SARF: A Simple, Adjustable, and Robust Fusion Method

Xiangchao Meng, *Member, IEEE*, Gang Yang, Feng Shao, *Member, IEEE*, Weiwei Sun, *Senior Member, IEEE*, Huanfeng Shen, *Senior Member, IEEE*, and Shutao Li, *Fellow, IEEE*

***Abstract*— Pansharpening aims to sharpen a low spatial resolution (LR) multispectral (MS) image using a high spatial resolution (HR) panchromatic (PAN) image to obtain the HR MS image. Though large numbers of pansharpening methods have been proposed, and many advanced methods have shown high quantitative results, few of them are widely used in real applications. This may be attributed to their instability for different images with different ground surface features, or the complexity to be implemented and the time-consuming process for some state-of-the-art methods. In this letter, we proposed a simple, adjustable, and robust fusion (SARF) method. In the proposed method, a spatial-spectral coenhanced strategy was proposed, and several details of the proposed fusion model were specifically designed for the "simple, adjustable, robust" features. It was tested and verified by four-band and eight-band MS images based on reduced resolution (RR) and full resolution (FR) experiments. The experimental results demonstrated the promising spatial visuality of the proposed method, and the spectral fidelity was more robust than most of component substitution (CS)-based and multiresolution analysis (MRA)-based methods.**

***Index Terms*— Fusion, multispectral (MS), panchromatic (PAN), pansharpening, remote sensing.**

I. Introduction

DUE to the technical limitations of sensors and other factors, the existing remote sensing observations generally feature a tradeoff between spatial and spectral resolutions [1]. For example, most of remote sensing satellites provide bundled high spatial resolution (HR) panchromatic (PAN) image with a low spectral resolution, and low spatial resolution (LR) multispectral (MS) image with a relatively higher spectral resolution, rather than an HR MS image. Fortunately, pansharpening, i.e., PAN/MS fusion, can overcome this limitation to obtain the images with both high spatial and spectral resolutions.

Pansharpening originated during the 1980s, and it has been developed for 40 years. The development of pansharpening methods can be divided into three major stages [2]: the initial stage in the 1980s; the rapid development stage in the 1990s; and the blowout development stage after 2000. To date, there have proposed large numbers of pansharpening methods, and these methods have been classified in several different ways. They were classified into two major categories in [3], i.e., the component substitution (CS)-based and the multiresolution analysis (MRA)-based methods. Kwan *et al.* [4], [5] classified existing pansharpening methods based on whether the point spread function (PSF) was used or not. Meng *et al.* [1], [2] summarized them into four family: CS-, MRA-, variational optimization (VO)-, and deep learning (DL)-based methods. Among them, the VO-based methods [6], [7] are based on the iterative optimization solution of a variational energy functional. They generally show higher quantitative evaluation results; however, they are time-consuming, especially for images in large dimensions. This has seriously hindered their engineering applications. The DL-based methods [8], [9] are regarded as a new generation of pansharpening methods, and they have attracted ever-increasing attention in recent years. The basic idea is to learn a network between the fused image and the observations based on DL theory, with all parameters automatically learned under the supervision of large-scale training samples. However, the performance of DL-based methods depends on the training data set, and the network training is also generally required to be retrained for different tasks, such as different remote sensing satellites or different thematic scenarios. The CS- and MRA-based methods have been developed from traditional understanding with rigid three-step processing ("forward projection transformation—CS—inverse projection transformation" for the CS-based methods, and "decomposition—fusion—reconstruction" for the MRA-based methods), to the current new general understanding featured by a unified fusion framework without forward and backward transformation. They are fast and easy to be implemented. Accordingly, some methods, such as the principal component analysis (PCA) method, the Brovey method, the Gram–Schmidt (GS) method, and so on, have been commercialized and widely used in real applications. However, most of the CS- and MRA-based methods do not have a robust performance for different data and scenarios. Besides, different applications may have different requirements for more spectral fidelity or more spatial

Manuscript received December 22, 2020; revised January 25, 2021 and February 15, 2021; accepted February 15, 2021. Date of publication February 24, 2021; date of current version December 27, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 41801252, in part by the Fellowship of China Postdoctoral Science Foundation under Grant 2020M672490, in part by the Natural Science Foundation of Ningbo City under Grant 2019A610098, and in part by the K. C. Wong Magna Fund in Ningbo University. *(Corresponding authors: Gang Yang; Shutao Li.)*

Xiangchao Meng is with the Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China, and also with the College of Electrical and Information Engineering, Hunan University, Changsha 410082, China (e-mail: mengxiangchao@nbu.edu.cn).

