | 0.9832 | 0.9813 | 0.9822 | **0.9855** |
| | G | 0.9582 | 0.9610 | 0.9544 | 0.9585 | 0.9611 | **0.9665** |
| | R | 0.9517 | 0.9564 | 0.9511 | 0.9542 | 0.9564 | **0.9630** |
| | NIR | 0.9295 | 0.9247 | 0.9286 | 0.9270 | 0.9395 | **0.9443** |
| | Avg | 0.9552 | 0.9476 | 0.9544 | 0.9553 | 0.9598 | **0.9648** |
| SAM | | 2.0948 | 2.4099 | 1.8639 | 2.0943 | 2.0059 | **1.8164** |
| ERGAS | | 1.8621 | 2.0169 | 1.8120 | 1.6873 | 1.6090 | **1.4711** |
| $Q_4$ | | 0.8367 | 0.8130 | 0.8304 | 0.8089 | 0.8477 | **0.8718** |

regularization parameters were set as 16 384 for IKONOS and 8192 for QuickBird, respectively.

### B. Experiments With IKONOS Data

The IKONOS system simultaneously offers a four-band MS image with a 4-m resolution and a single-band PAN image with a 1-m resolution. Fig. 3(a) and (b) shows the LR MS image with a resolution of 16 m and the PAN image with a resolution of 4 m, respectively. The original MS image with a 4-m resolution was used as reference. In this study, we utilized a simulated IKONOS LR MS image with the size of $64 \times 64$ and a corresponding PAN image with the size of $256 \times 256$.

The GS algorithm was implemented in ENVI 5.1 software in mode 1. The experimental results of the six image fusion methods are shown in Fig. 4. To facilitate a comparison, the details of local regions of the images are exhibited in Fig. 5. By visually comparing the fused images with the original image, it can be seen that all the fusion methods can effectively fuse the LR MS image and the HR PAN image. However, it can be clearly observed that the GIHS and GS results suffer from serious spectral distortion compared with the original MS image. Although the AWLP and PN-TSSC methods are able to maintain the spectral information of the original image, some details are missing and the image is blurred to some degree. In addition, we can see from Fig. 5(e) that the fusion result of SparseFI suffers from serious blurring and aliasing distortion. Overall, the image fusion result by the proposed method is the closest to the original image. The proposed method not only provides a high-spatial resolution while preserving the image details, but also decreases the spectral distortion to a great extent.

The quantitative assessment indexes for the fusion results are shown in Table I, in which the best results for each quality index are labeled in bold, and the second-best results for each quality index are underlined. In this table, B, G, R, and near-infrared represent the results of the blue, green, red, and near-infrared bands, respectively. We use the average values of all the band indexes to indicate the image quality.

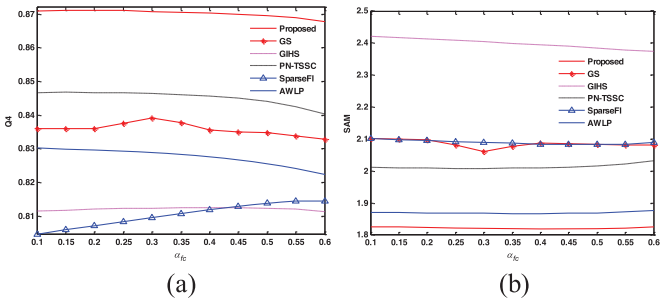



Fig. 9. Variation of the image fusion performance (Q4 and SAM) as a function of $\alpha_{cf}$. (a) Q4. (b) SAM.

It can be seen from Table I that the proposed method acquires the best evaluation results, in terms of CC, RMSE, ERGAS, SSIM, and $Q_4$, which suggests that the fusion result of the proposed method is the most closely correlated to the original MS image. In terms of the SAM index, AWLP and the proposed method perform much better than the other methods from the aspect of spectral information protection. On the whole, the quantitative assessment results are consistent with the visual evaluation, and the proposed method achieves a better fusion result than the other methods.

