Pansharpening Based on Variational Fractional-Order Geometry Model and Optimized Injection Gains

Yong Yang, *Senior Member, IEEE*, Hangyuan Lu, Shuying Huang, *Member, IEEE*, Weiguo Wan, and Luyi Li

***Abstract*—Pansharpening techniques fuse the complementary information from panchromatic (PAN) and multispectral (MS) images to obtain a high-resolution MS image. However, the majority of existing pansharpening techniques suffer from spectral distortion owing to the low correlation between the MS and PAN images, and difficulties in obtaining appropriate injection gains. To address these issues, this article presents a novel pansharpening method based on the variational fractional-order geometry (VFOG) model and optimized injection gains. Specifically, to improve the correlation between the PAN and MS images, the VFOG model is constructed to generate a refined PAN image with a similar spatial structure to the MS image, while maintaining the gradient information of the original PAN image. Furthermore, to obtain accurate injection gains, and considering that the vegetated and nonvegetated regions should be dissimilar, an optimized adaptive injection gain based on the normalized differential vegetation index is designed. The final pansharpened image is obtained by an injection model using the refined PAN image and optimized injection gains. Extensive experiments on various satellite datasets demonstrate that the proposed method offers superior spectral and spatial fidelity compared to existing state-of-the-art algorithms.**

***Index Terms*—Detail injection scheme, injection gain, pansharpening, variational fractional-order geometry model.**

I. Introduction

REMOTE sensing images, which are obtained by detection instruments without contacting the target scene, can provide relevant information regarding the structure and contents of the earth surface. Especially, a high-resolution multispectral (HRMS) image can not only be applied to the scientific research such as image classification [1], image segmentation [2], spectral unmixing [3], and super-resolution mapping [4], but also to the public service such as land resource management and environmental monitoring [5]. However, owing to the limitations of satellite technology, current satellites cannot provide HRMS images directly [6]. Instead, earth observation satellites such as GeoEye-1, QuickBird, and Pleiades can simultaneously capture multispectral (MS) images with high spectral resolution but low spatial resolution and panchromatic (PAN) images with high spatial resolution but low spectral resolution [7], [8]. To obtain the HRMS image, a feasible method is to fuse the information of the PAN and MS images using suitable algorithms. This process is known as pansharpening, which aims to improve the spatial resolution of the MS images by means of the PAN images. Furthermore, pansharpening is considered as a powerful postprocessing technique and has become a significant issue in the remote sensing field [9].

Various pansharpening methods have emerged over the past several decades. Vivone *et al.* [11] specifically analyzed and compared the performance and characteristics of pansharpening methods based on the means of generating the geometric details. These methods can be roughly divided into the following types.

1) *Component substitution (CS) based methods:* CS-based methods generate geometric details by estimating the difference between the synthetic intensity images and PAN images. Subsequently, the geometric details are injected into the upsampled MS images to obtain pansharpened images. Intensity-hue-saturation (IHS) transform [12], Gram-Schmidt transform [13], and partial replacement adaptive CS (PRACS) [14] are mainstream CS-based methods. These methods offer the advantages of simplicity, computational efficiency, and high spatial quality. Moreover, effects such as artifacts, aliasing, and texture blurring are minimized in the fusion process. However, such approaches usually suffer from varying degrees of spectral distortion in the pansharpened images [15].
2) *Multiresolution analysis (MRA) based methods:* MRA-based methods generate geometric details by estimating the difference between the original PAN images and low-pass-filtered PAN images. Subsequently, these details are injected into the upsampled MS images to generate pansharpened images. MRA-based methods include Laplace transform [16], the generalized Laplacian pyramid (GLP) [17], à trous wavelet transform [18], discrete Wavelet transform [19], and nonsubsampled Shearlet transform [20]. The main advantage of these methods is that they

Manuscript received July 11, 2021; revised October 11, 2021, December 21, 2021, and January 16, 2022; accepted February 18, 2022. Date of publication February 28, 2022; date of current version March 9, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 62072218 and Grant 61862030, in part by the Natural Science Foundation of Jiangxi Province under Grant 20192ACB20002 and Grant 20192ACBL21008, in part by the Natural Science Foundation of Zhejiang Province under Grant LY22F020017, and in part by the Talent Project of Jiangxi Thousand Talents Program under Grant jxsq2019201056. *(Yong Yang and Hangyuan Lu contributed equally to this work.) (Corresponding author: Shuying Huang.)*

Yong Yang and Shuying Huang are with the School of Computer Science and Technology, Tiangong University, Tianjin 300387, China (e-mail: greatyangy@126.com; shuyinghuang2010@126.com).

Hangyuan Lu is with the College of Information Engineering, Jinhua Polytechnic, Jinhua 321007, China (e-mail: lhyhziee@163.com).

Weiguo Wan is with the School of Software and Internet of Things Engineering, Jiangxi University of Finance and Economics, Nanchang 330032, China (e-mail: wanwgplus@163.com).

Luyi Li is with the School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330032, China (e-mail: ada_016@foxmail.com).

Digital Object Identifier 10.1109/JSTARS.2022.3154642

can preserve the spectral information in the pansharpened image more effectively. However, MRA-based methods usually produce several unsatisfactory artifacts [21].

3) *Variation-based methods:* Variation-based methods are generally defined by several optimal energy functions along with the specified spatial and spectral constraints. The pansharpened images are generated by minimizing the energy functions [22], which comprise techniques such as P+XS [23], variational wavelet pansharpening (VWP) [24], and sparsity-based methods [25]. These methods can obtain high-quality results, but generally require much higher computational costs compared to CS- or MRA-based methods.
4) *Deep learning (DL) based methods:* Numerous DL-based methods have recently been developed for pansharpening owing to their strong feature-learning characteristics. For example, Yang *et al.* [26] proposed a progressive cascade deep residual network with two residual subnetworks for pansharpening. Ozcelik *et al.* [27] considered the pansharpening as a colorization problem and proposed a PanColorization generative adversarial network (PanColorGAN) framework. To improve the spatial quality, Deng *et al.* [28] combined a deep convolutional neural network with an injection model. To take advantages of the variational model and deep learning, Yang *et al.* [29] proposed a variational network for hyperspectral-multispectral image fusion, where the model can be implicitly learned from the training data. Although these methods can obtain superior fusion quality, they usually require numerous learning samples and computing resources for training.

In recent years, CS- and MRA-based methods have generally been defined and realized by injection models. In particular, the CS-based injection model is used extensively owing to its high spatial quality and efficiency [30]. However, it is difficult to obtain satisfactory spectral quality with this model [28]. To address the abovementioned problem, Rahmani *et al.* [31] proposed an adaptive IHS (AIHS) method, which can improve the spectral fidelity by adaptively optimizing the edge information. Leung *et al.* [32] modified the AIHS method by combining the edges of the MS and PAN images and proposed an improved AIHS method. However, two deficiencies remain that may affect the fusion quality. The first is that the low correlation between the PAN and MS images may cause spectral distortion [30]. The second is that vegetated and nonvegetated regions are rarely differentiated when determining the injection gains, which will cause spatial and spectral distortion [33].

Linear histogram matching is generally performed in advance to minimize the spectral distortion caused by the low correlation between the PAN and MS images, so that the PAN image exhibits similar characteristics to the MS image. However, only a small part of the spectral distortion can be eliminated in this manner, which means that the difference between the histogram-matched PAN image and MS image remains large. To address this problem, in this article, the variational fractional-order geometry (VFOG) model is proposed to estimate a refined PAN image that is highly correlated to the MS image. The spatial structure of the refined PAN image should be closer to that of the MS image without losing the gradient features of the original PAN image. Most variational methods, such as P+XS, VWP, and total variation-based methods [22]–[24], characterize the geometry of the PAN images using the gradient features. However, they may easily produce blocky artifacts. To solve this problem, Liu *et al.* [34], [35] proposed the Hessian-feature-guided models to keep Hessian features consistent between the PAN image and the fused image. These integer-order computations are essential in the contrast enhancement, but there are still shortages such as staircase effects and ringing artifacts [36]. Instead, the fractional-order geometry (FOG) can maintain low frequency contours and enhance high frequency information, and has been proven as superior in extracting edges and textural details [37], [38]. Therefore, the FOG feature is utilized in the proposed variational model to represent the gradient information of the PAN images.

Moreover, the injection gains, which determine how many details to be injected, are an important factor affecting the performance of pansharpening methods. However, few existing injection gain approaches consider the fact that the injection gains of the vegetated regions should be unlike those of the nonvegetated regions, and spectral distortion mainly occurs in the vegetated regions [39]. As the normalized differential vegetation index (NDVI) is a superior index that can quantify the biophysical characteristics of vegetation [40], we employ the NDVI to design new injection gains. Furthermore, because certain regions of the injection gains are close to zero or even negative [41], fewer details are injected into the upsampled MS images. Therefore, the sigmoid function is defined to minimize the large dynamic range of the original injection gains into the range from 0 to 1. The proposed method was validated on multiple satellite datasets, including QuickBird, Pleiades, WorldView-2, and GeoEye-1. The experimental results of subjective and objective comparisons of the fused images obtained using the proposed method and several existing mainstream methods confirmed the effectiveness of the proposed method.

The main contributions of this article are summarized as follows.

1) To improve the correlation between the PAN and MS images without losing gradient information of the original PAN image, a VFOG model is designed for estimating a refined PAN image that can minimize the spectral distortion in the pansharpened image.
2) To further improve the spectral quality of the vegetation areas, the injection gains based on the NDVI is proposed. Besides, a sigmoid function is designed to reduce the dynamic range of the injection gains.
3) The VFOG model and optimized injection gains are mathematically combined in the injection scheme, the results of which verify the effectiveness and robustness of the proposed method in maintaining the spectral and spatial fidelity in various satellite datasets.

The rest of this article is organized as follows. Related work on the CS-based injection scheme is briefly reviewed in Section II. The proposed method is introduced in Section III. Section IV presents the experimental data settings, quality assessment indexes, experimental results, and analysis. Further experimental

discussions are provided in Section V. Finally, Section VI concludes this article.

II. Related Work

A. Notations

The auxiliary symbols and definitions are first introduced to simplify the observation. Furthermore, the notations in this article are presented in the following.

