Sparsity Constrained Fusion of Hyperspectral and Multispectral Images

Xiyou Fu, *Member, IEEE*, Sen Jia, *Senior Member, IEEE*, Meng Xu, *Member, IEEE*, Jun Zhou, *Senior Member, IEEE*, and Qingquan Li

***Abstract*— Fusing a Hyperspectral image (HSI) and a multispectral image (MSI) from different sensors is an economic and effective approach to get an image with both high spatial and spectral resolution, but localized changes between the multiplatform images can have negative impacts on the fusion. In this letter, we propose a novel sparsity constrained fusion method (SCFus) to fuse multiplatform HSIs and MSIs based on matrix factorization. Specifically, we imposed ℓ_1 norm on the residual term of the MSI to account for the localized changes between the hyperspectral and MSIs. Furthermore, we plugged a state-of-the-art denoiser, namely block-matching and 3-D filtering (BM3D), as the prior of the subspace coefficients by exploiting the plug-and-play framework. We refer to the proposed method as SCFus for hyperspectral and MSIs. Experimental results suggest that the proposed fusion method is more effective in fusing hyperspectral and MSIs than the competitors.**

***Index Terms*— Hyperspectral image (HSI), image changes, image fusion, multiplatform data.**

I. Introduction

THE high spectral resolution of hyperspectral images (HSIs) allows the precise recognition and identification of the materials present in the images and promotes remarkable applications in remote sensing, such as land cover classification, spectral unmixing [1], and anomaly detection [2]. However, the relative low spatial resolution hinders the further application of HSIs. An economic and effective approach is to fuse a low-spatial-resolution HSI (LR-HSI) and a high-spatial-resolution MSI (HR-MSI) to produce a high-spatial-resolution HSI (HR-HSI). Although there emerge some platforms carrying both hyperspectral and multispectral imaging sensors, the number of this kind of platforms is still limited [3]. Thus, fusing multiplatform hyperspectral and multispectral images (MSIs) has great potentials for practical applications.

In the past decades, many methods have been proposed to fuse hyperspectral and MSIs, such as matrix factorization-based methods [4]–[6], tensor factorization-based methods [7], [8], deep learning-based methods [9]–[11], and pan-sharpening methods extended for fusion of HSIs and MSIs [3], [12]. Most of the fusion methods work under an assumption that the observed images were acquired at almost the same observing conditions, implying that the LR-HSI and the HR-MSI can be viewed as spatially and spectrally downsampled versions of a unique underlying HR-HSI, respectively [13]. However, this assumption is not always met in practical applications, especially when fusing multiplatform data. It may happen that the observed HSI and MSI were acquired at long time intervals and different point of view. In this sense, the LR-HSI and the HR-MSI cannot be simply viewed as spatially and spectrally downsampled versions of a unique underlying HR-HSI because localized spatial and spectral changes will occur during the acquisition of the observed HSI and MSI, which will have negative impacts on the fusion results and have been neglected by many methods. Recently, Borsoi *et al.* [13] proposed a matrix factorization-based fusion method accounting for spectral changes. Then, a coupled tensor approximation fusion method [14] accounting for localized image changes was proposed. The performance improvements of these two methods highlight the importance of taking the inter-image changes into account when fusing multiplatform data.

In this letter, we propose a novel sparsity constrained method to fuse hyperspectral and MSIs, considering the presence of localized image changes between the observed HSIs and MSIs. We summarized the main contributions of this letter as follows.

1) To better fuse multiplatform hyperspectral and MSIs with the presence of localized image changes, we proposed a sparsity constrained fusion method (SCFus) based on matrix factorization by imposing the ℓ_1 norm on the residual term of the MSI to promote the sparsity, which is different from [15], [16].
2) The spectral low-rankness and spatial self-similarity of the underlying HR-HSI are exploited by using a subspace representation and plug-and-play technique, respectively. We plugged the block-matching and

