Pansharpening Based on Adaptive High-Frequency Fusion and Injection Coefficients Optimization

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

***Abstract*—The purpose of pansharpening is to fuse a multispectral (MS) image with a panchromatic (PAN) image to generate a high spatial-resolution multispectral (HRMS) image. However, the traditional pansharpening methods do not adequately take consideration of the information of MS images, resulting in inaccurate detail injection and spectral distortion in the pansharpened results. To solve this problem, a new pansharpening approach based on adaptive high-frequency fusion and injection coefficients optimization is proposed, which can obtain an accurate injected high-frequency component (HFC) and injection coefficients. First, we propose a multi-level sharpening model to enhance the spatial information of the MS image, and then extract the HFCs from the sharpened MS image and PAN image. Next, an adaptive fusion strategy is designed to obtain the accurate injected HFC by calculating the similarity and difference of the extracted HFCs. Regarding the injection coefficients, we propose injection coefficients optimization scheme based on the spatial and spectral relationship between the MS image and PAN image. Finally, the HRMS image is obtained through injecting the fused HFC into the upsampled MS image with the injection coefficients. Experiments with simulated and real data are performed on IKONOS and Pléiades datasets. Both subjective and objective results indicate that our method has better performance than state-of-the-art pansharpening approaches.**

***Index Terms*—High-frequency fusion, injection coefficients, multilevel sharpening, pansharpening.**

I. Introduction

NOWADAYS, the demand for high-precision satellite images is increasing in the applications of environmental monitoring [1], spectral unmixing [2], water quality assessment [3], etc. Nevertheless, limited by the sensor, the high spatial-resolution multispectral (HRMS) image is hard to obtain through a single sensor. This problem can be solved by fusing a high spatial-resolution panchromatic (PAN) image and a high spectral-resolution multispectral (MS) image to obtain a HRMS image. This process is also called pansharpening.

During the last few years, many pansharpening methods have been developed, which can be roughly divided into two categories: tradition methods and deep learning-based methods. The tradition methods include component substitution (CS) based methods, multiresolution analysis (MRA) based methods, and variational optimization (VO) based methods.

In the CS-based method, the MS image is first projected into another space. Then, the PAN image is adopted to replace the spatial information of the transformed components. Finally, the HRMS image is achieved by using the inverse transformation. The representative CS-based methods consist of the intensity-hue-saturation [4], [5], the principal component analysis [6], [7], and the Gram–Schmidt adaptive approach [8]. The advantages of these methods are that they are excellent in terms of spatial information and have low computational complexity. However, the disadvantage is that they are prone to spectral distortion [9].

MRA-based methods decompose the source images into low-frequency components (LFCs) and high-frequency components (HFCs) at different scales, and then these components are fused by some certain fusion rules. Finally, the fused components are inversely transformed to obtain the HRMS image. Common MRA-based methods include the Laplace transform [10], [11], à trous wavelet transform (ATWT) [12], [13], and smoothing filter based-intensity modulation [14]. This kind of method can preserve the spectral information well; however, there are still some defects in the enhancement of spatial information [15].

The pansharpening method based on VO includes two steps: the design of the energy function and its optimization solution [16]. The energy function is constructed using the methods such as the observation model and sparse representation, and then the optimization algorithm is used to tackle this problem to achieve a pansharpened image. Common VO-based methods include P+XS [17], regularized solutions of inverse problems [18], [19], coupled nonnegative matrix factorization [20], and sparse representation [21]. Although the results of the VO-based methods have high accuracy, the computational costs are relatively high [22].

Manuscript received 11 July 2022; revised 12 December 2022; accepted 21 December 2022. Date of publication 26 December 2022; date of current version 2 January 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 62072218, Grant 61862030, and Grant 62261025, in part by the Natural Science Foundation of Jiangxi Province under Grant 20192ACB20002 and Grant 20192ACBL21008, in part by the Talent project of Jiangxi Thousand Talents Program under Grant jxsq2019201056, and in part by the Postdoctoral Research Projects of Jiangxi Province under Grant 2020KY44. *(Corresponding author: Shuying Huang.)*

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

Chenxu Wan is with the School of Management Science and Engineering, Harbin Engineering University, Harbin 150006, China (e-mail: wanchenxu@126.com).

Shuying Huang is with the School of Software, Tiangong University, Tianjin 300387, China (e-mail: shuying-huang2010@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).

Digital Object Identifier 10.1109/JSTARS.2022.3232145

Recently, deep learning has impressive success in the field of remote sensing image processing. The deep learning-based methods solve the pansharpening problem by learning the features of the remote sensing images. In 2016, Masi et al. [23] developed a pansharpening approach based on the convolutional neural network. Since then, more deep learning-based pansharpening approaches have been developed, including the deep residual network [24], multiscale and multidepth convolutional neural network [25], and deep convolutional neural networks [26]. This type of method is good at learning complex features from images and obtain good fusion results. However, the disadvantages are that a large amount of training samples and an immense amount of training time are required, and the optimal parameters are difficult to adjust [27].

Based on above analysis, the article proposes a pansharpening method based on adaptive high-frequency fusion and injection coefficients optimization to obtain high spatial-resolution MS images, namely, HRMS images. First, the MS image is sharpened by a multilevel sharpening model, and the HFCs of the sharpened image (SI) and the PAN image are obtained by ATWT and the guided filter, respectively. Next, an adaptive fusion strategy is designed according to the difference and similarity between the high-frequencies of the PAN image and those of the sharpened MS image. Then, the adaptive injection coefficients based on the spatial and spectral information of source images are proposed. Finally, the HRMS image is achieved by injecting the fused HFC into the upsampled MS (UPMS) image with the designed injection coefficients. Experiments on IKONOS and Pléiades datasets show that compared with existing advanced pansharpening methods, the performance of our method is superior, both subjectively and objectively, which proves the performance of our method. The main contributions of this article are as follows.

1) An effective pansharpening approach based on adaptive high-frequency fusion and injection coefficients optimization is presented, which can effectively preserve the spatial and spectral information in the pansharpened image.
2) A multilevel sharpening scheme for MS images is designed, which can enhance the spatial information of MS images well.
3) An adaptive fusion strategy is proposed according to the difference and similarity between the HFCs of the PAN and sharpened MS images. This strategy can better retain the high-frequency information of the MS and PAN images.
4) Adaptive injection coefficients are defined based on the spectral and spatial information of the source images, which can effectively control the injection of the fused HFC to obtain the final HRMS image.

II. Related Works

With the advancement of pansharpening research, injection models are widely used in the traditional pansharpening approaches. Furthermore, the methods based on the injection model have gradually replaced the traditional pansharpening methods. In this section, we present the notations used in the article and concisely discuss the detail injection models of CS and MRA.

A. Notation

Take $\tilde{M} \in R^{M^*N^*K}$ as the UPMS image of the same size as the PAN image, where *M* and *N* denote the row and column of the matrix, respectively, and *K* denotes the number of spectral bands. $\hat{M} \in R^{M*N*K}$ denotes the HRMS image after pansharpening. $I \in R^{M*N}$ denotes the intensity component, which is obtained by a linear transformation from $\tilde{M}$. $P \in R^{M*N}$ is calculated by matching the histogram of the PAN image with *I*.

B. CS-based Detail Injection Model

The traditional CS-based pansharpening method is implemented by substituting the spatial component of the MS image with the PAN image. However, Tu et al. [28] found that this method can be realized in another way, which obtains the spatial detail of the PAN image by calculating the difference between the PAN image and the intensity component of the MS image. Thus, the HRMS image is obtained by injecting the extracted spatial details into the UPMS image with the injection coefficients. The CS-based detail injection model can be implemented without spatial transformation and is advocated in [7], [29]; it is defined as follows:

$$\hat{M}_k = \tilde{M}_k + g_k (P - I), \qquad k = 1, 2, 3 \ldots K \tag{1}$$

where g_k denotes the injection coefficients of the *k*th band. *I* is obtained by a linear combination of MS images [30], which is obtained by

$$I = \sum_{k=1}^{n} \alpha_k \tilde{M}_k \tag{2}$$

where α_k denotes the weight coefficients of the *k*th band.

C. MRA-based Detail Injection Model

Unlike the CS-based detail injection model, the MRA-based detail injection model obtains the HFC through the difference between the PAN image and its LFC. The formula of the MRA-based detail injection model is defined as follows:

$$\hat{M}_k = \tilde{M}_k + g_k (P - P_L), \qquad k = 1, 2, 3 \ldots n \tag{3}$$

where P_L denotes the LFC of *P*. Choi et al. [31] found that the injected HFC does not completely depend on the PAN image. Moreover, the HFC of an MS image is also related to the HFC of an HRMS image, which is difficult to obtain from the PAN image. However, the HFC of an MS image is important for subsequent results. Therefore, the new definition of HRMS is defined as follows:

$$\begin{aligned} \hat{M}_k &= \mathrm{LF}(\hat{M}_k) + \mathrm{HF}(\hat{M}_k) \\ &\approx \tilde{M}_k + g_k \mathrm{HF}(P, I) \end{aligned} \tag{4}$$

where *LF*(•) represents the LFC, and *HF*(•) represents the HFC.