Let $P, P_{\text{hist}} \in \Re^{s\times s}$ denote the original PAN image and histogram-matched PAN image, respectively. Furthermore, $s \times s$ denotes the size of the P or P_{hist} image. $m = (m_R, m_G, m_B, m_{\text{NIR}}) \in \Re^{s/c\times s/c\times 4}$ denotes the MS image, where m_R, m_G, m_B, and m_{NIR} are the four bands of the MS image, namely the red (R), green (G), blue (B), and near-infrared (NIR) bands, respectively. c is the ratio of spatial resolution between the MS and PAN images. Moreover, $F = (F_R, F_G, F_B, F_{\text{NIR}}) \in \Re^{s\times s\times 4}$ denotes the pansharpened image. $M = (M_R, M_G, M_B, M_{\text{NIR}}) \in \Re^{s\times s\times 4}$ represents the upsampled MS image, which is produced by bicubic interpolation. $\|\cdot\|_F$ and $(\cdot)$ denote the Frobenius norm and transpose operator, respectively.

B. CS-Based Injection Scheme

In general, the pansharpened image that is estimated by the CS-based injection scheme can be determined by

$$F_i = M_i + g_i\left(P - I\right), i = 1, \ldots, N \tag{1}$$

where g_i denotes the injection gains of the ith band and I is the intensity component of the upsampled MS image. The spatial quality of pansharpened images obtained by the CS-based method is usually superior to that of images obtained by the MRA-based method [41]. Therefore, we selected the CS-based injection scheme in this article.

As noted in [21] and [42], CS-based algorithms take advantage of combination coefficients to measure the I image. This process can be expressed as follows:

$$I = \sum_{i=1}^{N} \gamma_i M_i \tag{2}$$

where γ_i is the ith band combination coefficient, which is estimated by the linear relationship between P and each band of the M image, and is defined as

$$\min_{\gamma_1,\cdots,\gamma_N} \left\| P_{\text{hist}} - \sum_{i=1}^{N} \gamma_i M_i \right\|^2 s.t. \gamma_1 \geq 0, \ldots, \gamma_N \geq 0 \tag{3}$$

where N is the total number of bands of the M image. A classical method for minimizing the spectral distortion caused by the dissimilarity between the M and P images is preliminarily performing histogram matching to force the P image to have a similar mean and variance to the I image [21], [42]. Thus, P_{hist} can be defined as

$$P_{\text{hist}} = (P - \mu_P) \cdot \frac{\sigma_I}{\sigma_P} + \mu_I \tag{4}$$

where μ and σ are the mean and standard deviation, respectively.

Furthermore, the injection gains g_i play an important role in maintaining the spatial and spectral fidelity. Yang *et al.* [42] defined joint discriminant injection gains, which comprehensively consider the relation between the edge information of the P and M images, and are defined as follows:

$$g_i = \frac{M_i}{\frac{1}{N}\sum_{i=1}^{N} M_i} \left(\eta_i E_{P_{\text{hist}}} + (1 - \eta_i) E_{M_i}\right) \tag{5}$$

where η_i is the i-band tradeoff parameter, and $E_{P_{\text{hist}}}$ and E_{M_i} are the edge detection matrices that are estimated by the edges of the P_{hist} image and ith band of the M image, respectively. The edge detection matrix is defined as follows:

$$E_Q = \exp\left(-\frac{\lambda}{|\nabla Q|^4 + \varepsilon}\right) \tag{6}$$

where Q represents the input image, λ and ε are the tuning parameters, and the symbol ∇ denotes the gradient operator. By replacing Q with P_{hist} or M_i, $E_{P_{\text{hist}}}$ or E_{M_i} can be obtained, respectively.

Moreover, η_i in (5) is defined as

$$\min_{\chi_1,\ldots,\chi_N} \left\| E_{P_{\text{hist}}} - \sum_{i=1}^{N} \chi_i E_{M_i} \right\|^2 \text{s.t.} \chi_1 \geq 0, \ldots, \chi_N \geq 0 \tag{7}$$

$$\eta_i = \max\left(\chi_i, \frac{COV\left(E_{P_{\text{hist}}}, E_{M_i}\right)}{\sigma\left(E_{P_{\text{hist}}}\right) \cdot \sigma\left(E_{M_i}\right)}\right) \tag{8}$$

where χ_i represents the original weight parameter, and $COV(\cdot)$ and $\sigma(\cdot)$ are the covariance and standard deviation, respectively.

III. Proposed Method

A. Overall Framework

As CS-based injection schemes generally suffer from spectral distortion, numerous researchers have attempted to improve these methods using the P_{hist} image. As opposed to such methods, in this article, to minimize the spectral distortion in existing CS-based injection methods, the VFOG model is designed to estimate a refined PAN image P_{VFOG} that is highly relevant to the upsampled MS image. Moreover, to maintain the spectral and spatial fidelity in the pansharpened image more effectively, novel injection gains that consider the NDVI are defined. Fig. 1 presents the framework of the proposed method. The specific procedures of the proposed framework are listed as follows.

1) The I image and P_{hist} image can be obtained according to (2)–(4).
2) The P_{VFOG} image is estimated by employing the VFOG model. Thus, the high-frequency details are determined by the difference between the P_{VFOG} and I images.
3) The injection gains g_i are calculated by comprehensively considering the NDVI, and the edge information of P_{VFOG} and M. Here, G_i is the initial injection gain defined in formula (16), and b_i is a coefficient defined in formula (18).
4) Once g_i is obtained, the final pansharpened image can be achieved by linearly injecting the spatial details into each band of the M image. Here, HD represents the high-frequency details.

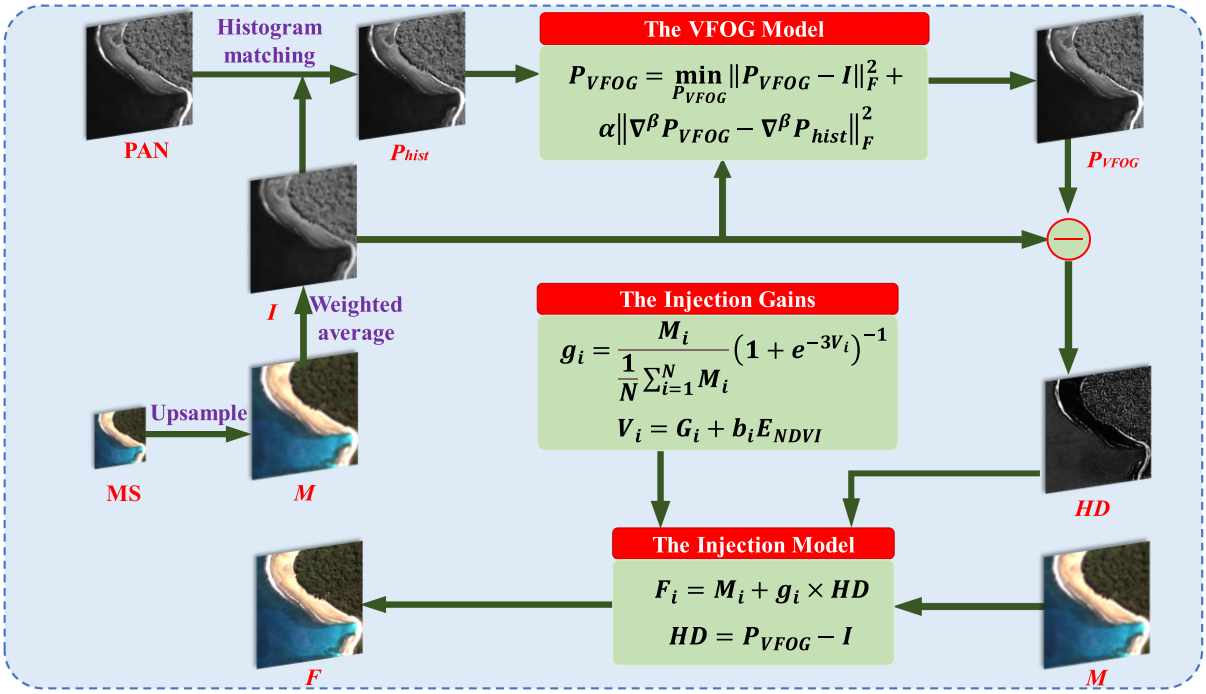

Fig. 1. Overall framework of the proposed method.

B. VFOG Model

To overcome the deficiencies of the CS-based injection scheme, we attempt to estimate a refined PAN image P_{VFOG} to minimize the spectral distortion further. Inspired by the superiority of variation-based methods; i.e., the contribution to minimizing the spectral distortion [38], we design the novel VFOG model in a unified variational framework to estimate the P_{VLFOG} image. In general, variation-based methods, which use the gradient features of the P image, suffer from staircase effects and the loss of several image textures. Therefore, the fractional-order geometric feature, which is an effective means of retaining the image textures and reducing the staircase effects, is used to represent the geometric information in the VFOG model.

To improve the correlations between the P_{VFOG} and M images, and to guarantee that P_{VFOG} preserves the geometric information from P_{hist}, two constraint terms are implemented in the VFOG model. The first term enforces P_{VFOG} to contain a similar spatial structure to the intensity component of the MS image. The second term is the spatial geometry information fidelity term. Specifically, the second term constrains the fractional-order gradient feature of the P_{VFOG} image to be consistent with that of P_{hist}, thereby ensuring that P_{VFOG} maintains spatial fidelity to P_{hist}. Mathematically, the energy function is designed as follows:

$$P_{\mathrm{VFOG}} = \min_{P_{\mathrm{VFOG}}} \|P_{\mathrm{VFOG}} - I\|_F^2 + \alpha \left\|\nabla^\beta P_{\mathrm{VFOG}} - \nabla^\beta P_{\mathrm{hist}}\right\|_F^2 \tag{9}$$

where α is the tradeoff parameter, $\nabla^\beta : \Re^{K\times K} \to \Re^{2\times K\times K}$ $(1 < \beta < 2)$ denotes the discrete fractional-order gradient operator, β represents the fractional order, and $[\nabla^\beta P_{\mathrm{hist}}]_k = ([D_h^\beta P_{\mathrm{hist}}]_k, [D_v^\beta P_{\mathrm{hist}}]_k)^T \in \Re^2$. $[D_h^\beta P_{\mathrm{hist}}]_k$ and $[D_v^\beta P_{\mathrm{hist}}]_k$ represent the kth columns of the horizontal and vertical discrete fractional-order gradient matrices, respectively. $D_h^\beta : \Re^{K\times K} \to \Re^{1\times K\times K}$ and $D_v^\beta : \Re^{K\times K} \to \Re^{1\times K\times K}$ are the horizontal and vertical discrete fractional-order gradient operators [38], which are formulated as

$$\left[D_h^\beta P_{\mathrm{hist}}\right]_k = \sum_{t=0}^{K-1} \omega_t^{(\beta)} P_{\mathrm{hist}}(r+t, s) \tag{10}$$

$$\left[D_v^\beta P_{\mathrm{hist}}\right]_k = \sum_{t=0}^{K-1} \omega_t^{(\beta)} P_{\mathrm{hist}}(r, s+t) \tag{11}$$

where (r, s) is the coordinate of a pixel and t represents the distance from the pixel. $\omega_t^{(\beta)} = (-1)^t \Gamma(\beta+1)/[\Gamma(t+1)\Gamma(\beta-t+1)]$ is the generalized binomial coefficient [37] and $\Gamma(\cdot)$ denotes the Gamma function, with $K \geq 3$.