Manuscript received October 12, 2021; revised January 9, 2022; accepted January 20, 2022. Date of publication January 25, 2022; date of current version February 9, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 41971300, Grant 61901278, and Grant 62001303; in part by the Key Project of Department of Education of Guangdong Province under Grant 2020ZDZX3045; in part by the Natural Science Foundation of Guangdong Province under Grant 2021A1515011413; in part by the China Postdoctoral Science Foundation under Grant 2021M692162; and in part by the Shenzhen Scientific Research and Development Funding Program under Grant 20200803152531004. *(Corresponding author: Sen Jia.)*

Xiyou Fu, Sen Jia, and Meng Xu are with the College of Computer Science and Software Engineering and the Key Laboratory for Geo-Environmental Monitoring of Coastal Zone of the Ministry of Natural Resources, Shenzhen University, Shenzhen 518060, China (e-mail: fuxiyou@qq.com; senjia@szu.edu.cn; m.xu@szu.edu.cn).

Jun Zhou is with the School of Information and Communication Technology, Griffith University, Nathan, QLD 4111, Australia (e-mail: jun.zhou@griffith.edu.au).

Qingquan Li is with the Key Laboratory for Geo-Environmental Monitoring of Coastal Zone of the Ministry of Natural Resources, Shenzhen University, Shenzhen 518060, China (e-mail: liqq@szu.edu.cn).

Digital Object Identifier 10.1109/LGRS.2022.3146248

3-D filtering (BM3D) denoiser [17] as image prior in the model, which proved to be effective in regularizing the subspace coefficients.

Extensive experiments were performed on three datasets in comparison with seven state-of-the-art fusion methods to fully evaluate the performance of the proposed method. We refer to the proposed method as SCFus for hyperspectral and MSIs.

II. Proposed Method

A. Observation Model

Let $\mathbf{Z} \in \mathbb{R}^{L\times WH}$ denote the HR-HSI, where W, H, and L denote the number of rows, columns, and bands in the HR-HSI, respectively. $\mathbf{Y}_h \in \mathbb{R}^{L\times wh}$, $\mathbf{Y}_m \in \mathbb{R}^{l\times WH}$ denote the matrices of LR-HSI, and HR-MSI, respectively, where w, h denote the number of rows, columns in LR-HSI, and l denote the number bands in HR-MSI, respectively. We have $W > w$, $H > h$, and $L > l$.

The LR-HSI $\mathbf{Y}_h$ can be represented as the spatially downsampled version of HR-HSI, $\mathbf{Z}$, as

$$\mathbf{Y}_h = \mathbf{ZBS} + \mathbf{N}_h \tag{1}$$

where $\mathbf{B} \in \mathbb{R}^{WH\times WH}$ is a blur kernel, $\mathbf{S} \in \mathbb{R}^{WH\times wh}$ is the downsampling matrix, and $\mathbf{N}_h \in \mathbb{R}^{L\times wh}$ is the additive Gaussian noise.

The HR-MSI, $\mathbf{Y}_m$, is the spectral downsampled version of $\mathbf{Z}$ and can be represented as

$$\mathbf{Y}_m = \mathbf{RZ} + \mathbf{N}_m \tag{2}$$

where $\mathbf{R} \in \mathbb{R}^{l\times L}$ is the spectral response function, and $\mathbf{N}_m \in \mathbb{R}^{l\times WH}$. If there is no image changes [13], [14] occurring between the acquisition of the observed LR-HSI and HR-MSI, $\mathbf{N}_m$ represents additive Gaussian noise. Otherwise, $\mathbf{N}_m$ accounts for not only additive Gaussian noise but also image changes. We obtained intensity maps of $\mathbf{N}_m$ for three pairs of real multiplatform images, i.e., Paris data, Lake Tahoe data, and Ivanpah Playa data, by computing $\mathbf{1}_l^T \cdot abs(\mathbf{N}_m)$,[1] followed by reshaping the row vectors into $W \times H$ matrices, as shown in the third column of Fig. 1. Pixels of large values in the third column of Fig. 1 indicate significant changes between HR-HSI and HR-MSI. We can see that localized changes exists in three datasets, presenting sparsity in varying degrees. The Paris scene exhibits very few changes. The Lake Tahoe scene and Ivanpah Playa scene show significant changes. We present the histograms of intensity maps of $\mathbf{N}_m$ for three datasets in the final column of Fig. 1. It can be clearly seen that most values are zero or near zero. Due to the presence of localized changes, there exist some pixels with large values, showing sparsity in varying degrees. This motivates us to take advantage of the sparsity characteristics of $\mathbf{N}_m$ to improve the fusion performance.

