

# A Unified Pansharpening Model Based on Band-adaptive Gradient and Detail Correction

Hangyuan Lu, Yong Yang, *Senior Member, IEEE*, Shuying Huang, *Member, IEEE*,

Wei Tu, and Weiguo Wan

**Abstract**—Pansharpening is used to fuse a panchromatic (PAN) image with a multispectral (MS) image to obtain a high-spatial-resolution multispectral (HRMS) image. Traditional pansharpening methods face difficulties in obtaining accurate details and have low computational efficiency. In this study, a unified pansharpening model based on the band-adaptive gradient and detail correction is proposed. First, a spectral fidelity constraint is designed by keeping each band of the HRMS image consistent with that of the MS image. Then, a band-adaptive gradient correction model is constructed by exploring the gradient relationship between a PAN image and each band of the MS image, so as to adaptively obtain an accurate spatial structure for the estimated HRMS image. To refine the spatial details, a detail correction constraint is defined based on the parameter transfer by designing a reduced-scale parameter acquisition model. Finally, a unified model is constructed based on the gradient and detail corrections, which is then solved by an alternating direction multiplier method. Both reduced-scale and full-scale experiments are conducted on several datasets. Compared with state-of-the-art pansharpening methods, the proposed method can achieve the best results in terms of fusion quality and has high efficiency. Specifically, our method improves the SAM and ERGAS metrics by 17.6% and 21.2% respectively compared to the traditional approach with the best average values, and improves these two metrics by 4.3% and 10.3% respectively compared to the learning-based approach with the best average values.

**Index Terms**—Pansharpening, band-adaptive, gradient correction, detail correction, parameter transfer.

## I. INTRODUCTION

With the rapid development of the satellite industry, remote sensing images have been widely used in natural resource monitoring, territorial management, and ecological protection [1]. An accurate high-spatial-resolution multispectral (HRMS) image provides important data for geographic information systems. However, due to the limitations of the existing satellites technique in bandwidth, payload, and so on,

most satellites cannot directly provide HRMS images [2]. They usually carry two different types of sensors: a panchromatic (PAN) sensor and a multispectral (MS) sensor. The PAN sensor has a high spatial resolution but contains little spectral information, whereas the MS sensor has a high spectral resolution but with low spatial resolution [3]. Through an appropriate algorithm, the geometric details of a PAN image and the spectral information of an MS image can be fused to obtain an HRMS image. This fusion process is also called pansharpening, and is an initial step of the image enhancement in many remote sensing tasks, such as target classification and semantic segmentation [4], [5].

Up to now, numerous pansharpening methods have been developed, which can be categorized into four groups including component substitution (CS)-based, multi-resolution analysis (MRA)-based, model-based, and deep learning (DL)-based methods [2], [6]. A CS-based method transforms an MS image to the spectral and spatial components, then uses the PAN image to replace the spatial component, and finally applies inverse transformation to obtain the fused image. Typical CS-based methods include principal component analysis (PCA) [7], gram Schmidt adaptive (GSA) [8], and the intensity–hue–saturation (IHS) transform [9]. These methods can effectively obtain high spatial quality of the fused images and are widely used in pansharpening, but they can easily cause spectral distortion, especially when the correlation between the PAN image and MS image is low [10].

Both the MRA-based and CS-based methods belong to the traditional approaches. The main difference between them is how they extract spatial details. For MRA-based methods, the spatial details of a fused image are obtained by computing the difference between a PAN image and its low-resolution version that is obtained by a pyramid transform or wavelet at multiple scales. Popular MRA-based methods include the wavelet transform [11], contourlet transform [12], additive wavelet luminance proportional [13], and à trous wavelet transform [14]. To improve the spatial quality, Kaplan *et al.* [15] proposed a bilateral filter luminance proportional (BFLP) method using an edge-preserving decomposition to obtained more satisfactory results. However, the effect of BFLP heavily depends on the accuracy of parameters setting. Thus, an adaptive multiscale bilateral filtering is proposed to further improve the generalization performance of BFLP [16]. These MAR-based methods usually perform well in spectral aspect. However, the use of multiscale transformation increases the computational complexity, and easily causes spatial distortion [2], [17].

The model-based methods learn the spectral and spatial information by making use of the assumed relations between pansharpened and observed MS and PAN images [18]. The

This work is supported by the National Natural Science Foundation of China (No.62072218 and No.61862030) and by the Natural Science Foundation of Jiangxi Province (No.20192ACB20002 and No.20192ACBL21008), and by the Talent project of Jiangxi Thousand Talents Program (No. jxsq2019201056).

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

Y. Yang and W. Tu are with the School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330032, China (e-mail: greatyangy@126.com; ncsytuwei@163.com).

S. Huang and W. Wan are with the School of Software and Communication Engineering, Jiangxi University of Finance and Economics, Nanchang 330032, China (e-mail: shuyinghuang2010@126.com; wanwgplus@163.com).

main process involved in this kind of approach is generally based on a variational optimization problem [19], and is comprised of two parts: (1) the construction of an energy function; and (2) the optimization solution of the function. The early works constructed the fusion models and restored the signal based on compressive sensing and dictionary learning [20], [21]. Later, Ghahremani *et al.* [22] improved the compressive sensing-based method by learning a new dictionary and seeking the sparse coefficient vectors with alternating direction multiplier method (ADMM). However, the dictionary learning process is usually time consuming. Instead, Yang *et al.* [23] proposed matting model and multiscale transform (MMMT) to improve the spectral quality and efficiency. Vivone *et al.* [24] further improved the efficiency by proposing the full-scale regression-based injection coefficients (FSRIC). But the effectiveness of these methods is affected by the correctness of the constructed model, and the accuracy of the spatial details of the fusion result needs to be improved [25]. Thus, an improved and robust band-dependent spatial details (BDSD) method is proposed by combining the BDSD with the removal of outliers and bisquare regression [26]. Another model-based approach was proposed to combine the low-rank fuzzy fusion (LRFF) model with the adaptive detail supplement model to improve the injected details [27]. These methods focus on improving the details and can obtain impressive spatial quality, but they do not fully consider spectral constraints, which may cause spectral distortion in the fusion results.

In recent years, deep learning (DL) methods have attracted increasing interest due to their strong ability to learn image features [28]. Masi *et al.* [29] applied the DL-based method used in the super-resolution, i.e., the super-resolution convolution neural network (SRCNN), to pansharpening, and obtained satisfactory results. To make full use of the details in the PAN image, Zhang *et al.* [30] proposed an end-to-end bi-directional pyramid network for pansharpening, and Deng *et al.* [31] exploited the combination of CNN with detail injection scheme to address the pansharpening problem. Meanwhile, Ozcelik *et al.* [32] proposed a PanColorization generative adversarial network (PanColorGAN) to further reduce the loss of detail information. DL-based methods show strong pansharpening ability, but they usually need a large number of samples and many computing resources for training; moreover, their generalizability needs to be improved [25], [33].

In this paper, aiming at the problems mentioned above, a novel pansharpening model based on gradient and detail correction is first proposed to maintain consistency between the degraded HRMS image and the up-sampled MS image (UPMS). Due to the diverse spatial information in each band of the HRMS image, we explore the gradient relationship between the MS and PAN images, and adaptively correct the gradient information for each band of the fused image. Moreover, since there is no reference image in the fusion process, keeping the structural consistency of the fused image with PAN image will cause inaccurate detail injection. To further refine the details, we take advantages of the CS and MRA methods, and construct a parameter acquisition model with a synthetic reference image. The obtained parameters are used for correcting the details in

each band of the fused image. In addition, a total variational regular term is defined to reduce the number of artifacts in the fused image. To sum up, the contributions of this work are as follows:

- (1) A gradient correction prior constraint is proposed to adaptively adjust the gradient information in each band of the fused image.
- (2) To further refine the injected details, a parameter transfer-based detail correction constraint is designed by constructing a parameter acquisition model.
- (3) A new pansharpening framework is presented by embedding the proposed constraint terms into a unified model that can obtain a high-quality fused image.

## II. TRADITIONAL DETAIL INJECTION MODEL

Both traditional CS- and MRA-based methods are image fusion methods based on the injection model. The main idea is to extract the detail information from a PAN image, and then inject this information into the MS image by using an injection coefficient. The model can combine different fusion methods to take advantages of these technologies. For an MS image with  $B$  bands  $\mathbf{M} \in \mathbb{R}^{m \times n \times B}$  and a PAN image  $\mathbf{P} \in \mathbb{R}^{M \times N}$ , the framework of the injection model can be formulated as follows [1]:

$$\widehat{\mathbf{M}}_b = \widetilde{\mathbf{M}}_b + g_b (\mathbf{P} - \mathbf{P}_b^L) \quad (1)$$

where  $\widehat{\mathbf{M}}_b$  and  $\widetilde{\mathbf{M}}_b$  represent the  $b$ th-band of the HRMS and UPMS images, respectively,  $g_b$  represents the  $b$ th-band injection coefficient, and  $\mathbf{P}_b^L$  denotes the  $b$ th-band low-resolution version of the PAN image. The evaluation accuracy of  $\mathbf{P}_b^L$  in model (1) directly affects the quality of the fused image. For the MRA-based method,  $\mathbf{P}_b^L$  can be obtained by applying a low-pass filter  $\mathbf{H}_b^L$  on  $\mathbf{P}$ , i.e.,

$$\mathbf{P}_b^L = \mathbf{H}_b^L \mathbf{P} \quad (2)$$

This kind of method can usually achieve a good spectral quality, but may cause spatial distortion [2], [34].

For the CS-based method,  $\mathbf{P}_b^L$  is replaced with the intensity component of the UPMS image ( $\mathbf{I}_{UP}$ ) after applying the IHS transform.  $\mathbf{I}_{UP}$  can be obtained by a linear combination of each band of the UPMS image:

$$\mathbf{I}_{UP} = \sum_{b=1}^B \alpha_b \widetilde{\mathbf{M}}_b \quad (3)$$

where  $\alpha_b$  represents the band weight. In this way, while a high spatial quality of the fusion result can be obtained, the spectral quality may be unsatisfactory [2]. To obtain  $\alpha_b$ , Yang *et al.* [23] proposed that the I component can be adaptively represented by a linear combination of each channel, that is,  $\alpha_b$  can be obtained by solving the following optimization problem:

$$\min_{\alpha_1, \dots, \alpha_B} \left\| \mathbf{P} - \sum_{b=1}^B \alpha_b \widetilde{\mathbf{M}}_b \right\|^2 \quad \text{s.t. } \alpha_1 \geq 0, \dots, \alpha_B \geq 0. \quad (4)$$

The injection model has good generalizability [31], and can

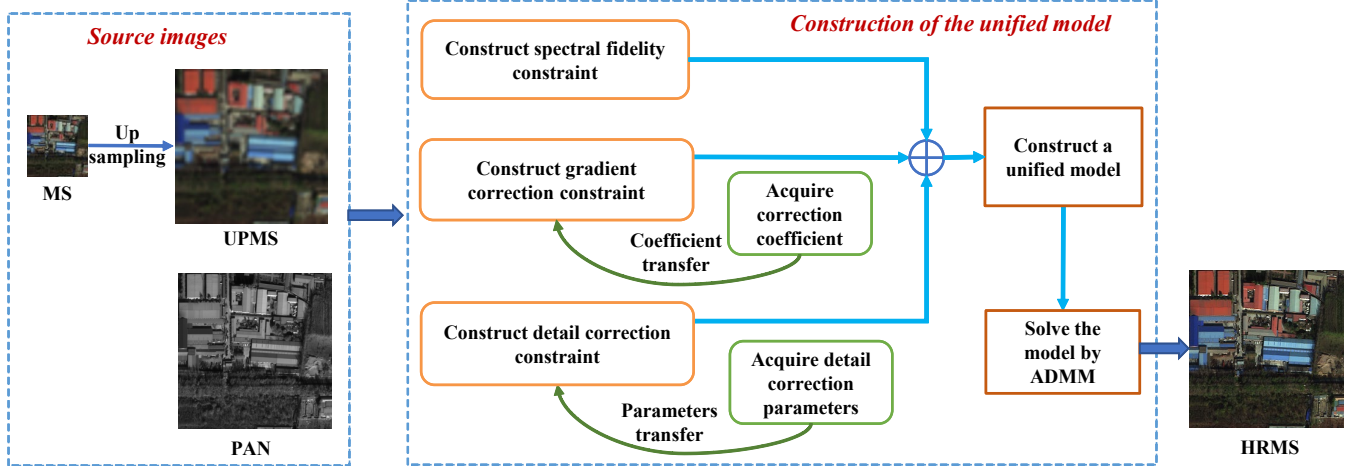


Fig. 1. Flowchart of the proposed pansharpening model.

