Unsupervised Hyperspectral Pansharpening by Ratio Estimation and Residual Attention Network

Jinyan Nie, Qizhi Xu, and Junjun Pan

***Abstract*— Most deep learning-based hyperspectral pansharpening methods use the hyperspectral images (HSIs) as the ground truth. Training samples are usually obtained by blurring and downsampling the panchromatic image and HSI. However, the blurring and downsampling operation lose much spatial and spectral information. As a result, the model parameters trained by these reduced-resolution samples are unsuitable for fusing full-resolution images. To tackle this problem, we propose an unsupervised hyperspectral pansharpening method via ratio estimation (RE) and residual attention network (RE-RANet). The spatial and spectral information of the fused image are derived from the original panchromatic and HSI rather than reduced-resolution images. At first, we generate the initial ratio image using the ratio enhancement method. The initial ratio image is fine-tuned by the residual attention network (RANet) to generate a multichannel ratio image. Then, we inject the multichannel ratio image that contains spatial detail information into the HSI. Finally, the generated hyperspectral image is constrained by the spatial constraint loss and the spectral constraint loss. Experiments on the EO-1 and Chikusei datasets verify the effectiveness of the proposed method. Compared with other state-of-the-art approaches, our method performs well in qualitative visual effects and quantitative evaluation indicators.**

***Index Terms*— Deep learning, hyperspectral pansharpening, ratio estimation (RE), residual attention network (RANet).**

I. Introduction

Hyperspectral remote sensing can obtain rich spectral information of ground objects, which can be used for target detection, land cover classification, artificial interpretation, and so on. Due to the limitation of physical conditions, there is a trade-off between spectral resolution and spatial resolution of hyperspectral images (HSIs). Hyperspectral pansharpening aims to fuse the HSI and panchromatic image (PAN) to improve the spatial resolution of the HSI, and various methods have been proposed [1].

Traditional hyperspectral pansharpening methods can be roughly divided into four main branches: component substitution (CS), multiresolution analysis (MRA), Bayesian, and matrix factorization (MF). The CS-based pansharpening methods contain principal component analysis (PCA) [2], Gram–Schmidt (GS) [3], and intensity–hue–saturation (IHS) [4], and so on. The CS-based pansharpening methods retain high fidelity of spatial details and are easy to be implemented. However, the spectral mismatch between the PAN and HSI will result in spectral distortion. The classic MRA-based pansharpening methods include decimated wavelet transform (DWT) [5], smoothing filter-based intensity modulation (SFIM) [6], guided filter [7], and Laplacian pyramid [8]. The MRA-based pansharpening methods keep spectral consistency. However, the design and implementation of spatial filters are complex. Because fusion problems are usually ill-posed, Bayesian [9] methods provide a convenient way to regularize the problem by defining an appropriate prior distribution for the scenarios of interest. The MF-based methods [1] depend on the unmixing process. The coupled nonnegative MF (CNMF) [10] is a typical representative of this kind of method.

In recent years, deep learning technology succeed in computer vision and image processing. Naturally, many deep learning-based methods have been introduced for HSI pansharpening tasks. Zheng *et al.* [11] proposed a hyperspectral pansharpening method based on guided filter and deep residual learning. They generated the initial HSI by enhancing the spatial information and then mapping the residual between the initial HSI and the reference HSI to improve fusion accuracy. He *et al.* [12] designed a spectral-fidelity convolutional neural network for hyperspectral pansharpening. They focused on the decomposability of HSI details and introduced a loss function of spectral fidelity. Zheng *et al.* [13] introduced a hyperspectral pansharpening method based on deep prior and dual attention residual network. They upsampled the HSI to the scale of PAN through a deep hyperspectral prior algorithm. Then, they designed a dual-attention residual network to learn spectral and spatial information adaptively. Xie *et al.* [14] proposed a 3-D generative adversarial network (GAN) for hyperspectral pansharpening. They used adversarial learning to search for the optimal high-resolution HSI to fool the discriminator network. Their loss functions contain global constraint, spectral constraint, and spatial constraint. Dong *et al.* [15] proposed a Laplacian pyramid dense network for hyperspectral pansharpening. The subband residuals were extracted from PAN and were injected into the upsampled HSI to reconstruct the high-resolution HSI step by step. This method simplifies the pansharpening problem into several pyramid-level learning issues.