### C. Experiments With QuickBird Data

In this section, we analyze another group of experimental results of image fusion with QuickBird data to further reveal the performance of the proposed algorithm. The QuickBird data set provides a four-band 2.8-m resolution MS image and a 0.7-m resolution PAN image. Fig. 6(a) and (b) shows a pair of simulated QuickBird images with resolutions of 11.2 and 2.8 m, respectively. The size of the LR MS image in this simulated

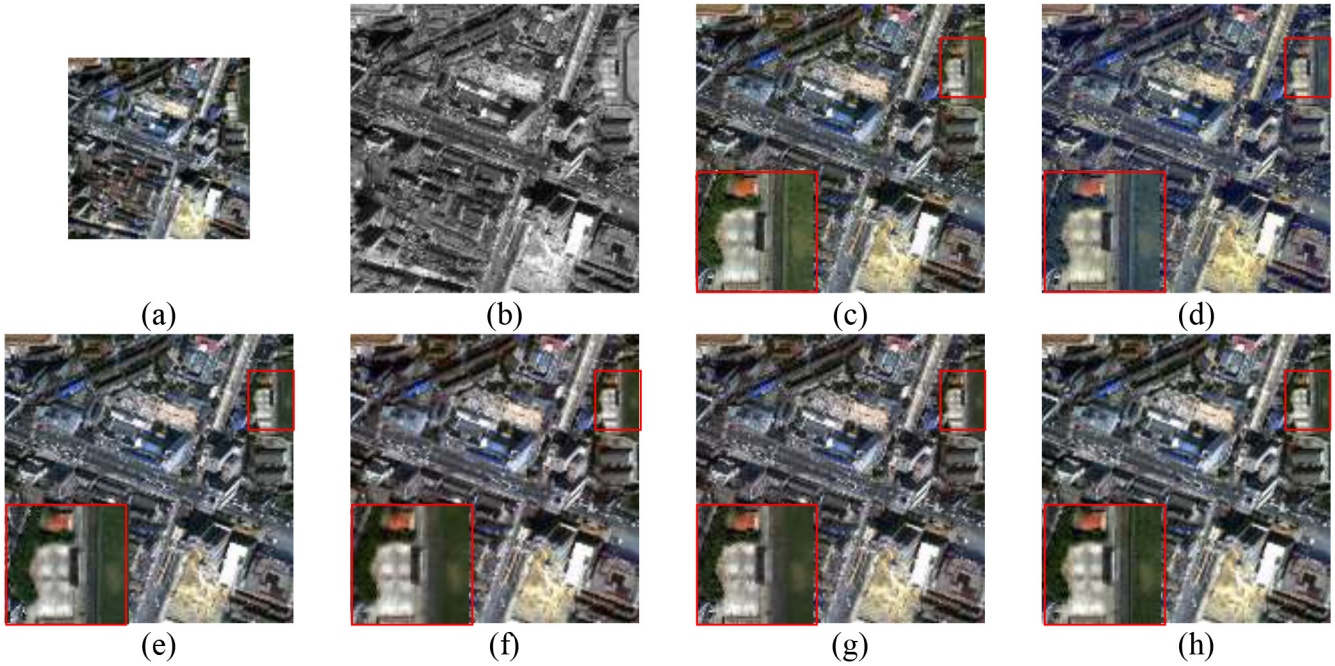



Fig. 10. Real-data experiment with IKONOS data. (a) MS image at a 2.8-m spatial resolution. (b) PAN image at a 0.7-m spatial resolution. (c) GS. (d) GIHS. (e) AWLP. (f) SparseFI. (g) PN-TSSC. (h) Proposed method.

experiment was $150 \times 150$, and the corresponding PAN image was sized $600 \times 600$, as shown in Fig. 6.

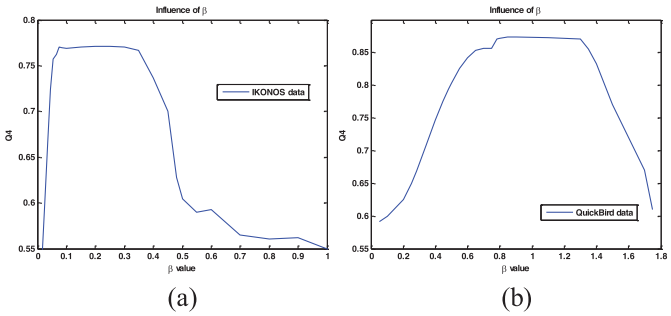
The results of the different image fusion methods are shown in Fig. 7, and the enlarged images for the same position are shown in Fig. 8. It can be clearly seen from Fig. 7(b) and (d) and the enlarged images in Fig. 8(b) and (d) that the GS and GIHS methods produce serious spectral distortion when compared with all the other methods. The image fusion result by the AWLP method is acceptable with regard to the spectral characteristic, but it has a little blurring compared with the original MS images.