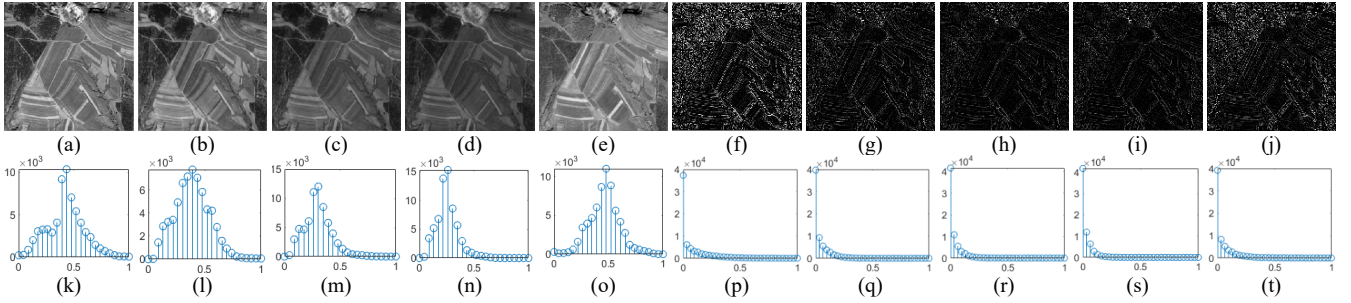


Fig. 2. The comparison of spatial information between a PAN image and a reference HRMS image. (a) is a PAN image. (b)-(e) are the four bands of the HRMS image such as red (R), green (G), blue (B), and near infrared (NIR) band. (f)-(j) are the corresponding gradient map of the images in (a)-(e). (k)-(t) are the corresponding histograms of the images in (a)-(j).

be combined with other fusion technology to improve the fusion quality.

### III. PROPOSED METHOD

In this study, a novel pansharpening method based on the band-adaptive gradient and detail correction is proposed. By exploring the spectral and spatial relationships between the PAN image and MS image, a unified pansharpening model is constructed to ensure the quality of the fusion image, as shown in Fig. 1. In the model, two band-adaptive prior terms are designed based on the spectral consistency and gradient correction to ensure the high fidelity of spectral and structural information between the fusion image and source images. Additionally, a detail correction constraint term is designed based on the reduced-scale parameter transfer to obtain more accurate details for the fused image. Finally, the optimal fusion result is obtained by using the ADMM algorithm to solve the proposed model. The detailed process is described in the following subsections.

#### A. Spectral Fidelity and Gradient Correction Prior Term

A proper pansharpening model should be able to obtain an HRMS image with good spectral and spatial fidelity to the source images. Therefore, it is very important to establish the precise relationship between the HRMS image and the source images for the construction of the model. According to [35], the low-resolution version of the HRMS image (L-HRMS) filtered

by a Gaussian filter can be considered as the UPMS image. Vivone *et al.* [36] argued that the Gaussian filter should match the modulation transfer function (MTF), which is the frequency domain form of the sensor point spread function. Therefore, in order to ensure the similarity between the L-HRMS image and UPMS image, we design a spectral fidelity prior term for the HRMS and UPMS images. First, an MTF-matched Gaussian filter  $H_b$  is applied to each band of the estimated HRMS image  $\hat{M}_b$  to achieve the L-HRMS, and then we make L-HRMS close to the UPMS image  $\tilde{M}_b$  to ensure spectral consistency between the HRMS image and MS image. The spectral fidelity prior term can be defined as follows.

$$\min_{\hat{M}_b} \frac{1}{2} \|H_b \hat{M}_b - \tilde{M}_b\|^2 \quad (5)$$

In addition to the spectral consistency, the HRMS image should also have spatial structure consistency with the PAN image. Previous pansharpening methods usually assumed that a PAN image can be considered as a linear combination of the multiple bands of the HRMS image, as shown in the following equation [2]:

$$P = \sum_{b=1}^B \hat{\alpha}_b \hat{M}_b \quad (6)$$

where  $\hat{\alpha}_b$  is the coefficient of  $b$ th-band of the HRMS image.  $\hat{\alpha}_b$  is unknown and usually set as an empirical value, which leads to uncertainty in the expression of the linear model. In

addition, the linear combination of the bands of an MS image cannot accurately express the spatial information in a PAN image [37]. Therefore, Eq. (6) cannot express the spatial differences among the bands of an MS image. In fact, the spatial information of different bands is widely diverse, which causes the correlation between each band of the HRMS image and PAN image to be very different. However, the geometric structure distribution of images in the same scene should be similar. Since the geometric structure can be expressed by the gradient domain [17], and the Laplacian operator can obtain sharpened boundaries and lines in any direction, we select the Laplacian operator to get the gradient information. To explicitly illustrate this, we take a group of source images from the IKONOS dataset as an example and analyze the spatial and structural relationship between PAN and HRMS images. Figs. 2 (a)–(e) show the PAN image and the different bands of HRMS image, Figs. 2 (f)–(i) show the corresponding gradient maps of (a)–(e), and the histograms of (a)–(j) are shown in Figs. 2 (k)–(t). We can see from the figure that the spatial information of the different bands of the HRMS image varies greatly. However, compared with the gradient histogram of the PAN image, all gradient histograms of the HRMS bands have almost the same shape as those of the PAN image, i.e., the gradient values of the four HRMS bands are proportional to those of the PAN image. Therefore, based on the above analysis, we construct a priori term about the spatial structure similarity between the PAN and HRMS images. Denoting the Laplacian operator as  $\mathbf{L}$ , the  $b$ th-band gradient map of the HRMS image is expressed as  $\widehat{\mathbf{LM}}_b$ , and the gradient map of the PAN image is expressed as  $\mathbf{LP}$ . With a correction coefficient, the two gradient maps should be close to each other. We thus define a band-adaptive gradient correction term to retain more accurate structure information as follows:

$$\min_{\omega_b} \frac{u}{2} \left\| \omega_b \widehat{\mathbf{LM}}_b - \mathbf{LP} \right\|^2 \quad (7)$$

where  $u$  is penalty parameter, and  $\omega_b$  is a correction coefficient.

In order to determine the band-adaptive correction coefficient  $\omega_b$  in Eq. (7), an HRMS image  $\widehat{\mathbf{M}}_b$  is required; however, such an image is unknown in real experiments. Therefore, the original PAN image is down-sampled to the same scale as the original MS image, and the original MS image is taken as a reference HRMS image. Then,  $\omega_b$  can be obtained by the constructed regression model as follows:

$$\min_{\omega_b} \frac{1}{2} \left\| \mathbf{LP}^d - \omega_b \mathbf{LM}_b \right\|^2 \quad s.t. \omega_b \geq 0. \quad (8)$$

where  $\mathbf{P}^d$  is the down-sampled PAN image, and  $\mathbf{M}_b$  is the  $b$ th-band of the original MS image. With the Lagrange multiplier method,  $\omega_b$  can be obtained by minimizing the energy function as follows:

$$\omega_b = \arg \min_{\omega_b} \left( \frac{1}{2} \left\| \mathbf{LP}^d - \omega_b \mathbf{LM}_b \right\|^2 + \frac{\delta}{2} \left\| \max(0, -\omega_b) \right\|^2 \right) \quad (9)$$

where  $\delta$  is a penalty parameter. The second term is a non-negative constraint. Through employing the gradient descent method, the discretized partial differential equation can be defined as

$$\frac{\omega_b^{t+1} - \omega_b^t}{\Delta t} = \mathbf{M}_b^T \mathbf{LP}^d - \omega_b^t \mathbf{M}_b^T \mathbf{LM}_b + \delta \max(0, -\omega_b^t) \quad (10)$$

where  $\Delta t$  is the step size of the gradient descent method. The  $t+1$  step of  $\omega_b$  can be obtained by

$$\omega_b^{t+1} = \omega_b^t + \Delta t (\mathbf{M}_b^T \mathbf{LP}^d - \omega_b^t \mathbf{M}_b^T \mathbf{LM}_b + \delta \max(0, -\omega_b^t)) \quad (11)$$

The final correction coefficient  $\omega_b$  can be obtained when the iterations meet the stop condition. The coefficient is then transferred to Eq. (7), and the spatial correction model is constructed. By combining equations (5) and (7), the spectral fidelity and structure correction model is defined as follows:

$$\min_{\widehat{\mathbf{M}}_b} \frac{1}{2} \left\| \mathbf{H}_b \widehat{\mathbf{M}}_b - \widetilde{\mathbf{M}}_b \right\|^2 + \frac{u}{2} \left\| \omega_b \widehat{\mathbf{LM}}_b - \mathbf{LP} \right\|^2 \quad (12)$$

### B. Detail Correction Constraint

Model (12) aims to keep the spectral fidelity of the UPMS image and the spatial fidelity of the PAN image. However, owing to the lack of a reference image, maintaining the structural consistency of the fused image with PAN image may still result in excessive or insufficient detail injection, which will affect both spectral and spatial quality. To ensure the accuracy of the pansharpening model, it is necessary to add a detail constraint to obtain proper details. According to formula (1), we can have

$$\widehat{\mathbf{M}}_b - \widetilde{\mathbf{M}}_b = g_b (\mathbf{P} - \mathbf{P}_b^L) \quad (13)$$

The left side of equation (13) represents the details injected into the UPMS image, and the right side of the equation represents the details extracted from the PAN image. To ensure the consistency of these two kinds of details in each band, we design a band-adaptive detail correction constraint based on formula (13):

$$\min_{\widehat{\mathbf{M}}_b} \frac{\lambda}{2} \left\| \widehat{\mathbf{M}}_b - \widetilde{\mathbf{M}}_b - g_b (\mathbf{P} - \mathbf{P}_b^L) \right\|^2 \quad (14)$$

where  $\lambda$  is a penalty parameter. In (14),  $\mathbf{P}_b^L$  plays an important role in determining the  $b$ th-band details, and  $g_b$  is a weight factor used to control the degree of detail injection. Both  $\mathbf{P}_b^L$  and  $g_b$  are still uncertain, and they are independent. Thus, to determine  $\mathbf{P}_b^L$  and  $g_b$ , we first find a solution for  $\mathbf{P}_b^L$ , and then one for  $g_b$ . As mentioned in Eqs. (2) and (3),  $\mathbf{P}_b^L$  can be obtained by CS- or MRA-based methods. In this study, a new approach is proposed by combining both CS- and MRA-based methods to take advantage of their respective merits.

Vivone *et al.* [35] argue that the blur function for remote sensing should have a Gaussian-like shape. Thus, a Gaussian filter  $\mathbf{H}_G$  is defined as the low-pass filter used in the MRA-based method.  $\mathbf{P}_b^L$  can be obtained by designing the following linear model:

$$\mathbf{P}_b^L = \beta_{b,1} \mathbf{I}_{UP} + \beta_{b,2} \mathbf{H}_G \mathbf{P} \quad s.t. \beta_{b,1} \geq 0, \beta_{b,2} \geq 0. \quad (15)$$

where  $\beta_{b,1}$  and  $\beta_{b,2}$  are band-adaptive weight coefficients.

Once  $\beta_{b,1}$ ,  $\beta_{b,2}$ , and  $\mathbf{H}_G$  are determined,  $\mathbf{P}_b^L$  can be obtained.  $\mathbf{H}_G$  is a distribution with  $N(0, \sigma^2)$ , and thus the only parameter that needs to be determined is the standard deviation  $\sigma$ , which can be estimated by building the relationship between  $\mathbf{I}_{UP}$  and  $\mathbf{P}$ .

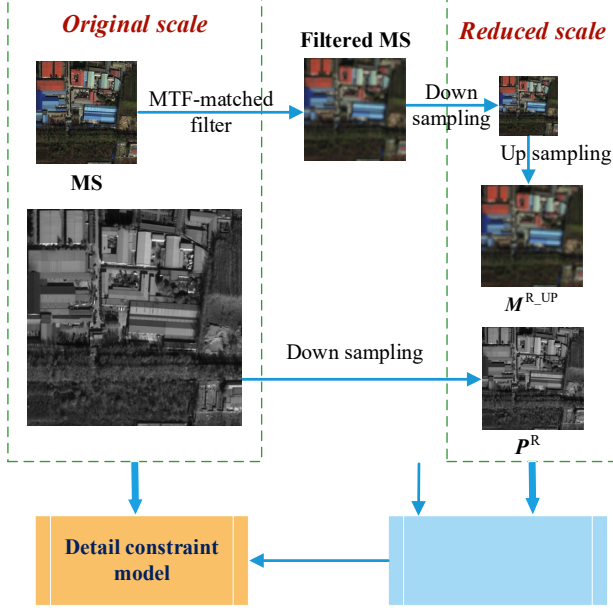


Fig.3 Flowchart of constructing detail constraint model based on the parameter transfer. The MS image is shown in RGB bands for better visual effect.

Denoting  $\hat{\mathbf{I}}$  as the I component of the required  $\hat{\mathbf{M}}$  image, the blurred version of  $\hat{\mathbf{I}}$  after applying  $\mathbf{H}_G$  on  $\hat{\mathbf{I}}$  should be close to  $\mathbf{I}_{UP}$ . Because  $\hat{\mathbf{I}}$  contains similar spatial information as the PAN image [36], [38], we use  $\mathbf{P}$  instead of  $\hat{\mathbf{I}}$  to establish a correlation function between  $\mathbf{P}$  and  $\mathbf{I}_{UP}$ . Thus,  $\mathbf{H}_G$  can be obtained by searching the optimal  $\sigma$  that makes  $\mathbf{H}_G \mathbf{P}$  highly correlated with  $\mathbf{I}_{UP}$ , i.e.,

$$\hat{\sigma} = \max_{\sigma} \text{corr}(\mathbf{H}_{G \sim N(0, \sigma)} \mathbf{P}, \mathbf{I}_{UP}) \quad (16)$$

where  $\hat{\sigma}$  represents the estimated  $\sigma$ , and  $\text{corr}(\cdot)$  is the correlation function. By setting an initial value of  $\sigma$ , and then iteratively computing the correlation according to Eq. (16) as  $\sigma$  changes, the optimal  $\sigma$  as well as  $\mathbf{H}_G$  can be obtained.