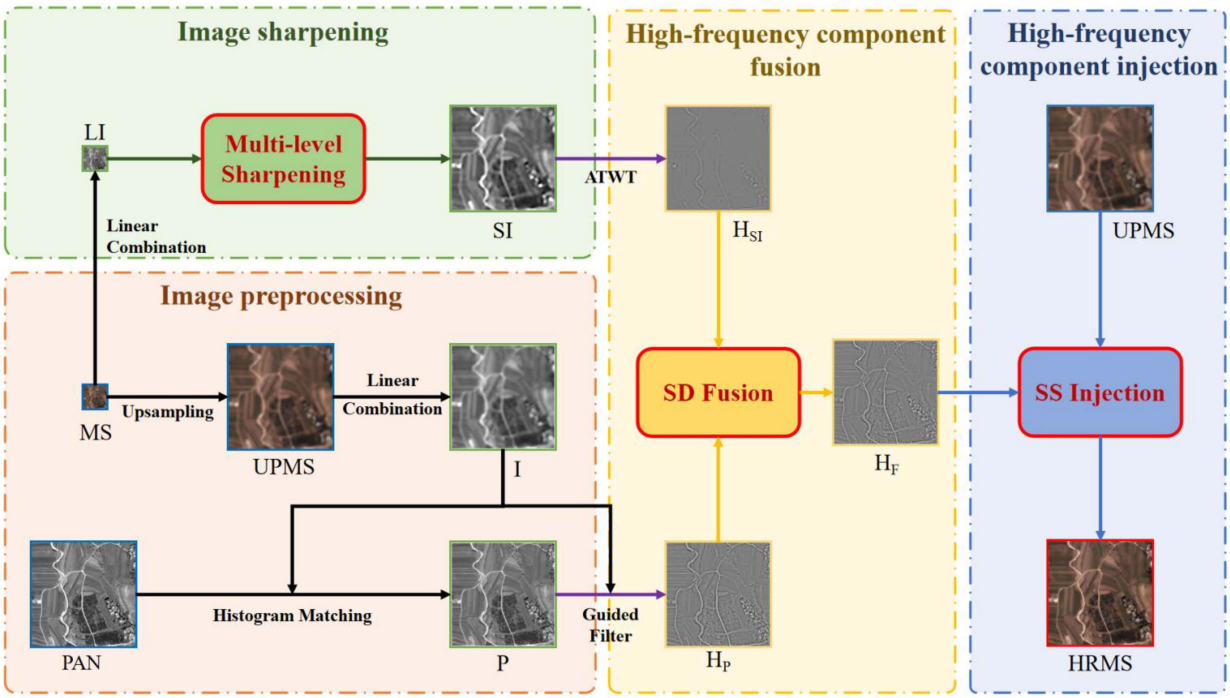

Fig. 1. Framework of the proposed method.

III. Proposed Method

The traditional MRA-based injection model generally fuses the high-frequency details of the source images, and then directly injects them into the UPMS image. Moreover, unlike other pansharpening approaches, the MRA-based approach not only preserves the spatial and spectral information well, but also has relatively low computational complexity. However, as the spatial information in the MS image is poor and blurry, the direct fusion of the decomposed HFCs causes artifacts and blurring in the final result. Therefore, how to effectively extract and fuse the HFCs of source images requires careful consideration. In addition, the injection coefficients also need to be carefully considered. This article proposes a new pansharpening method to solve the above problems, and the framework of our method is displayed in Fig. 1, whose process is described as follows.

1) The MS image and its upsampling version are preprocessed to obtain the linear combination components (*LI* and *I*), respectively, and the *P* component is achieved by using histogram matching on the PAN image and the *I* component.
2) A multilevel sharpening method is proposed to sharpen the *LI* component to obtain a SI.
3) The HFCs (H_{SI} and H_P) are extracted from SI and *P* through ATWT and the guided filter, respectively.
4) A high-frequency fusion rule is designed based on the similarity and difference between the high-frequency of SI (H_{SI}) and H_P to obtain the fused high-frequency (H_F), denoted as SD fusion.
5) The injection coefficients are defined based on the spatial and spectral relationship between the MS image and PAN image, which is called SS injection.
6) The fused high-frequency (H_F) is injected into the UPMS image by using the obtained injection coefficients to achieve the final HRMS image.

A. Multilevel Sharpening

MS images contain poor spatial information, and the HFC directly decomposed from the MS image is rough. Therefore, the fusion and injection of the HFC cause blur and artifacts in the final result. A feasible way to address the problem is to enhance the spatial information of MS images. Thus, to reduce the artifacts and enhance the spatial information while preserving the original spatial information, we propose a multilevel sharpening method to enhance the spatial information of the MS image.

Before the multilevel sharpening operation, the MS image is preprocessed to obtain the *I* and *LI* components. First, the MS image is upsampled to obtain $\tilde{M}$. Then, the *I* component is obtained by the linear combination of $\tilde{M}$. Similarly, *LI* is obtained by the linear combination of MS images. Finally, SI is obtained by multi-level sharpening of *LI*.

To sharpen the MS image, the spatial structure of the MS image is extracted by a convolution operator, and then the SI is obtained by adding the extracted spatial structure with the original image. A 3×3 operator kernel *W* is designed to sharpen the center information according to the surrounding information of the image. Because the size of the PAN image is four times that of the MS image, and it is difficult to obtain the satisfied sharpening results by directly upsampling the MS image with a factor of 4. Inspired by the image pyramid, we divide an upscale factor of 4 into two upscale factors of 2. Therefore, we propose a multi-level sharpening method. The specific steps

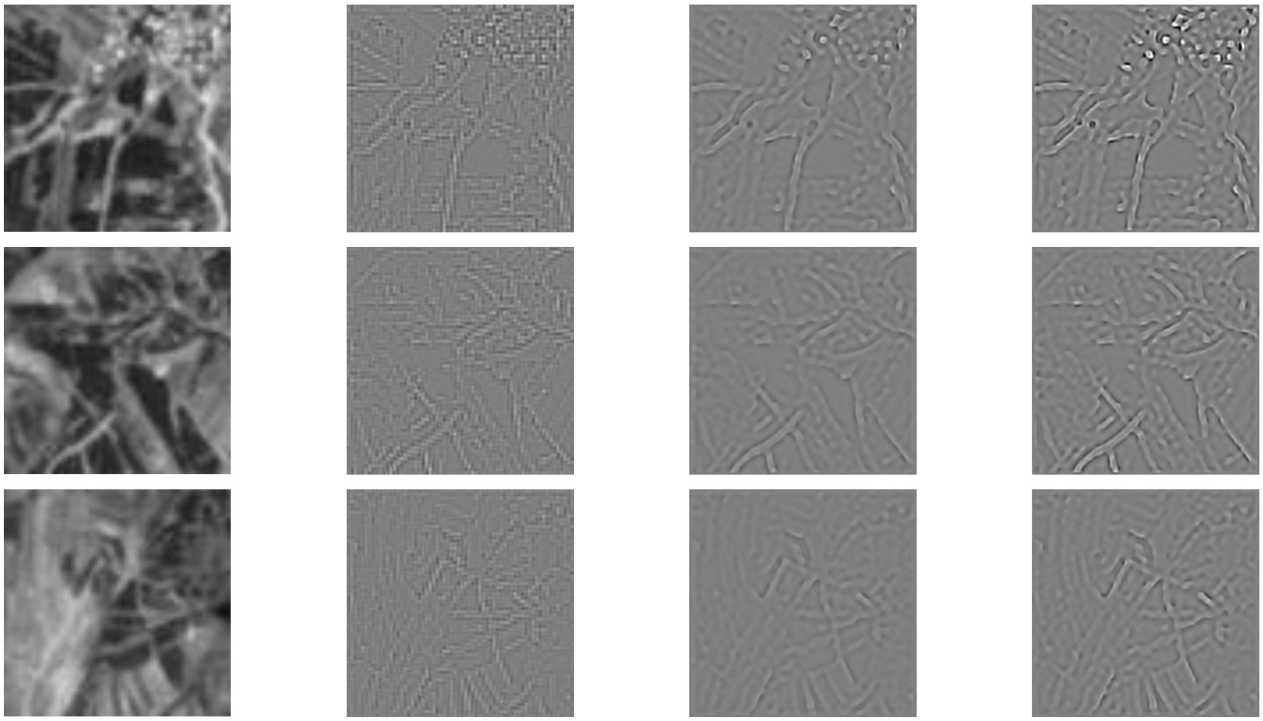

Fig. 2. HFCs of different sharpening methods. (a) LI. (b) Unsharpened HFCs. (c) Sharpened HFCs by direct upsampling four times. (d) Sharpened HFCs by the proposed multilevel sharpening model.

Algorithm 1: Multilevel Sharpening Method.

