HYPERSPECTRAL AND MULTISPECTRAL IMAGE FUSION VIA TENSOR SPARSITY REGULARIZATION

Jize Xue[1,2]*, Yongqiang Zhao*[1]*, Wenzhi Liao*[2]*, Wilfried Philips*[2]

[1]Research & Development Institute of Northwestern Polytechnical University in Shenzhen, Shenzhen 518057, China
[2]Department of Telecommunications and Information Processing, Ghent University-TELIN-IMEC, Ghent, Belgium

ABSTRACT

Hyperspectral image (HSI) super-resolution scheme based on HSI and multispectral image (MSI) fusion has been a prevalent research theme in remote sensing. However, most of the existing HSI-MSI fusion (HMF) methods adopt the sparsity prior across spatial or spectral domains via vectorizing hyperspectral cubes along a certain dimension, which results in the spatial or spectral informations distortion. Moreover, the current HMF works rarely pay attention to leveraging the nonlocal similar structure over spatial domain of the HSI. In this paper, we propose a new HSI-MSI fusion approach via tensor sparsity regularization which can encode essential spatial and spectral sparsity of an HSI. Specifically, we study how to utilize reasonably the sparsity of tensor to describe the spatial-spectral correlation hidden in an HSI. Then, we resort to an efficient optimization strategy based on the alternative direction multiplier method (ADMM) for solving the resulting minimization problem. Experimental results on Pavia University data verify the merits of the proposed HMF algorithm.

Index Terms— Hyperspectral and multispectral image fusion, tensor sparsity regularization.

1. INTRODUCTION

Hyperspectral image (HSI) are recorded by simultaneously capturing the information over two spatial and one spectral dimensions. The abundant spatial-spectral informations provide more accurate and reliable signature features on distinct materials, which contributes to various applications such as scene classification [1] and agriculture applications [2], etc. However, since the photon collected by the HSI sensors is spread over many spectral bands, which will greatly reduce spatial resolution of an HSI. Compared with HSI sensors, the MSI sensors provide much wider bandwidth and higher spatial resolution [3]. Therefore, it has become an important technology to fuse a low resolution (LR) HSI with a high resolution (HR) MSI to improve the spatial resolution of HSI [4, 5].

Recently, many approaches have been applied to fuse HSI and MSI effectively [6, 7, 8, 9, 10, 11]. In the light of the prior informations of the latent HSI used in HSI-MSI fusion, we can roughly classify these fusion approaches into two categories. The most commonly exploited research line for HMF problem is the spectral unmixing or matrix factorization based methods [6, 7, 8], which always concentrate to unmix the latent HSI into endmember and abundance with suitable constraints. For instance, given an LR HSI and an HR MSI, Kawakami et al. [6] first employed the unmixing scheme to perform HSI-MSI fusion, which is considered as sparse matrix factorization problem for the input HSI with l_1-norm constraint. Naoto et al. [7, 8] exploited a coupled nonnegative matrix factorization algorithm to alternately unmix a couple of HSI and MSI, the endmember of HSI and abundance matrices of MSI can be achieved simultaneously. Another research line for HMF problem is sparse representation based methods. The core issue of these methods is how to sparsely represent the latent HSI on an appropriate spectral dictionary learned from the input HSI. Specifically, Akhtar et al. [9] used K-SVD technique to learn the spectral dictionary, and then the learned spectral dictionary is employed to jointly code the latent HSI and the input MSI. Lately, they proposed a Bayesian sparse representation scheme to infer the spectral dictionary and sparse coefficient [10]. Recently, Dong et al. [11] introduced a non-negative spectral dictionary learning algorithm.

Although works reported in [6, 7, 8, 9, 10, 11] considered the prior information across the spatial or spectral structures of the latent HSI, few of them investigated the tensor sparsity prior, not to mention exploiting the multilinear sparsity structures over the spatial and spectral domains in the latent HSI. In [12, 13], the reasonable usage of GCS (i.e., global correlation across spectrum) and NSS (i.e., nonlocal self-similarity

This work was supported in part by the Science, Technology and Innovation Commission of Manicipality under Grants JCYJ20170815162956949, by the National Natural Science Foundation of China under Grants 61771391 and 61371152, by the innovation Foundation for Doctor Dissertation of Northwestern Polytechnical University (CX201917), by the Fund for Scientific Research in Flanders (FWO) project G037115N Data fusion for image analysis in remote sensing. Wenzhi Liao, a postdoctoral fellow of the Research Foundation Flanders (FWO-Vlaanderen), acknowledges its support.