As for  $\beta_{b,1}$  and  $\beta_{b,2}$ , they are critical for obtaining an appropriate  $\mathbf{P}_b^L$ . Due to the lack of an HRMS image  $\hat{\mathbf{M}}_b$ , a reduced-scale parameter acquisition model is first designed to calculate  $\beta_{b,1}$  and  $\beta_{b,2}$ . Then, the obtained parameters are transferred back into (15) to obtain  $\mathbf{P}_b^L$ , which is then substituted into (14) for constructing the detail correction constraint. The procedure of parameter acquisition is shown in Fig.3. To obtain the accurate parameters as well as the details with high efficiency, the original images are degraded into the reduced scale. In the degradation process, the MS image is first filtered by an MTF-matched filter, then down-sampled to a reduced scale, and finally up-sampled to the original scale to construct an image  $\mathbf{M}^{R-UP}$ ; this follows the Wald protocol [39]. The original MS image is regarded as a reference HRMS image. Meanwhile, the PAN image is also degraded and down-sampled to the same scale as the MS image. Through using the reference image and the reduced-scale images, the

---

**Algorithm 1** The overall process of proposed method

---

**Input:** PAN image  $\mathbf{P}$ , I component of UPMS ( $\mathbf{I}_{UP}$ )

**Initialize:**  $\tau=1$ ,  $\mathbf{A}^0=\mathbf{1}$ ,  $t=0$ ,  $\mathbf{X}^0=0$ ,  $\rho=1.01$ .

$\mathbf{H}_G$  is obtained by formula (16),

**For**  $b = 1:B$

$\omega_b$  is obtained by formula (11),

$\beta_{b,1}$  and  $\beta_{b,2}$  are obtained by formula (20)

$\mathbf{P}_b^L$  is obtained by formula (15),

$\mathbf{g}_b$  is obtained by formula (21)

**While do:**

Update  $\hat{\mathbf{M}}_b^{t+1}$  by formula (24),

Update  $\mathbf{X}^{t+1}$  by formula (27),

Update  $\mathbf{A}^{t+1}$  by formula (30),

$\tau = \rho\tau$ ,  $t=t+1$

**Until:**  $\frac{\|\hat{\mathbf{M}}_b^{t+1} - \hat{\mathbf{M}}_b^t\|}{\|\hat{\mathbf{M}}_b^t\|} < \zeta$

**End For**

**Output:** estimated HRMS image  $\hat{\mathbf{M}}$

---

parameters can be obtained with the parameter acquisition model. Initializing  $\mathbf{g}_b$  to 1, the parameter acquisition model is designed according to Eq. (14) as follows:

$$\min_{\mathbf{P}_b^{R-L}} \frac{1}{2} \|\mathbf{M}_b - \mathbf{M}_b^{R-UP} - \mathbf{P}^R + \mathbf{P}_b^{R-L}\|^2 \quad (17)$$

where the subscript  $b$  represents the  $b$ th-band,  $\mathbf{P}^R$  represents a reduced-scale PAN image, and  $\mathbf{P}_b^{R-L}$  represents the reduced-scale  $\mathbf{P}_b^L$ . Similar to formula (15),  $\mathbf{P}_b^{R-L}$  can be designed as follows:

$$\mathbf{P}_b^{R-L} = \beta_{b,1} \mathbf{I}^{R-UP} + \beta_{b,2} \mathbf{H}_G^R \mathbf{P}^R \quad \text{s.t. } \beta_{b,1} \geq 0, \beta_{b,2} \geq 0. \quad (18)$$

where  $\mathbf{I}^{R-UP}$  represents the I component of  $\mathbf{M}^{R-UP}$ .  $\mathbf{H}_G^R$  represents a Gaussian filter in reduced scale, and can be acquired by replacing  $\mathbf{I}_{UP}$  and  $\mathbf{P}$  with  $\mathbf{I}^{R-UP}$  and  $\mathbf{P}^R$ , respectively in Eq. (16). The variables in Eqs. (17) and (18) can be defined as follows:

$$\begin{cases} \mathbf{Z} = \mathbf{M}_b - \mathbf{M}_b^{R-UP} - \mathbf{P}^R \\ \boldsymbol{\beta}_b = [\beta_{b,1}, \beta_{b,2}] \\ \mathbf{Q} = [\mathbf{I}^{R-UP}, \mathbf{H}_G^R \mathbf{P}^R]^T \end{cases} \quad (19)$$

By substituting Eqs. (18) and (19) into Eq. (17), Eq. (17) can be simplified as follows:

$$\min_{\boldsymbol{\beta}_b} \frac{1}{2} \|\mathbf{Z} + \boldsymbol{\beta}_b \mathbf{Q}\|^2 \quad \text{s.t. } \boldsymbol{\beta}_b \geq 0. \quad (20)$$

$\boldsymbol{\beta}_b$  can be acquired by a gradient decent method, and the solution of (20) is similar to that of (8).

After acquiring  $\boldsymbol{\beta}_b$ ,  $\mathbf{P}_b^{R-L}$  can be acquired by formula (17).

After denoting  $\mathbf{D}_M = \mathbf{M}_b - \mathbf{M}_b^{R-UP}$  and  $\mathbf{D}_P = \mathbf{P}^R - \mathbf{P}_b^{R-L}$ ,  $\mathbf{g}_b$  can be obtained by:

$$\min_{\mathbf{g}_b} \frac{1}{2} \|\mathbf{D}_M - \mathbf{g}_b \mathbf{D}_P\|^2 \quad \text{s.t. } \mathbf{g}_b \geq 0. \quad (21)$$

The solution of formula (21) is also similar to that of formula (8), and  $g_b$  can be acquired by solving (21). Upon transferring the band-adaptive parameters  $\beta_b$  and  $g_b$  to (15) and (14), we can obtain the final detail correction constraint.

### C. Unified Model with Sparse Constraint

In order to reduce the number of artifacts and keep the sparsity in the gradient domain, a sparse constraint is designed based on a total variational (TV) prior. Together with formulas (12) and (14), the energy function can be defined as follows:

$$E_b = \frac{1}{2} \left\| \mathbf{H}_b \widehat{\mathbf{M}}_b - \widetilde{\mathbf{M}}_b \right\|^2 + \frac{u}{2} \left\| \omega_b \mathbf{L} \widehat{\mathbf{M}}_b - \mathbf{L} \mathbf{P} \right\|^2 + \frac{\lambda}{2} \left\| \widehat{\mathbf{M}}_b - \widetilde{\mathbf{M}}_b - g_b (\mathbf{P} - \mathbf{P}_b^L) \right\|^2 + \gamma \left\| \mathbf{L} \widehat{\mathbf{M}}_b \right\| \quad (22)$$

where  $\gamma$  is a penalty parameter. The first term is the spectral fidelity constraint, the second term is the structure correction constraint, the third term is the detail correction constraint, and the last term is the sparse constraint. Then,  $\widehat{\mathbf{M}}_b$  can be obtained by minimizing the energy function as follows:

$$\widehat{\mathbf{M}}_b = \arg \min_{\mathbf{M}_b} E_b \quad (23)$$

### D. The solution

To solve the model (22), we replace  $\mathbf{L} \widehat{\mathbf{M}}_b$  with  $\mathbf{X}$ , and apply the augmented Lagrangian function as:

$$E_b = \frac{1}{2} \left\| \mathbf{H}_b \widehat{\mathbf{M}}_b - \widetilde{\mathbf{M}}_b \right\|^2 + \frac{u}{2} \left\| \omega_b \mathbf{L} \widehat{\mathbf{M}}_b - \mathbf{L} \mathbf{P} \right\|^2 + \frac{\lambda}{2} \left\| \widehat{\mathbf{M}}_b - \widetilde{\mathbf{M}}_b - g_b (\mathbf{P} - \mathbf{P}_b^L) \right\|^2 + \gamma \left\| \mathbf{X} \right\| + \left\langle \mathbf{A}, \mathbf{X} - \mathbf{L} \widehat{\mathbf{M}}_b \right\rangle + \frac{\delta}{2} \left\| \mathbf{X} - \mathbf{L} \widehat{\mathbf{M}}_b \right\|^2 \quad (24)$$

where  $\mathbf{A}$  is a Lagrangian multiplier, and  $\delta$  is a penalty parameter. The augmented Lagrangian function can be solved by the ADMM method.

#### 1) Updating $\widehat{\mathbf{M}}_b$

By fixing  $\mathbf{X}$  and  $\mathbf{A}$ , we can proceed with updating  $\widehat{\mathbf{M}}_b$  through partial deviation as  $\partial E / \partial \widehat{\mathbf{M}}_b = 0$ . Then, we have

$$\begin{aligned} \partial E / \partial \widehat{\mathbf{M}}_b &= 0 \\ \Rightarrow (\mathbf{H}_b^T \mathbf{H}_b + u \omega_b^2 \mathbf{L}^T \mathbf{L} + \lambda + \delta \mathbf{L}^T \mathbf{L}) \widehat{\mathbf{M}}_b \\ &= \mathbf{H}_b^T \widetilde{\mathbf{M}}_b + u \omega_b \mathbf{L}^T \mathbf{L} \mathbf{P} + \lambda \widetilde{\mathbf{M}}_b + \lambda g_b (\mathbf{P} - \mathbf{P}_b^L) + \mathbf{L}^T \mathbf{A} + \delta \mathbf{L}^T \mathbf{X} \end{aligned} \quad (25)$$

Through applying the fast Fourier transform (FFT, denoted as  $\mathcal{F}(\cdot)$ ) and inverse FFT (IFFT, denoted as  $\mathcal{F}^{-1}(\cdot)$ ) to Eq.

(25), the  $t+1$  step of  $\widehat{\mathbf{M}}_b$  ( $\widehat{\mathbf{M}}_b^{t+1}$ ) is obtained as shown in Eq. (26).

#### 2) Updating $\mathbf{X}$

After fixing the obtained  $\widehat{\mathbf{M}}_b^{t+1}$  and  $\mathbf{A}$ , the  $t+1$  step of  $\mathbf{X}$

( $\mathbf{X}^{t+1}$ ) can be updated by

$$\mathbf{X}^{t+1} = \arg \min_{\mathbf{X}} \frac{\gamma}{\delta} \left\| \mathbf{X} \right\| + \frac{1}{2} \left\| \mathbf{X} - (\mathbf{L} \widehat{\mathbf{M}}_b^{t+1} - \frac{1}{\delta} \mathbf{A}) \right\|^2 \quad (27)$$

Formula (27) can be efficiently solved by the soft thresholding strategy as follows:

$$\mathbf{X}^{t+1} = \text{shrink}(\mathbf{L} \widehat{\mathbf{M}}_b^{t+1} - \frac{1}{\delta} \mathbf{A}, \frac{\gamma}{\delta}) \quad (28)$$

where

$$\text{shrink}(x, y) = \text{sgn}(x) \cdot \max(|x| - y, 0) \quad (29)$$

#### 3) Updating $\mathbf{A}$

By fixing  $\widehat{\mathbf{M}}_b^{t+1}$  and  $\mathbf{X}^{t+1}$ , the updating of  $\mathbf{A}$  at the  $t+1$  step ( $\mathbf{A}^{t+1}$ ) can be obtained by the gradient decent method:

$$\mathbf{A}^{t+1} = \mathbf{A}^t + \tau^{t+1} (\mathbf{X}^{t+1} - \mathbf{L} \widehat{\mathbf{M}}_b^{t+1}) \quad (30)$$

where  $\tau$  is the step length, and we set  $\tau^{t+1} = \rho \tau^t$ ,  $\rho > 1$  to accelerate the iteration process.

The proposed method is summarized in Algorithm 1. By initializing the parameters and the filter obtained by Algorithm 1,  $\widehat{\mathbf{M}}_b$  can be computed iteratively until the stop condition is met.

## IV. EXPERIMENTAL RESULTS

In this section, a series of experiments are conducted to verify the effectiveness of the proposed method. First, the experimental datasets and the comparison methods are detailed in Section IV-A. Then, the setting of the optimal parameters is presented in section IV-B. An ablation study is conducted in Section IV-C to testify the effectiveness of the proposed constraints. Furthermore, a convergence analysis is carried out to show the robustness of the proposed method in Section IV-D. Finally, comparison experiments on reduced-scale and full-scale images are implemented in sections IV-E and IV-F, respectively, and an efficiency analysis is conducted in Section IV-G.