Input: *LI*, convolution kernel W, $\lambda = 2$.
Stage I:
1) Sharpen the image
C_{LI} is obtained by the convolution:
$$C_{LI} = LI * W$$
Fuse the images before and after convolution:
$$F_{IL} = \alpha \cdot LI + \beta \cdot C_{LI}$$
where α and β denote weight parameters.
2) Up-sample the fused image with an upscale factor of λ
$$U_{LI} = UP(F_{LI}, \lambda)$$
Stage II:
3) Progressive sharpening and upsampling
LI=U_{LI}, and repeat Stage I to obtain the U_{LI}.
Stage III:
4) Denoising and sharpening
G_{LI} is obtained by performing Gaussian filtering on U_{LI};
LI=G_{LI}, and repeat 1) to obtain the sharpened image F_{IL};
SI is obtained by performing guided filtering on F_{IL}.
Output: The sharpened image *SI*.

of multilevel sharpening are shown in Algorithm 1. Through the algorithm, the SI is obtained by sharpening the *LI* image step by step, which improves the spatial information of the MS image in SI.

Furthermore, to illustrate the effectiveness of the multi-level sharpening model, our method is compared to the method without the sharpening process and the method with direct upsampling four times. Fig. 2 shows the comparison results of the unsharpening and sharpening methods on three MS images. Fig. 2(a) shows the component LI of each MS image. Fig. 2(b)–(d) shows the HFCs of the results of each method. It can be seen that the details of the sharpened HFC are clearer, while the details of the unsharpened HFC are blurred and rough. Moreover, the effect of the multi-level sharpening model is better than that of using direct upsampling four times, which indicates that our sharpening method has superiority performance in enhancing the spatial information.

B. High-frequency Fusion Based on Similarity and Difference

Before fusing the HFCs, the HFCs of the source images need to be extracted separately. This article employs ATWT [12] and guided filtering [32] to filter the MS image and PAN image, respectively, to obtain the corresponding HFCs. ATWT is obtained by improving the discrete wavelet transform, which has many advantages, including non-orthogonality, shift invariance, non-decimation, and redundancy [13]. Therefore, to better extract the HFC from the MS image, we use ATWT to obtain the HFC (H_{SI}). Guided filtering is similar to bilateral filtering in that it is also an edge-preserving filtering algorithm. In addition, it can constrain the input image according to the changing trend of the guided map, which can solve the problem of gradient inversion [32]. In order to make the extracted HFC similar to the

MS image, we use I as the guided map and filter the image (P) to obtain the HFC (H_P).

Most pansharpening methods focus on the spatial structures of the PAN image but ignore some important spatial structures in the MS image, causing the final result to be different from the reference image. This problem can be solved by fusing the high-frequency of the source images to enhance the similarity between the injected high-frequency and MS image [15]. Actually, the source images are captured in the same scene but have a different emphasis on storing information. Therefore, the spatial information of the source images has high similarity, but at the same time, there are certain differences between them. Their HFCs also have the same characteristics; that is, the overall structure of the HFCs is similar, and there are some differences in local details. Based on this analysis, to retain the PAN image information well while supplementing the missing information of MS images, a fusion weight based on the similarity and difference of the HFCs is proposed. Structural similarity (SSIM) is an index to measure the similarity of the two images, which consists of mean value, standard deviation, and covariance, representing brightness, contrast, and SSIM, respectively [33]. The root-mean-square error (RMSE) can well express the difference between the two images [34]. Therefore, we use SSIM to measure the similarity of HFCs and use RMSE to evaluate the difference between HFCs. Before calculation, the pixel value is normalized between 0 and 1 so that the values calculated by SSIM and RMSE are between 0 and 1. Therefore, the proposed fusion weight is defined as follows:

$$\begin{aligned}\theta &= \frac{1}{2}(\mathrm{SSIM}_N + \mathrm{RMSE}_N)\\ &= \frac{1}{2} * \frac{(2\mu_P\mu_{SI} + C_1)(2\sigma_{PSI} + C_2)}{(\mu_P^2 + \mu_{SI}^2 + C_1) + (\sigma_P^2 + \sigma_{SI}^2 + C_2)}\\ &\quad + \frac{1}{2}\sqrt{\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}(P(i,j) - SI(i,j))^2} \end{aligned} \tag{5}$$

where SSIM_N and RMSE_N denote the normalized SSIM and RMSE, respectively. μ_P and μ_{SI} represent the mean values of P and SI, respectively. σ_P and σ_{SI} represent the variance of P and SI, respectively. σ_{PSI} represents the covariance of P and SI. $C_1 = (k_1L)^2$ and $C_2 = (k_2L)^2$ represent the constants used to keep stability, and L represents the dynamic range of pixel values [33]. Equation (5) can be divided into two parts: the first half and the second half, which represent the similarity and difference between the HFCs, respectively. To enable the fused HFC to better retain the information of HFCs (H_{SI}) and H_P, we set θ as the weight of H_{SI} and (1-θ) as the weight of H_P. Therefore, with the proposed fusion weight, the fusion strategy is defined as

$$H_F = \theta * H_{\mathrm{SI}} + (1-\theta) * H_P \tag{6}$$

where H_F is the fused HFC.

C. Injection Coefficients Based on the Spatial and Spectral Relationship

In the injection model, the injection coefficients play a key role in the process of injecting the fused HFC into the UPMS image. The good injection coefficients can improve the accuracy of the information injection. If the injection coefficients are too large or too small, spatial or spectral distortion occurs in the final pansharpened result. Therefore, it is important to find the suitable injection coefficients. The adaptive IHS (AIHS) [35] obtains the injection coefficients by calculating the edge information of the source image, which can be represented as follows:

$$g_k^E = \frac{\tilde{M}_k}{\frac{1}{K}\sum_{k=1}^{K}\tilde{M}_k}\left(\beta_k W_{\tilde{M}_k} + (1-\beta_k)W_P\right) \tag{7}$$

where β_k denotes the kth tradeoff coefficient of the MS image. $W_{\tilde{M}_k}$ and W_P denote the edge detection matrices of MS and PAN images, respectively. $gE\ k$ denotes the kth injection coefficients calculated by the AIHS method. Unfortunately, although this method is based on spatial information, it ignores the relationship between the spectra of the MS image. Therefore, the injection of spatial information may cause a change in the original ratio between the spectra of the MS image. A practical way to deal with this problem is to utilize spectral information to constrain the injection coefficients. Hence, in the article, we propose to combine the spatial and spectral information to design better injection coefficients. Based on AIHS, a spectral information constraint is presented to optimize the injection coefficients by considering the relationship between the PAN image and the various spectral bands of the MS image. Thus, the SSIM of the PAN image and each spectral band of the MS image is multiplied by the ratio of the standard deviation of the PAN image and each spectral band of the MS image, which is designed as

$$W_k^{os} = \mathrm{SSIM}\left(\tilde{M}_k, P\right) * \left(\frac{\mathrm{std}\left(\tilde{M}_k\right)}{\mathrm{std}\,(P)}\right) \tag{8}$$

where W_k^{os} denotes the spectral relationship of the kth spectral band, SSIM(•) denotes the SSIM between images, and std(•) means the standard deviation function. Subsequently, the spectral relationship $Wos\ k$ is used to constrain $gE\ k$ to obtain the spectral constrained coefficients. The $gE\ k$ and spectral constraint coefficients are then combined to obtain the final injection coefficient g_k

$$g_k = \frac{1}{2}\left(g_k^E + \left(\frac{W_k^{os}}{\sum_{i=1}^{K}W_i^{os}} * \sum_{j=1}^{K} g_j^E\right)\right). \tag{9}$$

Finally, the fusion high-frequency (H_F) is injected into $\tilde{M}_k$through the injection coefficients g_k, and thus the pansharpened result ($\hat{M}_k$) is obtained by

$$\hat{M}_k = \tilde{M}_k + g_k * H_F. \tag{10}$$

IV. Experimental Results and Analysis

A. Datasets

This article is tested and verified on two widely used datasets, including IKONOS and Pléiades. The IKONOS satellite is a commercial satellite, and the dataset used in this experiment includes urban and farmland areas. The Pléiades consists of Pléiades-1 and Pléiades-2 satellites, and the dataset used in this experiment covers the area of Shenzhen in China. In this experiment, the two datasets each contain 60 sets of MS images and PAN images. MS images of four bands (red, green, blue, and near-infrared) were used in this experiment, and its sizes is $256\times256\times4$. The sizes of PAN images are $1024\times1024\times1$. To objectively verify the effectiveness of our method, the test experiments include simulated data evaluation and real data evaluation. The simulated data assessment is to downsample the original data according to Wald's protocol [36], and use the original data as ground truth (GT). The simulated MS images are obtained through the MTF [10] operation and the downsampling operation with a sampling factor of 4, and they are used as input images. The real data assessment uses the original data to perform experiments as the input images to the experiment.