It is obvious that (9) is a typical convex optimization problem that can be solved by the sliding block technique and least square method. Denote p_{VFOG}, m_{b}, and p_{hist} as the vectorized blocks of the P_{VFOG}, I, and P_{hist} images, respectively. Then, the vectorized block version of (9) can be expressed as follows:

$$p_{\mathrm{VFOG}} = \min_{P_{\mathrm{VFOG}}} \|p_{\mathrm{VFOG}} - m_{\mathrm{b}}\|_F^2 + \alpha \left\|\nabla^\beta p_{\mathrm{VFOG}} - \nabla^\beta p_{\mathrm{hist}}\right\|_F^2. \tag{12}$$

To solve (12), we set the partial deviation of p_{VFOG} to zero, then we can obtain the equation as

$$\begin{aligned} &p_{\mathrm{VFOG}} - m_{\mathrm{b}} + \alpha(\nabla^\beta)^T(\nabla^\beta p_{\mathrm{VFOG}} - \nabla^\beta p_{\mathrm{hist}}) = 0 \\ &\Rightarrow p_{\mathrm{VFOG}} - m_{\mathrm{b}} + \alpha(\nabla^\beta)^T\nabla^\beta p_{\mathrm{VFOG}} - \alpha(\nabla^\beta)^T\nabla^\beta p_{\mathrm{hist}} = 0 \\ &\Rightarrow (H + \alpha(\nabla^\beta)^T\nabla^\beta)p_{\mathrm{VFOG}} = m_{\mathrm{b}} + \alpha(\nabla^\beta)^T\nabla^\beta p_{\mathrm{hist}} \end{aligned}$$

$$\Rightarrow p_{\mathrm{VFOG}} = (H + \alpha(\nabla^{\beta})^T \nabla^{\beta})^{-1}(m_{\mathrm{b}} + \alpha(\nabla^{\beta})^T \nabla^{\beta} p_{\mathrm{hist}})$$
$$= (H + \alpha A)^{-1}(m_{\mathrm{b}} + \alpha A p_{\mathrm{hist}}) \tag{13}$$

where

$$A = (\nabla^{\beta})^T \nabla^{\beta} = ((D_h^{\beta}, D_v^{\beta})^T)^T (D_h^{\beta}, D_v^{\beta})^T$$
$$= D_h^{\beta}(D_h^{\beta})^T + D_v^{\beta}(D_v^{\beta})^T \tag{14}$$

and H represents an identity matrix.

According to the designed VFOG model in (9), the difference between the generated P_{VFOG} image and M image is smaller than that between the P_{hist} image and M image. Furthermore, the fractional-order gradient is applied to preserve the spatial features of the P_{hist} image more effectively.

C. Injection Gains Based on NDVI

The injection gains g_i determine how many high-frequency details are added into each band of the M image, which plays a decisive role in the effectiveness of fused algorithms. The spectral dissimilarities between the P and M images may vary significantly according to different land cover types, such as vegetated areas, water, and buildings. Moreover, the main spectral distortion usually occurs within vegetated areas. Xu *et al.* [43] verified that the injection gains based on vegetated and nonvegetated areas are dissimilar. The NDVI, which is an important indicator that quantifies the state of vegetation growth, has been established as a practicable and effective means of estimating the injection gains [39]. The NDVI is defined as follows:

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - R}{\mathrm{NIR} + R} \tag{15}$$

where R and NIR are the pixel values of R and NIR bands of the M image, respectively.

Choi *et al.* [39] confirmed that a high or moderate correlation is usually reflected between the injection gains and NDVI. Therefore, we consider not only the edge detection matrices of M_i and P_{VFOG}, but also the edge detection matrix of the NDVI. By replacing Q with P_{VFOG} and the NDVI, the edge detection matrices $E_{P_{\mathrm{VFOG}}}$ and E_{NDVI}, respectively, can be obtained using (6).

At present, the majority of injection gains in existing injection models consider the edge detection matrices that are guided by the edges of a PAN image and the different bands of an MS image [21]. Inspired by this concept, in this article, we define the initial injection gains, denoted as G_i, as follows:

$$G_i = a_i E_{P_{\mathrm{VFOG}}} + (1 - a_i) E_{M_i} \tag{16}$$

where a_i is the weight parameter, which is obtained by measuring the relationship between the P_{VFOG} images and different bands of the MS images [32]. As mentioned in Section II, a_i can be obtained by substituting $E_{P_{\mathrm{hist}}}$ with $E_{P_{\mathrm{VFOG}}}$ in (7) and (8).

It is obvious that the injection gains based on different bands of the M image, denoted as V_i, should not only consider the tradeoff between E_{M_i} and $E_{P_{\mathrm{VLFOG}}}$, but should also take E_{NDVI} into account. This process can be expressed as follows:

$$V_i = G_i + b_i E_{\mathrm{NDVI}} \tag{17}$$

where b_i is the discrimination coefficient, the purpose of which is to modulate the relation between V_i and E_{NDVI}. Choi *et al.* [39] verified that the NDVI was negatively correlated with the injection gains of the visible bands and positively correlated with the injection gains of the NIR band. That is, when the correlation coefficients are greater than 0, the information that is being processed is in the NIR band. Moreover, given that vegetation has very strong reflectivity in the NIR band, it has more details in the NIR band. Thus, the degree of injection should be improved for the vegetation information in the NIR band. Therefore, b_i is defined as

$$b_i = \begin{cases} \frac{\sigma^2(G_i)}{2\sigma^2(E_{\mathrm{NDVI}})}, & \text{if } \mathrm{CR}(G_i, \mathrm{NDVI}) > 0 \\ 0, & \text{if } \mathrm{CR}(G_i, \mathrm{NDVI}) < 0 \end{cases} \tag{18}$$

where σ^2 denotes the variance and $\mathrm{CR}(\cdot)$ is the function for calculating the correlation coefficients, which is defined as

$$\mathrm{CR}(S_1, S_2) = \frac{COV(S_1, S_2)}{\sigma(S_1) \cdot \sigma(S_2)} \tag{19}$$

where S_1 and S_2 are the two input source images.

Given that the injection gains are close to 0 or even negative in certain areas, and the sigmoid function can reduce the dynamic range of the injection gains [44], the sigmoid function is used to optimize V_i preliminarily, which can be expressed as follows:

$$OV_i = \left(1 + e^{-3V_i}\right)^{-1} \tag{20}$$

where OV_i is the initial optimized V_i. Furthermore, it is particularly important that the proportion between different bands of the MS image remain unchanged to maintain the spectral information in the M image [32]. Therefore, the final injection gains g_i can be obtained by

$$g_i = \frac{M_i}{\frac{1}{N}\sum_{i=1}^{N} M_i} \times OV_i. \tag{21}$$

In general, the designed g_i comprehensively consider the relations between E_{NDVI}, E_{M_i}, and $E_{P_{\mathrm{VLFOG}}}$. Moreover, the sigmoid function and normalization method are applied to optimize the injection gains further. With the aid of the defined g_i, the high-frequency details HD from P_{VFOG} can be injected into the MS image more accurately. Therefore, the fused image F can be obtained using the injection model as follows:

$$\begin{cases} F_i = M_i + g_i \times HD \\ HD = P_{\mathrm{VLFOG}} - I \end{cases}. \tag{22}$$

IV. Experimental Results and Analysis

A. Experimental Datasets and Settings

In this article, we collected the datasets obtained from four different satellites, namely Pleiades, QuickBird, GeoEye-1, and WorldView-2. The characteristics of the four satellites are listed in Table I. In our work, each dataset contained 90 sets of images. These satellite images, which included various land cover types, such as lakes, buildings, vegetation, and mountainous areas, were used to verify the performance of our algorithm comprehensively.

To assess the effectiveness of our algorithm more thoroughly, two categories of experiments, namely simulated and real data experiments, were conducted. In simulated data experiments, the original MS images are usually regarded as reference images

TABLE I
SPECIFICATIONS OF THE PLEIADES, QUICKBIRD, GEOEYE-1, AND WORLDVIEW-2 SATELLITES

Sensor		Pleiades	QuickBird	GeoEye-1	WorldView-2
MS resolution/size		2 m 256 × 256	1.84 m 256 × 256	1.65 m 256 × 256	1.84 m 256 × 256
PAN resolution/size		0.5 m 256 × 256	0.46 m 256 × 256	0.41 m 256 × 256	0.46 m 256 × 256
Wavelength	PAN	450-900 nm	470-830 nm	450-800 nm	450-800 nm
	blue	450-520 nm	430-550 nm	450-510 nm	450-510 nm
	green	520-600 nm	500-620 nm	510-580 nm	510-580 nm
	red	630-690 nm	590-710 nm	655-690 nm	630-690 nm
	NIR1	760-900 nm	740-940 nm	780-920 nm	770-895 nm