### A. Experiment Setting

Three datasets, IKONOS, Pléiades, and WorldView-3, are collected from <http://www.kosmos-images.com>. The MS images of IKONOS and Pléiades datasets contain 4 bands: red (R), green (G), blue (B), and near infrared (NIR), while the MS images of WorldView-3 dataset contain 8 bands including R, G, B, near-infrared1, near-infrared2, coastal blue, yellow, and red edge. Besides, TianGong-2 dataset which contains the MS image with 14 visible NIR bands is collected from <http://www.msadc.cn/sjfwgth/> to further testify the performance of the proposed method.

Two kinds of experiments are conducted. The first one is a simulated experiment. In this experiment, the MS and PAN images are degraded and down-sampled by a factor of 4, and the original MS image is treated as a reference image; this follows the Wald protocol [39]. This simulated experiment is also called a reduced-scale experiment. The sizes of MS image

---


$$\widehat{\mathbf{M}}_b^{t+1} = \mathcal{F}^{-1} \left( \frac{\mathcal{F}(\mathbf{H}_b^T \widetilde{\mathbf{M}}_b + u \omega_b \mathbf{L}^T \mathbf{L} \mathbf{P} + \lambda \widetilde{\mathbf{M}}_b + \lambda g_b (\mathbf{P} - \mathbf{P}_b^L) + \mathbf{L}^T \mathbf{A} + \delta \mathbf{L}^T \mathbf{X})}{\mathcal{F}(\mathbf{H}_b^T \mathbf{H}_b + \mu \mathbf{L}^T \mathbf{L} + \delta \mathbf{L}^T \mathbf{L})} \right) \quad (26)$$



TABLE I  
ILLUSTRATION OF THE FUSION QUALITY METRICS

	FULL NAME	DESCRIPTION	RS/FS	OPTIMAL
UIQI↑[40]	Universal Image Quality Index	Representing a global fusion quality	RS	1
SAM↓[41]	Spectral Angle Mapper	Measuring the spectral distortion	RS	0
ERGAS↓[36]	Erreur Relative Global Adimensionnelle De Synthèse	Evaluating the spatial and spectral quality	RS	0
SCC↑[42]	Spatial Correlation Coefficient	Indicating a spatial quality	RS	1
Q4/Q8↑[43]	Q4 for a 4-band image, Q8 for an 8-band image	A vector extension of the Q-index to evaluating the global quality	RS	1
RMSE↓[17]	Root Mean Squared Error	Indicating the errors between fused image and reference image	RS	0
D <sub>λ</sub> ↓[44]	-	Indicating spectral distortion	FS	0
D <sub>s</sub> ↓[44]	-	Indicating spatial distortion	FS	0
QNR↑[44]	Quality with No Reference	Comprehensive metric combining $D_s$ and $D_r$	FS	1

↑ indicates the optimal value is 1, ↓ indicates the optimal value is 0. RS represents reduced scale, and FS represents full scale.

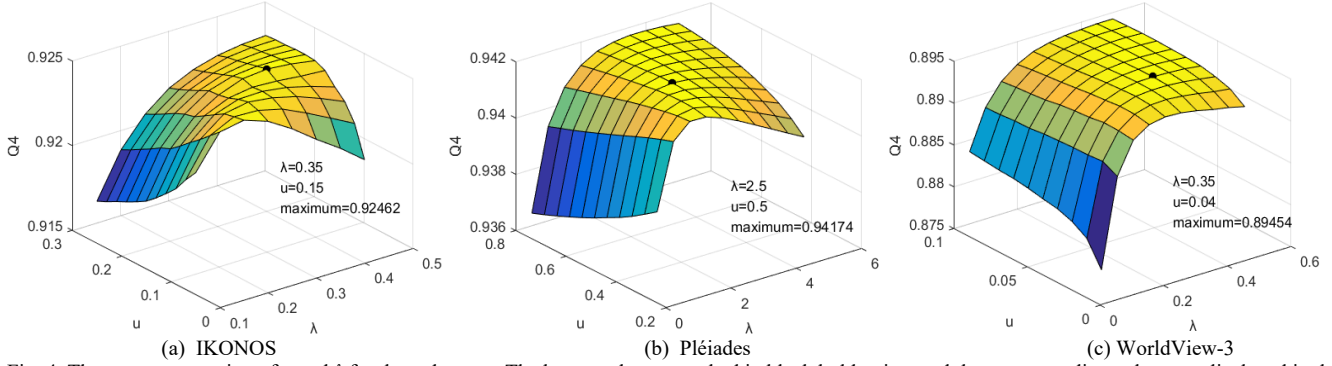


Fig. 4. The parameter setting of  $u$  and  $\lambda$  for three datasets. The best results are marked in black bold points and the corresponding values are displayed in the down-right corner. The maximum value means the best value of Q4.

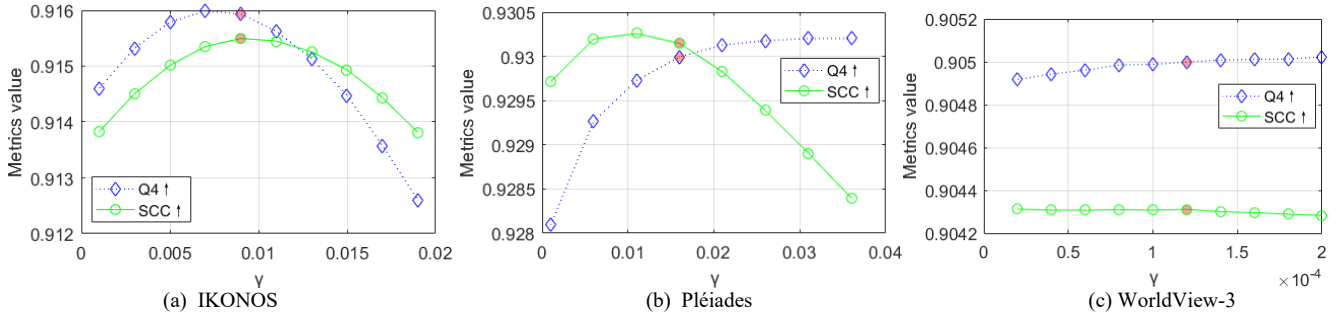


Fig. 5. The parameter setting of  $\gamma$  for three datasets. The best results are marked in red points.

are  $64 \times 64 \times 4$  in IKONOS and Pléiades datasets,  $64 \times 64 \times 8$  in WorldView-3 dataset, and  $32 \times 32 \times 14$  in TianGong-2 dataset. The sizes of PAN images are  $256 \times 256$  in IKONOS, Pléiades, and WorldView-3 datasets, and  $128 \times 128$  in TianGong-2 dataset. The second experiment is a real experiment, and is also called a full-scale experiment; in this experiment, the sizes of MS and PAN images are 4 times larger than those in the reduced scale experiment. Six widely used quality metrics are employed to evaluate the reduced scale results, and three full-scale metrics are employed to evaluate the real image fusion results. The detailed information is shown in Table I.

Some traditional and state-of-the-art pansharpening methods are introduced as comparative methods in these experiments. For example, the traditional methods include CS-based approach GSA [8], and MRA-based approach BFLP [15]. The model-based methods include MMT [23], FSRIC [24], fusion via compressive sensing (FCS) [22], and LRFF [27]. Furthermore, fusion with deep convolutional neural networks (Fusion-Net) [31] which belongs to DL-based method is included for comparison. Since both the model-based and DL-based methods learn spectral and spatial information using

energy functions, we consider them as the learning-based methods. Besides, an up-sampling method with a polynomial kernel, also known as EXP [45], is used as a benchmark rather than a comparative pansharpening method. All the codes of the comparison methods are provided by the authors, and all the parameters are set by their default values of the original papers.

### B. Parameters Setting

Three penalty parameters ( $u$ ,  $\lambda$ , and  $\gamma$  in formula (22)) need to be determined.  $u$  represents the weight of spatial sharpness, and  $\lambda$  represents the importance of accuracy of detail injection. These two parameters play critical roles in the overall spatial and spectral quality, and interact with each other. They can be determined together by the grid search method.  $\gamma$  is a penalty parameter on the sparse regular term, and is relatively independent of the other two parameters. It is thus determined after fixing the parameters  $u$  and  $\lambda$ .

In order to determine the values of  $u$  and  $\lambda$ , the Q4 metric is employed since it considers both radiometric and spectral distortions and can represent the overall fusion quality [26]. Ten remote sensing image pairs are randomly collected from

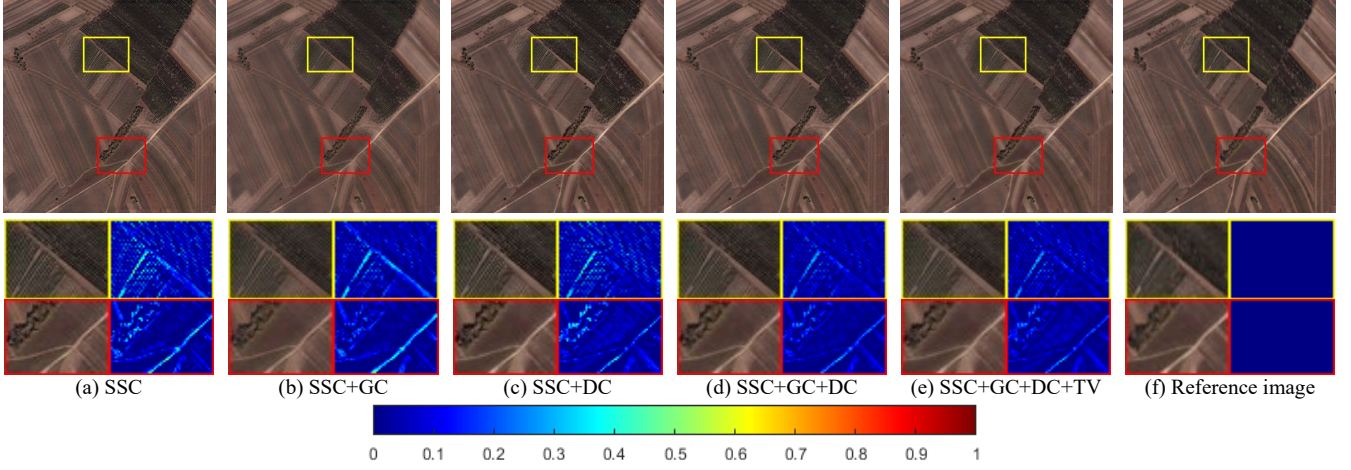


Fig. 6. Fusion results and AEMs of an image pair by different ablation models.

TABLE II  
OBJECTIVE EVALUATION OF DIFFERENT COMBINATIONS OF THE MODELS IN THE ABLATION STUDY

Models	Spectral/spatial constraint	Gradient correction	Detail constraint	Total variation	UIQI↑	SAM↓	ERGAS↓	SCC↑	Q4↑	RMSE↓
SSC	✓	×	×	×	0.8861	<b>3.9706</b>	3.2691	0.8873	0.8850	0.0434
SSC+GC	✓	✓	×	×	0.8977	<u>4.0801</u>	3.0342	0.8985	0.8984	0.0381
SSC+DC	✓	×	✓	×	0.8946	4.4010	3.2819	0.9011	0.8939	0.0414
SSC+GC+DC	✓	✓	✓	×	<u>0.9159</u>	4.3330	<u>2.8044</u>	0.9124	<u>0.9169</u>	0.0380
SSC+GC+DC+TV	✓	✓	✓	✓	<b>0.9185</b>	4.3013	<b>2.7496</b>	<b>0.9141</b>	<b>0.9202</b>	<b>0.0369</b>

each dataset; an image pair includes an MS image and a PAN image. The average Q4 value is obtained from ten fusion results on each dataset. The results are shown in Fig. 4. The points with the highest Q4 values represent the optimal values of  $u$  and  $\lambda$ , which are displayed in the lower right corner of Fig. 4.

To determine the values of  $\gamma$  for each dataset, the SCC metric considering the spatial fidelity is added to find an optimal parameter value, because the sparse constraint may affect the spatial quality. The obtained optimal  $\gamma$  values are shown in Fig. 5. We can infer that the best  $\gamma$  for the IKONOS dataset is 0.009, that for Pléiades is 0.015, and that for WorldView-3 is  $1.2 \times 10^{-4}$ .

### C. Ablation Study

To testify the effectiveness of the band-adaptive gradient correction constraint, detail correction constraint, and the total variation constraint in the proposed method, an ablation study is conducted on the three datasets. Fifteen image pairs are collected from the three datasets. Taking an image pair from the IKONOS dataset as an example, the fused results obtained by different ablation models are shown in Fig. 6. The average objective evaluation is shown in Table II with best ones shown in bold and the second-best ones underlined. Here, SSC is the original spectral and spatial fidelity constraints, GC is the gradient correction as in Eq. (8), DC is the detail correction as in Eq. (14), and TV is the total variation.

In Fig. 6, the upper half part displays the fusion results, and the bottom half part displays the enlarged yellow and red rectangles as well as the corresponding absolute error maps (AEMs) [31]. The less residues in AEMs, the better the fusion results. From the figure, we can see that the result of SSC obviously contains the most residues. With the addition of GC, the residues of SSC+GC are reduced. However, there are still some obvious residues left in the road areas. Replacing GC with

DC, the residues of SSC+DC in road areas are much less than those of SSC+GC. However, the residues of SSC+DC in the vegetation areas are increased. Taking advantages of GC and DC, the result of SSC+GC+DC contains much less residues both in road and vegetation areas. With the additional TV prior, the result of the final model is improved slightly compared with that of SSC+GC+DC.