The above-mentioned deep learning pansharpening methods take the original HSI as the ground truth, resulting from a lack of high-resolution HSI for supervision in the training process. And the reduced-resolution PAN and HSI are used

Manuscript received December 22, 2021; accepted January 31, 2022. Date of publication February 4, 2022; date of current version March 1, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 61972021 and Grant 61672076. *(Corresponding author: Qizhi Xu.)*

Jinyan Nie and Junjun Pan are with the State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, China.

Qizhi Xu is with the School of Mechatronical Engineering, Beijing Institute of Technology, Beijing 100181, China (e-mail: qizhi@bit.edu.cn).

Digital Object Identifier 10.1109/LGRS.2022.3149166

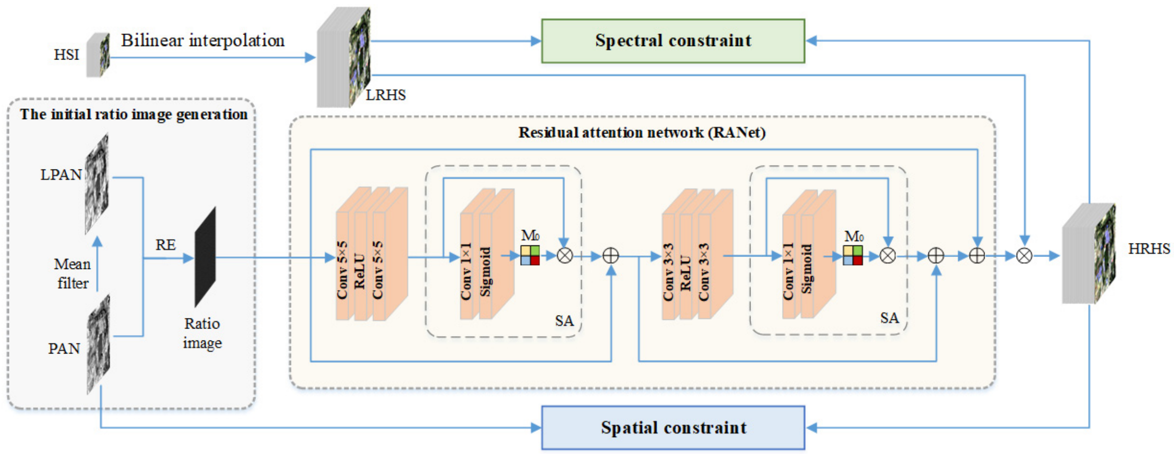

Fig. 1. Flowchart of RE-RANet. ⊕ represents the elementwise summation and ⊗ represents the elementwise multiplication.

as the training samples according to Wald protocol [16]. The relationship between high-resolution and low-resolution HSIs cannot be simply simulated via blurring and downsampling. To tackle these problems, we propose an unsupervised hyperspectral pansharpening method by ratio estimation (RE) and residual attention network (RE-RANet). Our method preserves rich spatial and spectral information by pansharpening on the original HSI and PAN. First, the initial ratio image is generated by the ratio enhancement method and then sent to the residual attention network (RANet) for fine-tuning. Second, the spatial detail information and spectral information are combined by multiplying the upsampled HSI with the generated multichannel ratio image. Third, spatial constraint loss and spectral constraint loss are designed to guide network optimization.

The rest of this letter is organized as follows. The RE-RANet is described in Section II. Experiments and analysis are given in Section III. The conclusion is in Section IV.