It can be observed that the fusion results of the SparseFI and PN-TSSC methods in Fig. 8(e) and (f) contain serious artifacts. Outlier pixels in some regions with small objects or fine details appear in these two images. Overall, it can be seen that the fusion result by the proposed method is the closest to the original image. To sum up, the proposed method provides higher quality spatial details and decreases the spectral distortion, from a visual point of view.

The visual interpretation can only show the quality of three bands of the fusion results. Therefore, we also evaluated the fusion results for all the MS bands via a comprehensive quantitative evaluation. The quantitative assessment results are shown in Table II, where the best results of each index are marked in bold. Here, it can be clearly observed that the proposed method obtains the best evaluation results for all the bands, in terms of both CC and SSIM, which illustrates that the proposed method maintains the highest correlation and the least radiometric distortion between the fusion result and the original MS image. For the ERGAS index and the Q4 index, the proposed method again performs better than the other fusion methods, as expected. As for the SAM index, the proposed method again provides the best results. For all the quantitative indexes in this table, the proposed method performs better than the other image fusion methods. Overall, the conclusion can be drawn that the proposed method is suitable for the image fusion of QuickBird data, and it performs better than the other state-of-the-art methods.

TABLE III
QUALITY MEASURES FOR THE DIFFERENT IMAGE FUSION APPROACHES WITH REAL IKONOS DATA

| | GS | GIHS | AWLP | SparseFI | PN-TSSC | **Proposed** |
|---|---|---|---|---|---|---|
| $D_\lambda$ | 0.0366 | 0.0731 | 0.0552 | 0.0632 | 0.0359 | **0.0298** |
| $D_s$ | 0.1192 | 0.1362 | 0.1018 | 0.1054 | 0.0624 | **0.0398** |
| QNR | 0.8486 | 0.8006 | 0.8486 | 0.8381 | 0.9039 | **0.9316** |




(a) (b)

Fig. 11. Plot of the influence of $\beta$ for the IKONOS and QuickBird data sets (Q4). (a) Q4 with IKONOS data. (b) Q4 with QuickBird data.

### D. Robustness to the Imagery System MTF

The MTF of the imagery system was investigated in [11], [13], [62], and [64]. It corresponds to the imagery system

TABLE IV
QUANTITATIVE ASSESSMENT RESULTS OF THE PROPOSED METHOD WITH DIFFERENT VALUES OF $\lambda$ ($\lambda$ IS FROM $3 \times 10^5$ TO $3 \times 10^{11}$) FOR THE IKONOS DATA

| $\lambda$ | $3\times10^5$ | $5\times10^6$ | $3\times10^6$ | $3\times10^7$ | $8\times10^7$ | $3\times10^8$ | $3\times10^9$ | $3\times10^{10}$ | $3\times10^{11}$ |
|---|---|---|---|---|---|---|---|---|---|
| $CC_{Avg}$ | 0.9231 | 0.0.9467 | 0.9477 | 0.9483 | 0.9485 | **0.9486** | 0.9484 | 0.9485 | 0.9484 |
| $SSIM_{Avg}$ | 0.9213 | 0.9433 | 0.9444 | 0.9453 | 0.9455 | **0.9456** | 0.9455 | 0.9455 | 0.9454 |
| ERGAS | 5.5377 | 4.6698 | 4.6359 | 4.5867 | 4.5756 | **4.5701** | 4.5778 | 4.5727 | 4.5807 |
| SAM | 6.2832 | 7.8791 | 7.7662 | 7.6715 | 7.6346 | **7.6327** | 7.6380 | 7.6317 | 7.6417 |
| $Q_4$ | 0.5958 | 0.7673 | 0.7689 | 0.7695 | **0.7697** | 0.7695 | 0.7696 | 0.7610 | 0.7273 |
| $RMSE_{Avg}$ | 21.8748 | 18.7518 | 18.5468 | 18.3731 | 18.3186 | **18.2983** | 18.3299 | 18.3093 | 18.3427 |