B. Benchmark and Quality Indexes

Some advanced pansharpening methods are used as benchmarks: the Gram–Schmidt adaptive (GSA) method [8]; the partial replacement adaptive CS (PRACS) method [31]; the bilateral filtering-based luminance proportional (BFLP) method [37]; the context-based decision (CBD) method [10]; the matting model and multiscale transform (MMMT) based method [38]; the full scale regression-based injection coefficients (FSRIC) method [39]; the robust band-dependent spatial-detail (BDSD-PC) method [40], the pansharpening neural network (PNN) [23]; and the fusion with deep convolutional neural networks (Fusion-Net) [26]. Among them, GSA, PRACS, and BDSD-PC are based on CS; BFLP, CBD, MMMT, and FSRIC are based on MRA; PNN and Fusion-Net are deep-learning-based methods. All of the methods used in this article are derived from publicly sourced codes provided by the corresponding authors. In the experiment, all the methods are tested in the environment of MATLAB2019B, and the hardware device is a computer with a 3.2GHz CPU and 16GB RAM.

For the simulated dataset and the real dataset, we use two different sets of indexes for the objective assessment. The simulated data assessment indexes include the correlation coefficient (CC), hypercomplex quality assessment $Q2^n$ (Q4 for four bands) [41], spectral angle mapper (SAM) [42], erreur relative global adimensionnelle de synthèse (ERGAS) [43], and peak signal-to-noise ratio (PSNR) [44]. For the real data assessment, the hybrid quality with no reference (HQNR) index [45] is used for the objective assessment. It includes the spectral distortion index (D_λ^K) [46] and the spatial distortion index (D_S) [47].

C. Simulated Datasets Assessment

In this section, we introduce in detail the assessment of pansharpened results on simulated dataset. All of the pansharpened results of the contrast methods as well as the proposed method are analyzed and evaluated subjectively and objectively. Fig. 3 displays the simulated experiment results on Pléiades dataset. Fig. 3(a) shows the GT, and Fig. 3(b) shows the simulated MS image. Fig. 3(c)–(l) shows the pansharpened results of the comparison methods. It can be observed clearly that the results of BFLP and Fusion-Net have spectral distortion, but it is difficult to find obvious differences for other methods. In order to make observation more convenient, we subtract the pansharpened results from the GT to obtain the residual images, and the results are displayed in Fig. 4. Fig. 4(a) is an ideal residual image, and Fig. 4(b)–(l) show the residual images of other methods. It is obvious that the residual result of BFLP has some large residual blocks, and the residual results of GSA, FSRIC, and BDSD-PC have many residues, which indicates that spatial information of the four methods is different from that of GT. The residual images of MMMT, FS, and PNN contain color information, which indicates that the results of these methods have large differences in spectra compared to GT. Observing the residual images, we can find that our result is closest to the ideal residual image, which indicates that the proposed method performs the best in the subjective assessment. Table I gives the objective evaluation results of Fig. 3. In the table, the best value is bolded and the second-best value is underlined. We can see that our method obtains the best values in all the quality assessment metrics except SAM.

TABLE I
Quantitative Assessment of Results in Fig. 3

	CC↑	$Q2^n$↑	SAM↓	ERGAS↓	PSNR↑
GSA	0.9005	0.7802	7.5063	6.1803	24.2845
PRACS	0.9252	0.8429	5.5732	4.7024	26.2867
BFLP	0.9237	0.7871	10.5218	7.5200	22.3262
CBD	0.9079	0.7930	6.8261	5.6271	24.9439
MMMT	0.9168	0.8180	5.8596	5.0551	25.7230
FSRIC	0.9130	0.8039	6.5233	5.3780	25.2820
BDSD-PC	0.9171	0.8129	5.9455	5.3379	25.3602
PNN	0.9079	0.8234	5.7142	5.3996	25.2582
Fusion-Net	<u>0.9359</u>	<u>0.8589</u>	**4.4581**	<u>4.4562</u>	<u>26.8425</u>
Proposed	**0.9398**	**0.8727**	<u>4.6130</u>	**4.1461**	**27.3643**

Fig. 5 displays the pansharpened images of the degraded experiments on IKONOS. Fig. 5(a)–(b) shows GT and the degraded MS image, respectively. Fig. 5(c)–(l) shows the pansharpened results of each comparison method. The results of BFLP and PNN exhibit some spectral distortion. There are some colored noises in the result of the Fusion-Net. In order to show the differences of the compared methods more conveniently, we display the residual images in Fig. 6. Fig. 6(a) shows the ideal residual image, and Fig. 6(b)–(l) shows the residual results of other methods. From Fig. 6, it can be observed that the results of the BFLP and MMMT methods have obvious color patches, which indicates that the two methods have spectral differences compared with the ideal residual image. Also, there are many residues and color information in the residual results of PNN and Fusion-Net, which indicates that the results of these two methods are different from GT, both in space and spectrum.

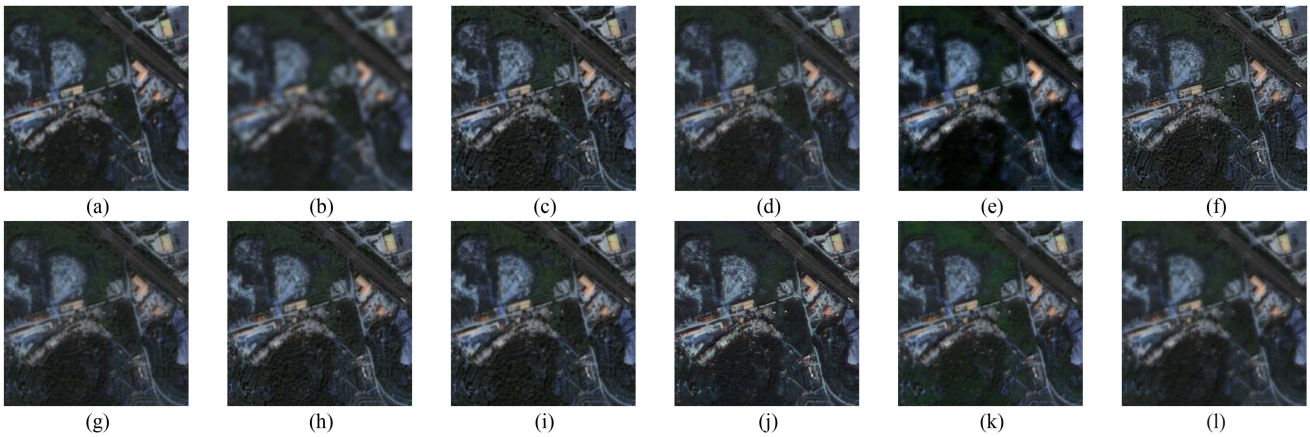

Fig. 3. Fusion results of Pléiades simulated dataset. (a) Ground truth. (b) Upsampled MS. (c) Gram–Schmidt adaptive. (d) Partial replacement adaptive CS. (e) Bilateral filtering-based luminance proportional. (f) Context-based decision. (g) Matting model and multiscale transform. (h) Full scale regression-based injection coefficient. (i) Band-dependent spatial-detail. (j) Pansharpening neural network. (k) Fusion-Net. (l) Proposed.

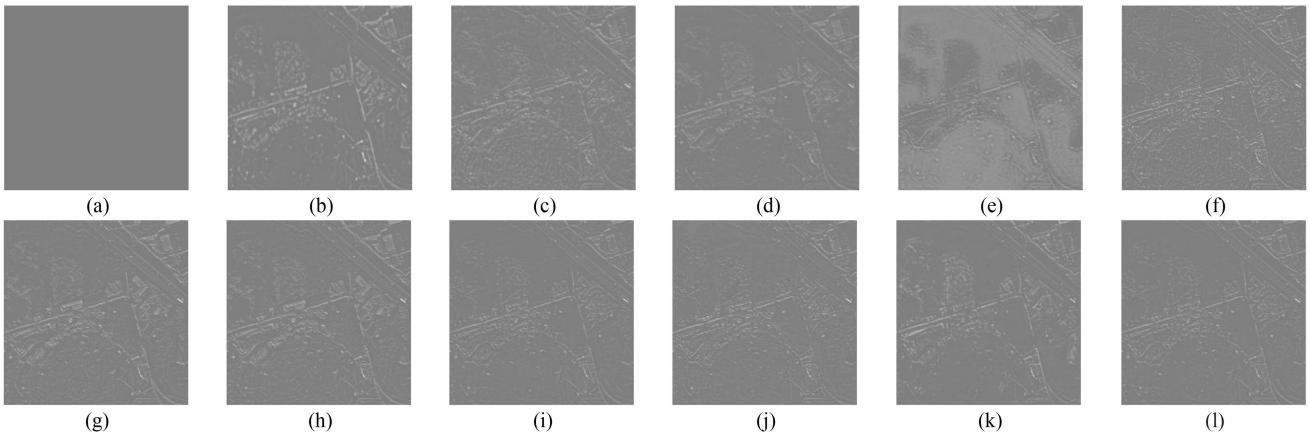

Fig. 4. Residual images estimated by the difference between GT and the pansharpened image of each algorithm in Fig. 3. (a) Ideal image. (b) Upsampled MS. (c) Gram–Schmidt adaptive. (d) Partial replacement adaptive CS. (e) Bilateral filtering-based luminance proportional. (f) Context-based decision. (g) Matting model and multiscale transform. (h) Full scale regression-based injection coefficient. (i) Band-dependent spatial-detail. (j) Pansharpening neural network. (k) Fusion-Net. (l) Proposed.