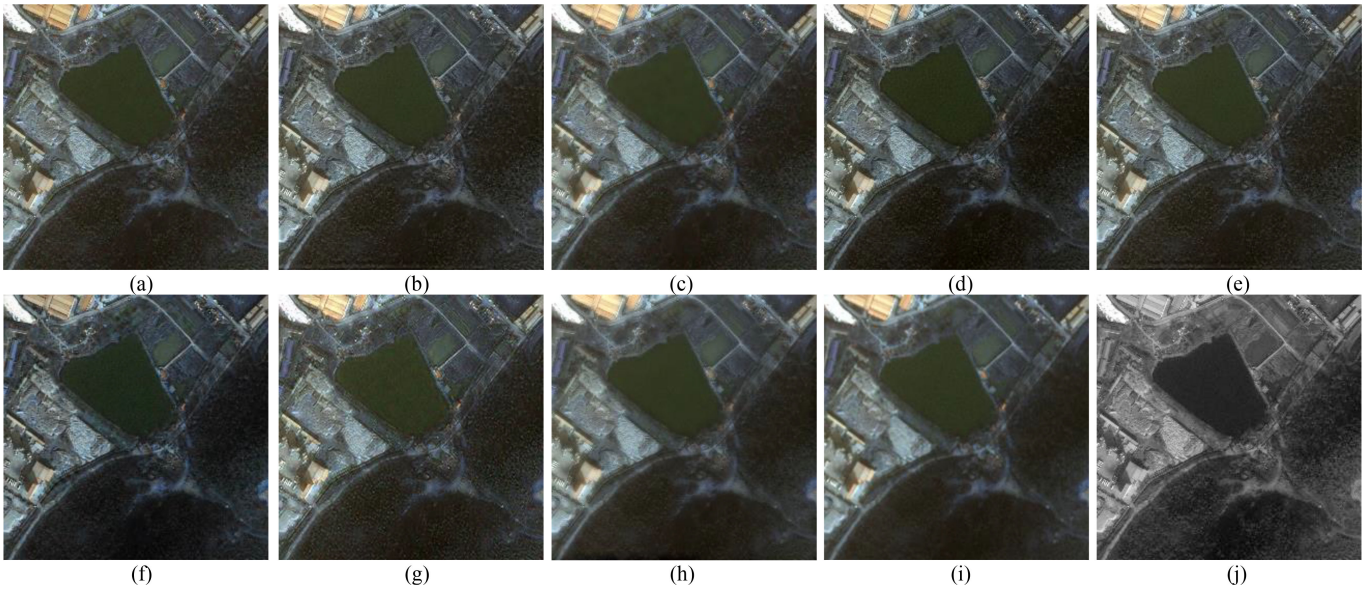

Fig. 2. Fused results by different methods on the Pleiades dataset. (a) GS. (b) MTF-GLP. (c) PRACS. (d) BFLP. (e) FSRIC. (f) PanColorGAN. (g) Fusion-Net. (h) Proposed. (i) Reference. (j) PAN.

according to Wald's protocol [45]. Specifically, the original MS and PAN images were degraded and down-sampled with a decimation factor of 4. Thus, the down-sampled MS and PAN images were used as the input images for the simulated data validation. Furthermore, seven state-of-the-art pansharpening algorithms were used for comparison to verify the performance of the proposed algorithm: GS [13], Generalized Laplacian Pyramid with MTF-matched filter (GLP-MTF) [46], PRACS [14], the bilateral filter luminance proportional (BFLP) method [47], the full-scale regression-based injection coefficient (FS-RIC) method [48], PanColorGAN [27], and Fusion-Net [28]. All of the methods used for comparison were public source codes provided by the corresponding authors, and tested on the computer with Intel(R) Core (TM) i5 @ 2.60 GHz CPU and 16 GB RAM.

B. Quality Evaluation Indexes

To further quantitatively verify the performance of the proposed method, we adopted two types of metrics based on the different experiments; i.e., the simulated and real data experiments. In particular, the metrics used for the simulated data experiments include correlation coefficient (CC) [38], universal image quality index (UIQI) [11], root-mean-square error (RMSE) [26], peak signal-to-noise ratio (PSNR) [49], relative average spectral error (RASE) [50], erreur relative global adimensionnelle de synthese (ERGAS) [21], and spectral angle mapper (SAM) [51]. The metrics used for the real data experiments include spectral distortion index (D_λ), spatial distortion index (D_s), and hybrid quality with no reference (HQNR) [52], [53].

C. Experimental Results

The simulated data experiments were conducted on the images acquired from the Pleiades and QuickBird sensors, whereas the images obtained from the WorldView-2 and GeoEye-1 sensors were employed in the real data experiments. Both subjective and objective evaluations were the essential aspects for verifying the effectiveness of the pansharpening algorithm. Four groups of images from the Pleiades, QuickBird, WorldView-2, and GeoEye-1 datasets are presented as examples to observe the performance of the proposed algorithm as well as the comparison methods subjectively. For convenience of comparison, only the R, G, and B bands of the pansharpened images are included.

Fig. 2 displays the pansharpened images that were generated by the different methods on the Pleiades dataset. Moreover, to

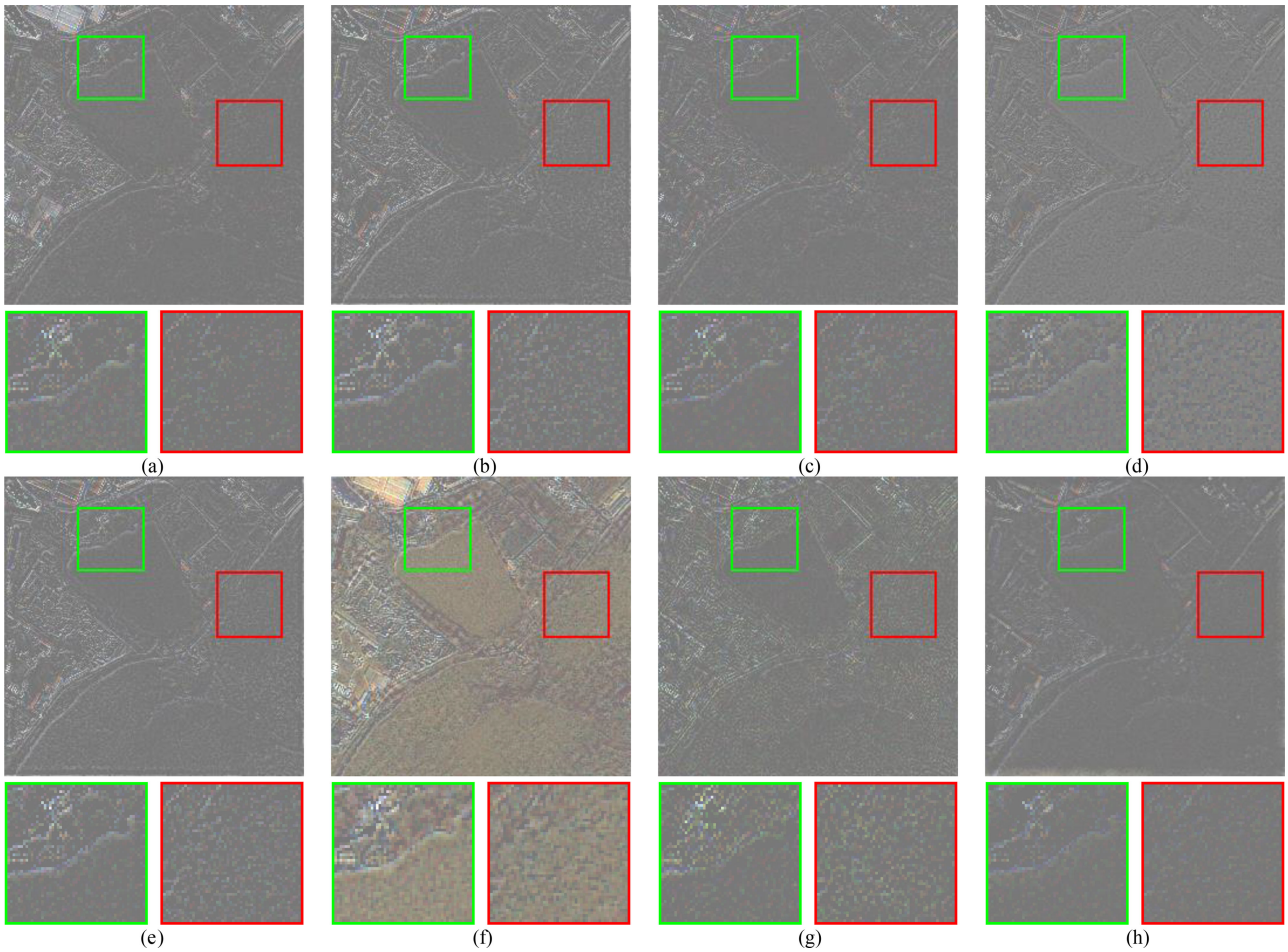

Fig. 3. Residual images estimated by the difference between the reference image and the fused results by different methods in Fig. 2. (a) GS. (b) MTF-GLP. (c) PRACS. (d) BFLP. (e FSRIC. (f) PanColorGAN. (g) Fusion-Net. (h) Proposed.

distinguish the differences of the pansharpened results (fused results) obtained by the different methods more effectively, the residual images estimated by the differences between the reference image and fused results are presented in Fig. 3. The red and green rectangular regions within the residual images are magnified and displayed at the bottom of the residual images. As can be observed from Fig. 3, the residual image of PanColorGAN produced serious spectral distortion and undesired artifacts. Although the GS, MTF-GLP, PRACS, BFLP, FSRIC, and Fusion-Net algorithms could reduce some of the spectral distortion, several undesired artifacts remained in the residual images, particularly in the water and forest areas marked by the green and red rectangles. In contrast, the proposed algorithm obtained less residual information, indicating that our method could eliminate the spectral distortion more effectively, retain more textures, and achieve a better fusion effect.

Table II displays the objective results of Fig. 2 and the average objective results of the Pleiades dataset including 90 groups of images. For clearer observation, the upward arrow in Table II indicates that a larger measured result of the indicator means a better effect, whereas the downward arrow indicates the opposite. Furthermore, the best results of each indicator are displayed in bold, and the second-best ones are underlined. As indicated in Table II, the proposed algorithm achieved the optimal results in all indexes for the results of Fig. 2. It also performed the best in all indexes for the average evaluation except SAM index. This further confirms the superiority of the proposed algorithm on the Pleiades dataset.

Fig. 4 presents the pansharpened images that were generated by the comparison and proposed algorithms on the QuickBird dataset. It is evident that the BFLP and PanColorGAN algorithms produced serious spectral distortion, particularly in the plain areas, and the result of PRACS was blurred. Moreover, Fig. 5 depicts the residual images for a more intuitive comparison. According to Fig. 5, the residual images that were generated by the GS, MTF-GLP, and FSRIC algorithms exhibited spatial and spectral distortion, especially in the vegetation areas marked by the green rectangles. In comparison, the PRACS and Fusion-Net algorithms could eliminate some spatial distortion and retain more spectral information. However, when comparing the residual image of the proposed method with those of other pansharpening methods, it is clear that our method could reduce the spectral distortion and improve the spatial fidelity more effectively, thereby demonstrating the superiority of the proposed method.