B. Sparsity Constrained Fusion Model

The HR-HSI can be represented by the subspace and its corresponding coefficients, which can be written as

$$\mathbf{Z} = \mathbf{DA} \tag{3}$$

[1]Function $abs(\mathbf{N}_m)$ outputs the absolute values of $\mathbf{N}_m$.

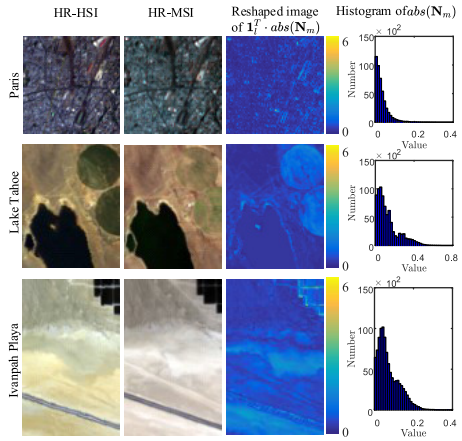

Fig. 1. Illustration of the existence of localized image changes between observed images.

where $\mathbf{D} \in \mathbb{R}^{L\times k}$, $\mathbf{A} \in \mathbb{R}^{k\times WH}$ are the subspace and the corresponding coefficients, respectively. Since LR-HSI retains most of the spectral information of HR-HSI, the subspace can be estimated from LR-HSI using subspace learning methods, i.e., singular value decomposition (SVD) [5].

Then, the objective function of the proposed group sparsity constrained fusion method is written as

$$\min_{\mathbf{A}} \frac{1}{2}\|\mathbf{Y}_h - \mathbf{DABS}\|_F^2 + \frac{\lambda}{2}\|\mathbf{Y}_m - \mathbf{RDA}\|_1 + \beta\phi(\mathbf{A}) \tag{4}$$

where $\phi(\cdot)$ is the regularizer defined to further enhance the spatial correlation of subspace coefficients $\mathbf{A}$, $\lambda > 0$ and $\beta > 0$ are parameters to regularize the group sparsity property of $(\mathbf{Y}_m - \mathbf{RDA})$ and the image prior of $\mathbf{A}$. The ℓ_1 norm, $\|\cdot\|_1$, is introduced to promote the sparsity of $(\mathbf{Y}_m - \mathbf{RDA})$.

After introducing two auxiliary variables, i.e., $\mathbf{V}_1 = \mathbf{Y}_m - \mathbf{RDA}$, $\mathbf{V}_2 = \mathbf{A}$, we can obtain the augmented Lagrangian function

$$\begin{aligned} &L(\mathbf{A}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{G}_1, \mathbf{G}_2) \\ &= \frac{1}{2}\|\mathbf{Y}_h - \mathbf{DABS}\|_F^2 + \frac{\lambda}{2}\|\mathbf{V}_1\|_1 \\ &\quad + \beta\phi(\mathbf{V}_2) + \frac{\mu}{2}\left\|\mathbf{V}_1 - \mathbf{Y}_m + \mathbf{RDA} + \frac{\mathbf{G}_1}{\mu}\right\|_F^2 \\ &\quad + \frac{\mu}{2}\left\|\mathbf{V}_2 - \mathbf{A} + \frac{\mathbf{G}_2}{\mu}\right\|_F^2 \end{aligned} \tag{5}$$

where μ denotes the penalty parameter and $\mathbf{G}_1$, $\mathbf{G}_2$ are the Lagrangian multipliers. By referring to the alternative method of multipliers (ADMMs), we can split (5) into several subproblems and solve each subproblem alternatively.