Gang Yang and Weiwei Sun are with the Department of Geography and Spatial Information Techniques, Ningbo University, Ningbo 315211, China (e-mail: yanggang@nbu.edu.cn; sunweiwei@nbu.edu.cn).

Feng Shao is with the Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China (e-mail: shaofeng@nbu.edu.cn).

Huanfeng Shen is with the School of Resources and Environmental Sciences, Wuhan University, Wuhan 430079, China (e-mail: shenhf@whu.edu.cn).

Shutao Li is with the College of Electrical and Information Engineering, Hunan University, Changsha 410082, China (e-mail: shutao_li@hnu.edu.cn).

Digital Object Identifier 10.1109/LGRS.2021.3060095

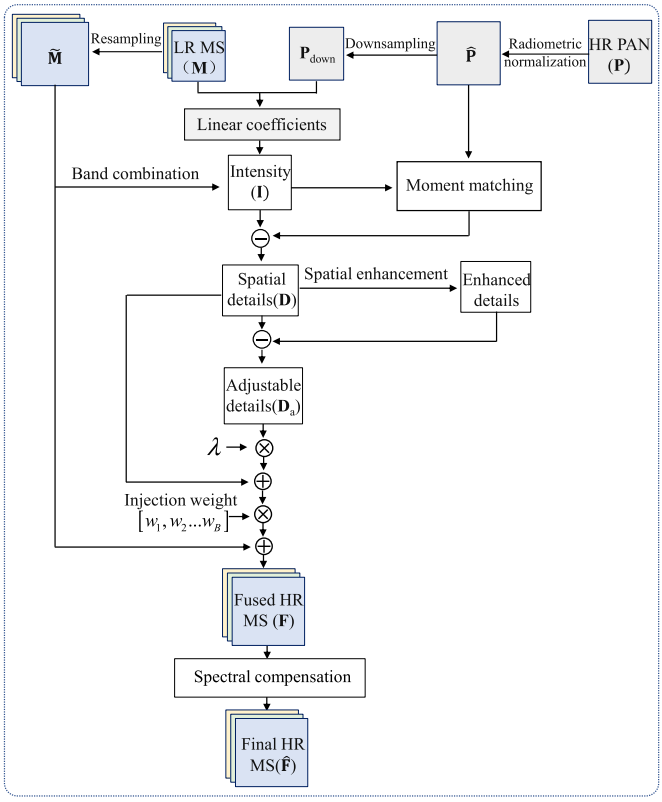

Fig. 1. Flowchart of the proposed method.

enhancement; hence, adjustable pansharpening methods are necessary.

To solve these problems, we proposed a simple, adjustable, and robust pansharpening method for remote sensing images. There are two main contributions in this letter.

1) We propose a simple, adjustable, and robust fusion (SARF) model. In the method, a spatial-spectral coenhanced fusion strategy is proposed by comprehensively considering "simple, adjustable, robust" features. A spatial enhancement strategy by considering noise disturbance is proposed to satisfy the demands for sharper spatial information in some real applications. In addition, a spectral compensation is further developed to ensure the spectral robustness to different satellites and scenarios.
2) The proposed SARF is simple yet effective. It has better visual perception and robust performance than most of CS- and MRA-based methods, and a higher computational efficiency than VO- and DL-based methods. In addition, it has robust performance for different scenarios and misregistration alignments.

II. SARF

In this letter, a simple, adjustable, and robust pansharpening method is proposed. It consists of three major steps, i.e., 1) radiometric normalization; 2) spatial detail extraction and enhancement; and 3) fused image initialization and spectral compensation. The flowchart of the proposed method is shown in Fig. 1.

A. Radiometric Normalization

In this letter, considering possible radiometric difference between PAN and MS, a moment matching is performed on PAN for the radiometric normalization, represented as

$$\hat{\mathbf{P}} = \frac{\left(\tilde{\mathbf{M}}_{\mathbf{I}}\right)_{\text{std}}}{(\mathbf{P})_{\text{std}}} \times \left[\mathbf{P} - (\mathbf{P})_{\text{mean}}\right] + \left(\tilde{\mathbf{M}}_{\mathbf{I}}\right)_{\text{mean}} \tag{1}$$

where $\mathbf{P}$ denotes the original PAN, $\tilde{\mathbf{M}}_{\mathbf{I}}$ is the average intensity component of the resampled MS image $\tilde{\mathbf{M}}$, $\hat{\mathbf{P}}$ is the normalized PAN, and the subscripts $(\cdot)_{\text{std}}$ and $(\cdot)_{\text{mean}}$ denote the standard deviation and mean value, respectively.