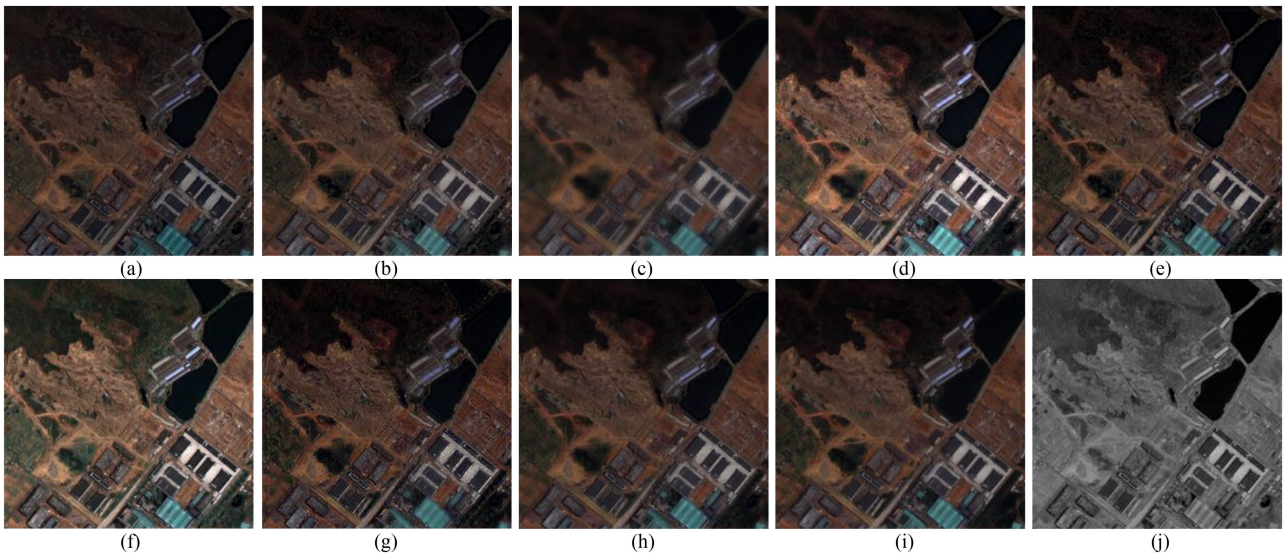

Fig. 4. Fused results by different methods on the QuickBird dataset. (a) GS. (b) MTF-GLP. (c) PRACS. (d) BFLP. (e) FSRIC. (f) PanColorGAN. (g) Fusion-Net. (h) Proposed. (i) Reference. (j) PAN.

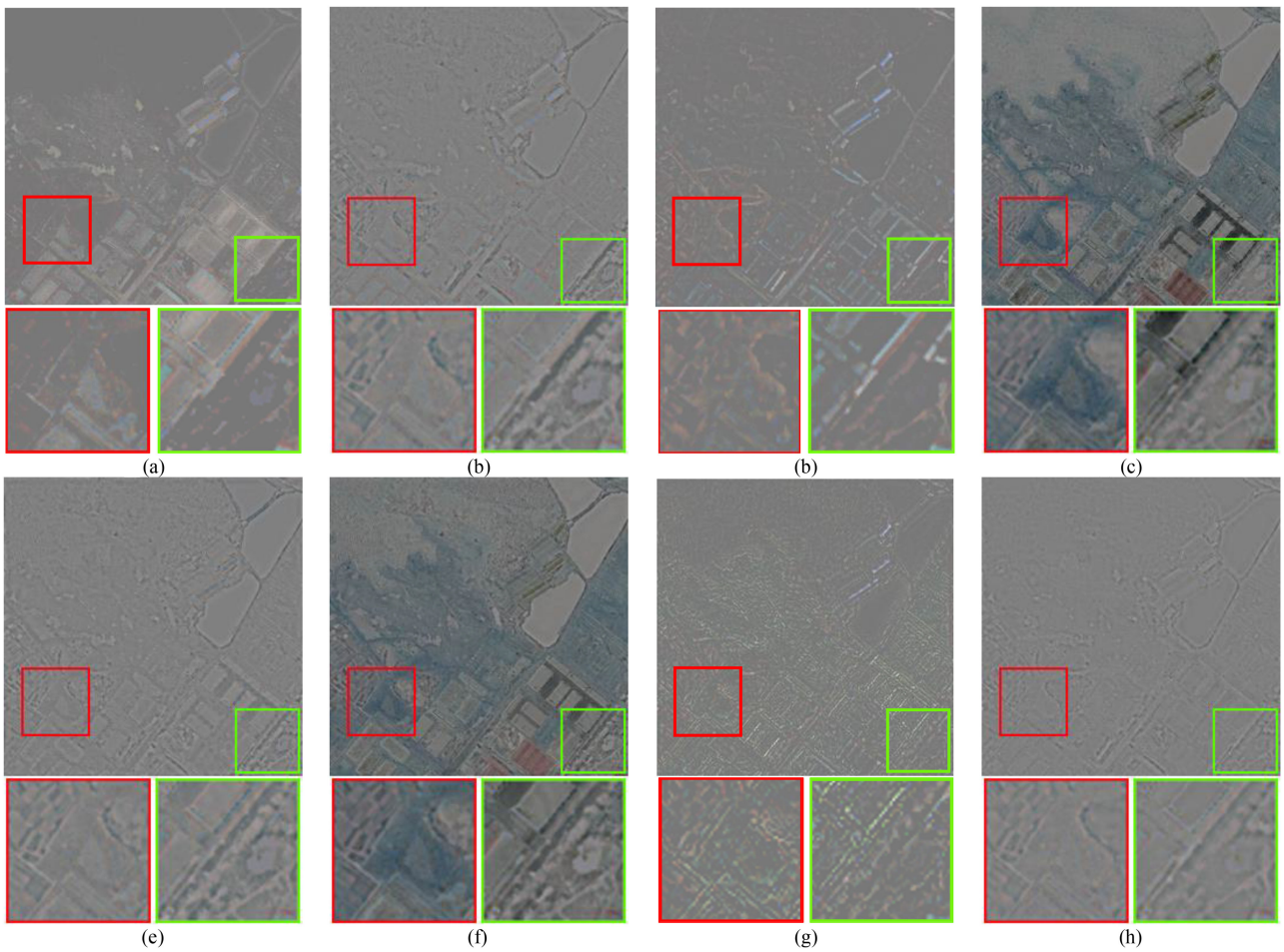

Fig. 5. Residual images estimated by the difference between the reference image and the fused results by different methods in Fig. 4. (a) GS. (b) MTF-GLP. (b) PRACS. (c) BFLP. (e) FSRIC. (f) PanColorGAN. (g) Fusion-Net. (h) Proposed.

TABLE II
QUANTITATIVE RESULTS ON THE PLEIADES DATASET

	Indexes	PSNR ↑	CC ↑	UIQI ↑	RASE ↓	RMSE ↓	SAM ↓	ERGAS ↓
The objective results of Fig. 2	GS [13]	26.9262	0.9604	0.9347	16.7280	11.6280	3.5312	4.1868
	MTF-GLP [46]	24.7565	0.9575	0.9390	17.9667	12.4695	3.7759	4.5088
	PRACS [14]	27.7588	0.9704	0.9435	13.8717	8.9505	3.6676	3.2230
	BFLP [47]	24.4023	0.9669	0.9387	20.4867	14.2290	4.4364	5.8700
	FSRIC [48]	24.0376	0.9362	0.9168	22.5176	15.6315	8.8907	5.7159
	PanColorGAN [27]	23.9787	0.9580	0.9240	23.3662	16.2331	7.2312	6.9755
	Fusion-Net [28]	27.6307	0.9633	0.9431	16.6178	11.5448	4.1170	4.1889
	Proposed	**28.0580**	**0.9799**	**0.9562**	**12.0199**	**8.3385**	**3.2348**	**3.0194**
The average objective results of the Pleiades dataset	GS [13]	24.8482	0.9230	0.9093	15.2788	12.3930	3.1777	3.8578
	MTF-GLP [46]	24.7182	0.9213	0.9121	16.3115	13.1325	3.3754	4.1182
	PRACS [14]	26.2405	0.9455	0.9272	13.5401	10.9395	**3.1760**	3.4103
	BFLP [47]	23.0365	0.9398	0.8635	31.1656	24.6330	4.3707	6.9830
	FSRIC [48]	24.8060	0.8968	0.9036	23.0713	15.1400	9.4747	6.2855
	PanColorGAN [27]	21.5511	0.9134	0.8735	26.6102	22.0507	7.0564	8.0575
	Fusion-Net [28]	27.0591	0.9286	0.9156	15.1501	12.1635	3.5231	3.8165
	Proposed	**30.4054**	**0.9620**	**0.9646**	**11.7831**	**7.7557**	4.0165	**3.0544**

TABLE III
QUANTITATIVE RESULTS ON THE QUICKBIRD DATASET

	Indexes	PSNR ↑	CC ↑	UIQI ↑	RASE ↓	RMSE ↓	SAM ↓	ERGAS ↓
The objective results of Fig. 4	GS [13]	25.1073	0.9297	0.8081	21.9085	11.0670	7.7587	5.9953
	MTF-GLP [46]	25.9482	0.9384	0.8545	16.7049	10.8405	6.4225	4.5238
	PRACS [14]	26.5438	**0.9683**	0.8532	15.0216	**7.5735**	5.0628	3.8970
	BFLP [47]	22.6415	0.9285	0.9104	24.5053	19.5503	3.6086	6.0015
	FSRIC [48]	22.8954	0.8843	0.8932	22.9337	18.6722	7.6902	5.8481
	PanColorGAN [27]	18.0317	0.9405	0.6706	70.8366	35.7502	8.2287	11.0931
	Fusion-Net [28]	**28.8108**	0.9346	0.8317	24.0465	12.1359	7.2768	5.8164
	Proposed	27.6871	0.9447	**0.9573**	**13.0794**	10.7527	**2.7766**	**3.2834**
The average objective results of the QuickBird dataset	GS [13]	23.8959	0.8649	0.7970	26.4317	16.9575	8.7022	6.8904
	MTF-GLP [46]	24.2922	0.8587	0.7246	25.8085	14.1015	8.8162	8.9039
	PRACS [14]	25.7437	**0.8920**	0.7947	25.0926	**9.5880**	5.7194	6.7451
	BFLP [47]	21.3780	0.8421	0.8366	41.0340	21.7591	6.6655	8.7390
	FSRIC [48]	21.3791	0.7607	0.8036	41.0285	21.7562	10.5275	10.5052
	PanColorGAN [27]	24.3111	0.8741	0.7939	27.8035	18.2155	9.5529	7.2466
	Fusion-Net [28]	25.4337	0.8805	0.7497	25.6721	14.5548	6.0043	6.3273
	Proposed	**26.0085**	0.8788	**0.9082**	**24.0780**	12.7678	**5.6853**	**6.1194**

Table III presents the objective results of Fig. 4 and the average objective results of the QuickBird dataset including 90 groups of images. As indicated in Table III, the proposed algorithm obtained the best fused results on most indexes, which further demonstrates the superiority of the proposed algorithm.