TABLE V
QUANTITATIVE ASSESSMENT RESULTS OF THE PROPOSED METHOD WITH DIFFERENT VALUES OF $\lambda$ ($\lambda$ IS FROM $3 \times 10^3$ TO $3 \times 10^7$) FOR THE QUICKBIRD DATA

| $\lambda$ | $3\times10^3$ | $8\times10^3$ | $4\times10^4$ | $8\times10^4$ | $3\times10^5$ | $6\times10^5$ | $3\times10^6$ | $6\times10^6$ | $3\times10^7$ |
|---|---|---|---|---|---|---|---|---|---|
| $CC_{Avg}$ | 0.9126 | 0.9373 | 0.9639 | **0.9647** | **0.9647** | 0.9579 | 0.9420 | 0.5606 | 0.2850 |
| $SSIM_{Avg}$ | 0.9024 | 0.9300 | **0.9616** | 0.9611 | 0.9607 | 0.9546 | 0.9586 | 0.8867 | 0.8276 |
| ERGAS | 2.4349 | 2.0650 | **1.5303** | 1.5815 | 1.5889 | 1.6958 | 1.9354 | 7.8806 | 18.6517 |
| SAM | 2.2144 | 2.0392 | **1.8275** | 1.8670 | 1.8732 | 2.0783 | 2.9283 | 3.5370 | 4.9979 |
| $Q_4$ | 0.6180 | 0.7247 | 0.8528 | 0.8732 | **0.8746** | 0.8628 | 0.8727 | 0.7873 | 0.7310 |
| $RMSE_{Avg}$ | 27.3442 | 23.2838 | 16.9518 | **17.7782** | 17.8844 | 19.3393 | 22.4210 | 94.4426 | 223.8258 |

module of the Fourier transform of its PSF, and is also usually considered to obey a Gaussian distribution.

In practical cases, increasing the cutoff of the Nyquist frequency violates the Shannon's condition and causes aliasing. In contrast, decreasing the cutoff value reduces the amount of signal in the pass-band and leads to blurring. The tail of the MTF function also gives rise to aliasing. Therefore, the cutoff value must be a tradeoff between the maximum spatial resolution and minimum aliasing of the sampled signal.

We also studied the robustness of the proposed method to MTF variation of the imagery system. This experiment was implemented with the QuickBird data. The cutoff frequency was the Nyquist frequency, which is equal to (0.125), as shown in [13]. The simulated PAN and MS images were generated by applying Gaussian-tailored low-pass filters with the frequency response at the cutoff Nyquist frequency increasing from 0.1 to 0.6 in steps of 0.05 [65]. Fig. 9 shows the variation of the performance ($Q_4$ and SAM) of the different approaches as functions of $\alpha_{cf}$.

The performances of all the results improve as $\alpha_{cf}$ increases, and the proposed method provides the best result in terms of $Q_4$ and SAM. We can see from the diagram that SparseFI and GS are more sensitive to the aliasing increase than the other methods. Overall, the proposed method is more robust and obtains a better result than the other methods.

### E. Real-Data Experiment

We also evaluated the proposed method in a real-data experiment. Fig. 10(a) and (b) shows the IKONOS image at a 2.8-m (MS) and 0.7-m (PAN) spatial resolution, respectively. The $D_\lambda$, $D_s$, and QNR quality assessment indexes were used to evaluate the fused image [63]. The results of GS, GIHS, AWLP, SparseFI, PN-TSSC, and the proposed method are shown in Fig. 10(c)–(h).

Here, it can be seen that the GIHS method generates spectral distortion, as shown in Fig. 10(d). We can also see from Fig. 10(c), (e), and (f) that the results of GS, AWLP, and SparseFI, respectively, contain a little blurring. PN-TSSC and the proposed method [Fig. 10(g) and (h)] generate an HR MS image with satisfactory spectral and spatial preservation. The quality assessment indexes are shown in Table III. Clearly, the advantage of PN-TSSC and the proposed method is that they result in very-low-spatial distortion ($D_s < 0.1$). However, the proposed method generates better $D_\lambda$, $D_s$, and QNR values than PN-TSSC.