II. Methodology

The proposed method mainly contains three parts: 1) generating the initialized ratio image; 2) fine-tuning the ratio image by RANet; and 3) spatial constraint and spectral constraint. The flowchart of the RE-RANet is given in Fig. 1.

A. Ratio Estimation

The RE strategy is inspired by the ratio enhancement pansharpening method [17]. It considers that the ratio of PAN to the degraded image is equal to the ratio of fused image to low-resolution HSI (LRHS). It can be expressed as

$$\frac{P(i,j)}{D(i,j)} = \frac{F_k(i,j)}{H_k(i,j)}, \quad k = 1, 2, \ldots, n \tag{1}$$

where (i, j) is the coordinate of the pixel in the image, P represents the PAN, D is the degraded image, F is the fusion result, H is the upsampled HSI, and k stands for band index.

For the fused images, it can be rewritten as

$$F_k(i,j) = \frac{P(i,j)}{D(i,j)} \times H_k(i,j), \quad k = 1, 2, \ldots, n \tag{2}$$

where n is the number of bands in the HSI.

The hypothesis is that the grayscale information of PAN can be eliminated by dividing the PAN from its degraded image. The ratio of PAN to degraded image preserved the spatial detail information, while the HSI contains spectral information. Multiplying these two can combine the spatial information with spectral information, effectively. Since the degraded image is unknown, the low-resolution PAN (LPAN) is calculated as the estimate of degraded image. To obtain LPAN, mean filter is implemented on PAN, as shown in follows:

$$L(i,j) = (P * M)(i,j) \tag{3}$$

where M represents a mean filter, and L represents the LPAN. The initial ratio image R_e can be expressed as

$$R_e(i,j) = \frac{P(i,j)}{L(i,j)}. \tag{4}$$

Then, the ratio image $R_e(i, j)$ is input into the RANet. The output of the network is a multichannel ratio image R', which has the same number of bands as the HSI. It can be expressed as

$$R'(i,j) = f(R_e(i,j);\ \theta) \tag{5}$$

where $f(\cdot)$ denotes the RANet, θ is the trainable parameters of the network. After obtaining the new multichannel ratio image, it is injected into the HSI to obtain the high-resolution HSI (HRHS), which can be expressed as

$$F_k(i,j) = R'_k(i,j) \times H_k(i,j), \quad k = 1, 2, \ldots, n. \tag{6}$$

B. Architecture of the RANet

The initial ratio image is coarse, so the RANet is designed to fine-tune it and the network is no need to be complex. The architecture of RANet consists of two cascading spatial residual attention blocks (SRAB) and a skip connection. Each SRAB consists of the convolution operation, spatial attention (SA) module, and skip connection. The kernel size of the first and second SRABs are 5×5 and 3×3, respectively. The corresponding number of extracted feature maps is 512 and 256. M_0 in the SA module refers to the spatial attention mask, and it can be obtained via 1×1 convolution and sigmoid layer. The edge and textured areas are rich in high-frequency information, and the smoother areas are rich in low-frequency components, correspondingly. The former is harder to sharpen than the latter. Therefore, the SA module is built to take full advantage of spatial relationships, and skip connection is adopted for feature reuse and improving training stability.

C. Loss Functions

The loss functions consist of spatial constraint loss and spectral constraint loss, which can be defined as

$$L = L_p + \alpha L_h \tag{7}$$

where L_p represents the spatial constraint loss and L_h denotes the spectral constraint loss, α is the tunable parameter to balance L_p and L_h.

The spatial constraint loss L_p can be expressed as

$$L_p = \frac{1}{N}\sum_{i=1}^{N}\left\|h\left(I_f^i\right) - h\left(I_p\right)\right\|_F^2 + \frac{1}{N}\sum_{i=1}^{N}\left\|\nabla\left(I_f^i\right) - \nabla\left(I_p\right)\right\|_F^2 \tag{8}$$

where N is the number of training data, I_f^i is the i^{th} HRHS, I_p denotes the PAN, and $\|\cdot\|_F$ stands for the matrix Frobenius norm. $h(\cdot)$ is the high-pass filter to extract high-frequency information of the image, and $\nabla(\cdot)$ is the gradient operator to obtain the gradient information of the image. The goal of our method is to inject the spatial information of the PAN into the HSI. Because there is no high-resolution HSI as a ground truth, we use the high-frequency information and gradient information of the PAN to constrain the spatial information of the fusion image.