1) Subproblem of $\mathbf{A}$*:* Subproblem of $\mathbf{A}$ can be solved by minimizing $L(\mathbf{A}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{G}_1, \mathbf{G}_2)$ with respect to $\mathbf{A}$, such that

$$\begin{aligned} \mathbf{A} = \arg\min_{\mathbf{A}} &\frac{1}{2}\|\mathbf{Y}_h - \mathbf{DABS}\|_F^2 \\ &+ \frac{\mu}{2}\left\|\mathbf{V}_1 - \mathbf{Y}_m + \mathbf{RDA} + \frac{\mathbf{G}_1}{\mu}\right\|_F^2 \\ &+ \frac{\mu}{2}\left\|\mathbf{V}_2 - \mathbf{A} + \frac{\mathbf{G}_2}{\mu}\right\|_F^2. \end{aligned} \tag{6}$$

By setting the derivative of (6) with respect to $\mathbf{A}$ to zero, we can finally acquire a Sylvester equation, which can be solved analytically and efficiently as done in [15].

2) Subproblem of $\mathbf{V}_1$*:* Subproblem of $\mathbf{V}_1$ can be solved by minimizing $L(\mathbf{A}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{G}_1, \mathbf{G}_2)$ with respect to $\mathbf{V}_1$, such that

$$\mathbf{V}_1 = \arg\min_{\mathbf{V}_1} \frac{\lambda}{2}\|\mathbf{V}_1\|_1 + \frac{\mu}{2}\left\|\mathbf{V}_1 - \mathbf{Y}_m + \mathbf{RDA} + \frac{\mathbf{G}_1}{\mu}\right\|_F^2 \quad (7)$$

which has a close-form solution by using the ℓ_1 norm minimization operator [18].

3) Subproblem of $\mathbf{V}_2$*:* Subproblem of $\mathbf{V}_2$ can be solved by minimizing $L(\mathbf{A}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{G}_1, \mathbf{G}_2)$ with respect to $\mathbf{V}_2$ as

$$\mathbf{V}_2 = \arg\min_{\mathbf{V}_2} \beta\phi(\mathbf{V}_2) + \frac{\mu}{2}\left\|\mathbf{V}_2 - \mathbf{A} + \frac{\mathbf{G}_2}{\mu}\right\|_F^2. \quad (8)$$

Based on the plug-and-play technique, the solution of the problem (8) can be obtained by using a denoising operator as

$$\mathbf{V}_2 \leftarrow \text{Denoiser}\left(\mathbf{A} - \frac{\mathbf{G}_2}{\mu}, \frac{\beta}{\mu}\right) \quad (9)$$

where Denoiser$(\cdot)$ is a plugged denoiser, $(\mathbf{A} - \mathbf{G}_2/\mu)$ is the input noisy data to be denoised, and β/μ is the input noise level of the denoiser. We use the BM3D [17] as the plugged denoiser to solve the problem (9). BM3D is a state-of-the-art image denoiser based on an enhanced sparse representation in transform domain. It is composed of grouping, collaborative filtering, and aggregation. The grouping is realized by block-matching and the collaborative filtering is accomplished by shrinkage in a 3-D transform domain.

4) Updating of $\mathbf{G}_1$*:* We update $\mathbf{G}_1$ in each iteration as

$$\mathbf{G}_1 = \mathbf{G}_1 + \mu(\mathbf{V}_1 - \mathbf{Y}_m + \mathbf{RDA}). \quad (10)$$

5) Updating of $\mathbf{G}_2$*:* We update $\mathbf{G}_2$ in each iteration as

$$\mathbf{G}_2 = \mathbf{G}_2 + \mu(\mathbf{V}_2 - \mathbf{A}). \quad (11)$$

III. Experiments

In this section, we present experiments to assess the performance of the proposed method for fusing HSIs and MSIs.

A. Experimental Datasets

We evaluated the performance of the proposed fusion method using three multiplatform datasets.