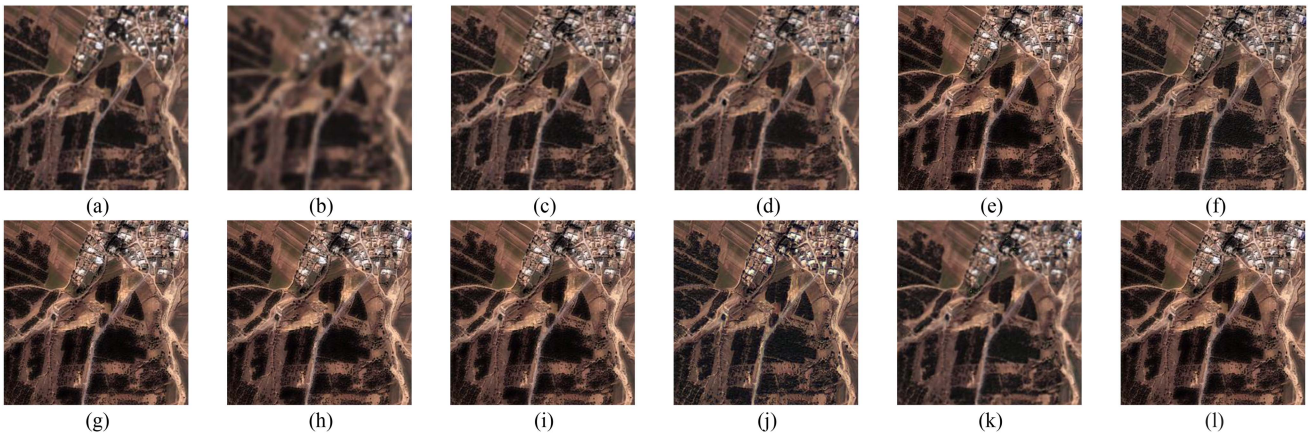

Fig. 5. Fusion results of IKONOS simulated dataset. (a) Ground truth. (b) Upsampled MS. (c) Gram–Schmidt adaptive. (d) Partial replacement adaptive CS. (e) Bilateral filtering-based luminance proportional. (f) Context-based decision. (g) Matting model and multiscale transform. (h) Full scale regression-based injection coefficient. (i) Band-dependent spatial-detail. (j) Pansharpening neural network. (k) Fusion-Net. (l) Proposed.

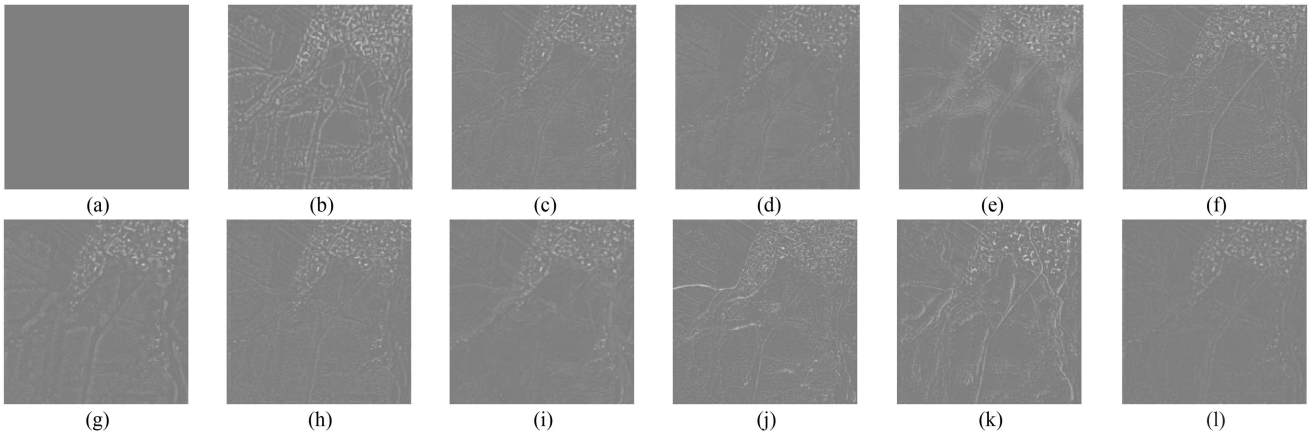

Fig. 6. Residual images estimated by the difference between GT and the pansharpened image of each algorithm in Fig. 5. (a) Ideal image. (b) Upsampled MS. (c) Gram–Schmidt adaptive. (d) Partial replacement adaptive CS. (e) Bilateral filtering-based luminance proportional. (f) Context-based decision. (g) Matting model and multiscale transform. (h) Full scale regression-based injection coefficient. (i) Band-dependent spatial-detail. (j) Pansharpening neural network. (k) Fusion-Net. (l) Proposed.

TABLE II
QUANTITATIVE ASSESSMENT OF RESULTS IN FIG. 5

	CC↑	$Q2^n$↑	SAM↓	ERGAS↓	PSNR↑
GSA	0.9428	0.9126	7.2255	4.5861	24.8720
PRACS	0.9470	0.9195	6.2327	4.2249	25.4339
BFLP	0.9396	0.8713	6.3960	5.5014	22.9288
CBD	0.9327	0.8870	5.5954	5.2275	23.4241
MMMT	0.9373	0.9024	6.8020	4.5387	24.7183
FSRIC	0.9424	0.9111	6.9172	4.4636	25.0148
BDSD-PC	0.9463	0.9178	7.3866	4.5410	24.9434
PNN	0.9373	0.8981	6.4922	4.8233	24.1327
Fusion-Net	0.9348	0.8972	**5.0898**	4.4977	24.7499
Proposed	**0.9528**	**0.9299**	5.6855	**3.8911**	**26.0479**

Compared with the residual results of other methods, it can be observed that the result of the proposed method has the least residues, which indicates that our approach can obtain the best result both spatially and spectrally. Furthermore, Table II gives the objective evaluation results of Fig. 5, and it shows that our method achieves the best values among all the quality assessment indexes except SAM.

To further demonstrate the validation of our method, we perform quantitative evaluations on Pléiades and IKONOS. The average quantitative results of the two datasets (each includes 60 sets of images) are given in Table III. From the table, it can be seen that for these two simulated datasets, all of the assessment indexes of our method are the best. Therefore, through the objective results presented here, the effectiveness of our method is verified.

D. Real Dataset Assessment

This section presents and analyzes the pansharpened results on the real data of the IKONOS and Pléiades. Fig. 7 shows the fusion results of IKONOS dataset. Fig. 7(a)–(b) shows the original source images. Fig. 7(c)–(l) displays the pansharpened results of the comparison methods. As can be seen from Fig. 7, the results of PNN and Fusion-Net exhibit spectral distortion, as indicated by the arrow area. To better differentiate the fusion results by different methods, we enlarge a part of the area (blue boxes) in Fig. 7, and the closeups are displayed in Fig. 8. From the red and green marked boxes in Fig. 8, we can observe that there are obvious color deviations in GSA, MMMT, FSRIC, and BDSD-PC. In the red area of Fig. 8(j), the color of the object is purple, which is different from the source PAN image showing that the result of PNN has spectral distortion. In Fig. 8(k), there is a large amount of colored noise, signifying that Fusion-Net is weak in retaining spectral information. Furthermore, Table IV gives the objective evaluation results of Fig. 7. The assessment result of our method is the best on the HQNR and D_S metrics.

In addition, the average objective evaluation results on the two real datasets are given in Table V. From the table, we can see that the comprehensive index HQNR of our method is the best, while D_S and D_λ^K are either the best or the second-best on these two datasets, which indicates the performance of our method.