over space) have led to quite powerful HSI denoising algorithms. In this paper, we are the first to exploit the GCS and NSS prior knowledge to construct the tensor structure sparsity of latent HSI that is a faithful sparsity form for HSI-MSI fusion task. Similar as in [14], we establish 3^{rd}-order tensors by stacking all their unfolding version of non-local similar full-band patches. For this way, it is natural for the formed tensors to provide a reliable description with the purpose of delivering both the GCS and NSS priors underlying an HSI, which can jointly represent the global spectral and local-nonlocal spatial correlations along different modes of this tensor.

2. PROBLEM FORMULATION

In the paper, we follow exactly the same description of definitions of tensor terminologies in [13]. A tensor of order N, is denoted as $\mathcal{X} \in \mathbb{R}^{I_1 \times \cdots \times I_n \times \cdots \times I_N}$. Denote $\|\mathcal{X}\|_F$ and $\|\mathcal{X}\|_1$ as the F-norm and l_1 norm of a tensor $\mathcal{X}$, respectively. The "unfold" and "fold" operations along the mode-n of a tensor $\mathcal{X}$ are defined as $\mathrm{unfold}_n(\mathcal{X}) := X_{(n)} \in \mathbb{R}^{I_n \times (I_1 \times \cdots \times I_{n-1} I_{n+1} \times \cdots \times I_N)}$ and $\mathrm{fold}_n(X_{(n)}) := \mathcal{X}$. The Kronecker product of matrices $A \in \mathbb{R}^{I \times J}$ and $B \in \mathbb{R}^{K \times L}$ is a matrix of size $IK \times JL$, denoted by $A \otimes B$. The multiplication of a tensor $\mathcal{X}$ with a matrix $Y \in \mathbb{R}^{I_k \times J_k}$ on mode-k is denoted by $\mathcal{X} \times_k Y = \mathcal{Z}$, which also can be defined in terms of mode-k unfolding as $Z_k = Y X_k$. The Tucker decomposition form of a tensor $\mathcal{X}$ is: $\mathcal{X} = \mathcal{G} \times_1 \mathrm{U}_1 \times_2 \cdots \times_N \mathrm{U}_\mathrm{N}$, where $\mathcal{G}$ is the core tensor and U_n is the factor matrix in each mode.

2.1. HSI-MSI Fusion Formulation

The target HR-HSI is denoted by 3^{rd}-order tensor $\mathcal{X} \in \mathbb{R}^{W \times H \times S}$ ($W \times H$ spatial resolution and S spectral bands). The LR-HSI $\mathcal{Y} \in \mathbb{R}^{w \times h \times S}$ and HR-MSI $\mathcal{Z} \in \mathbb{R}^{W \times H \times s}$ are corresponding to spatially and spectrally downsampled versions of the $\mathcal{X}$, repectively, i.e.,

$$\mathcal{Y}{=}\mathcal{X}\times_1 \mathrm{P}_1 \times_2 \mathrm{P}_2, \mathcal{Z}{=}\mathcal{X}\times_3 \mathrm{P}_3, \tag{1}$$

where $\mathrm{P}_1 \in \mathbb{R}^{w \times W}$ and $\mathrm{P}_2 \in \mathbb{R}^{h \times H}$ are two spatial downsampling matrixes, and $\mathrm{P}_3 \in \mathbb{R}^{s \times S}$ is the spectral downsampling matrix. Combining with the observation model in Eq. (1), the HSI-MSI fusion can be formulated as

$$\min_{\mathcal{X}} \|\mathcal{Y} - \mathcal{X}\times_1 \mathrm{P}_1 \times_2 \mathrm{P}_2\|_F^2 + \|\mathcal{Z} - \mathcal{X}\times_3 \mathrm{P}_3\|_F^2 + \lambda\phi(\mathcal{X}), \tag{2}$$

where λ is a parameter and $\phi(\mathcal{X})$ denotes a suitable regularization term of the $\mathcal{X}$. In the following, we will study the tensor sparsity term of $\mathcal{X}$, and then introduce it into the HSI-MSI fusion framework.