Spectral information of fused images is provided by HSI. The spectral constraint loss L_p can be expressed as

$$L_h = \frac{1}{N}\sum_{i=1}^{N}\left\|\downarrow g\left(I_f^i\right) - I_{hs}^i\right\|_F^2 \tag{9}$$

where $g(\cdot)$ is a Gaussian filter, $\downarrow$ represents the downsampling operation, and I_{hs} denotes the HSI. The purpose of blurring and downsampling the fused image is to degenerate the fused image into a low-resolution HSI, so that the spectral information of the HRHS is consistent with the original HSI.

III. Experiments and Analysis

A. Datasets and Implementation Details

To verify the effectiveness of the algorithm, experiments are conducted on two public datasets. The first one is collected by the Earth Observing-1 (EO-1) satellite. The spatial resolution of Hyperion HSI camera is 30 m, and the spectral resolution is 10 nm with 242 bands ranging from 400 to 2500 nm. The ALI camera produces PAN with a spatial resolution of 10 m. In the experiment, we used 162 bands without water vapor absorption bands and noise bands. The HSI and PAN use for training are $1200 \times 182 \times 162$ pixels and 3600×546 pixels, respectively. The HSI and PAN use for the test are $133 \times 133 \times 162$ pixels and 399×399 pixels, respectively. The second dataset is the Chikusei dataset [18]. It was collected by the Headwall Hyperspec-VNIR-C imaging sensor in Chikusei, Ibaraki, Japan. It provides a spatial resolution of 2.5 m, and 128 spectral bands ranging from 363 to 1018 nm. After noise bands elimination, 124 bands remain for the experiment. Since the Chikusei dataset does not contain PAN, we utilized the synthesized PAN and HSI. In training, the size of HSI and PAN are $150 \times 150 \times 124$ pixels and 450×450 pixels, respectively. In the test, the size of HSI is $100 \times 70 \times 124$ pixels, and the PAN is 300×210 pixels.

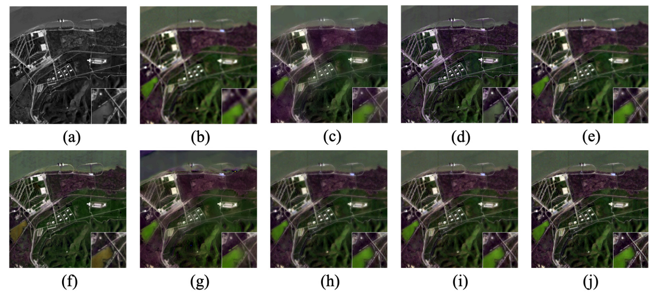

Fig. 2. Visual results of different methods in EO-1. (a) Pan, (b) upsampled HSI, (c) GFPCA, (d) HySure, (e) SFIM, (f) CNMF, (g) DRCNN, (h) DHP-DARN, (i) DRSAN, and (j) proposed method. The false color image is chosen for clear visualization (red: 29, green: 23, blue: 16).