1) Paris data is composed of an HSI captured by the Hyperion and an MSI obtained by the Advanced Land Imager instruments on board the Earth Observing-1 Mission satellite [5], as shown in Fig 1. Both images have a size of 72 × 72 pixels. The HSI has 128 spectral bands. The MSI has nine bands.
2) Lake Tahoe data are composed of an HSI acquired by the airborne visible infrared imaging spectrometer (AVIRIS) instrument and an MSI captured by the Sentinel-2A instrument [13], as shown in Fig. 1. Both images have a size of 100 × 80 pixels. The HSI has 173 spectral bands. The MSI has ten bands.
3) Ivanpah Playa data is composed of an HSI acquired by the AVIRIS instrument, and an MSI captured by the Sentinel-2A instrument [13], as shown in Fig. 1. Both images have a size of 128 × 80 pixels. The HSI has 173 spectral bands. The MSI has ten bands.

The LR-HSI was generated by applying a Gaussian filter, with its variance equal to 1, to the HR-HSI, followed by downsampling every two pixels. Then, white Gaussian noise (WGN) was added to the LR-HSI and HR-MSI to obtain the observed LR-HSI and HR-MSI with a 30 and 40 dB SNR, respectively. The spectral response function was obtained as a known prior as in [13].

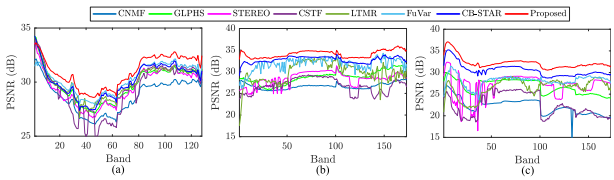

Fig. 2. Band-wise PSNR for three datasets. (a) Paris. (b) Lake Tahoe. (c) Ivanpah Playa.

B. Compared Methods

We compared the proposed method with seven state-of-the-art fusion methods. Generalized Laplacian pyramid hyperspharpening (GLPHS) [3] is an adaptation of the generalized Laplacian pyramid (GLP) method [12] for the fusion of HSIs and MSIs. Among the matrix factorization-based methods, we compared the coupled nonnegative matrix factorization (CNMF) [4], low tensor multi-rank (LTMR) [6], and fusion with spectral variability (FuVar) [13]. For tensor factorization-based methods, we considered super-resolution tensor-reconstruction (STEREO) [7], coupled sparse tensor factorization (CSTF) [8], and coupled tucker decompositions for hyperspectral super-resolution with variability (CB-STAR) [14]. Note that FuVar and CB-STAR were proposed recently accounts for spatially and spectrally localized changes into the fusion. The parameters of the compared methods were set according to their original references.

C. Experimental Results

We show the mean peak signal-to-noise ratio (MPSNR), root mean square error (RMSE), spectral angle mapper (SAM), Erreur Relative Globale Adimensionnelle de Synthèse (ERGAS), and mean universal image quality index (MUIQI) of the recovered HSIs for the three datasets in Table I, where the best results were highlighted in bold. In terms of the five evaluation indexes, it can be clearly seen that FuVar, CB-STAR, and the proposed method achieved better reconstruction performance than other methods because they take spatially and spectrally localized changes into consideration while fusing. Among them, the proposed method achieved considerable better overall reconstruction performance than FuVar and CB-STAR for the three datasets. The reason why the proposed method achieved much better fusion performance boils down to the introduction of ℓ_1 norm to incorporate the image changes between hyperspectral and MSIs and the introduction of the prior to effectively regularize the subspace coefficients.

The peak signal to noise ratio (PSNR) values of each restoration band for the three datasets are presented in Fig. 2. It can be clearly observed from Fig. 2 that the PNSR values of the proposed method are significant higher than that of the competitors at almost all the bands.

For a visual comparison, we present the reconstructed images and corresponding error images of the 20th band for Paris data, tenth band for Lake Tahoe data, and 30th band