B. Spatial Detail Extraction and Enhancement

High spatial details and adjustable spatial structures are extracted. First, an optimal intensity of the MS is calculated

$$\mathbf{I} = c_b\tilde{\mathbf{M}}_b + c_{b+1}\tilde{\mathbf{M}}_{b+1} + \ldots + c_Q\tilde{\mathbf{M}}_Q \quad \text{with } b \geq 1,\ Q \leq B \tag{2}$$

where $\mathbf{I}$ is the intensity image, c_b is the coefficient for the bth MS band, and B denotes the total number of MS bands. Most of the existing solutions of the coefficients can be concluded into the following several ways.

1) The coefficients are obtained by spectral transformation or the correlation between PAN and MS. For example, $c_b = \text{cov}(\tilde{\mathbf{M}}_b, \tilde{\mathbf{M}}_I)/\text{var}(\tilde{\mathbf{M}}_I)$ for GS fusion, where $\text{cov}(\cdot)$ and $\text{var}(\cdot)$ denote the covariance and variance, respectively, $\tilde{\mathbf{M}}_b$ is the bth band.
2) The coefficients are determined based on spectral response function (SRF) of satellite sensors. However, in fact, the SRF may be different from the responses of the observed data sets due to atmospheric effects, illumination conditions, and so on. Furthermore, if the SRF is not obtained, then the method will not work.
3) The coefficients are calculated by regression analysis. In our proposed method, the coefficients $\{c_b, c_{b+1}, \ldots, c_Q\}$ mainly contribute to the fitting of the intensity image. Therefore, the coefficients are calculated based on the unconstrained least square linear regression. In addition, considering the computational efficiency of the algorithm, the coefficients are calculated on the spatial scale of LR MS $\mathbf{M}$

$$\underset{\{c_b, c_{b+1}, \ldots, c_Q\}}{\text{argmin}} \left\{ \left\| \mathbf{P}_{\text{down}} - \sum c_b\mathbf{M}_b \right\|_2^2 \right\} \quad \text{with } 1 \leq b \leq Q \tag{3}$$

where $\mathbf{P}_{\text{down}}$ is the downsampled PAN image.

Second, the spatial details of HR PAN are extracted

$$\mathbf{D} = f\left(\hat{\mathbf{P}}, \mathbf{I}\right) - \mathbf{I} \tag{4}$$

where $\mathbf{D}$ is the extracted spatial details and $f(\cdot)$ denotes the moment matching between $\hat{\mathbf{P}}$ obtained by (1) and $\mathbf{I}$ calculated by (2).

Finally, the extracted spatial detail $\mathbf{D}$ is further enhanced, to obtain the adjustable spatial structures of the fused image

$$\mathbf{D}_a = g_e(g_w(\mathbf{D})) - \mathbf{D} \tag{5}$$

where $\mathbf{D}_a$ is the adjustable spatial structures and $g_w(\cdot)$ denotes the Wiener filter operation to avoid noise disturbance during spatial enhancement. $g_e(\cdot)$ is the enhancement filter, represented as

$$g_e = \frac{1}{a+1}\begin{bmatrix} -a & a-1 & -a \\ a-1 & a+5 & a-1 \\ -a & a-1 & -a \end{bmatrix}. \tag{6}$$

In this letter, the default parameter of $a = 0.2$ was utilized.