Fig. 6 displays the fused images that were produced by the comparison and proposed algorithms on the WorldView-2 dataset. We provide the zoomed details for a more intuitive comparison. It is evident from Fig. 6 that the BFLP algorithm generated great serious spectral distortion in the area marked by the red rectangle, although it achieved satisfactory spatial performance. The GS, FSRIC, and Fusion-Net algorithms produced excessive sharpening, especially in the zoomed areas. The MTF-GLP and PRACS algorithms showed some improvements in the spatial quality. The BFLP algorithm achieved progress in the spectral preservation but produced several undesired artifacts. In contrast, the proposed algorithm produced superior pansharpened images in terms of the spectral and spatial quality.

Table IV displays the objective assessments of Fig. 6 and the average objective results of the WorldView-2 dataset including 90 groups of images. Both the objective results of Fig. 6 and the average results of the WorldView-2 dataset demonstrate that the proposed algorithm achieved the best results on HQNR indexes, and the second-best on D_s indexes. In general, the proposed algorithm exhibited outstanding performance with respect to the spatial and spectral fidelity.

Fig. 7 presents the fused images that were obtained by the comparative and proposed algorithms on the GeoEye-1 dataset. The visual comparison demonstrates that the fused images obtained by the PanColorGAN methods exhibited serious spectral distortion. Although the results obtained by the GS, MTF-GLP, BFLP, and FSRIC methods exhibited certain improvements, they were unsatisfactory in the spatial aspect, particularly in the zoomed areas. The result of Fusion-Net suffered from over detail injection. The fused images produced by the PRACS and the proposed algorithms had a high visual quality. However, the

Fig. 6. Fused results by different methods on the WorldView-2 dataset. (a) GS. (b) MTF-GLP. (c) PRACS. (d) BFLP. (e) FSRIC. (f) PanColorGAN. (g) Fusion-Net. (h) Proposed. (i) MS. (j) PAN.

TABLE IV
QUANTITATIVE RESULTS ON THE WORLDVIEW-2 DATASET

	The objective results of Fig. 6			The average objective results of the WorldView-2 dataset		
Indexes	$D_\lambda \downarrow$	$D_S \downarrow$	HQNR↑	$D_\lambda \downarrow$	$D_S \downarrow$	HQNR↑
GS [13]	0.1227	0.1253	0.7674	0.1488	0.1314	0.7424
MTF-GLP [46]	**0.0355**	0.1512	0.8187	**0.0444**	0.1491	<u>0.8151</u>
PRACS [14]	<u>0.0773</u>	0.1096	<u>0.8216</u>	<u>0.0921</u>	0.1176	0.8032
BFLP [47]	0.1941	0.1230	0.7067	0.1843	0.1254	0.7125
FSRIC [48]	0.2125	0.1485	0.6705	0.1917	0.1443	0.6912
PanColorGAN [27]	0.5451	0.1099	0.4049	0.4261	0.1218	0.5018
Fusion-Net [28]	0.1874	**0.0617**	0.7625	0.1754	**0.0592**	0.7754
Proposed	0.0724	<u>0.0964</u>	**0.8381**	0.0753	<u>0.1090</u>	**0.8239**

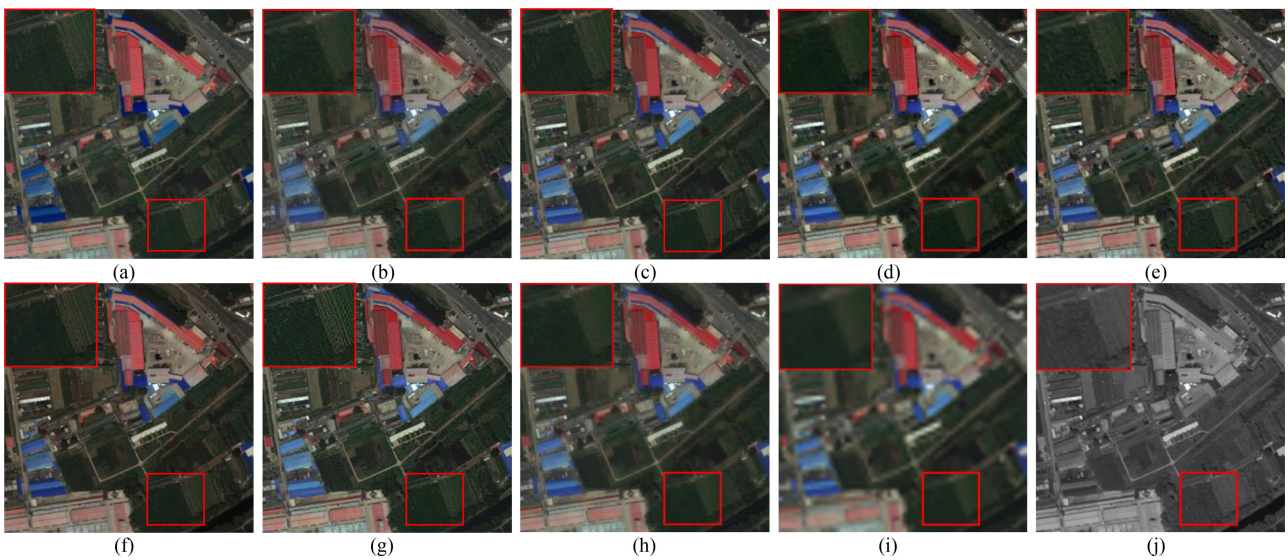

Fig. 7. Fused results by different methods on the GeoEye-1 dataset. (a) GS. (b) MTF-GLP. (c) PRACS. (d) BFLP. (e) FSRIC. (f) PanColorGAN. (g) Fusion-Net. (h) Proposed. (i) MS. (j) PAN.

TABLE V
QUANTITATIVE RESULTS ON THE GEOEYE-1 DATASET

Indexes	The objective results of Fig. 7			The average objective results of the GeoEye-1 dataset		
	$D_\lambda \downarrow$	$D_S \downarrow$	HQNR↑	$D_\lambda \downarrow$	$D_S \downarrow$	HQNR↑
GS [13]	0.1619	0.1348	0.7251	0.1664	0.1427	0.7180
MTF-GLP [46]	**0.0571**	0.1605	0.7915	**0.0796**	0.1729	0.7644
PRACS [14]	0.0884	0.0589	0.8579	0.0921	0.1176	0.8032
BFLP [47]	0.3088	0.1390	0.5951	0.2904	0.1536	0.5970
FSRIC [48]	0.3155	0.1244	0.5994	0.2855	0.1456	0.6077
PanColorGAN [27]	0.5024	0.1426	0.4266	0.4540	0.1488	0.4623
Fusion-Net [28]	0.2638	0.0478	0.7010	0.2289	0.0674	0.7181
Proposed	0.0929	**0.0476**	**0.8630**	0.1473	**0.0445**	**0.8147**

TABLE VI
AVERAGE RUNNING TIME OF THE PROPOSED METHOD AND COMPARISON METHODS.

Algorithms	GS [13]	MTF-GLP [46]	PRACS [14]	BFLP [47]	FSRIC [48]	PanColorGAN [27]	Fusion-Net [28]	Proposed
Time (s)	0.05	0.28	0.12	1.70	0.09	4.15	9.75	1.4225

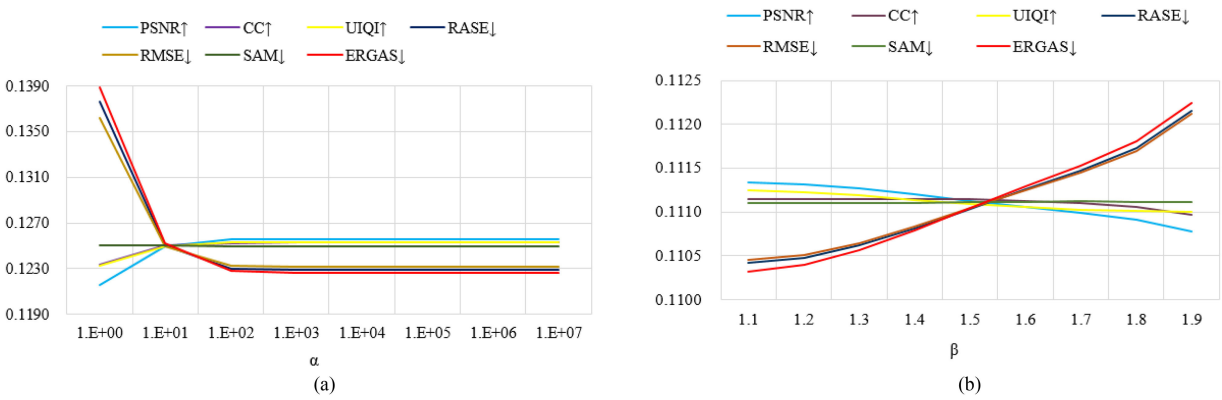

Fig. 8. Different normalized values of indexes corresponding to different values of (a) α and (b) β.

objective assessment results presented in Table V indicate that the proposed algorithm performed the best on D_s and HQNR indexes. Therefore, the proposed algorithm could achieve the optimal performance in terms of the overall quality compared to the other algorithms.

In addition to abovementioned subjective and objective evaluations, we calculated the average running time of the proposed and comparison methods on the four datasets including 360 groups of images, as indicated in Table VI. It can be observed that the BFLP, PanColorGan, and Fusion-Net methods required a relatively long time, whereas the GS, MTF-GLP, PARCS, and FSRIC methods required a relatively short time. As the VFOG model defined in the proposed algorithm is solved using sliding block technology, the proposed algorithm could not achieve optimal efficiency but was at an intermediate level among all algorithms. However, overall, according to the previous experimental results on the four datasets, the proposed algorithm still outperformed the other algorithms.