TABLE I
QUANTITATIVE INDEXES OF THE PROPOSED METHOD AND COMPARED METHODS ON TESTED DATASETS

Data	Indexes	CNMF	GLPHS	STEREO	CSTF	LTMR	FuVar	CB-STAR	Proposed
Paris	**MPSNR**	28.71	29.82	29.56	29.35	29.96	30.33	30.12	**30.99**
	RMSE	11.74	10.31	10.58	11.05	10.09	9.56	9.92	**8.92**
	SAM	3.05	2.80	3.30	3.29	2.85	2.72	2.78	**2.56**
	ERGAS	6.32	5.57	5.69	5.86	5.47	5.19	5.37	**4.81**
	MUIQI	0.84	0.88	0.87	0.87	0.88	0.90	0.89	**0.91**
Lake Tahoe	**MPSNR**	22.25	26.76	27.09	22.70	26.51	28.17	30.54	**32.23**
	RMSE	21.70	12.99	14.06	22.69	14.20	11.08	8.10	**6.71**
	SAM	7.99	4.91	7.24	7.28	6.80	4.30	3.75	**2.26**
	ERGAS	9.55	5.65	5.70	8.74	5.86	4.76	3.56	**2.91**
	MUIQI	0.86	0.95	0.95	0.92	0.95	0.96	0.98	**0.99**
Ivanpah Playa	**MPSNR**	26.66	28.78	28.45	26.86	29.77	31.82	32.80	**34.13**
	RMSE	12.11	9.73	10.21	12.26	9.35	6.96	6.01	**5.16**
	SAM	1.05	1.49	2.00	1.79	2.16	1.54	1.15	**0.76**
	ERGAS	3.22	2.56	2.68	3.14	2.51	1.84	1.59	**1.36**
	MUIQI	0.76	0.85	0.85	0.84	0.88	0.94	0.95	**0.97**

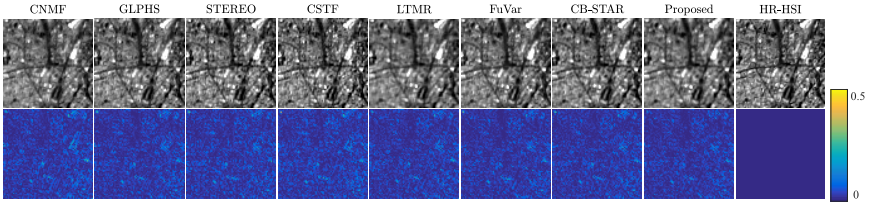

Fig. 3. Twentieth band of reconstructed images (first row) and the corresponding error images (second row) for Paris data.

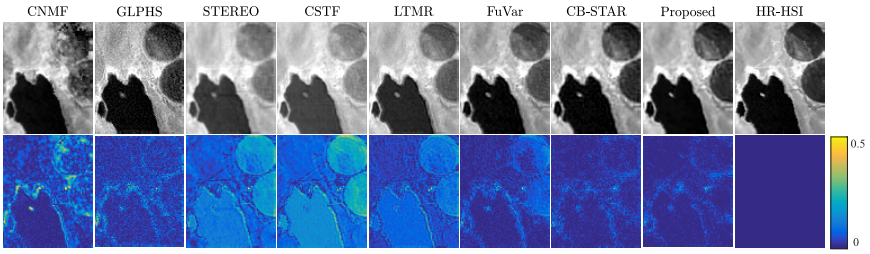

Fig. 4. Tenth band of reconstructed images (first row) and the corresponding error images (second row) for Lake Tahoe data.

for Ivanpah Playa data in Figs. 3–5, respectively. The error images were obtained by calculating the absolute values of the differencing images between ground truth and the estimated images. For Paris data, as few image changes occur, all the methods achieve comparable fusion results. The error image of the proposed method has fewer errors and is closer to the ground truth than other methods. For Lake Tahoe data, changes mainly occur at the crop circles and around the lake, as shown in the last column of Fig. 1. For the Ivanpah Playa data, changes mainly occur in the central part and lower left corner of the images, as indicated in the last column of Fig. 1. For Lake Tahoe and Ivanpah Playa data, the reconstructed images and error images of GLPHS are noisy. Many residuals remain in the error images of CNMF, STEREO, CSTF, and LTMR, in particular in the areas of changes. Fewer residuals left in the error image of FusVar, CB-STAR, and the proposed method, highlighting the importance of taking images changes into consideration while fusing. The error images of proposed method are more close to the ground truth on three datasets, suggesting that the proposed method is more effective than FuVar and CB-STAR, which are also methods accounting for spatially and spectrally localized changes while fusing.