C. Fused Image Initialization and Spectral Compensation

The preliminary fused image is obtained based on the injection of the extracted basic and adjustable spatial details

$$\mathbf{F}_b = \tilde{\mathbf{M}}_b + w_b(\mathbf{D} + \lambda \mathbf{D}_a) \tag{7}$$

where $\mathbf{F}_b$ denotes the bth band of the fused image, λ is the parameter, and w_b indicates the injection weight. On the one hand, in the spectral dimension, the weight w_b may be equal for all spectral bands, or determined by a band-dependent solution, such as the popular high-pass modulation (HPM) injection scheme applied in the additive wavelet luminance proportional (AWLP) [10] and Generalized Laplacian pyramid (GLP) using modulation transfer function (MTF), namely MTF-GLP-HPM [11] and so on. On the other hand, in the spatial dimension, it is determined by a global model, which is applied in most of the commercial methods, or a local model, such as the context-based decision (CBD) model [12], and the local model proposed by the band-dependent spatial detail (BDSD) [13], partial replacement adaptive CS (PRACS) [14], and so on. In a nutshell, the spatially local model generally shows relatively better performance on spectral fidelity; however, the robustness to spatial enhancement is somewhat discount, such as the PRACS [14], and it is also slightly time-consuming compared with a global model. Therefore, a simple yet effective band-dependent global injection model is proposed. Considering that the injected spatial details should be proportional to the spatial structures of original MS bands, i.e., the less of the quantity of spatial gradients of MS band, the less of the demands for the injected spatial details. The proposed weight is represented as: $w_b = (\mathbf{M}_b)_{\text{gra}}/(\sum_{b=1}^{B} \mathbf{M}_b/B)_{\text{gra}}$, where the subscript $(\cdot)_{\text{gra}}$ denotes the average gradient.

Then, spectral compensation is proposed to improve the spectral fidelity and robustness, represented as

$$\hat{\mathbf{F}} = \mathbf{F} + \hat{\mathbf{F}}_{\text{res}} \tag{8}$$

where $\hat{\mathbf{F}}$ is the final fused image and $\hat{\mathbf{F}}_{\text{res}}$ denotes the ideal spectral residual. It is obtained based on the original LR MS, due to inherent unavailable of the ideal fused image.

First, the spectral difference is calculated by original LR MS and the preliminary fused image, represented as

$$\hat{\mathbf{F}}'_{\text{res}} = \mathbf{M} - f_{\text{down}}(f_{\text{MTF}}(\mathbf{F})) \tag{9}$$

where $f_{\text{MTF}}(\cdot)$ and $f_{\text{down}}(\cdot)$ denote the MTF blurring and downsampling, respectively. $\hat{\mathbf{F}}'_{\text{res}}$ denotes the spectral difference between the original LR MS and fused images.

Then, $\hat{\mathbf{F}}'_{\text{res}}$ is upsampled to the spatial dimension of fused image for spectral compensation, represented as

$$\hat{\mathbf{F}} = \mathbf{F} + \left(\hat{\mathbf{F}}'_{\text{res}}\right)_{\text{up}} \tag{10}$$

where the subscript $(\cdot)_{\text{up}}$ denotes the upsampling operation, including MTF blurring and resampling.

III. Experiments

The reduced resolution (RR) and full resolution (FR) experiments were implemented on QuickBird and WorldView-2 images, respectively. The RR experiment was based on the Wald's protocol. To date, there are many quantitative indices, such as the radiometric and geometric index (RG index) [15], correlation coefficient (CC), relative dimensionless global error in synthesis (ERGAS), spectral angle mapper (SAM), and $Q2^n$ [16] (an extension of Q4 [17]). In this RR experiment, the ERGAS, SAM, and $Q2^n$ were employed. The FR experiment was directly performed by original HR PAN and LR MS images, and the popular nonreference quantitative evaluation indices of the quality with no reference (QNR) [18] and generalized QNR (GQNR) [19] were utilized.

TABLE I
QUANTITATIVE EVALUATION RESULTS IN THE RR EXPERIMENT

Methods	ERGAS	SAM	$Q2^n$
GS	2.2418	2.2055	0.9167
ATWT_M2	2.3298	2.1110	0.8990
AWLP	<u>1.8505</u>	<u>1.9221</u>	**0.9400**
MTF_GLP	1.9890	2.1240	0.9331
PNN	3.3384	2.9843	0.8585
TFNet	2.2533	2.2536	0.9043
SRCNN	3.2441	2.9334	0.8618
MSDCNN	2.3892	2.3134	0.9093
Proposed SARF	**1.7461**	**1.7838**	<u>0.9367</u>

The SARF was compared with several state-of-the-art pansharpening methods in [3], including the GS, à trous wavelet transform (ATWT) using Model 2 (ATWT-M2), AWLP, MTF-GLP, convolutional neural network (CNN)-based pansharpening neural network (PNN) [20], two-stream fusion network (TFNet) [21], super-resolution CNN (SRCNN) [22], and the multiscale and multidepth convolutional neural network (MSDCNN) [23]. To ensure the consistency and comparability, all the DL-based methods were trained based on the Wald's protocol by the same data set [2] in the RR and FR experiments, respectively. In the RR experiment, 5000, 2000, and 1000 pairs of QuickBird PAN (64×64) and MS ($16 \times 16 \times 4$) were employed as the training, validation, and testing data sets, respectively. In the FR experiment, 5000, 2000, and 1000 pairs of WorldView-2 PAN (64×64) and MS ($16 \times 16 \times 8$) images were used, respectively. The parameter is $\lambda = 0$.