### F. Parameter Analysis

The impact of parameters $\lambda$ (14) and $\beta$ (12) in the proposed method was also investigated. We set parameter $\lambda$ as a constant, and the trend graphs of $Q_4$ with respect to parameter $\beta$ for the IKONOS and QuickBird data sets are shown in Fig. 11(a) and (b), respectively. The results of the two groups of experiments show that the proposed method can reach a stable performance when parameter $\beta$ is set as [0.06, 0.36] and [0.8, 1.3] for the IKONOS and QuickBird data sets, respectively. The difference can be explained by the fact that the two data sets have different spectral characteristics. It can also be seen from Fig. 11 that the proposed method has a wide stable range with regard to parameter $\beta$.

We also fixed the optimal $\beta$ value and explored the influence of parameter $\lambda$. The influence of regularization parameter $\lambda$ on the IKONOS and QuickBird data sets is shown in Tables IV and V, respectively. For simplicity, the $\mathrm{CC_{Avg}}$, $\mathrm{SSIM_{Avg}}$ ($\mathrm{RMSE_{Avg}}$, $\mathrm{CC_{Avg}}$, and $\mathrm{SSIM_{Avg}}$ are the average of RMSE, CC, and SSIM, respectively), ERGAS, SAM, and $Q_4$ indexes are used to evaluate the performance. Here, it can be observed that all the indexes are stable over a wide range for both the IKONOS and QuickBird data sets in terms of $\lambda$.

### G. Time Cost

Except for the GS method, all the other methods were implemented in MATLAB 2012a. The personal computer used was a

TABLE VI
TIME COSTS OF THE DIFFERENT IMAGE FUSION METHODS FOR THE $256 \times 256$ IKONOS IMAGE

| Method | GS | GIHS | AWLP | SparseFI | PN-TSSC | Proposed |
|---|---|---|---|---|---|---|
| Time (s) | 0.03 | 0.07 | 0.03 | 0.2 | 0.6785 | 85 |

TABLE VII
TIME COSTS OF THE DIFFERENT IMAGE FUSION METHODS FOR THE $600 \times 600$ QUICKBIRD IMAGE

| Method | GS | GIHS | AWLP | SparseFI | PN-TSSC | Proposed |
|---|---|---|---|---|---|---|
| Time (s) | 2.08 | 0.21 | 0.48 | 21.85 | 46.89 | 720 |

DELL T1500. The central processing unit (CPU) was a dual-core Intel Core i3 540 at 3.07 GHz. The RAM was 6 GB and the operating system was Windows 7 64-bit. The running times of all the methods in the experiments with the IKONOS and QuickBird data sets are provided in Tables VI and VII, respectively.

From Tables VI and VII, it can be clearly seen that the sparse regularization based fusion methods are time-consuming due to the large computational complexity of the $\ell_1$-norm minimization problem. Because the proposed method takes more time than the other methods, parallel processing and other acceleration strategies will be investigated in our future work.

## V. CONCLUSION

This paper has presented a novel remote sensing image fusion method based on the analysis sparse model. With the PAN and MS image degradation model, the remote sensing image fusion task is formulated as an ill-posed inverse problem. An image fusion model which is based on a sparse regularization method is proposed to recover the HR MS image from the observed measurements. In our observation model, the MTF of the imagery system is taken into account, which makes the method more practical for real remote sensing applications. During the operator learning stage, the analysis operator for each band is trained by the GOAL algorithm with the upsampled LR MS image and the PAN image. These analysis operators are then combined into one operator for the LR MS image. With this analysis operator, the arising optimization task for the image fusion is solved by the conjugate gradient method. The proposed method was compared with the state-of-the-art image fusion techniques on both IKONOS and QuickBird data. The experimental results suggest that the proposed method is competitive in both spatial and spectral quality. The experimental results also show that the parameters of the proposed method are stable. However, the proposed scheme takes more time than the traditional methods. Therefore, in the future, we will focus on acceleration strategies to solve this problem.

## ACKNOWLEDGMENT

The authors would like to thank the editor and the anonymous reviewers, whose comments greatly helped in improving the presentation and quality of this paper. They are very grateful for the helpful discussion with A. Kallel.