TABLE I

Quantitative Assessment for EO-1 Dataset

Method	Q (↑)	ERGAS (↓)	SAM (↓)	CC (↑)	RMSE (↓)	Time(s) (↓)
GFPCA	0.7789	8.0542	7.3130	0.9172	11.3118	2.1498
HySure	0.8849	6.7296	5.4725	0.9011	8.6989	84.0886
SFIM	0.7674	8.8168	6.0191	0.8997	14.6828	2.3386
CNMF	0.8356	6.9783	5.5207	0.8849	10.6556	16.1294
DRCNN	0.7670	27.7391	15.5770	0.6462	27.4370	**1.1289**
DHP-DARN	0.8367	5.4264	6.0253	0.8793	7.9583	2.6937
DRSAN	0.8524	5.8395	5.9374	0.8532	8.2947	2.9343
Proposed	**0.9104**	**5.1973**	**5.2649**	**0.9221**	**7.3834**	2.1149

The size of HSI patches used in the training model is 30×30 pixels, and the PAN patches are 90×90 pixels. We set the batch size to 16 and the initialized learning rate to 0.001. The decay rate is set to 0.99 with decay step 10 000. The RMSProp optimizer is adopted. In loss functions, α is set to 1.

B. Evaluation Metrics and Compared Methods

To evaluate the fusion quality of different methods objectively, several quality evaluation metrics are used, including quality index (Q) [19], erreur relative globale adimensinnelle de synthèse (ERGAS) [20], spectral angle map (SAM) [21], cross correlation (CC) [22], root mean squared error (RMSE) [1], and time [15].

To verify the effectiveness of the proposed method, seven different hyperspectral pansharpening methods are used for comparison, including guided filter PCA (GFPCA) [23], HySure [24], SFIM [25], CNMF [26], deep residual convolutional neural network (DRCNN) [27], deep hyperspectral prior and dual-attention residual network (DHP-DARN) [13], and deep residual spatial attention network (DRSAN) [28]. In addition, the traditional pansharpening methods are per-

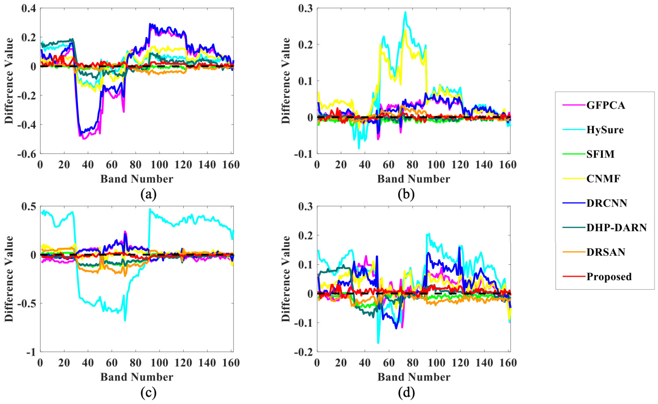

Fig. 3. Spectral reflectance difference values at four randomly selected locations on EO-1 dataset. Coordinate (a) (129, 226), (b) (138, 72), (c) (89, 90), and (d) (158, 76).

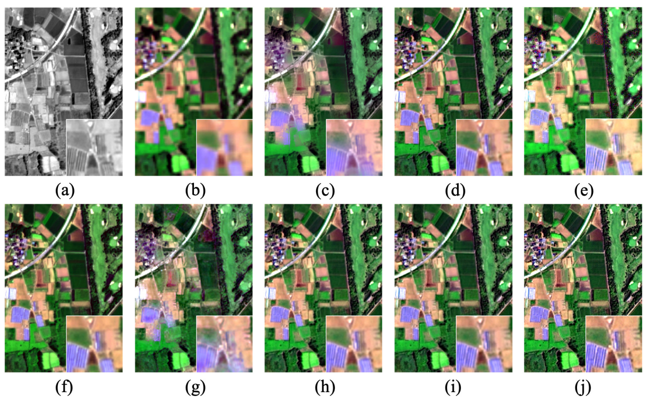

Fig. 4. Visual results of different methods in Chikusei dataset. (a) Pan, (b) upsampled HSI, (c) GFPCA, (d) HySure, (e) SFIM, (f) CNMF, (g) DRCNN, (h) DHP-DARN, (i) DRSAN, and (j) proposed method. The false color image is chosen for clear visualization (red: 60, green: 40, blue: 20).