A. RR Experiment

The RR experimental results based on QuickBird data are shown in Fig. 2. It is shown that GS fusion result has sharp spatial details; however, it has slight spectral distortion. The ATWT-M2 and the DL-based PNN, TFNet, SRCNN, and MSDCNN methods show good spectral fidelity; however, they are slightly blurring. The proposed fusion result has both high spectral fidelity and spatial enhancement. The quantitative evaluation results are shown in Table I. The best results are marked in bold, and the second is marked in underline. It shows that the SARF has the best spectral fidelity in terms of most of quantitative evaluation indices.

B. FR Experiment

The FR experimental results based on WorldView-2 are shown in Fig. 3. It shows that the GS fusion result has slightly spectral distortions. The DL-based PNN, TFNet, SRCNN, and MSDCNN methods show overall good performance. The proposed SARF shows competitive performance. The quantitative evaluation indices of D_λ (spectral evaluation), D_s (spatial evaluation), QNR (overall evaluation), and GQNR in Table II shows that the proposed SARF achieves competitive performance compared with the CS- and MRA-based methods, and it has similar performance with the DL-based methods.

C. Discussions

1) Advantage and Drawback: This letter has presented a simple, adjustable, and robust pansharpening method.

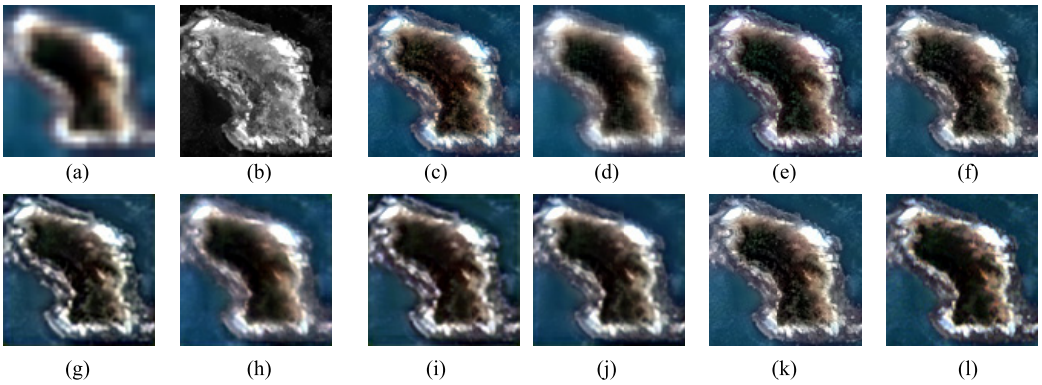

Fig. 2. RR experimental results based on QuickBird images. (a) Upsampling MS. (b) HR PAN. (c) GS. (d) ATWT-M2. (e) AWLP. (f) MTF-GLP. (g) PNN. (h) TFNet. (i) SRCNN. (j) MSDCNN. (k) Proposed SARF. (l) Original MS.