## REFERENCES

[1] Q. Liu, Y. Wang, Z. Zhang, and L. Liu, "Pan-sharpening based on weighted red black wavelets," *IET Image Process.*, vol. 8, no. 8, pp. 477–488, 2014.
[2] L. Zhang, H. Shen, W. Gong, and H. Zhang, "Adjustable model-based fusion method for multispectral and panchromatic images", *IEEE Trans. Syst. Man Cybern. Part B*, vol. 42, no. 6, pp. 1693–1704, 2012.
[3] 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.
[4] T. M. Tu, S. C. Su, H. C. Shyu, and P. S. Huang, "A new look at HIS like image fusion methods," *Inf. Fusion*, vol. 2, no. 3, pp. 177–186, Sep. 2001.
[5] V. P. Shah and N. H. Younan, "An efficient pan-sharpening method via a combined adaptive PCA approach and contourlets," *IEEE Trans. Geosci Remote Sens.*, vol. 46, no. 5, pp. 1323–1335, May 2008.
[6] 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.
[7] Y. Zhang, "Problems in the fusion of commercial high-resolution satellite images as well as Landsat-7 images, and initial solutions," in *Proc. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 2002, vol. 34, pp. 587–592.
[8] 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.
[9] X. Otazu, M. Gonzalez-Audicana, O. Forsand, and J. Nunez, "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.
[10] K. Amolins, Y. Zhang, and P. Dare, "Wavelet based image fusion techniques—An introduction, review and comparison," *ISPRS J. Photogramm. Remote Sens.*, vol. 62, no. 4, pp. 249–263, Sep. 2007
[11] 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.
[12] H. Song, S. Yu, L. Song, and X. Yang, "Fusion of multispectral and panchromatic satellite images based on contourlet transform and local average gradient," *Opt. Eng.*, vol. 46, p. 020502, Feb. 2007.
[13] A. Kallel, "MTF-adjusted pan sharpening approach based on coupled multiresolution decompositions," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 3124–3145, Jun. 2015.
[14] 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.
[15] Z. H. 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.
[16] K. A. Kalpoma and J. I. Kudoh, "Image fusion processing for IKONOS 1-m color imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3075–3086, Oct. 2007.
[17] 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.
[18] 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.
[19] S. Li, H. Yin, and L. Fang, "Remote sensing image fusion via sparse presentations over learned dictionaries," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 9, pp. 4779–4789, Sep. 2013.
[20] M. Guo, H. Zhang, J. Li, and L. Zhang, "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.
[21] Y. Tsaig and D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory.*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
[22] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
[23] A. Buades, B. Coll, and J. M. Morel, "A nonlocal algorithm for image denoising," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR'05)*, 2005, vol. 2, pp. 60–65.
[24] T. Peleg, Y. Eldar, and M. Elad, "Exploiting statistical dependencies in sparse representations for signal recovery," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2286–2303, May 2012.
[25] S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.