E. Research on Multi-Level Sharpening Parameters

As introduced in Section III-A, to improve the spatial quality of the HFC extracted from the MS image, we propose a multilevel sharpening operation by superimposing the extracted high-frequency on the MS image. The choice of kernels W and parameters α and β are discussed through experiments. Four kernels and the corresponding α and β values are set, as given in Table VI, and the HFCs of the sharpening results performed by them are shown in Fig. 9. Fig. 9(a) shows the HFC extracted from the result without the sharpening method, and Fig. 9(b)– (i) shows the HFCs extracted from the sharpening results using eight kinds of parameters corresponding to A1-A8. It can be seen that the results using the parameters of A4, A6, and A8 are oversharpened. The result using the parameters of A2 has the best sharpening effect, and has better texture details than those of other methods. Moreover, to better illustrate the effects of different methods, we objectively evaluate the final pansharpened results. The experiment is conducted on a simulated dataset

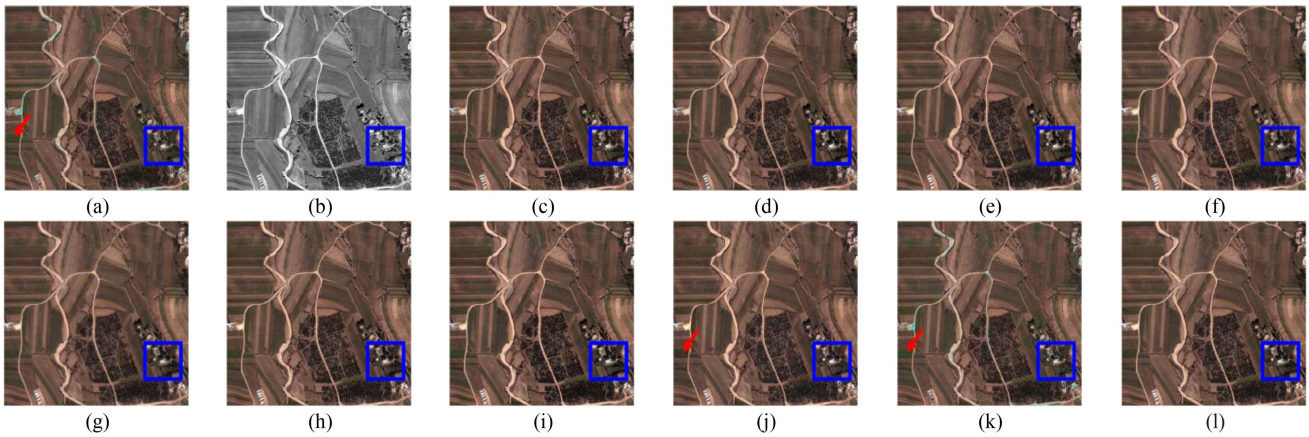

Fig. 7. Fusion results of the full resolution IKONOS dataset. (a) Multispectral. (b) Panchromatic. (c) Gram–Schmidt adaptive. (d) Partial replacement adaptive CS. (e) Bilateral filtering-based luminance proportiona. (f) Context-based decision. (g) Matting model and multiscale transform. (h) Full scale regression-based injection coefficient. (i) Band-dependent spatial-detail. (j) Pansharpening neural network. (k) Fusion-Net. (l) Proposed.

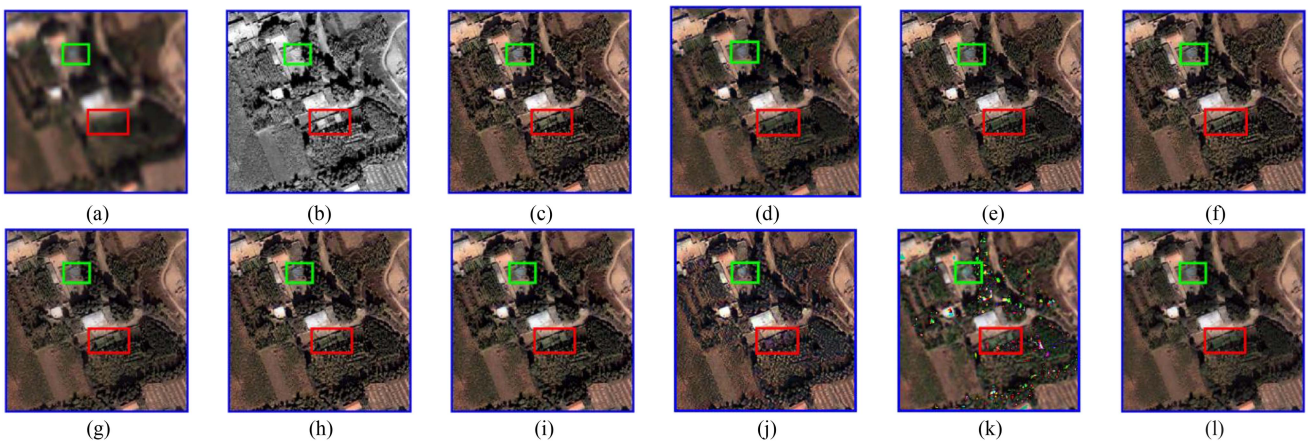

Fig. 8. Closeups of the fusion results in Fig. 7. (a) Multispectral. (b) Panchromatic. (c) Gram–Schmidt adaptive. (d) Partial replacement adaptive CS. (e) Bilateral filtering-based luminance proportional. (f) Context-based decision. (g) Matting model and multiscale transform. (h) Full scale regression-based injection coefficient. (i) Band-dependent spatial-detail. (j) Pansharpening neural network. (k) Fusion-Net. (l) Proposed.

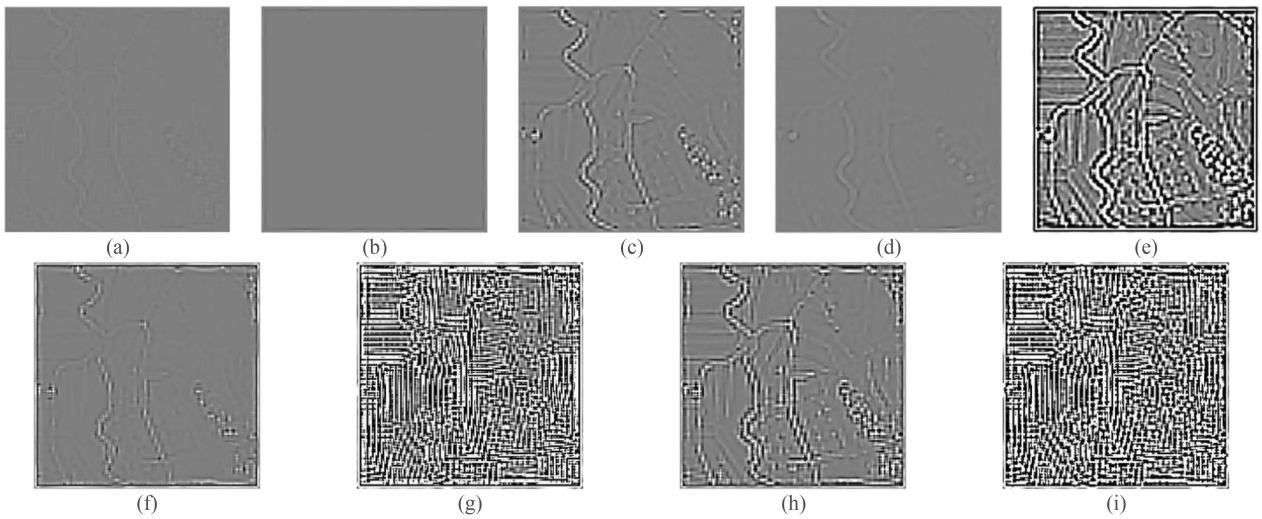

Fig. 9. HFCs of the SI. (a) Multispectral. (b) A1. (c) A2. (d) A3. (e) A4. (f) A5. (g) A6. (h) A7. (i) A8.

TABLE III
AVERAGE QUANTITATIVE RESULTS ON THE DEGRADED DATASETS

	Pléiades					IKONOS				
	CC↑	$Q2^n$↑	SAM↓	ERGAS↓	PSNR↑	CC↑	$Q2^n$↑	SAM↓	ERGAS↓	PSNR↑
GSA	0.9085	0.7615	5.9677	5.7634	26.1004	0.9249	0.8631	6.1391	4.6005	26.8180
PRACS	0.9322	0.8304	5.0589	4.6732	27.7948	0.9293	0.8680	5.2714	4.4245	27.5480
BFLP	0.9237	0.7382	6.0044	7.8789	23.5236	0.9303	0.7407	6.1553	7.3675	22.0691
CBD	0.9160	0.7738	5.6448	5.3183	26.7215	0.9236	0.8603	5.9959	4.5604	26.8878
MMMT	0.9234	0.8062	5.1875	4.9317	27.2985	0.9182	0.8510	6.0027	4.5160	27.0610
FSRIC	0.9205	0.7845	5.4817	5.1287	27.0085	0.9271	0.8665	5.8448	4.4372	27.1608
BDSD-PC	0.9224	0.7949	5.2467	5.1472	26.9434	0.9389	0.8843	5.7158	4.1619	27.6863
PNN	0.9140	0.8000	5.7479	5.4469	26.5870	0.9008	0.8220	6.7570	5.6850	25.5901
Fusion-Net	0.9355	0.8390	4.4591	4.5762	28.0374	0.8978	0.8300	5.1421	6.1494	26.6989
Proposed	**0.9438**	**0.8568**	**4.4184**	**4.1694**	**28.7234**	**0.9412**	**0.8873**	**4.8787**	**3.9096**	**28.4340**