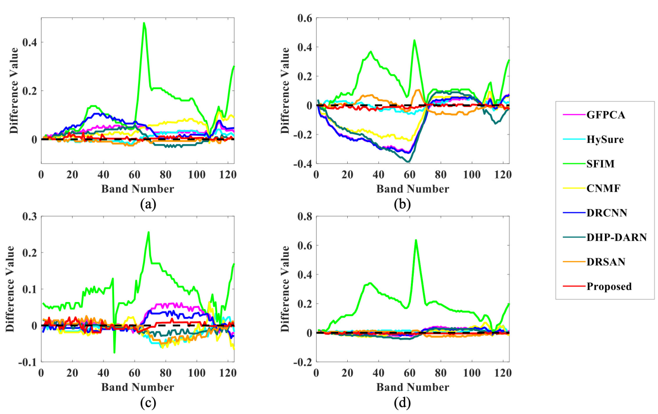

Fig. 5. Spectral reflectance difference values at four randomly selected locations on Chikusei dataset. Coordinate (a) (42, 189). (b) (58, 93), (c) (90, 147), and (d) (20, 149).

formed on Intel Core i7-8700 CPU at 3.20 GHz. The deep learning pansharpening methods are performed on NVIDIA GeForce GTX 1080Ti GPU.

C. Experiments and Discussion

The visual results on the EO-1 dataset are illustrated in Fig. 2. It is observed that the proposed method has good visual results. The GFPCA and SFIM have good color fidelity, but

TABLE II

QUANTITATIVE ASSESSMENT FOR CHIKUSEI DATASET

Method	Q (↑)	ERGAS (↓)	SAM (↓)	CC (↑)	RMSE (↓)	Time(s) (↓)
GFPCA	0.6996	7.7778	5.2997	0.8094	10.4624	0.6142
HySure	0.8201	4.3335	5.2672	0.8256	9.3447	28.1752
SFIM	0.8202	9.6464	4.1682	0.8439	15.1758	**0.4894**
CNMF	0.8946	6.2779	3.7672	0.9146	8.5365	3.8101
DRCNN	0.7630	8.4678	4.2339	0.8089	19.3449	0.5679
DHP-DARN	0.8736	4.3654	3.7864	0.8937	8.2941	0.6934
DRSAN	0.8683	4.5246	3.9532	0.8796	8.4892	0.7247
proposed	**0.9028**	**4.1904**	**3.6381**	**0.9298**	**7.3851**	0.6534

TABLE III

QUANTITATIVE RESULTS WITH DIFFERENT PARAMETERS

Data set	α	Q (↑)	ERGAS (↓)	SAM (↓)	CC (↑)	RMSE (↓)
EO-1	0.1	7.4749	7.0828	7.5784	0.8375	8.9346
	0.5	0.8895	5.2749	6.8053	0.9142	7.7804
	1	**0.9104**	**5.1973**	**5.2649**	**0.9221**	**7.3834**
	5	0.7949	6.7854	5.8923	0.8642	9.3585
Chikusei	0.1	0.8234	6.7384	5.9649	0.8197	9.5538
	0.5	0.9011	5.7876	4.8346	0.8879	8.6835
	1	**0.9028**	**4.1904**	**3.6381**	**0.9298**	**7.385**
	5	0.8532	6.9678	4.0896	0.8563	8.9463

the spatial details are blurred. The spatial detail of HySure and CNMF are clear, but their color fidelity are poor, especially in the lake area of zoom. The spatial details and color fidelity of DRCNN are mediocre. Especially on the shore, there is significant distortion. DHP-DARN and DRSAN are a little fuzzy compared to the proposed method. Quantitative assessments of different methods on the EO-1 dataset are given in Table I. It can be found that the proposed method obtains the most significant Q and CC values, and the smallest ERGAS, SAM, and RMSE values. In comparison to the traditional methods such as GFPCA, HySure, SFIM, and CNMF, the testing time of proposed method is faster. In comparison to the deep learning methods, the proposed algorithm is more time-consuming than DRCNN, but significantly improves the fusion effect. The spectral reflectance difference values of different methods are compared to verify the spectral preservation capability. In each subfigure, the spectral reflectance difference closer to the baseline (0), the better the spectrum is preserved. On the EO-1 dataset, it can be seen intuitively from Fig. 3 that the spectral reflectance difference values of the proposed method is close to the baseline. Compared with other methods, the spectrum of the proposed algorithm is better preserved.