V. DISCUSSION

A. Analysis of Parameters

The main parameters used in the proposed method, namely α and β in (9), are generated in numerous experiments. In this article, for generalization, we used 30 groups of images from each satellite dataset to constitute a new dataset containing 120 groups of images. Thereafter, the optimal values of α and β were obtained through experiments performed on this new dataset. The RMSE, RASE, SAM, PSNR, CC, UIQI, and ERGAS indexes were employed to examine how these parameters affected the overall spatial and spectral performance of the proposed method, as illustrated in Fig. 8. It is worth noting that all of the results in Fig. 8 are the normalized average values. It is obvious from Fig. 8(a) that when $\alpha \geq +031.\text{E}$, the proposed method could realize the best tradeoff between these indexes. Moreover, Fig. 8(b) indicates that when $\beta = 1.1$, the proposed method could achieve the best performance. Therefore, we set $\alpha = 1.\text{E}$ and $\beta = 1.1$, in this article, to achieve better spectral and spatial performance.

B. Performance of VFOG Model

Most researchers have utilized the P_{hist} image to solve the spectral distortion problem in the CS-based injection scheme. However, this approach is not very effective owing to the low correlation between the P_{hist} and M images. Therefore, we designed the VFOG model to estimate the P_{VFOG} image with similar characteristics to the M image to minimize the spectral distortion. To verify whether the P_{VFOG} image was more related

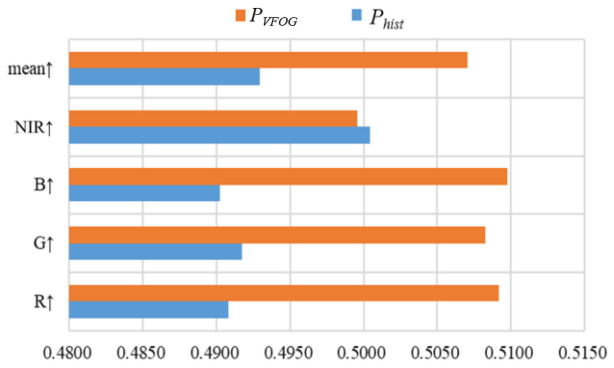

Fig. 9. Correlation between P_{hist} images, P_{VFOG} images and different bands of M images.

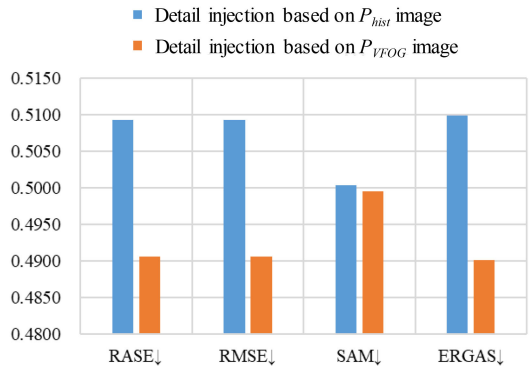

Fig. 10. Comparison of the performance of detail injection based on P_{hist} and P_{VFOG} images.

to the M image compared to the P_{hist} image, a comparison experiment was conducted based on the new dataset.

We compared the correlation between the P_{hist} and M images with the correlation between P_{VFOG} and M images, and the results are presented in Fig. 9. All of the measured results in Fig. 9 are the normalized average values of 120 groups of fusion results. As indicated in Fig. 9, the correlations between the R, G, and B bands of the M and P_{VFOG} images were significantly higher than the correlations between the R, G, and B bands of the M and P_{hist} images, whereas the correlation between the NIR band of the M and P_{VFOG} images was slightly lower than that between the NIR band of the M and P_{hist} images. However, according to the mean value, the correlation between the M and P_{VFOG} images was significantly higher than that between the M and P_{hist} images. Therefore, the correlation between the P_{VFOG} and M images was higher than that between the P_{hist} and M images, and the high-frequency details of the P_{VFOG} image had the approximate texture distribution of the M image, which led to a better spectral quality in the pansharpened image.

As mentioned previously, the high-frequency details that were extracted from the P_{VFOG} image could effectively reduce the spectral distortion of the pansharpened image. To verify this point, the high-frequency details from the P_{hist} and P_{VFOG} images were measured, following which these details were injected into the M images to obtain the fused images. The superior performance of the VFOG model could be proven by objectively assessing the degree of spectral distortion in the fused images. As illustrated in Fig. 10, the RASE, RMSE, SAM, and ERGAS indexes were applied to estimate the spectral fidelity of the fused images quantitatively. The results in Fig. 10 are the normalized average indexes of 120 groups of objective index results. It is obvious that, compared to the algorithm using the P_{hist} image, the algorithm using the P_{VFOG} image achieved the optimal results on all indexes, which indicates that the algorithm using the VFOG model achieved better spectral performance. Therefore, the VFOG model could effectively improve the spectral fidelity of the fused images.

C. Performance of Injection Gains

As vegetation has very strong reflectivity in the NIR band, more vegetation information exists in this band. Therefore, it is necessary to inject more vegetation information into the NIR band. To confirm this claim, comparative experiments were conducted based on the injection gains with E_{NDVI} and without E_{NDVI}. The PSNR, CC, UIQI, RASE, RMSE, and ERGAS indexes were used to compare the performance of the algorithms quantitatively. All of the measured results in Fig. 11 are the normalized average indexes of 120 groups of objective index results. It is evident from Fig. 11 that the algorithm with the injection gains considering E_{NDVI} could achieve the optimal results on all indexes. Thus, it was reasonable to inject more vegetation information into the NIR band. This further validates the necessity of introducing E_{NDVI} to determine the injection gains.

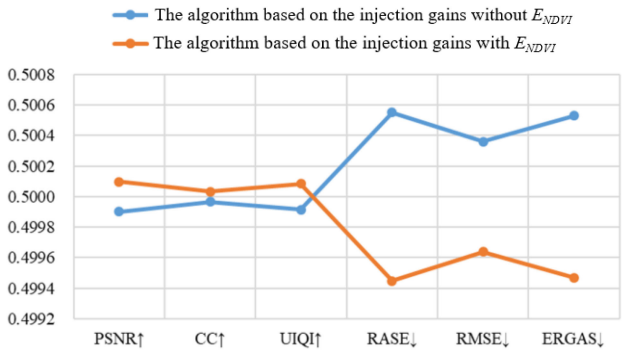

Fig. 11. Performance comparison of E_{NDVI}.

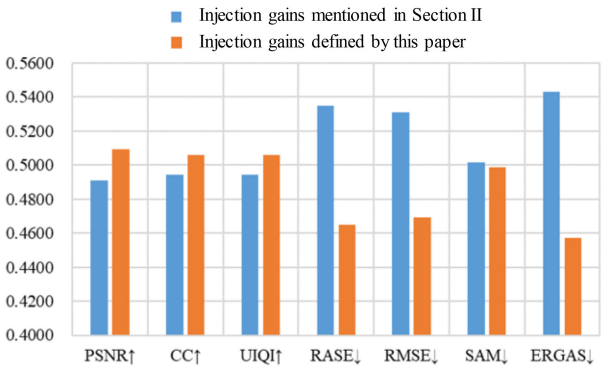

Fig. 12. Comparison of the performance of injection gains mentioned in Section II and defined by this article.

Moreover, comparative experiments were conducted to validate the performance of the injection gains designed in this article. Specifically, the defined injection gains and the injection gains mentioned in Section II were used to inject the high-frequency details into the M images to obtain the final fused images. Through quantitative comparison of the normalized average results of each index, the evaluation results are shown in Fig. 12. It is apparent that the algorithm with the defined injection gains achieved the optimal results on all indexes, which further validated the fact that the defined injection gains could effectively increase the performance of the algorithm.

VI. Conclusion

In this article, a novel pansharpened algorithm based on the VFOG model and optimized injection gains has been presented. The main purpose of the proposed algorithm is to minimize the spectral distortion caused by the dissimilarities between the PAN and MS images. Moreover, more accurate injection gains were determined to ensure the performance of the proposed method. Therefore, the VFOG model, which comprehensively considers the spatial structure of the MS image and the gradient characteristics of the PAN image, was designed to reduce the spectral distortion of the pansharpened images that are obtained by the proposed method. Moreover, as larger spectral dissimilarity usually occurs within vegetated areas, the NDVI was introduced to modify the injection gains and to improve the quality of the final fused images further. The experimental results demonstrate that the proposed algorithm can generate better fusion results in terms of spectral and spatial fidelity compared with various other pansharpening methods.

References

[1] B. Deng, S. Jia, and D. Shi, "Deep metric learning-based feature embedding for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 2, pp. 1422–1435, Feb. 2020.

[2] J. Nalepa, M. Antoniak, M. Myller, P. R. Lorenzo, and M. Marcinkiewicz, "Towards resource-frugal deep convolutional neural networks for hyperspectral image segmentation," *Microprocessors Microsyst.*, vol. 73, 2020, Art. no. 102994.

[3] L. Qi, J. Li, Y. Wang, Y. Huang, and X. Gao, "Spectral–Spatial-Weighted multiview collaborative sparse unmixing for hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 12, pp. 8766–8779, Dec. 2020.

[4] P. Wang, L. Wang, H. Leung, and G. Zhang, "Super-resolution mapping based on spatial–spectral correlation for spectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 3, pp. 2256–2268, Mar. 2021.

[5] N. Chang, K. Bai, S. Imen, C. Chen, and W. Gao, "Multisensor satellite image fusion and networking for all-weather environmental monitoring," *IEEE Syst. J.*, vol. 12, no. 2, pp. 1341–1357, Jun. 2018.

[6] B. Aiazzi, L. Alparone, S. Baronti, R. Carla, A. Garzelli, and L. Santurri, "Sensitivity of pansharpening methods to temporal and instrumental changes between multispectral and panchromatic data sets," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 1, pp. 308–319, Jan. 2017.

[7] P. Liu and L. Xiao, "A novel generalized intensity-hue-saturation (GIHS) based pan-sharpening method with variational hessian transferring," *IEEE Access*, vol. 6, pp. 46751–46761, 2018.

[8] S. Lolli, L. Alparone, A. Garzelli, and G. Vivone, "Haze correction for contrast-based multispectral pansharpening," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2255–2259, Dec. 2017.

[9] Q. Wang, W. Shi, and P. M. Atkinson, "Information loss-guided multi-resolution image fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 45–57, Jan. 2020.

[10] N. Chang, K. Bai, S. Imen, C. Chen, and W. Gao, "Multisensor satellite image fusion and networking for all-weather environmental monitoring," *IEEE Syst. J.*, vol. 12, no. 2, pp. 1341–1357, Jun. 2018.