TABLE IV
QUANTITATIVE ASSESSMENT OF RESULTS IN FIG. 7

	HQNR↑	D_λ^K ↓	D_S ↓
GSA	0.7551	0.0504	0.2048
PRACS	0.8363	0.0344	0.1340
BFLP	0.7854	0.0239	0.1953
CBD	0.7970	0.0265	0.1813
MMMT	0.8309	0.0392	0.1352
FSRIC	0.8027	**0.0213**	0.1798
BDSD-PC	0.7744	0.0596	0.1765
PNN	0.8012	0.0544	0.1527
Fusion-Net	0.8113	0.1114	0.0870
Proposed	**0.8899**	0.0264	**0.0859**

TABLE V
AVERAGE QUANTITATIVE RESULTS ON THE REAL DATASETS

	PLEIADES			IKONOS		
	HQNR↑	D_λ^K ↓	D_S ↓	HQNR↑	D_λ^K ↓	D_S ↓
GSA	0.4870	0.3129	0.2954	0.7340	0.0639	0.2165
PRACS	0.6950	0.1435	0.1893	0.8086	0.0555	0.1443
BFLP	0.7726	0.1035	0.1389	0.7165	0.0843	0.2190
CBD	0.7860	0.0547	0.1694	0.7817	0.0395	0.1863
MMMT	0.7668	0.0790	0.1674	0.8366	0.0450	0.1240
FSRIC	0.7876	0.0510	0.1708	0.7913	**0.0296**	0.1848
BDSD-PC	0.7395	0.1192	0.1612	0.7820	0.0774	0.1525
PNN	0.7285	0.1244	0.1696	0.7586	0.0717	0.1830
Fusion-Net	0.7515	0.1510	**0.1168**	0.8001	0.1281	**0.0841**
Proposed	**0.8415**	**0.0409**	0.1229	**0.8827**	0.0354	0.0850

composed of IKONOS and Pléiades, containing 35 sets of images. The results are evaluated by the assessment objective indexes, which are given in Table VII. From the table, the result using the parameters of A2 is the best in all of the indexes. Therefore, these parameters are selected to sharpen the MS image.

F. Ablation Study

To further illustrate the effectiveness of each component in our method as illustrated in Section III–C, we conduct an ablation experiment. As given in Table VIII, the ablation experiment includes three comparison methods: the method without multilevel sharpening (O_1), the method without similarity and dissimilarity fusion (O_2), and the method without spatial and spectral injection (O_3). The experiment is conducted on the same simulated dataset as Section III-E. The results are evaluated by the assessment metrics, which are shown in Fig. 10. From the figure, it can be observed that our method performs the best in all of the indexes, which further verifies the effectiveness of the proposed method.

G. Comparison of Time Consumption

To compare the computational cost of each approach, Table IX gives the average running time required for the computation of different methods. The Fusion-Net method is running on the GPU, while other methods are running on the CPU. Therefore, the Fusion-Net method does not participate in the comparison. Table IX gives the average time consumption of the comparison methods on the simulated and real datasets. Among the methods, GSA, PRACS, and BDSD-PC are the methods based on CS, which generally have less calculation time. The BFLP and MMMT methods take more time than the proposed methods. Although the proposed method consumes more time than some traditional methods, the gap is small. The computational cost of the proposed method is mainly spent on multi-level sharpening of MS images and obtaining HFCs. On the whole, our method can obtain better pansharpened results with relatively high efficiency.

V. DISCUSSION

In order to keep the final pansharpened result with accurate high-frequency information, the high-frequency information of the MS image is usually added into the injected HFC. However, it ignores the fact that the spatial information of the MS image is blurry and rough. When the high-frequency information of the MS image is fused with the high-frequency information of the PAN image, the final pansharpened result is blurry and has artifacts. These problems are caused by the introduction of coarse spatial information in the HFCs of MS images. To improve the precision of HFC injection, a feasible approach is to improve the spatial information of the MS image. Therefore, we propose a multilevel sharpening operation to improve the spatial information of the MS image without introducing

TABLE VI
SHARPENING PARAMETERS

	A1	A2	A3	A4	A5	A6	A7	A8
W	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}$
α	0.5	1	0.5	1	0.5	1	0.5	1
β	0.5	1	0.5	1	0.5	1	0.5	1

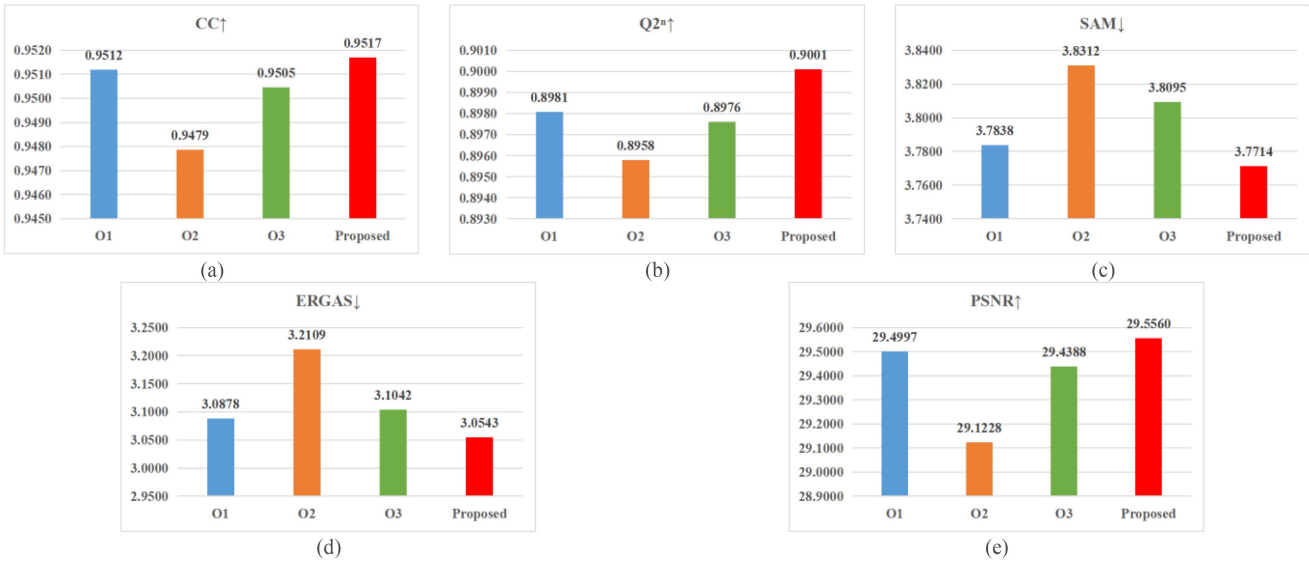

Fig. 10. Various indicators of the pansharpened results with different methods in Table VIII. (a) Correlation coefficient. (b) $Q2^n$. (c) Spectral angle mapper. (d) Erreur relative global adimensionnelle de synthèse. (e) Peak signal-to-noise ratio.

TABLE VII
AVERAGE QUANTITATIVE RESULTS OF DIFFERENT PARAMETERS

	CC↑	$Q2^n$↑	SAM↓	ERGAS↓	PSNR↑
A1	0.9509	0.8979	3.7900	3.1006	29.4663
A2	**0.9517**	**0.9001**	**3.7714**	**3.0543**	**29.5560**
A3	0.9514	0.8984	3.7820	3.0796	29.5208
A4	0.6340	0.4972	4.9904	12.5779	16.7217
A5	0.9393	0.8796	3.7920	3.3091	28.6674
A6	0.1895	0.0511	7.1717	35.1421	8.0058
A7	0.9384	0.8841	3.8238	3.4224	28.4170
A8	0.1565	0.0368	7.2455	40.6434	6.7840

TABLE VIII
COMPARISON METHODS FOR ABLATION STUDY

	Multi-level sharpening	Similarity and dissimilarity fusion	Spatial and spectral injection
O_1	×	√	√
O_2	√	×	√
O_3	√	√	×
Proposed	√	√	√

TABLE IX
AVERAGE CONSUMING TIME (SECOND) OF THE COMPARISON METHODS

	Degraded Data	Real Data
GSA	0.0475	0.6667
PRACS	0.0515	1.0672
BFLP	1.1844	9.7349
IMG	0.1406	1.7106
MMMT	8.6538	118.4500
FSRIC	0.0446	0.3969
BDSD-PC	0.0346	0.2362
PNN	0.1150	1.0339
Proposed	0.2106	2.4069

other irrelevant information. Through multi-level sharpening, the high-frequency of the sharpened MS and PAN images are extracted.

Moreover, the PAN image and the MS image are the shots of the same scene, but the main information in the two images is different. For their HFCs, their overall structure is similar, but their local details are different. In response to this phenomenon, we propose an adaptive fusion strategy based on the similarity and dissimilarity of the HFCs. The fusion is performed by calculating the similarity and difference between the HFCs of

the SI and that of PAN image. Then, the fused HFC is injected into the UPMS image. To optimize the injection coefficients, we design the injection coefficients based on the spatial and spectral relation between PAN and MS images. The injection coefficients based on spatial information are first obtained by calculating the edge matrices of the source images. Then, the SSIM between the various bands of the MS image and the PAN image are calculated, and they are normalized as the spectrum constraints to correct the obtained injection coefficients. To further enhance the spatial information and maintain the original spectral ratio between the bands, we combine the injection coefficients based on spatial information with the injection coefficients corrected by the spectral information to achieve the final injection coefficients. At the end, the HFC is injected into the UPMS image with the proposed injection coefficients, and the pansharpened result is obtained. The comparison with many of the latest pansharpening methods proves the effectiveness of our method.