We present the reconstructed spectral curves of a pixel for three datasets in Fig. 6. It can be observed from Fig. 6 that spectral curves recovered by the proposed method are more closer to the ground truth, which means that the proposed method can retain more accurate spectral features than the competitors.

TABLE II
RECONSTRUCTED MPSNR VALUES OF ABLATION EXPERIMENTS

Data	$\text{Method}_{\mathbf{F}}$	$\text{Method}_{\beta=0}$	Proposed
Paris	30.35	29.85	**30.99**
Lake Tahoe	27.96	30.00	**32.23**
Ivanpah Playa	31.91	31.14	**34.13**

Ablation studies were performed to verify the effectiveness of the contributions involved in the proposed method. First, we verified the effectiveness of the plugged BM3D denoiser as the regularizer by setting β to 0 (denoted as $\textbf{Method}_{\beta=0}$). Second, we verified the effectiveness of the proposed ℓ_1 norm on the residual term of the MSI by replacing the ℓ_1 norm of (4) with the Frobenius norm (denoted as $\textbf{Method}_{\mathbf{F}}$). From Table II, we can see that the MPSNR values of the proposed method are higher than the other two methods on all the datasets, demonstrating the effectiveness of the contributions involved in the proposed method.

D. Parameters Analysis

We show reconstructed MPSNRs on all the datasets with various parameter settings of k, λ, and β in Fig. 7. It can be

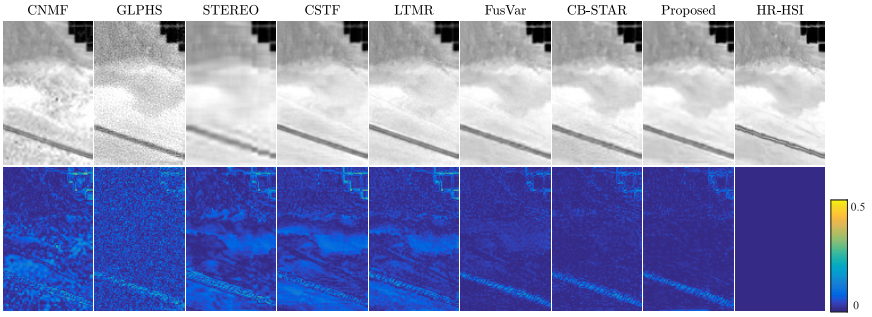

Fig. 5. Thirtieth band of reconstructed images (first row) and the corresponding error images (second row) for Ivanpah Playa data.

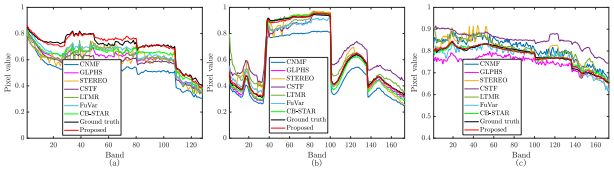

Fig. 6. Reconstructed spectral curves of a pixel in (a) Paris data, (b) Lake Tahoe data, and (c) Ivanpah Playa data.

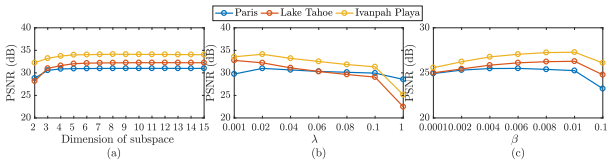

Fig. 7. MPSNR values of the proposed method with the changes of parameters. (a) k. (b) λ. (c) β.

seen from Fig. 7(a) that the MPSNR values of all the datasets are practically constant when the dimension of subspace is greater than 5. In this letter, we fixed dimension of subspace to 8 on all the datasets. From Fig. 7(b), we can see that the MPSNRs of the Lake Tahoe and Ivanpah Playa data decrease gradually when λ varies from 0.02 to 0.1, whereas the MPSNRs of the Paris data remain relatively stable. By setting λ to 0.02, we could obtain good results for all the datasets. From Fig. 7(c), we can see that the MPSNRs of the Lake Tahoe and Ivanpah Playa data increase gradually when β varies from 0.002 to 0.01, whereas the MPSNRs of the Paris data remain relatively stable. In this letter, we set β to 0.006 for all the datasets.