[11] G. Vivone, L. Alparone, and J. Chanussot, "A critical comparison among pansharpening algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2565–2586, May 2015.

[12] M. Chikr El-Mezouar, N. Taleb, K. Kpalma, and J. Ronsin, "An IHS-Based fusion for color distortion reduction and vegetation enhancement in IKONOS imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 5, pp. 1590–1602, May 2011.

[13] C. A. Laben and B. V. Brower, "Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening," U.S. Patent 6011875, Jan. 2000.

[14] J. Choi, K. Yu, and Y. Kim, "A new adaptive component-substitution-based satellite image fusion by using partial replacement," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 295–309, Jan. 2011.

[15] Y. Yang, H. Lu, S. Huang, Y. Fang, and W. Tu, "An efficient and high-quality pansharpening model based on conditional random fields," *Inf. Sci.*, vol. 553, pp. 1–18, Apr. 2021.

[16] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "An MTF-based spectral distortion minimizing model for pan-sharpening of very high resolution multispectral images of urban areas," in *Proc. 2nd GRSS/ISPRS Joint Workshop Remote Sens. Data Fusion Over Urban Areas*, Berlin, Germany, 2003, pp. 90–94.

[17] G. Vivone, S. Marano, and J. Chanussot, "Pansharpening: Context-based generalized laplacian pyramids by robust regression," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 9, pp. 6152–6167, Sep. 2020.

[18] J. Nunez, X. Otazu, O. Fors, A. Prades, V. Pala, and R. Arbiol, "Multiresolution-based image fusion with additive wavelet decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1204–1211, May 1999.

[19] S. Li, J. T. Kwok, and Y. Wang, "Using the discrete wavelet frame transform to merge landsat TM and SPOT panchromatic images," *Inf. Fusion*, vol. 3, no. 1, pp. 17–23, 2002.

[20] D. Lv, F. Li, Q. Guo, X. Wang, and T. Chen, "Unsupervised change detection in remote sensing image based on image fusion in nonsubsampled shearlet transform domain and fuzzy k-means clustering," in *Proc. IEEE 3rd Adv. Inf. Technol. Electron. Automat. Control Conf.*, 2018, pp. 1568–1573.

[21] Y. Yang, H. Lu, S. Huang, and W. Tu, "Remote sensing image fusion based on fuzzy logic and salience measure," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 11, pp. 1943–1947, Nov. 2020.

[22] L. Meysam and G. Hassan, "A new variational model in texture space for pansharpening," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 8, pp. 1269–1273, Aug. 2018.

[23] 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.

[24] M. Möeller, T. Wittman, A. L. Bertozzi, and M. Burger, "A variational approach for sharpening high dimensional images," *SIAM J. Imag. Sci.*, vol. 5, no. 1, pp. 150–178, 2012.

[25] X. X. Zhu and R. Bamler, "A sparse image fusion algorithm with application to pan-sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2827–2836, May 2013.

[26] Y. Yang, W. Tu, S. Huang, and H. Lu, "PCDRN: Progressive cascade deep residual network for pansharpening," *Remote Sens.*, vol. 12, no. 4, Feb. 2020, Art. no. 676.

[27] F. Ozcelik, U. Alganci, E. Sertel, and G. Unal, "Rethinking CNN-Based pansharpening: Guided colorization of panchromatic images via GANs," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 4, pp. 3486–3501, Jul. 2020.

[28] L.-J. Deng, G. Vivone, C. Jin, and J. Chanussot, "Detail injection-based deep convolutional neural networks for pansharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 8, pp. 6995–7010, Aug. 2021.

[29] J. Yang, L. Xiao, Y.-Q. Zhao, and J. C.-W. Chan, "Variational regularization network with attentive deep prior for hyperspectral–multispectral image fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5508817.

[30] Y. Yang, H. Lu, S. Huang, and W. Tu, "Pansharpening based on joint-guided detail extraction," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 389–401, Jan. 2021.

[31] S. Rahmani, M. Strait, D. Merkurjev, M. Moeller, and T. Wittman, "An adaptive IHS pan-sharpening method," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 746–750, Oct. 2010.

[32] Y. Leung, J. Liu, and J. Zhang, "An improved adaptive intensity-hue-saturation method for the fusion of remote sensing images," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 5, pp. 985–989, May 2014.
[33] H. Zhang, J. Ma, C. Chen, and X. Tian, "NDVI-Net: A fusion network for generating high-resolution normalized difference vegetation index in remote sensing," *ISPRS J. Photogramm. Remote Sens.*, vol. 168, pp. 182–196, Oct. 2020.
[34] P. Liu, L. Xiao, and S. Tang, "A new geometry enforcing variational model for pan-sharpening," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 12, pp. 5726–5739, Dec. 2016.
[35] P. Liu, L. Xiao, J. Zhang, and B. Naz, "Spatial-Hessian-Feature-Guided variational model for pan-sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 2235–2253, Apr. 2016.
[36] Y.-F. Pu *et al.*, "A fractional-order variational framework for retinex: Fractional-order partial differential equation-based formulation for multi-scale nonlocal contrast enhancement with texture preserving," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1214–1229, Mar. 2018.
[37] Z. Jun and W. Zhihui, "A class of fractional-order multi-scale variational models and alternating projection algorithm for image denoising," *Appl. Math. Model.*, vol. 35, no. 5, pp. 2516–2528, 2011.
[38] P. Liu, L. Xiao, and T. Li, "A variational pan-sharpening method based on spatial fractional-order geometry and spectral-spatial low-rank priors," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1788–1802, Mar. 2018.
[39] J. Choi, G. Kim, N. Park, H. Park, and S. Choi, "A hybrid pansharpening algorithm of VHR satellite images that employs injection gains based on NDVI to reduce computational costs," *Remote Sens.*, vol. 9, no. 10, pp. 976–997, 2017.
[40] A. Garzelli, B. Aiazzi, L. Alparone, S. Lolli, and G. Vivone, "Multispectral pansharpening with radiative transfer-based detail-injection modeling for preserving changes in vegetation cover," *Remote Sens.*, vol. 10, no. 8, Aug. 2018, Art. no. 1308.
[41] J. Choi, P. Honglyun, and D. Seo, "Pansharpening using guided filtering to improve the spatial clarity of VHR satellite imagery," *Remote Sens.*, vol. 11, no. 6, pp. 633–652, 2019.
[42] Y. Yang, L. Li, S. Huang, Y. Zhang, and H. Lu, "Remote sensing image fusion with convolutional sparse representation based on adaptive dictionary learning," *J. Signal Process.*, vol. 36, no. 1, pp. 125–138, 2020.
[43] Q. Xu, Y. Zhang, B. Li, and L. Ding, "Pansharpening using regression of classified MS and pan images to reduce color distortion," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 1, pp. 28–32, Jan. 2015.
[44] J. Choi, J. Yeom, A. Chang, Y. Byun, and Y. Kim, "Hybrid pansharpening algorithm for high spatial resolution satellite imagery to improve spatial quality," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 3, pp. 490–494, May 2013.
[45] 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.
[46] 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.
[47] N. H. Kaplan and I. Erer, "Bilateral filtering-based enhanced pansharpening of multispectral satellite images," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 11, pp. 1941–1945, Nov. 2014.
[48] G. Vivone, R. Restaino, and J. Chanussot, "Full scale Regression-based injection coefficients for panchromatic sharpening," *IEEE Trans. Image Process.*, vol. 27, no. 7, pp. 3418–3431, Jul. 2018.
[49] T. Wang, F. Fang, F. Li, and G. Zhang, "High-Quality Bayesian pansharpening," *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 227–239, Jan. 2019.
[50] Y. Yang, L. Wu, S. Huang, J. Sun, W. Wan, and J. Wu, "Compensation details-based injection model for remote sensing image fusion," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 5, pp. 734–738, May 2018.
[51] R. H. Yuhas, A. F. H. Goetz, and J. W. Boardman, "Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm," in *Proc. Summaries 3rd Annu. JPL Airborne Geosci. Workshop*, 1992, pp. 147–149.
[52] G. Vivone, R. Restaino, and J. Chanussot, "A Bayesian procedure for full-resolution quality assessment of pansharpened products," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4820–4834, Aug. 2018.
[53] G. Vivone, M. Dalla Mura, A. Garzelli, and F. Pacifici, "A benchmarking protocol for pansharpening: Dataset, preprocessing, and quality assessment," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 6102–6118, Jun. 2021.

Yong Yang (Senior Member, IEEE) received the Ph.D. degree in biomedical engineering from Xi'an Jiaotong University, Xi'an, China, in 2005.

From 2009 to 2010, he was a Postdoctoral Research Fellow with Chonbuk National University, Jeonju, South Korea. He is currently a Distinguished Professor with the School of Computer Science and Technology, Tiangong University, Tianjin, China. His current research interests include image fusion, image super-resolution reconstruction, medical image processing and analysis, and pattern recognition.

Prof. Yang is an Associate Editor of the IEEE ACCESS and an Editor of *KSII Transactions on Internet and Information Systems*.

Hangyuan Lu received the Ph.D. degree in management science and engineering from Jiangxi University of Finance and Economics, Nanchang, China, in 2021.

He is currently an Associate Professor with Jinhua Polytechnic, Jinhua, China. His current research interests include image fusion, sparse representation, and deep learning.

Shuying Huang (Member, IEEE) received the Ph.D. degree in computer application technology from the Ocean University of China, Qingdao, China, in 2013.

She is currently an Associate Professor with the School of Computer Science and Technology, Tiangong University, Tianjin, China. Her current research interests include image and signal processing, and pattern recognition.

Weiguo Wan received the B.S. degree in mathematics and applied mathematics from Jiangxi Normal University, Nanchang, China, in 2014, and the Ph.D. degree in computer science and engineering from Jeonbuk National University, Jeonju, South Korea, in 2020.

He is currently a Lecturer with the School of Software and Internet of Things Engineering, Jiangxi University of Finance and Economics, Nanchang, China. His current research interests include computer vision, deep learning, face sketch synthesis and recognition, and remote sensing image fusion.

Luyi Li received the B.S. degree in soft engineering from Minnan Normal University, Fujian, China, in 2017, and the M.S. degree in computer technology from the Jiangxi University of Finance and Economics, Nanchang, China, in 2020.

Her current research interests include image fusion, sparse representation, and deep learning.