[26] J. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.
[27] M. Elad, "Optimized projections for compressed sensing," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5695–5702, Dec. 2007
[28] E. J. Candes, Y. C. Eldar, D. Needell, and P Randall, "Compressed sensing with coherent and redundant dictionaries," *Appl. Comput. Harmon. Anal.*, vol. 31, pp. 59–73, May 2010.
[29] Q. Ning, K. Chen, L. Yi, C. Fan, Y. Lu, and J. Wen, "Image super-resolution via analysis sparse prior," *IEEE Signal Process. Lett.*, vol. 20, no. 4, pp. 399–402, Dec. 2013.
[30] M. Elad, P. Milanfar, and R. Rubinstein, "Analysis versus synthesis in signal priors," *Inverse Prob.*, vol. 23, no. 3, pp. 947–968, Jun. 2007.
[31] S. Nam, M. E. Davies, M. Elad, and R. Gribonval, "The cosparse analysis model and algorithms," *Appl. Comput. Harmon. Anal.*, vol. 34, no. 1, pp. 30–56, Jan. 2013.
[32] M. Yaghoobi, S. Nam, R. Gribonva, and M. E. Davies, "Constrained overcomplete analysis operator learning for cosparse signal modeling," *IEEE Trans. Signal Process.*, vol. 61, no. 9, pp. 23–41, May 2013.
[33] J. S. Turek, I. Yavneh, and M. Elad, "On MAP and MMSE estimators for the co-sparse analysis model," *Digit. Signal Process.*, vol. 28, pp. 57–74, May 2014.
[34] C. Nieuwenhuis, D. Cremers, S. Hawe, and M. Kleinsteuber, "Co-sparse textural similarity for image segmentation," ar Xiv preprint arXiv:1312.4746, 2013.
[35] S. Hawe, M. Kleinsteuber, and K. Diepold, "Analysis operator learning and its application to image reconstruction," *IEEE Trans. Image Process.*, vol. 22, no. 6, pp. 2138–2150, Jun. 2013.
[36] R. Giryes, S. Nam, M. Elad, R. Gribonval, and M. E. Davies, "Greedy-like algorithms for the cosparse analysis model," *Linear Algebra Appl.*, vol. 441, pp. 22–60, Jan. 2014.
[37] S. Vaiter, G. Peyre, C. Dossal, and J. Fadili, "Robust sparse analysis regularization," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2001–2016, Apr. 2013.
[38] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*. Hoboken, NJ, USA: Wiley, 1977.
[39] G. Subham and R. Yoram, "Application of L1-norm regularization to epicardial potential solution of the inverse electrocardiography problem," *Ann. Biomed. Eng.*, vol. 37, no. 5, pp. 902–912, May 2009.
[40] D. L. Donoho and E. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via $l_1$ minimization," *Proc. Natl. Acad. Sci.*, vol. 100, no. 5, pp. 2197–2202, 2003.
[41] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *J. Roy. Statist. Soc. B*, vol. 58, no. 1, pp. 267–288, 1996.
[42] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, 2009.
[43] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM J. Comput.*, vol. 24, no. 2, pp. 227–234, 1995.
[44] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proc. IEEE*, vol. 98, no. 6. pp. 948–958, Jun. 2010.
[45] R. Rubinstein, T. Faktor, and M. Elad, "Analysis K-SVD: A dictionary learning algorithm for the analysis sparse model," *IEEE Trans. Signal Process.*, vol. 61, no. 3, pp. 661–677, Feb. 2013.
[46] M. R. Garey and D. S. Johnson, "A guide to the theory of NP-completeness," in *Computers and Intractability*. San Francisco, CA, USA: Freeman, 1990.
[47] R. Giryes, S. Nam, R. Gribonval, and M. E. Davies, "Iterative cosparse projection algorithms for the recovery of cosparse vectors," in *Proc. 19th Eur. Signal Process. Conf. (EUSIPCO'11)*, Barcelona, Spain, 2011, pp. 1460–1464.
[48] R. Giryes and M. Elad, "CoSaMP and SP for the cosparse analysis model," in *Proc. 20th Eur. Signal Process. Conf. (EUSIPCO'12)*, Bucharest, Romania, 2012.
[49] J. Worman, S. Hawe, and M. Kleinsteuber, "Analysis based blind compressive sensing," *IEEE Signal Process. Lett.*, vol. 20, no. 5, pp. 491–494, May 2013.
[50] Y. Chen, R. Ranftl, and T. Pock, "Insights into analysis operator learning: A view from a higher-order filter-based MRF model," *IEEE Trans. Image Process.*, vol. 23, no. 3, pp. 1060–1072, Mar. 2014.
[51] D. Zorn and Y. Weiss, "From learning models of nature image patches to whole image restoration," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV'11)*, Barcelona, Spain, 2011.
[52] Y. H. Dai and Y. Yuan, "An efficient hybrid conjugate gradient method for unconstrained optimization," *IMA J. Numer. Anal.*, vol. 103, pp. 33–47, 2001.
[53] M. Yaghoobi, S. Nam, R. Gribonval, and M. E. Davies, "Analysis operator learning for overcomplete cosparse representations," in *Proc. Eur. Signal Process. Conf. (EUSIPCO'11)*, 2011, pp. 1470–1474.
[54] M. Kiechle, T. Habigt, S. Hawe, and M. Kleinsteuber, "A bimodal co-sparse analysis model for image processing," *Int. J. Comput. Vis.*, vol. 114, no. 2, pp. 233–247, Nov. 2014.
[55] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted l1 minimization," *J. Fourier Anal. Appl.*, vol. 14, no. 5–6, pp. 877–905, 2008.
[56] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton, NJ, USA: Princeton Univ. Press, 2008.
[57] L. Wald, T. Ranchin, and M. Mangolini, "Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images," *Photogramm. Eng. Remote Sens.*, vol. 63, no. 6, pp. 691–699, 1997.
[58] M. M. Khan, L. Alparone, and J. Chanussot, "Pan-sharpening quality assessment using the modulation transfer functions of instruments," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3880–3891, Nov. 2009.
[59] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. M. Bruce, "Comparison of pansharpening algorithms: Outcome of the 2006 GRSS data fusion contest," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3012–3021, Oct. 2007.
[60] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error measurement to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[61] 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.
[62] L. Alparone, S. Baronti, A. Garzelli, and F. Nencini, "A global quality measurement of pan-sharpened multispectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 313–317, Oct. 2004.
[63] 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.
[64] J. M. Delvit, D. Leger, S. Roques, and C. Valorge, "Modulation transfer function and noise assessment," *IEEE Trans. Geosci. Remote Sens.*, vol. 7, no. 10, pp. 4500–4502, Jul. 2003.
[65] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "Advantages of laplacian pyramids over "a trous wavelet transforms for pansharpening of multispectral images," *Proc. SPIE*, vol. 8537(04), pp. 1–10, 2012.