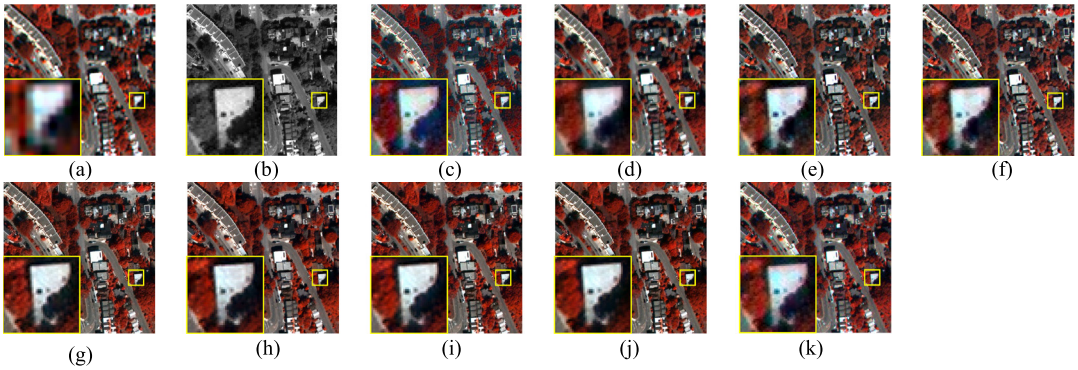

Fig. 3. FR experimental results based on WorldView-2 images. (a) Upsampling MS. (b) HR PAN. (c) GS. (d) ATWT-M2. (e) AWLP. (f) MTF-GLP. (g) PNN. (h) TFNet. (i) SRCNN. (j) MSDCNN. (k) Proposed SARF.

Fig. 4. Proposed method performs better for moving targets. (a) Resampled LR MS. (b) HR PAN. (c) GS. (d) ATWT-M2. (e) AWLP. (f) MTF-GLP. (g) Proposed SARF with $\lambda = 0$. (h) Proposed SARF with $\lambda = 0.3$.

TABLE II
QUANTITATIVE EVALUATION RESULTS IN THE FR EXPERIMENT

Methods	D_s	D_λ	QNR	GQNR
GS	0.1273	0.0426	0.8355	0.1981
ATWT_M2	0.1336	0.0714	0.8045	0.3702
AWLP	0.1265	0.0820	0.8019	0.4151
MTF_GLP	0.1358	0.0911	0.7855	0.4475
PNN	0.0736	0.0418	0.8877	0.2359
TFNet	0.0736	0.0259	**0.9024**	0.1296
SRCNN	**0.0679**	0.0430	0.8920	0.2029
MSDCNN	0.0925	0.0398	0.8714	0.1652
Proposed_SARF	0.1168	**0.0239**	0.8620	**0.1123**

The "simple" indicates the easy implement. The proposed SARF, CS-, and MRA-based methods have approximate computational efficiency, such as a subsecond computational time for the QuickBird HR PAN (100×100) and LR MS ($25 \times 25 \times 4$) images in the RR experiment. The DL-based methods have competitive advantage on spectral fidelity; however, they are generally time-consuming in the network training. In addition, existing methods generally require retraining for different tasks. The "adjustable" denotes the characteristics of additional adjustable spatial details according to different demands for more spectral fidelity or spatial sharpness. The "robust" represents the advantage of robust performance for different ground surface features. Although the SARF may not achieve the best performance in some cases, such as in terms of some quantitative evaluation indices. It shows more robustness for different surface features. For example, as shown in Fig. 4, based on IKONOS images, the GS, ATWT-M2, AWLP, and MTF-GLP methods perform poorly for moving objects, despite their possible better quantitative evaluation results. The proposed fusion results in Fig. 4(g) and (h) show more robust performance. This is because, on the one hand, the proposed SARF is established on the CS-based fusion framework that is generally robust to the misregistration errors; on the other hand, the proposed spatial detail injection weight has competitive advantage on the spatial enhancement.

In addition, there are some limitations to be further solved. Though the proposed method achieves acceptable and robust spectral information preservation, the spectral fidelity should be further tested and verified for some quantitative remote sensing applications, such as fused images for water parameter inversions, and so on. Besides, the proposed SARF can be

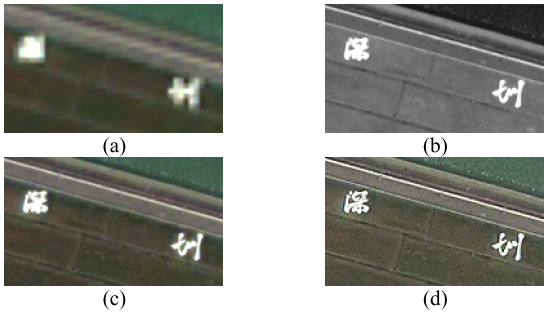

Fig. 5. Influence of the parameter λ on the fused result. (a) Resampled LR MS. (b) HR PAN. (c) Proposed SARF with $\lambda = 0$. (d) Proposed SARF with $\lambda = 0.8$.