IV. Conclusion

In this letter, we propose a novel SCFus to fuse multiplatform hyperspectral and MSIs based on matrix factorization. By imposing the ℓ_1 norm on the reconstruction error of the MSI term to promote sparsity, the proposed fusion method can account for the case in which localized changes occur between the images obtained from multiplatforms. Meanwhile, we used a plugged in prior, i.e., BM3D, to regularize the subspace coefficients effectively. Experimental results suggest that the proposed method is effective in fusing multiplatform hyperspectral and MSIs.

References

[1] X. Xu, B. Pan, Z. Chen, Z. Shi, and T. Li, "Simultaneously multiobjective sparse unmixing and library pruning for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 4, pp. 3383–3395, Apr. 2021.

[2] X. Fu, S. Jia, L. Zhuang, M. Xu, J. Zhou, and Q. Li, "Hyperspectral anomaly detection via deep plug-and-play denoising CNN regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 11, pp. 9553–9568, Nov. 2021.

[3] N. Yokoya, C. Grohnfeldt, and J. Chanussot, "Hyperspectral and multispectral data fusion: A comparative review of the recent literature," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 2, pp. 29–56, Jun. 2017.

[4] 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.

[5] 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.

[6] R. Dian and S. Li, "Hyperspectral image super-resolution via subspace-based low tensor multi-rank regularization," *IEEE Trans. Image Process.*, vol. 28, no. 10, pp. 5135–5146, Oct. 2019.

[7] C. I. Kanatsoulis, X. Fu, N. D. Sidiropoulos, and W.-K. Ma, "Hyperspectral super-resolution: A coupled tensor factorization approach," *IEEE Trans. Signal Process.*, vol. 66, no. 24, pp. 6503–6517, Dec. 2018.

[8] S. Li, R. Dian, L. Fang, and J. M. Bioucas-Dias, "Fusing hyperspectral and multispectral images via coupled sparse tensor factorization," *IEEE Trans. Image Process.*, vol. 27, no. 8, pp. 4118–4130, Aug. 2018.

[9] R. Dian, S. Li, A. Guo, and L. Fang, "Deep hyperspectral image sharpening," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 11, pp. 5345–5355, Nov. 2018.

[10] X. Han, J. Yu, J. Luo, and W. Sun, "Hyperspectral and multispectral image fusion using cluster-based multi-branch BP neural networks," *Remote Sens.*, vol. 11, no. 10, p. 1173, May 2019.

[11] W. Xie, Y. Cui, Y. Li, J. Lei, Q. Du, 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.

[12] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, "MTF-tailored multiscale fusion of high-resolution MS and PAN imagery," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 5, pp. 591–596, May 2006.

[13] R. A. Borsoi, T. Imbiriba, and J. C. M. Bermudez, "Super-resolution for hyperspectral and multispectral image fusion accounting for seasonal spectral variability," *IEEE Trans. Image Process.*, vol. 29, pp. 116–127, 2020.

[14] R. A. Borsoi, C. Prevost, K. Usevich, D. Brie, J. C. M. Bermudez, and C. Richard, "Coupled tensor decomposition for hyperspectral and multispectral image fusion with inter-image variability," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 3, pp. 702–717, Jan. 2021.

[15] R. Dian, S. Li, and X. Kang, "Regularizing hyperspectral and multispectral image fusion by CNN denoiser," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 3, pp. 1124–1135, Mar. 2021.

[16] X. Fu, S. Jia, M. Xu, J. Zhou, and Q. Li, "Fusion of hyperspectral and multispectral images accounting for localized inter-image changes," *IEEE Trans. Geosci. Remote Sens.*, early access, Nov. 1, 2021, doi: 10.1109/TGRS.2021.3124240.

[17] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[18] L. Zhuang and M. K. Ng, "Hyperspectral mixed noise removal by ℓ_1-norm-based subspace representation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 1143–1157, 2020.