Although the proposed method is better than many popular methods, it still has room for improvement. In the practical application of remote sensing, there are certain requirements for time consumption. Although the proposed method has a relatively fast calculation speed, it still does not meet the real-time requirements.

VI. Conclusion

A novel pansharpening approach based on adaptive high-frequency fusion and injection coefficients optimization has been proposed to improve the quality of pansharpened images. To preserve more spatial information from the MS image in the pansharpened result, we have proposed a multilevel sharpening method to sharpen the MS image step by step. The HFC is obtained by fusing the HFCs of the sharpened MS and PAN images according to their similarity and dissimilarity. Finally, the injection coefficients are designed based on the spatial and spectral relationship between the PAN image and MS image. The fused HFC is injected into the UPMS image with the designed injection coefficients to obtain the pansharpened result. Experimental results illustrate that our method has better performance in subjective and objective assessments compared to some advanced pansharpening approaches.

References

[1] D. Hong, N. Yokoya, J. Chanussot, and X. X. Zhu, "An augmented linear mixing model to address spectral variability for hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1923–1938, Apr. 2019.

[2] X. Zhuge, X. Zou, and Y. Wang, "Determining AHI cloud-top phase and intercomparisons with MODIS products over north pacific," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 1, pp. 436–448, Jan. 2021.

[3] D. Hong, X. Wu, P. Ghamisi, J. Chanussot, N. Yokoya, and X. X. Zhu, "Invariant attribute profiles: A spatial-frequency joint feature extractor for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 6, pp. 3791–3808, Jun. 2020.

[4] W. Carper, T. Lillesand, and R. Kiefer, "The use of intensity-hue-saturation transformations for merging spot panchromatic and multispectral image data," *Photogrammetric Eng. Remote Sens.*, vol. 56, no. 4, pp. 459–467, 1990.

[5] J. A. Malpica, "Hue adjustment to IHS pan-sharpened IKONOS imagery for vegetation enhancement," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 1, pp. 27–31, Jan. 2007.

[6] V. K. Shettigara, "A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution data set," *Photogramm. Eng. Remote Sens.*, vol. 58, no. 5, pp. 561–567, 1992.

[7] D. Bueso, M. Piles, and G. Camps-Valls, "Nonlinear PCA for Spatio-Temporal analysis of Earth observation data," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 8, pp. 5752–5763, Aug. 2020.

[8] B. Aiazzi, S. Baronti, and M. Selva, "Improving component substitution pansharpening through multivariate regression of MS+Pan Data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3230–3239, Oct. 2007.

[9] C. Chen, Y. Li, W. Liu, and J. Huang, "SIRF: Simultaneous satellite image registration and fusion in a unifified framework," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4213–4224, Nov. 2015.

[10] 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, 2006.

[11] 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.

[12] 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.

[13] G. Vivone, L. Alparone, A. Garzelli, and S. Lolli, "Fast reproducible pansharpening based on instrument and acquisition modeling: AWLP revisited," *Remote Sens.*, vol. 11, no. 5, pp. 2315, 2019.

[14] J. G. Liu, "Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details," *Int. J. Remote Sens.*, vol. 21, no. 18 pp. 3461–3472, 2000.

[15] Y. Yang, C. Wan, S. Huang, H. Lu, and W. Wan, "Pansharpening based on low-Rank fuzzy fusion and detail supplement," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 5466–5479, 2020.

[16] H. Lu, Y. Yang, S. Huang, W. Tu, and W. Wan, "A unified pansharpening model based on band-adaptive gradient and detail correction," *IEEE Trans. Image Process.*, vol. 31, pp. 918–933, 2022.

[17] 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.

[18] X. He, L. Condat, J. M. Bioucas-Dias, J. Chanussot, and J. Xia, "A new pansharpening method based on spatial and spectral sparsity priors," *IEEE Trans. Image Process.*, vol. 23, no. 9, pp. 4160–4174, Sep. 2014.

[19] L. Deng, G. Vivone, W. Guo, M. D. Mura, and J. Chanussot, "A variational pansharpening approach based on reproducible kernel Hilbert space and heaviside function," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4330–4344, Sep. 2018.

[20] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 528–537, Feb. 2012.

[21] 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.

[22] F. D. Javan, F. Samadzadegan, S. Mehravar, A. Toosi, R. Khatami, and A. Stein, "A review of image fusion techniques for pan-sharpening of high-resolution satellite imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 171, pp. 101–117, 2021.

[23] G. Masi, D. Cozzolino, V. erdoliva, and L. G. Scarpa, "Pansharpening by convolutional neural networks," *Remote Sens.*, vol. 8, no. 7, pp. 594:1–594:22, Jul. 2016.

[24] Y. Yang, W. Tu, S. Huang, and H. Lu, "PCDRN: Progressive cascade deep residual network for pansharpening," *Remote Sens.*, vol. 12, no. 4, pp. 1–20, 2020.

[25] Q. Yuan, Y. Wei, X. Meng, H. Shen, and L. Zhang, "A multiscale and multidepth convolutional neural network for remote sensing imagery pansharpening," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 978–989, Mar. 2018.

[26] 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.

[27] X. Fu, Z. Lin, Y. Huang, and X. Ding, "A variational pan-sharpening with local gradient constraints," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 10257–10266.

[28] T. Tu, S. Su, H. Shyu, and P. Huang, "A new look at IHS-like image fusion methods," *Inf. Fusion*, vol. 2, no. 3, pp. 177–186, 2001.

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

[30] G. Vivone et al., "A critical comparison among pansharpening algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2565–2586, May 2015.
[31] 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.
[32] K. He, J. Sun, and X. Tang, "Guided image fifiltering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1397–1409, Jun. 2013.
[33] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[34] Y. Yang, S. Tong, S. Huang, and P. Lin, "Multifocus image fusion based on NSCT and focused area detection," *IEEE Sensors J.*, vol. 15, no. 5, pp. 2824–2838, May 2015.
[35] 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.
[36] 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.
[37] N. H. Kaplan and I. Erer, "Bilateral fifiltering-based enhanced pansharpening of multispectral satellite images," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 11, pp. 1941–1945, Nov. 2014.
[38] Y. Yang, W. Wan, S. Huang, P. Lin, and Y. Que, "A novel pan-sharpening framework based on matting model and multiscale transform," *Remote Sens.*, vol. 9, no. 4, pp. 1–21, 2017.
[39] G. Vivone, R. Restaino, and J. Chanussot, "Full scale regression-based injection coeffificients for panchromatic sharpening," *IEEE Trans. Image Process.*, vol. 27, no. 7, pp. 3418–3431, Jul. 2018.
[40] G. Vivone, "Robust band-dependent spatial-detail approaches for panchromatic sharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6421–6433, Sep. 2019.
[41] A. Garzelli and F. Nencini, "Hypercomplex quality assessment of multi/hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 662–665, Oct. 2009.
[42] 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.
[43] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. M. Bruce, "Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 3012–3021, Oct. 2007.
[44] Z. H. Nezhad, A. Karami, R. Heylen, and P. Scheunders, "Fusion of hyperspectral and multispectral images using spectral unmixing and sparse coding," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 6, pp. 2377–2389, Jun. 2016.
[45] G. Vivone et al., "A new benchmark based on recent advances in multispectral pansharpening: Revisiting pansharpening with classical and emerging pansharpening methods," *IEEE Geosci. Remote Sens. Mag.*, vol. 9, no. 1, pp. 53–81, Mar. 2021.
[46] M. M. Khan, L. Alparone, and J. Chanussot, "Pansharpening quality assessment using the modulation transfer functions of instruments," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3880–3891, Nov. 2009.
[47] L. Alparone, B. Aiazzi, S. Baronti, A. Garzelli, F. Nencini, and M. Selva, "Multispectral and panchromatic data fusion assessment without reference," *Photogrammetric Eng. Remote Sens.*, vol. 74, no. 2, pp. 193–200, Feb. 2008.

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.

Dr. Yang is an Associate Editor for IEEE ACCESS and an Editor for *KSII Transactions on Internet and Information Systems*.

Chenxu Wan received the M.S. degree in computer science and technology from Jiangxi University of Finance and Economics, Nanchang, China, in 2021. He is currently working toward the Ph.D. degree in management science and engineering with Harbin Engineering University, Harbin, China.

His research interests include machine learning and data mining.

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

She is currently a Professor with the School of Software, Tiangong University, Tianjin, China. Her current research interests include image and signal processing, and pattern recognition.

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.

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.