From Table II, we can see that the objective evaluation is consistent with the subjective evaluation. The proposed model and SSC+GC+DC model obviously obtain better results than the model without DC or GC. Note that compared with other models, the SSC and SSC+GC models seem to have more satisfactory SAM values, which indicates less spectral distortion. This is because the detail injection will inevitably cause the spectral distortion. However, aside from the SAM metric, other metric values of SSC and SSC+GC are much worse than those of the proposed model. Therefore, the ablation study verifies the effectiveness of the proposed method.

### D. Convergence Analysis

In order to verify the robustness of the proposed model, a convergence analysis of the solution is conducted. The energy functions for updating  $\hat{\mathbf{M}}_b$ ,  $\mathbf{X}$ , and  $\mathbf{A}$  are convex with lower bounds, and the solutions for the updating steps (Eqs. (24), (27), and (30)) are exact. Therefore, the respective energy functions decrease in these steps. In addition, model (22) is proper and closed, and thus will converge as iteration progresses. To further verify the convergence of the proposed model, a series of remote sensing images from the three datasets are randomly collected and processed. We employ two indicators to show the convergence. The first indicator is the relative error between iteration steps of  $\hat{\mathbf{M}}_b$  [28], which is defined as:



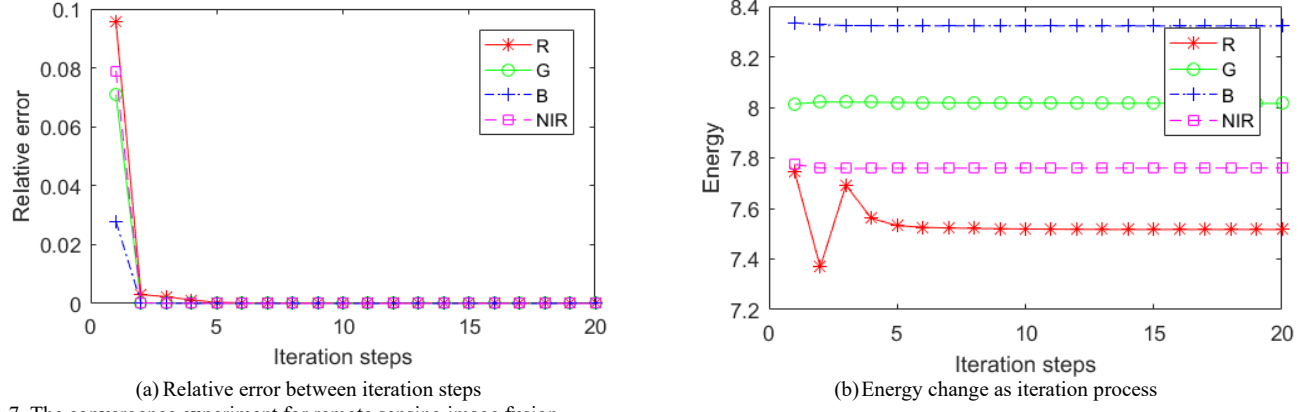


Fig. 7. The convergence experiment for remote sensing image fusion.

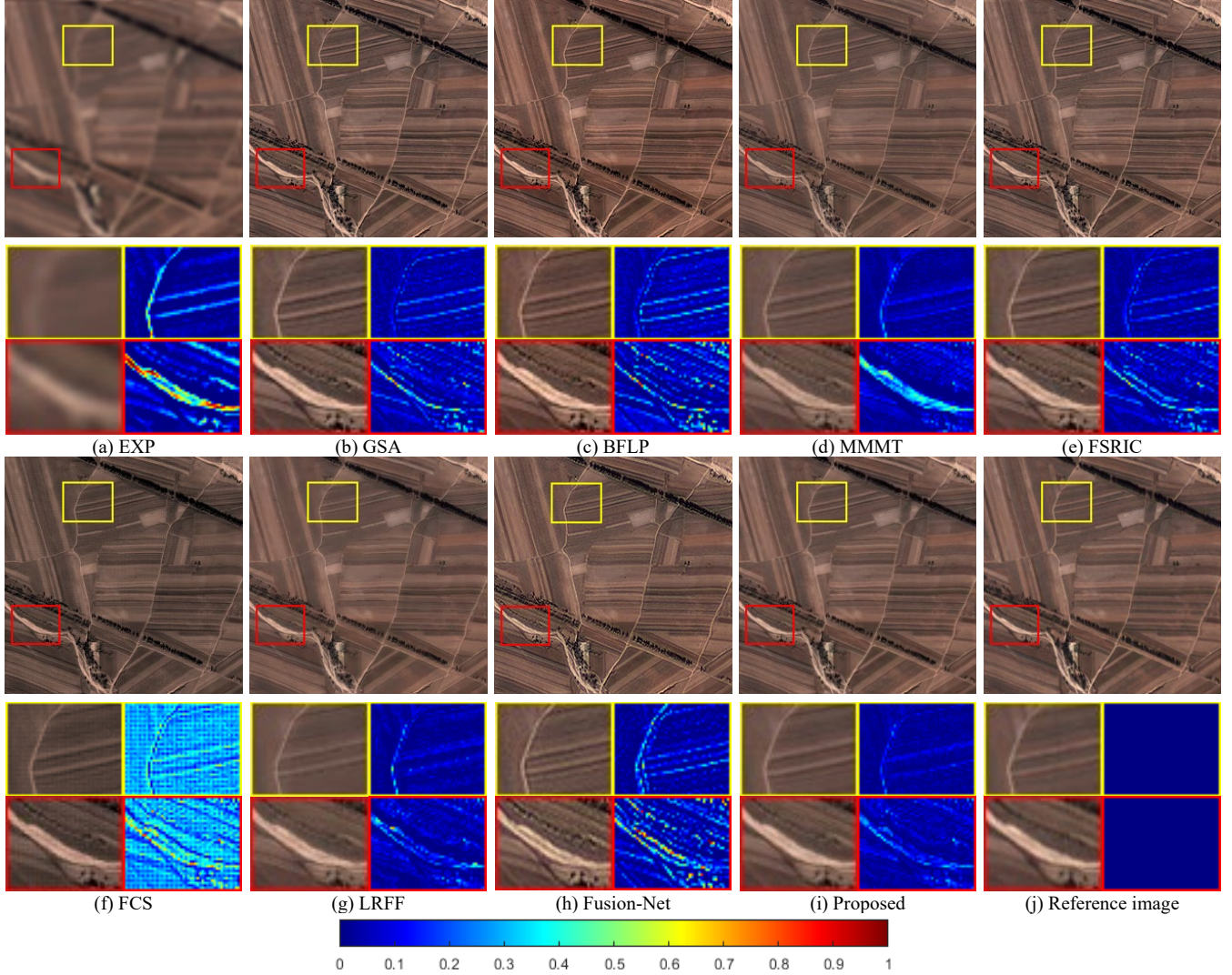


Fig. 8. Fusion results and the corresponding AEMs of the reduced scale images from IKONOS dataset.

$$\frac{\|\widehat{\mathbf{M}}_b^{t+1} - \widehat{\mathbf{M}}_b^t\|}{\|\widehat{\mathbf{M}}_b^t\|} \quad (31)$$

Taking an image from the Pléiades dataset as an example, the convergence results of R, G, B, and NIR bands are shown in Fig. 7 (a). The second indicator is the energy function in formula

(22), and the corresponding convergence results are shown in Fig. 7 (b). We can see from the figure that the proposed model converges quickly, in less than 5 steps for every band. The relative error of each band almost converges to 0 as iteration progresses. From Fig. 7 (b), we can see that the energy function quickly converges to an optimal value for each band, especially for G, B, and NIR bands. The energy of the R band fluctuates

TABLE III  
QUANTITATIVE EVALUATION OF FUSION RESULTS IN FIG. 8, AND AVERAGE QUANTITATIVE EVALUATION ON IKONOS DATASET

METHOD	RESULTS IN FIG. 8						AVERAGE					
	UIQI↑	SAM↓	ERGAS↓	SCC↑	Q4↑	RMSE↓	UIQI↑	SAM↓	ERGAS↓	SCC↑	Q4↑	RMSE↓
EXP [45]	0.7213	3.0146	3.2290	0.6384	0.7295	13.3062	0.7105	3.9215	4.3046	0.6195	0.7091	13.8195
GSA [8]	0.8838	3.9962	2.6410	0.8610	0.8844	10.0041	0.8682	5.2195	3.4350	0.8581	0.8691	10.2998
BFLP [15]	0.8666	3.8683	3.1696	0.8733	0.8583	12.6888	0.8146	6.3128	7.3887	0.8667	0.7062	23.4323
MMMT [23]	0.8766	3.6946	2.3809	0.8658	0.8855	9.5130	0.8624	4.9050	3.1512	0.8521	0.8710	9.8830
FSRIC [24]	0.8890	3.8782	2.4575	0.8604	0.8929	9.5313	0.8700	5.0983	3.3101	0.8557	0.8730	10.0763
FCS [22]	0.8779	5.7793	5.0986	0.8956	0.7412	22.067	0.8561	6.7116	5.4626	0.8836	0.6927	18.2609
LRFF [27]	0.9129	3.1211	1.9471	0.9005	0.9140	7.7365	0.8971	4.0399	2.7120	0.8933	0.8975	8.5297
Fusion-Net [31]	0.8720	3.6084	2.8416	0.8536	0.8620	10.9425	0.8388	4.9502	3.9453	0.8369	0.8296	11.6590
Proposed	<b>0.9346</b>	<b>2.8672</b>	<b>1.7987</b>	<b>0.9199</b>	<b>0.9330</b>	<b>7.1224</b>	<b>0.9149</b>	<b>3.8086</b>	<b>2.4658</b>	<b>0.9089</b>	<b>0.9151</b>	<b>7.8549</b>

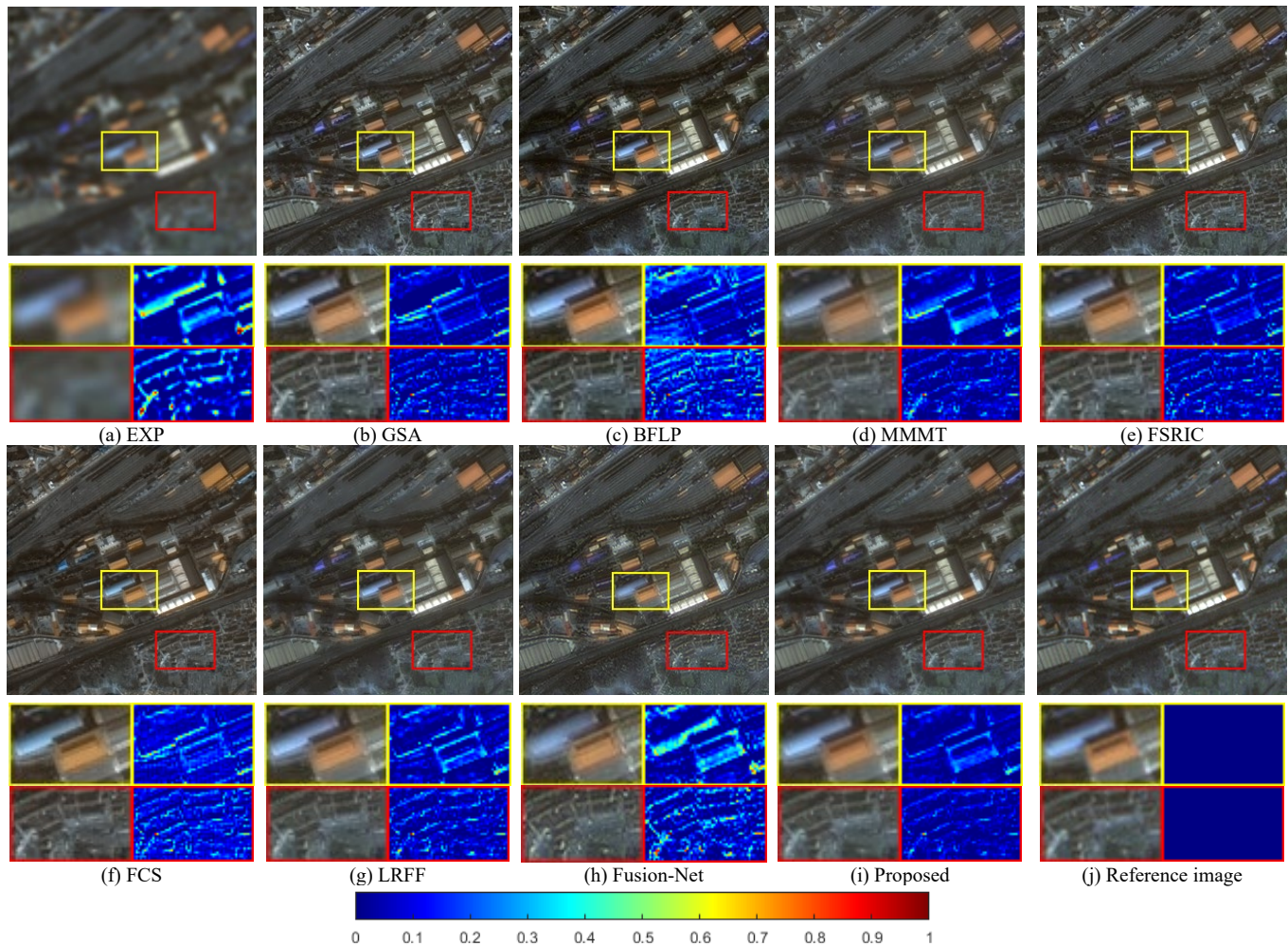


Fig. 9. Fusion results and the corresponding AEMs of the reduced scale images from Pléiades dataset.

slightly at the primary stage, but also converges in less than 5 steps. The reason for this is that the R band is the first band involved in the iteration process. When the iteration converges for the R band, the relevant variables also achieve relatively optimal values. Therefore, bands G, B, and NIR are much more stable and converge more quickly. This kind of result has been observed in all our datasets. Overall, the designed model achieves good convergence performance.