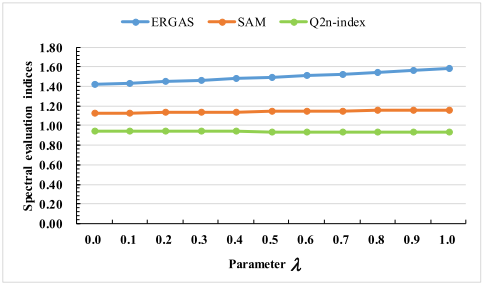

Fig. 6. Influence of the parameter λ on spectral fidelity.

further improved by combining the parallel computing to improve computational efficiency, especially for images with large dimensions.

2) Parameter Analysis: The sole parameter λ is analyzed and discussed. The experiment was implemented on real QuickBird images. Fig. 5 shows the results. It is shown that with the increasing of λ, the fused image is sharper. In addition, the quantitative evaluation results with different λ are shown in Fig. 6. It is shown that with the increasing of the parameter λ, the spectral fidelity is decreasing. Therefore, the adjustable parameter λ should be selected according to different demands for more spectral fidelity or more spatial sharpness. In most cases, the parameter λ can be set between 0 and 0.3.

IV. Conclusion

This letter has presented a simple, adjustable, and robust fusion method. The proposed SARF is based on the CS-based framework, and a spatial-spectral coenhanced strategy is proposed. On the one hand, a spatial enhancement strategy by considering noise disturbance is proposed to provide adjustable details for different applications. On the other hand, a spectral compensation based on the residual of original MS is further proposed to ensure the robustness to different satellites and scenarios. The proposed SARF was tested and verified by both four-band and eight-based satellite images based on RR and FR experiments. The experimental results show the competitive performance of the proposed SARF.

References

[1] 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. 101–113, Mar. 2019.

[2] X. Meng *et al.*, "A large-scale benchmark data set for evaluating pansharpening performance: Overview and implementation," *IEEE Geosci. Remote Sens. Mag.*, early access, Apr. 29, 2020, doi: 10.1109/MGRS.2020.2976696.

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

[4] C. Kwan, J. H. Choi, S. Chan, J. Zhou, and B. Budavari, "Resolution enhancement for hyperspectral images: A super-resolution and fusion approach," presented at the IEEE Int. Conf. Acoust., Speech Signal Process., New Orleans, LA, USA, 2017.

[5] J. Zhou, C. Kwan, and B. Budavari, "Hyperspectral image super-resolution: A hybrid color mapping approach," *J. Appl. Remote Sens.*, vol. 10, no. 3, Sep. 2016, Art. no. 035024.

[6] S. Li and B. Yang, "A new pan-sharpening method using a compressed sensing technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 738–746, Feb. 2011.

[7] X. Meng, H. Shen, Q. Yuan, H. Li, L. Zhang, and W. Sun, "Pansharpening for cloud-contaminated very high-resolution remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 5, pp. 2840–2854, May 2019.

[8] W. Huang, L. Xiao, Z. Wei, H. Liu, and S. Tang, "A new pan-sharpening method with deep neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 5, pp. 1037–1041, May 2015.

[9] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, Jun. 2016.

[10] X. Otazu, M. González-Audícana, O. Fors, and J. Núñez, "Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 10, pp. 2376–2385, Oct. 2005.

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

[12] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. M. Bruce, "Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3012–3021, Oct. 2007.

[13] A. Garzelli, F. Nencini, and L. Capobianco, "Optimal MMSE pan sharpening of very high resolution multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 1, pp. 228–236, Jan. 2008.

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

[15] M. Hasanlou and M. R. Saradjian, "Quality assessment of pan-sharpening methods in high-resolution satellite images using radiometric and geometric index," *Arabian J. Geosci.*, vol. 9, no. 1, pp. 45:1–45:10, Jan. 2016.

[16] 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.

[17] L. Alparone, S. Baronti, A. Garzelli, and F. Nencini, "A global quality measurement of pan-sharpened multispectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 313–317, Oct. 2004.

[18] 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.

[19] C. Kwan, B. Budavari, A. C. Bovik, and G. Marchisio, "Blind quality assessment of fused WorldView-3 images by using the combinations of pansharpening and hypersharpening paradigms," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1835–1839, Oct. 2017.

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

[21] X. Liu, Q. Liu, and Y. Wang, "Remote sensing image fusion based on two-stream fusion network," *Inf. Fusion*, vol. 55, pp. 1–15, Mar. 2020.

[22] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2016.

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