Fig. 4 shows the visual results of different methods on the Chikusei dataset. We can see that the GFPCA, SFIM, and DRCNN have poor spatial detail preservation and color fidelity. Both HySure, CNMF, DHP-DARN, and DRSAN and the proposed method have fine spatial detail and color fidelity. However, the detailed reconstruction of the proposed algorithm is best. Quantitative assessments of different methods on the Chikusei dataset are given in Table II. It can be seen that our method obtains the most significant Q and CC values, and the smallest ERGAS, SAM, and RMSE values. In addition, the testing time of the proposed algorithm is above average compared with other methods. The comparisons of spectral reflectance difference values on Chikusei are given in Fig. 5.

It can be found that the fluctuation of the spectral reflectance difference values of the proposed method is smaller than that of other methods. That is to say, the spectral distortion of the proposed method is smaller than other methods.

In order to explore the influence of parameters in (7) on the fusion effect, different parameters are selected for experiments, as shown in Table III. It can be seen that the evaluation index of $\alpha = 1$ is better than other parameters. Therefore, $\alpha = 1$ is set in the experiment to balance spatial constraint loss and spectral constraint loss.

IV. Conclusion

In this letter, an unsupervised HSI pansharpening method based on RE-RANet is proposed. Our method directly fuses on the original images rather than on the reduced-resolution images. The main contribution of our method is fine-tuning the coarse ratio image via the RANet and injecting spatial details into HSI. Additionally, a spatial constraint loss and a spectral constraint loss are proposed to preserve the spatial and spectral information in the fused images. The experimental results demonstrated that RE-RANet achieved a satisfactory fusion performance in visual comparison and quantitative evaluation. This method uses bilinear interpolation to upsample HSI, which will blur the edges of HSI and reduce the fusion accuracy. In the future, we plan to use the deep image prior technique in the upsampling process and reduce the spectral and spatial distortion before fusion.

Acknowledgment

The authors gratefully acknowledge Space Application Laboratory, Department of Advanced Interdisciplinary Studies, The University of Tokyo, for providing the Chikusei data.

References

[1] L. Loncan *et al.*, “Hyperspectral pansharpening: A review,” *IEEE Trans. Geosci. Remote Sens.*, vol. 3, no. 3, pp. 27–46, Sep. 2015.

[2] J. Qu, Y. Li, and W. Dong, “Guided filter and principal component analysis hybrid method for hyperspectral pansharpening,” *J. Appl. Remote Sens.*, vol. 12, no. 1, p. 015003, Jan. 2018.

[3] Z. Huang, Q. Chen, Q. Chen, and X. Liu, “Variational pansharpening for hyperspectral imagery constrained by spectral shape and Gram–Schmidt transformation,” *Sensors*, vol. 18, no. 12, p. 4330, 2018.

[4] M. Ghahremani and H. Ghassemian, “Nonlinear IHS: A promising method for pan-sharpening,” *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 11, pp. 1606–1610, Nov. 2016.

[5] T. Zaveri, I. Makwana, and M. Zaveri, “A fuzzy based hybrid multispectral image fusion method using DWT,” in *Proc. 10th Int. Conf. Hybrid Intell. Syst.*, Aug. 2010, pp. 13–18.

[6] B. Han and Y. Zhao, “An improved smoothing filter-based intensity modulation algorithm for hyperspectral image fusion,” *Remote Sens. Inf.*, vol. 27, no. 5, pp. 44–47, Oct. 2012.