**Chang Han** received the B.S. degree in electronic science and technology from Nanchang Hangkong University, Nanchang, China, in 2006, the M.S. degree in physical electronics from Fuzhou University, Fuzhou, China, in 2009. He is currently pursuing the Ph.D. degree from the School of Automation, Huazhong University of Science and Technology, Wuhan, China.

His research interests include image fusion, sparse representation, pattern recognition, computer vision, and technology of electronics.

**Hongyan Zhang** (M'13) received the B.S. degree in geographic information system and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2005 and 2010, respectively.

He is currently an Associate Professor with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University. He has authored or co-authored more than 40 research papers. His research interests include image reconstruction, hyperspectral image processing, sparse representation, and low-rank methods for sensing image imagery.

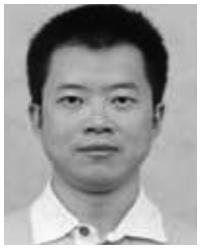
Dr. Zhang is a Reviewer of more than 10 international academic journals.

**Changxin Gao** received the B.S. degree in optical information science and technology and the Ph.D. degree in pattern recognition and intelligent systems from Huazhong University of Science and Technology, Wuhan, China, in 2005 and 2010, respectively.

He is currently an Assistant Professor with the School of Automation, Huazhong University of Science and Technology. From 2012 to 2013, he was a Postdoctoral Researcher with Université Catholique de Louvain, Louvain-la-Neuve, Belgium. His research interests include pattern recognition and computer vision.

**Cheng Jiang** received the B.S. degree in remote sensing science and technology and the Ph.D. degree in photogrammertry and remote sensing from Wuhan University, Wuhan, China, in 2010 and 2015, respectively.

He is currently a Post-Doctor with Institut de Planètologie et d'Astrophysique de Grenoble (IPAG), Saint-Martin-d'Hères, France. His research interests include image fusion, denoising, super-resolution in remote sensing images, and high accuracy DEM for Mars data.



**Nong Sang** received the B.E. degree in computer science and engineering, the M.S. degree in pattern recognition and intelligent control, and the Ph.D. degree in pattern recognition and intelligent systems from Huazhong University of Science and Technology, Wuhan, China, in 1990, 1993, and 2000.

He is currently a Professor with the School of Automation, Huazhong University of Science and Technology. His research interests include computational modeling of biological vision perception and applications in computer vision, image analysis, and object recognition based on statistical learning, medical image processing and analysis, interpretation of remote sensing images, and intelligent video surveillance.

**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 of 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, 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 authored more than 410 research papers. 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, an Executive Member (Board of Governor) of the China National Committee of the International Geosphere-Biosphere Programme, and an Executive Member of the China Society of Image and Graphics. 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 of 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*, the *Journal of Remote Sensing*, and the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He was the recipient of the 2010 Best Paper Boeing Award and the 2013 Best Paper ERDAS Award of American Society of Photogrammetry and Remote Sensing (ASPRS), respectively.