#### E. Reduced-scale Experiments

##### 1) IKONOS Dataset

Fusion experiments are conducted on 60 image pairs from the IKONOS dataset. Taking one fused image as an example,

the fusion results by different methods and the corresponding AEMs are displayed in Fig. 8 for visualization. We can see from the figure that the result of BFLP is brighter than the reference image overall, while the result of FCS is apparently darker. From AEMs, we can see that the results of Fusion-Net, BFLP, MMT, and FCS obviously contain more residues than those of the proposed method, while the results of GSA, FSRIC, and LRFF contain slightly more residues than those of the proposed method.

To further verify the effectiveness of all comparison methods, an objective evaluation is conducted. The results of Fig. 8 are tabulated in the left part of Table III, and the average evaluation results of 60 images from the IKONOS dataset are tabulated in



TABLE IV  
QUANTITATIVE EVALUATION OF FUSION RESULTS IN FIG. 9, AND AVERAGE QUANTITATIVE EVALUATION ON PLÉIADES DATASET

METHOD	RESULTS IN FIG. 9						AVERAGE					
	UIQI↑	SAM↓	ERGAS↓	SCC↑	Q4↑	RMSE↓	UIQI↑	SAM↓	ERGAS↓	SCC↑	Q4↑	RMSE↓
EXP [45]	0.8015	3.9133	4.9888	0.6612	0.7790	14.2572	0.7898	2.9110	4.0986	0.6810	0.7819	13.2158
GSA [8]	0.8944	4.0455	3.8649	0.8245	0.8776	11.1192	0.8888	3.1965	3.3250	0.8621	0.8829	10.4437
BFLP [15]	0.8938	4.5780	4.9065	0.8205	0.7589	14.1215	0.8639	3.5301	6.1081	0.8531	0.7459	18.8575
MMMT [23]	0.8953	4.3541	3.7137	0.8258	0.8875	9.5761	0.8931	3.1682	3.0394	0.8658	0.8905	9.5761
FSRIC [24]	0.9049	3.9971	3.5694	0.8258	0.8908	10.3888	0.8954	3.1060	3.0942	0.8596	0.8911	9.9257
FCS [22]	0.9071	4.6792	3.8631	0.8386	0.8813	11.0960	0.8915	3.8340	3.4024	0.8661	0.8677	10.7684
LRFF [27]	0.9265	3.9724	3.1082	0.8551	0.9171	8.9208	0.9133	3.0133	2.7034	0.8753	0.9106	8.6478
Fusion-Net [31]	0.8799	4.2262	4.0104	0.7749	0.8622	11.4299	0.8698	3.3373	3.343	0.8148	0.8633	10.5348
Proposed	<b>0.9318</b>	<b>3.9122</b>	<b>3.0254</b>	<b>0.8676</b>	<b>0.9244</b>	<b>8.6491</b>	<b>0.9260</b>	<b>2.9752</b>	<b>2.4876</b>	<b>0.8922</b>	<b>0.9241</b>	<b>8.1286</b>

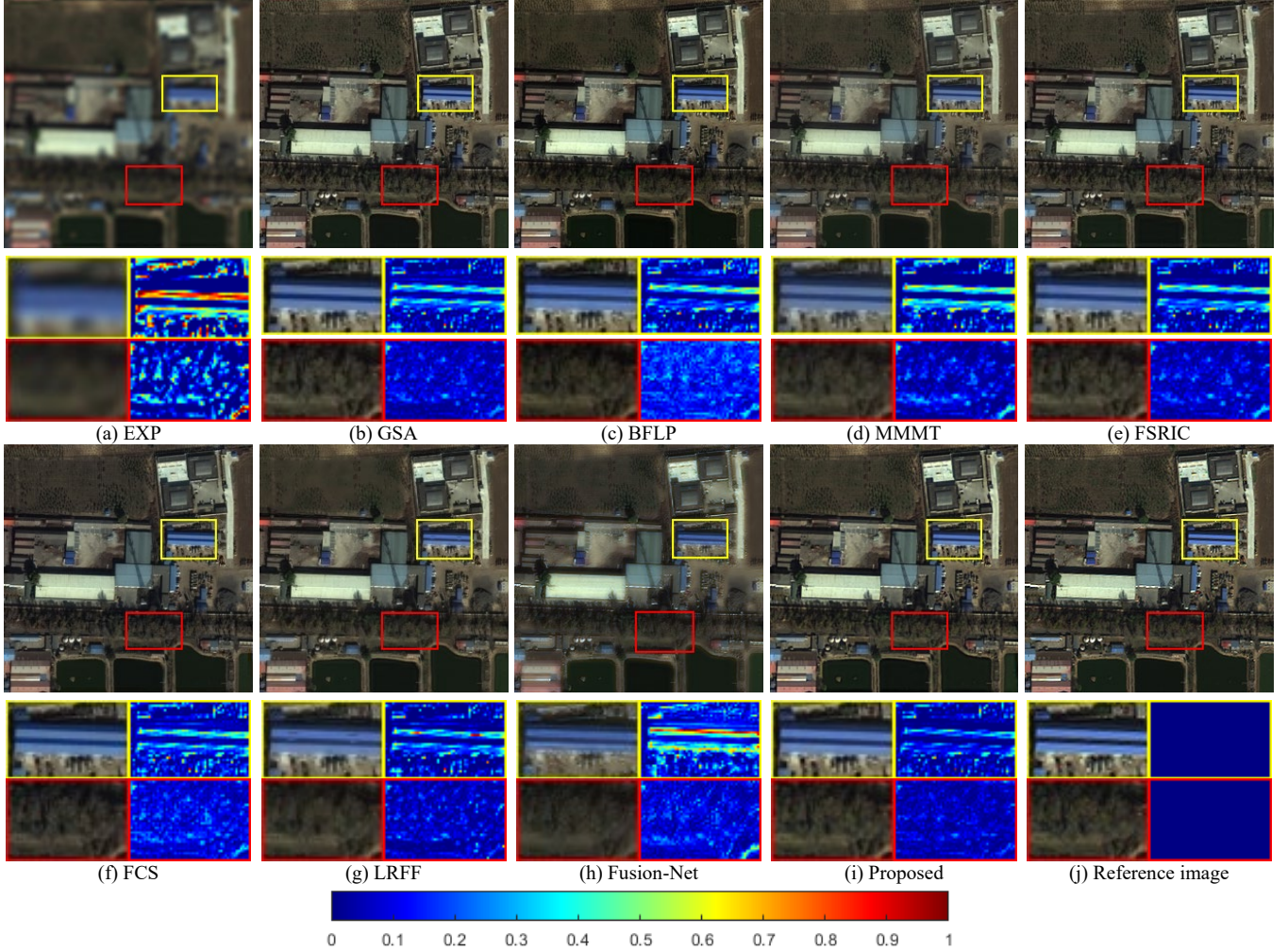


Fig. 10. Fusion results and the corresponding AEMs of the reduced scale images from WorldView-3 dataset with 8 bands.

the right part of Table III. We can see that the proposed method performs the best in all metrics, both in regard to the results presented in Fig. 8 and the average results. Moreover, the improvement of our method over other methods is significant.

## 2) Pléiades Dataset

Fusion experiments are also conducted on 60 image pairs from the Pléiades dataset. Taking a fused image as an example, the results of different methods and the corresponding AEMs are shown in Fig. 9. Overall, we can see that the result of BFLP is too bright and thus it may suffer from over injection; in contrast, the result of MMT is blurred, and thus it may suffer

from insufficient injection. Meanwhile, the result of FCS suffers from spectral distortion, since the color of some roofs turns light blue instead of remaining dark blue (as in the reference image), and the result of Fusion-Net suffers from spatial distortion because of introducing extra artifact. From the enlarged AEMs in yellow and red rectangles, we can see that the results of the proposed method contain less residues than those of other comparison methods.

An objective evaluation is carried out for the fusion results obtained on the Pléiades dataset. The results of Fig. 9 are tabulated in the left part of Table IV, and the average fusion

TABLE V  
QUANTITATIVE EVALUATION OF FUSION RESULTS IN FIG.10, AND AVERAGE QUANTITATIVE EVALUATION ON WORLDVIEW-3 DATASET

METHOD	RESULTS IN FIG.10						AVERAGE					
	UIQI↑	SAM↓	ERGAS↓	SCC↑	Q8↑	RMSE↓	UIQI↑	SAM↓	ERGAS↓	SCC↑	Q8↑	RMSE↓
EXP [45]	0.6697	6.2313	7.6323	0.5840	0.6932	21.3977	0.6690	5.0665	6.3078	0.6019	0.6673	16.7565
GSA [8]	0.9121	6.9638	3.9831	0.9017	0.9284	11.1053	0.8673	6.0949	4.1814	0.8335	0.8794	10.9186
BFLP [15]	<u>0.9212</u>	6.8542	4.2552	<u>0.9115</u>	0.8933	11.8524	0.8672	6.8722	6.4605	<u>0.8599</u>	0.8138	17.2254
MMMT [23]	0.8834	6.6724	4.5789	0.8984	0.8983	12.6770	0.8421	5.5929	4.3891	0.8409	0.8489	11.6147
FSRIC [24]	0.9058	6.8136	4.1050	0.9014	0.9227	11.6807	0.8664	5.9076	4.0927	0.8358	0.8794	10.9103
FCS [22]	0.9145	6.6823	4.0182	0.9099	0.9229	11.2761	<u>0.8743</u>	6.2206	4.3651	0.8558	0.8530	11.2896
LRFF [27]	0.9092	<u>6.2597</u>	3.9904	0.8993	0.9200	11.1161	0.8720	<u>5.1252</u>	<u>3.8997</u>	0.8506	0.8753	<u>10.3238</u>
Fusion-Net [31]	0.8822	8.164	5.0018	0.8772	0.8982	13.8858	0.8336	7.6868	5.2346	0.8254	0.8357	13.0084
Proposed	<b>0.9331</b>	<b>6.0618</b>	<b>3.5946</b>	<b>0.9193</b>	<b>0.9386</b>	<b>10.1790</b>	<b>0.8967</b>	<b>4.9456</b>	<b>3.7528</b>	<b>0.8645</b>	<b>0.8985</b>	<b>10.0334</b>