[7] J. Qu, Y. Li, and W. Dong, “Hyperspectral pansharpening with guided filter,” *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 11, pp. 2152–2156, Nov. 2017.

[8] P. Addesso, R. Restaino, and G. Vivone, “An improved version of the generalized Laplacian pyramid algorithm for pansharpening,” *Remote Sens.*, vol. 13, no. 17, p. 3386, Aug. 2021.

[9] 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.

[10] C.-H. Lin, F. Ma, C.-Y. Chi, and C.-H. Hsieh, “A convex optimization-based coupled nonnegative matrix factorization algorithm for hyperspectral and multispectral data fusion,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1652–1667, Mar. 2018.

[11] Y. Zheng, J. Li, and Y. Li, “Hyperspectral pansharpening based on guided filter and deep residual learning,” in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2019, pp. 616–619.

[12] L. He, J. Zhu, J. Li, D. Meng, J. Chanussot, and A. Plaza, “Spectral-fidelity convolutional neural networks for hyperspectral pansharpening,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 5898–5914, Sep. 2020.

[13] Y. Zheng, J. Li, Y. Li, G. Jie, X. Wu, and J. Chanussot, “Hyperspectral pansharpening using deep prior and dual attention residual network,” *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 11, pp. 8059–8076, Nov. 2020.

[14] W. Xie, Y. Cui, Y. Li, J. Lei, and J. Li, “HPGAN: Hyperspectral pansharpening using 3-D generative adversarial networks,” *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 1, pp. 463–477, Jan. 2021.

[15] W. Dong, T. Zhang, J. Qu, S. Xiao, J. Liang, and Y. Li, “Laplacian pyramid dense network for hyperspectral pansharpening,” *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–13, 2022.

[16] L. Wald, T. Ranchin, and M. Mangolini, “Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images,” *Photogramm. Eng. Remote Sens.*, vol. 63, no. 6, pp. 691–699, 1997.

[17] Q. Xu, Y. Zhang, B. Li, and L. Ding, “Pansharpening using regression of classified MS and pan images to reduce color distortion,” *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 1, pp. 28–32, Jan. 2015.

[18] N. Yokoya and A. Iwasaki, “Airborne hyperspectral data over chikusei,” Space Appl. Lab., Univ. Tokyo, Tokyo, Japan, Tech. Rep. SAL-2016-05-27, May 2016. [Online]. Available: http://park.itc.u-tokyo.ac.jp/sal/hyperdata/TechRepSAL20160527.pdf

[19] Z. Wang and A. C. Bovik, “A universal image quality index,” *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.

[20] L. Wald, *Data Fusion: Definitions and Architectures: Fusion of Images of Different Spatial Resolutions*. Paris, France: Presses des MINES, 2002.

[21] R. H. Yuhas, A. F. H. Goetz, and J. W. Boardman, “Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm,” in *Proc. JPL, Summaries 3rd Annu. JPL Airborne Geosci. Workshop, AVIRIS Workshop*, vol. 1, 1992.

[22] 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, 1998.

[23] L. Wenzhi *et al.*, “Processing of multiresolution thermal hyperspectral and digital color data: Outcome of the 2014 IEEE GRSS data fusion contest,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2984–2996, Jun. 2015.

[24] M. Simoes, J. Bioucas-Dias, L. B. Almeida, and J. Chanussot, “A convex formulation for hyperspectral image superresolution via subspace-based regularization,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 6, pp. 3373–3388, Jun. 2015.

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

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

[27] Y. Zheng, J. Li, Y. Li, K. Cao, and K. Wang, “Deep residual learning for boosting the accuracy of hyperspectral pansharpening,” *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 8, pp. 1435–1439, Aug. 2020.

[28] Y. Zheng, J. Li, Y. Li, Y. Shi, and J. Qu, “Deep residual spatial attention network for hyperspectral pansharpening,” in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Sep. 2020, pp. 2671–2674.