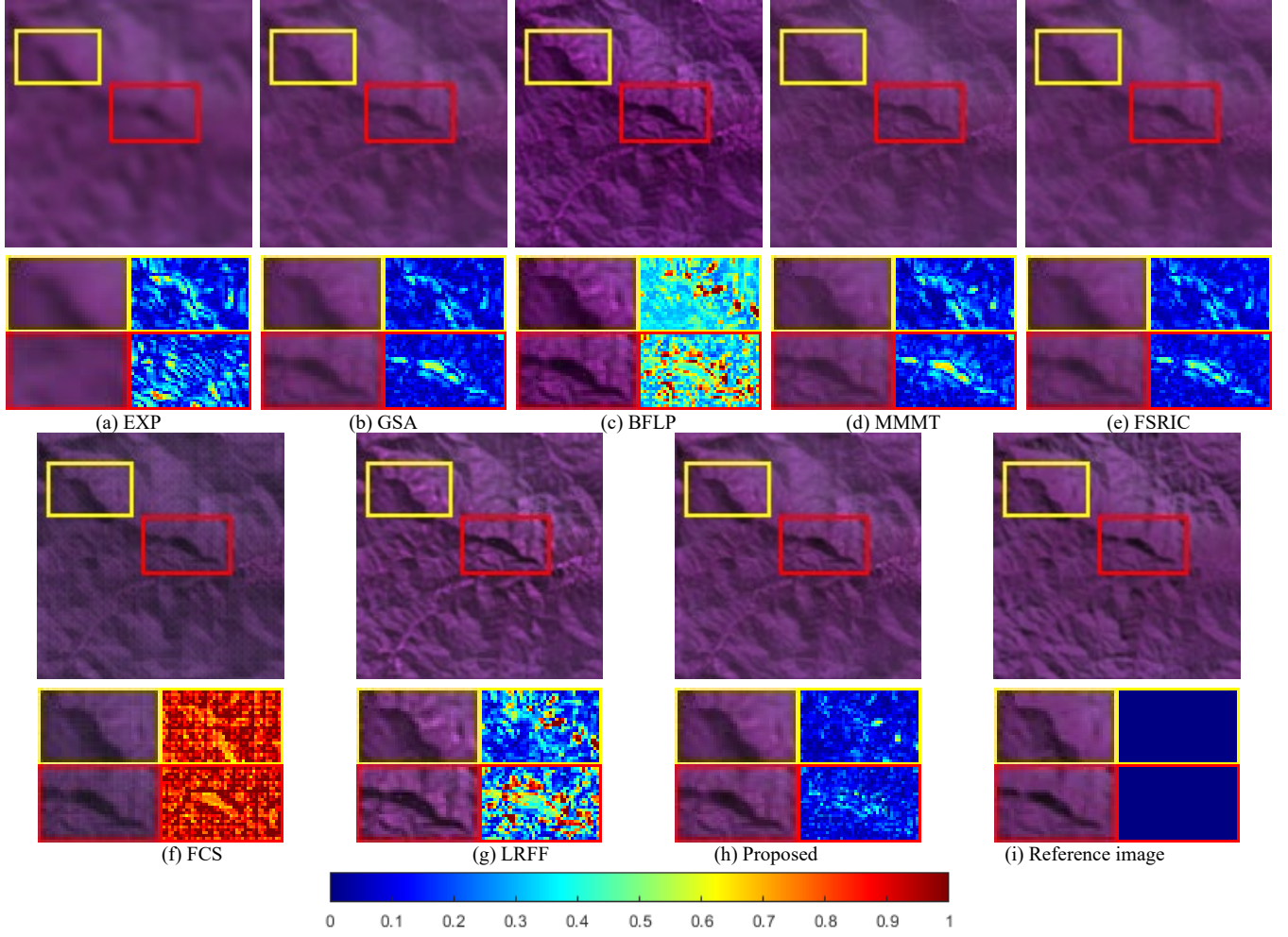


Fig. 11. Fusion results and the corresponding AEMs of the reduced scale images from TianGong-2 dataset with 14 bands.

results are tabulated in the right part of Table IV. We can see that the objective evaluation is consistent with the subjective evaluation, and the proposed method achieves the best results in all metrics both in regards to the results presented in Fig. 9 and the average results.

### 3) WorldView-3 Dataset

Additional fusion experiments are conducted on 60 image pairs in WorldView-3 dataset with 8 bands, and the fusion results of an image pair and the corresponding AEMs are shown in Fig. 10. We can see from the enlarged rectangles that the fused images of MMT, LRFF, and Fusion-Net are apparently blurred due to the insufficient detail injection. As observed by the AEMs in the rectangles, the results of other

comparison methods contain more residues than those of the proposed method. Overall, we can see that the result produced by the proposed method is closer to the reference image.

The objective evaluation results of Fig. 10 and the average evaluation results are tabulated in Table V. We can see from the table that the objective evaluation is consistent with the subjective assessment, and the proposed method has the best values in all metrics both in regards to the results in Fig. 10 and the average results. This further verifies the effectiveness of the proposed method.

### 4) TianGong-2 Dataset

To further verify the effectiveness of the proposed method, we conducted the experiments on TianGong-2 dataset with



TABLE VI  
QUANTITATIVE EVALUATION OF FUSION RESULTS IN FIG. 11, AND AVERAGE QUANTITATIVE EVALUATION ON TIANGONG-2 DATASET

METHOD	RESULTS IN FIG. 11					AVERAGE				
	UIQI↑	SAM↓	ERGAS↓	SCC↑	RMSE↓	UIQI↑	SAM↓	ERGAS↓	SCC↑	RMSE↓
EXP [45]	0.7104	2.813	2.6459	0.7891	10.965	0.7069	3.5345	3.518	0.7973	10.3785
GSA [8]	0.9128	2.4415	1.7171	0.9013	8.2365	0.9235	2.4025	1.9225	0.9334	6.4690
BFLP [15]	0.6974	7.6395	6.4419	0.8058	18.5895	0.6093	5.9151	17.5605	0.8253	45.339
MMMT [23]	0.8667	2.3881	1.9512	0.9055	8.4660	0.8462	2.9256	2.5830	0.9104	7.4970
FSRIC [24]	0.9108	3.0050	1.8733	0.9002	8.7465	0.9203	2.7803	2.0724	0.9312	6.5280
FCS [22]	0.4747	20.6073	17.4077	0.8581	39.2190	0.6191	18.3185	12.0454	0.8367	31.1865
LRFF [27]	0.6148	4.7526	4.5990	0.7719	14.9685	0.5649	15.0722	17.2226	0.6043	35.6065
Proposed	<b>0.9377</b>	<b>2.0763</b>	<b>1.5457</b>	<b>0.9228</b>	<b>7.7010</b>	<b>0.9321</b>	<b>2.3613</b>	<b>1.8870</b>	<b>0.9340</b>	<b>6.3750</b>

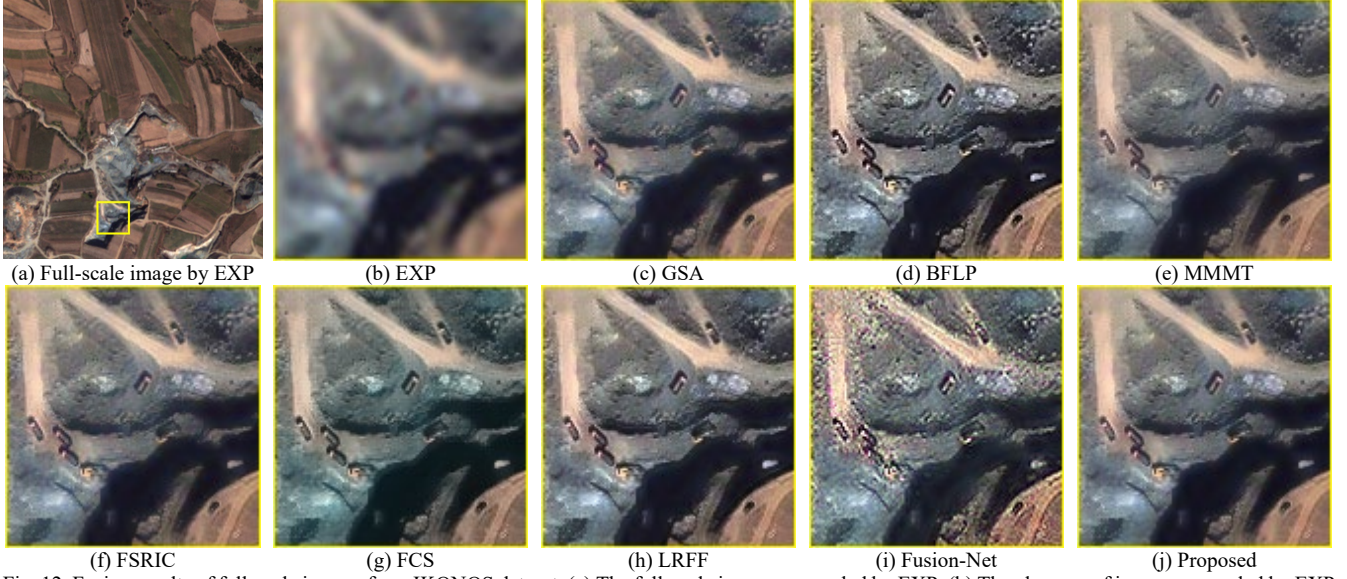


Fig. 12. Fusion results of full-scale images from IKONOS dataset. (a) The full-scale image upsampled by EXP. (b) The close-up of image upsampled by EXP. (c)-(j) The close-up of pansharpened images by different methods.

TABLE VII  
QUANTITATIVE EVALUATION OF FUSION RESULTS IN FIG. 12, AND AVERAGE QUANTITATIVE EVALUATION ON IKONOS DATASET

METHOD	RESULTS IN FIG. 12			AVERAGE		
	$D_s$ ↓	$D_r$ ↓	QNR↑	$D_s$ ↓	$D_r$ ↓	QNR↑
EXP [45]	0.0020	0.3001	0.6985	0.0023	0.3259	0.6725
GSA [8]	0.1488	0.1196	0.7494	0.1947	0.1728	0.6678
BFLP [15]	0.1933	0.1457	0.6892	0.2022	0.1607	0.6727
MMMT [23]	0.1594	0.0714	0.7806	0.1553	0.0667	0.7888
FSRIC [24]	0.1474	0.0906	0.7754	0.1889	0.1360	0.7022
FCS [22]	0.1818	0.0975	0.7384	0.1978	0.1106	0.7150
LRFF [27]	0.1519	0.0754	0.7842	0.1455	<b>0.0592</b>	<u>0.8043</u>
Fusion-Net [31]	<b>0.1011</b>	0.0939	<u>0.8145</u>	<b>0.1051</b>	0.1068	0.7988
Proposed	0.1099	<b>0.0498</b>	<b>0.8458</b>	0.1078	0.0647	<b>0.8348</b>

14-band MS images. The fusion results of TianGong-2 dataset are displayed in Fig. 11. Note that Fusion-Net is designed for the MS image with 4 and 8 bands, and is not suitable for the MS image with 14 bands. Therefore, Fusion-Net is excluded from the comparison for TianGong-2 dataset. Besides, since there are no R, G, and B bands in the MS image, the first three bands of the fused images are used to form the pseudo-color images, as shown in the upper half part of the figure. We can see that the results of GSA, MMT, FSRIC, and RBDSD are obviously blurred, while the results of LRFF and BFLP suffer from detail over-injection. FCS method seems produce serious spectral distortion. Overall, the result of the proposed method is closer to the reference image than other comparison methods. From AEMs, we can see that the proposed method evidently contains

less residues than other methods, which further verifies the effectiveness of our method.

The objective evaluation results of Fig. 11 and the average index values of 60 image pairs are tabulated in Table VI. Because the metric  $Q2^n$  is not suitable for the image with 14 bands, it is removed from the evaluation metrics. From the table, we can see that some methods are unstable, especially for LRFF and BFLP, which leads to much inferior results in some metrics. The proposed method performs the best in terms of all metrics.

#### F. Full-scale Experiments

Full-scale experiments are conducted on real images from the IKONOS dataset to acquire fused images with a size of  $1024 \times 1024 \times 4$ . Due to the lack of a reference image, the result of the EXP method is taken as a spectral reference, and we



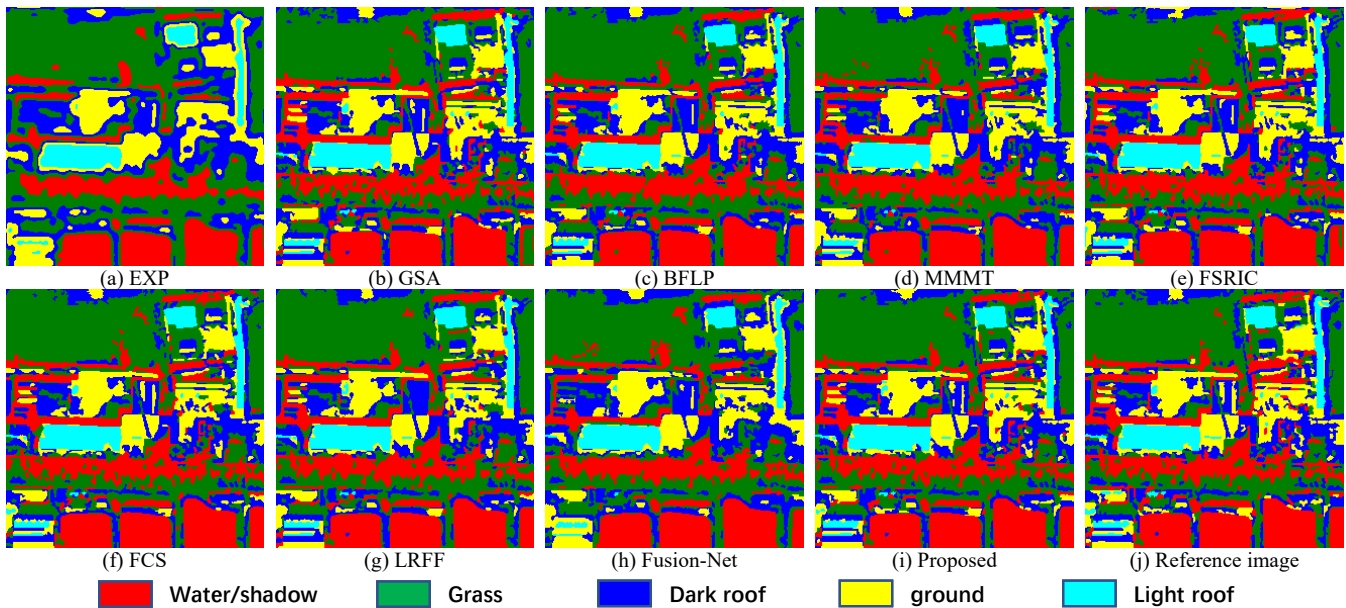


Fig. 13. Classification results of the images in Fig. 10.

TABLE VIII

OBJECTIVE EVALUATION OF CLASSIFICATION RESULTS

Method	OA $\uparrow$	K $\uparrow$	OE $\downarrow$	PA $\uparrow$
EXP [45]	0.7228	0.6285	0.2897	0.7103
GSA [8]	0.8732	0.8309	0.1276	0.8724
BFLP [15]	0.8618	0.8152	0.1433	0.8567
MMMT [23]	0.8498	0.7992	0.1553	0.8447
FSRIC [24]	0.8667	0.8222	0.1344	0.8656
FCS [22]	<u>0.8759</u>	<u>0.8340</u>	<u>0.1259</u>	<u>0.8741</u>
LRFF [27]	0.8653	0.8196	0.1413	0.8587
Fusion-Net [31]	0.8117	0.7487	0.1890	0.8110
Proposed	<b>0.8816</b>	<b>0.8420</b>	<b>0.1209</b>	<b>0.8791</b>

mainly consult the objective evaluation and use the QNR metric to represent the overall quality. Because the fused images are too large to display here, we only show one part of the area marked in the yellow rectangle. Taking an image pair from the dataset as an example, the fusion results are shown in Fig. 12. From the figure, we can see that the results of Fusion-Net and BFLP obviously introduce extra artifacts, and the trucks and mountains look unnatural. The result of MMT is blurred, and the result of FCS suffers from obvious spectral distortion since the overall color is changed compared with the reference provided by the EXP method. The results of other methods are difficult to differentiate. The objective evaluation of the results in Fig. 12 is tabulated in the left part of Table VII, and the average fusion results are shown in the right part of the table. We can see that the proposed method achieves the best results in  $D_s$  and QNR metrics, and the second-best result in  $D_i$  metric in regards to the results presented in Fig. 12. As for the average fusion results, our method obtains the best QNR metric, and the second-best  $D_i$  and  $D_s$  metrics. Specifically, our method has a big advantage in the critical metric QNR.

### G. Classification Application

To testify the practicality of the pansharpened images obtained by the proposed method in classification application, we conducted the classification experiments by ENVI

classification tool<sup>1</sup> on the pansharpened images. Taking the fused images in Fig. 10 as an example, their classification results are shown in Fig. 13. From the figure, we can see that the results of MMT and Fusion-Net appear some classification errors especially in the green and red classes. The result of FSRIC exists some obvious errors in the red class. Other results are hard to tell the difference visually.

To quantify the accuracy of the classification, four metrics provided by the ENVI tool are employed. They are the overall accuracy (OA) and kappa coefficient (K) indicating the pixel accuracy, the omission error (OE) indicating the classification error, and the producer accuracy (PA) indicating the accuracy of each class. The objective evaluation results of Fig. 13 are shown in Table VIII. From the table, we can see that the proposed method can achieve the best classification results in all metrics, which proves the effective performance of our fusion method.

### H. Efficiency Analysis

With the FFT and IFFT, the computation complexity for the solution of the proposed model is  $O(B \times s \times K \log_2 K)$ , where  $B$  is the number of bands,  $s$  is the number of iterations, and  $K$  is the number of pixels in an image. Due to the fine-tuned parameters in the initialization step, the proposed model can converge quickly (see sections IV-B and IV-D). Note that the time for parameter settings is excluded from the efficiency analysis, because the parameters are preset and fixed during the fusion process. The average running time of the different methods on all datasets is shown in Table IX. We can see from the table that for reduced-scale experiments, the proposed method has high efficiency and the operation time is less than 1 second; meanwhile, the full-scale experiments consume much more time than the reduced-scale experiments. However, the proposed method is much faster than other optimization-based methods such as MMT, FCS, and LRFF, and faster than

<sup>1</sup> <https://www.13harrisgeospatial.com/Software-Technology/ENVI>

TABLE IX  
AVERAGE RUNNING TIME (S) FOR REDUCED-SCALE AND FULL-SCALE  
IMAGES

Method	Reduced scale	Full scale
EXP [45]	0.02	0.32
GSA [8]	0.07	1.51
BFLP [15]	1.71	17.03
MMMT [23]	14.5	194.24
FSRIC [24]	0.16	2.05
FCS [22]	35.6	529.69
LRFF [27]	7.21	80.67
Fusion-Net [31]	8.71	147.39
Proposed	0.64	11.29

DL-based Fusion-Net. GSA is a classical CS-based method, while FSRIC is a model-based method without iterative process, and thus these methods consume less time. However, in general, the proposed method obtains much better fusion quality and high efficiency.

## V. CONCLUSION

Aiming at the problems of the inappropriate construction of models and low efficiency in pansharpening approaches, a novel pansharpening model based on the band-adaptive gradient and detail correction is proposed. To maintain spectral fidelity, a spectral constraint is first proposed to maintain the spectral consistence of MS image bands. Due to the diversity of the gradient information in different bands, a gradient correction constraint is then designed to improve the structural quality. Additionally, to improve the accuracy of the spatial details, a detail correction constraint based on the parameter transfer is constructed. A total variational term is designed to further improve the fusion quality by reducing the number of artifacts in the fused image. Finally, a unified model is constructed based on the gradient and detail correction, and solved by the ADMM method. An ablation study and convergence analysis further verify the effectiveness of the proposed method. Compared with some classical and state-of-the-art pansharpening methods, our method achieves the best overall results and demonstrates high efficiency.

## REFERENCES

- [1] Y. Yang, L. Wu, S. Huang, W. Wan, W. Tu, and H. Lu, "Multiband Remote Sensing Image Pansharpening Based on Dual-injection Model," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 1888–1904, Mar. 2020.
- [2] X. Meng, H. Shen, H. Li, L. Zhang, and R. Fu, "Review of the pansharpening methods for remote sensing images based on the idea of meta-analysis: Practical discussion and challenges," *Inf. Fusion*, vol. 46, pp. 102–113, Mar. 2019.
- [3] H. Ghassemian, "A review of remote sensing image fusion methods," *Inf. Fusion*, vol. 32, pp. 75–89, Nov. 2016.
- [4] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, and B. A. Johnson, "Deep learning in remote sensing applications: A meta-analysis and review," *ISPRS J. Photogramm. Remote Sens.*, vol. 152, pp. 166–177, Jun. 2019.
- [5] L.-J. Deng, G. Vivone, W. Guo, M. Dalla 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.
- [6] G. Vivone, M. Dalla Mura, A. Garzelli, R. Restaino, G. Scarpa, M. O. Ulfarsson, L. Alparone, and J. Chanussot, "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.
- [7] Z. Zhou, N. Ma, Y. Li, P. Yang, P. Zhang, and Y. Li, "Variational PCA fusion for Pan-sharpening very high resolution imagery," *Sci. China Inf. Sci.*, vol. 57, no. 11, pp. 1–10, Nov. 2014.
- [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, Sep. 2007.
- [9] S. Rahmani, M. Strait, D. Merkurjev, M. Moeller, and T. Wittman, "An Adaptive IHS Pan-Sharpener Method," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 746–750, Oct. 2010.
- [10] Y. Yang, H. Lu, S. Huang, and W. Tu, "Pansharpening Based on Joint-Guided Detail Extraction," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 14, pp. 389–401, Oct. 2020.
- [11] 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.
- [12] M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2091–2106, Dec. 2005.
- [13] X. Otazu, M. Gonzalez-Audicana, O. Fors, and J. Nunez, "Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 10, pp. 2376–2385, Oct. 2005.
- [14] Y. Kim, C. Lee, D. Han, Y. Kim, and Y. Kim, "Improved Additive-Wavelet Image Fusion," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 2, pp. 263–267, 2011.
- [15] 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.
- [16] N. H. Kaplan, I. Erer, O. Ozcan, and N. Musaoğlu, "MTF driven adaptive multiscale bilateral filtering for pansharpening," *Int. J. Remote Sens.*, vol. 40, no. 16, pp. 6262–6282, Aug. 2019.
- [17] 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.
- [18] S. Li and B. Yang, "A New Pan-Sharpener Method Using a Compressed Sensing Technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 738–746, 2011.
- [19] F. Fang, F. Li, C. Shen, and G. Zhang, "A Variational Approach for Pan-Sharpener," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2822–2834, Jul. 2013.
- [20] X. X. Zhu and R. Bamler, "A Sparse Image Fusion Algorithm With Application to Pan-Sharpener," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2827–2836, May 2013.
- [21] S. Li, H. Yin, and L. Fang, "Remote Sensing Image Fusion via Sparse Representations Over Learned Dictionaries," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 9, pp. 4779–4789, 2013.
- [22] M. Ghahremani, Y. Liu, P. Yuen, and A. Behera, "Remote sensing image fusion via compressive sensing," *ISPRS J. Photogramm. Remote Sens.*, vol. 152, pp. 34–48, Jun. 2019.
- [23] Y. Yang, W. Wan, S. Huang, P. Lin, and Y. Que, "A Novel Pan-Sharpener Framework Based on Matting Model and Multiscale Transform," *Remote Sens.*, vol. 9, no. 4, p. 391, Apr. 2017.
- [24] G. Vivone, R. Restaino, and J. Chanussot, "Full Scale Regression-Based Injection Coefficients for Panchromatic Sharpener," *IEEE Trans. Image Process.*, vol. 27, no. 7, pp. 3418–3431, Jul. 2018.
- [25] E. Vargas, O. Espitia, H. Arguello, and J.-Y. Tournet, "Spectral Image Fusion From Compressive Measurements," *IEEE Trans. Image Process.*, vol. 28, no. 5, pp. 2271–2282, May 2019.
- [26] G. Vivone, "Robust Band-Dependent Spatial-Detail Approaches for Panchromatic Sharpener," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6421–6433, Sep. 2019.
- [27] Y. Yang, C. Wan, S. Huang, H. Lu, and W. Wan, "Pansharpening Based on Low-Rank Fuzzy Fusion and Detail Supplement," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 5466–5479, Sep. 2020.
- [28] 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.
- [29] G. Masi, D. Cozzolino, L. Verdoliva, and G. Scarpa, "Pansharpening by Convolutional Neural Networks," *Remote Sens.*, vol. 8, no. 7, p. 594, 2016.

- [30] Y. Zhang, C. Liu, M. Sun, and Y. Ou, "Pan-Sharpener Using an Efficient Bidirectional Pyramid Network," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5549–5563, Aug. 2019.
- [31] 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, 2021.
- [32] 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.
- [33] Y. Yang, H. Lu, S. Huang, and W. Tu, "Remote Sensing Image Fusion Based on Fuzzy Logic and Saliency Measure," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 11, pp. 1943–1947, Dec. 2019.
- [34] R. Restaino, G. Vivone, M. Dalla Mura, and J. Chanussot, "Fusion of Multispectral and Panchromatic Images Based on Morphological Operators," *IEEE Trans. Image Process.*, vol. 25, no. 6, pp. 2882–2895, Jun. 2016.
- [35] G. Vivone, P. Addesso, R. Restaino, M. Dalla Mura, and J. Chanussot, "Pansharpening Based on Deconvolution for Multiband Filter Estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 540–553, Jan. 2019.
- [36] G. Vivone, L. Alparone, J. Chanussot, M. Dalla Mura, A. Garzelli, G. A. Licciardi, R. Restaino, and L. Wald, "A Critical Comparison Among Pansharpening Algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2565–2586, May 2015.
- [37] X. Fu, Z. Lin, Y. Huang, and X. Ding, "A Variational Pan-Sharpener With Local Gradient Constraints," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, 2019, pp. 10257–10266.
- [38] Y. Yang, H. Lu, S. Huang, and W. Tu, "An Efficient Pansharpening Method Based On Conditional Random Fields," in *2020 IEEE International Conference on Multimedia and Expo (ICME)*, London, United Kingdom, 2020, pp. 1–6.
- [39] 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, pp. 691–699, 1997.
- [40] Zhou W and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, 2002.
- [41] R. H. Yuhas, A. F. H. Goetz, and J. W. Boardman, "Discrimination among semi-arid landscape endmembers using the Spectral Angle Mapper (SAM) algorithm," presented at the Summaries of the Third Annual JPL Airborne Geoscience Workshop, JPL Publ., 1992, vol. 1, pp. 147–149.
- [42] J. Zhou, D. L. Civco, and J. A. Silander, "A wavelet transform method to merge Landsat TM and SPOT panchromatic data," *Int. J. Remote Sens.*, vol. 19, no. 4, pp. 743–757, Jan. 1998.
- [43] L. Alparone, S. Baronti, A. Garzelli, and F. Nencini, "A global quality measurement of pan-sharpened multispectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 313–317, 2004.
- [44] L. Alparone, B. Aiazzi, S. Baronti, A. Garzelli, F. Nencini, and M. Selva, "Multispectral and Panchromatic Data Fusion Assessment Without Reference," *Photogramm. Eng. Remote Sens.*, vol. 74, no. 2, pp. 193–200, Feb. 2008.
- [45] B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli, "Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 10, pp. 2300–2312, Jan. 2002.