2.2. Tensor Sparsity Regularization

Some pioneering studys related with tensor sparsity were presented in [13, 14]. And the work of [14] has drawn attention to the tensor sparsity formulation of an HSI in HSI compressive sensing reconstruction. Along this research direction, we impose the tensor sparsity constraint on nonlocal patch group to characterize the spatial-spectral correlation in our work. For each patch group, it contains similar patches by k-NN clustering. When clustering, to avoid destroying the high spectral correlation, we unfold a series of 3D cubes into corresponding 2D matrices along the spectral modes and obtain a new 3^{rd}-order tensor by stacking a set of similar items. Such constructed 3^{rd}-order tensor simultaneously employ the local and nonlocal spatial sparsity, and spectral sparsity, which contribute to maximize the benefit from tensor sparsity representation. Thus each 3^{rd}-order tensor $\mathcal{R}_i\mathcal{Z}$ can be approximated by tensor sparsity problem:

$$\begin{aligned}\min_{\mathcal{G}_i,\mathrm{U}_{\mathrm{i},1},\mathrm{U}_{\mathrm{i},2},\mathrm{U}_{\mathrm{i},3}} & \lambda\sum\nolimits_i \|\mathcal{R}_i\mathcal{X} - \mathcal{G}_i\times_1 \mathrm{U}_{\mathrm{i},1}\times_2 \mathrm{U}_{\mathrm{i},2}\times_3 \mathrm{U}_{\mathrm{i},3}\|_F^2 \\ & + \beta\|\mathcal{G}_i\|_1,\end{aligned} \tag{3}$$

where $\mathcal{R}_i$ is the operation to extract similar patches within the i-th group of $\mathcal{X}$. We incorporate the tensor sparsity regularization into the HSI-MSI fusion framework in Eq. (2) as

$$\begin{aligned}&\min_{\mathcal{X},\mathcal{G}_i,\mathrm{U}_{\mathrm{i},1},\mathrm{U}_{\mathrm{i},2},\mathrm{U}_{\mathrm{i},3}} \|\mathcal{Y} - \mathcal{X}\times_1 \mathrm{P}_1 \times_2 \mathrm{P}_2\|_F^2 + \|\mathcal{Z} - \mathcal{X}\times_3 \mathrm{P}_3\|_F^2 \\ &+ \lambda\sum\nolimits_i \|\mathcal{R}_i\mathcal{X} - \mathcal{G}_i\times_1 \mathrm{U}_{\mathrm{i},1}\times_2 \mathrm{U}_{\mathrm{i},2}\times_3 \mathrm{U}_{\mathrm{i},3}\|_F^2 + \beta\|\mathcal{G}_i\|_1,\end{aligned} \tag{4}$$

2.3. Algorithm Optimization

For the proposed fusion model, we apply the ADMM [15], an effective strategy for solving large scale optimization problems, to optimize Eq. (4). First, we introduce two auxiliary variables $\mathcal{U}, \mathcal{S}$, and equivalently reformulated Eq. (4) as

$$\begin{aligned}&\min_{\mathcal{X},\mathcal{G}_i,\mathrm{U}_{\mathrm{i},1},\mathrm{U}_{\mathrm{i},2},\mathrm{U}_{\mathrm{i},3}} \|\mathcal{Y} - \mathcal{X}\times_1 \mathrm{P}_1 \times_2 \mathrm{P}_2\|_F^2 + \|\mathcal{Z} - \mathcal{S}\times_3 \mathrm{P}_3\|_F^2 \\ &+ \lambda\sum\nolimits_i \|\mathcal{R}_i\mathcal{U} - \mathcal{G}_i\times_1 \mathrm{U}_{\mathrm{i},1}\times_2 \mathrm{U}_{\mathrm{i},2}\times_3 \mathrm{U}_{\mathrm{i},3}\|_F^2 + \beta\|\mathcal{G}_i\|_1, \\ &\qquad \text{s.t. } \mathcal{X} = \mathcal{U}, \mathcal{X}{=}\mathcal{S},\end{aligned} \tag{5}$$

Then, we have following augmented Lagrangian function:

$$\begin{aligned}&\mathcal{L}(\mathcal{X},\mathcal{U},\mathcal{S},\mathcal{G}_i,\mathrm{U}_{\mathrm{i},1},\mathrm{U}_{\mathrm{i},2},\mathrm{U}_{\mathrm{i},3},\mathcal{V}_1,\mathcal{V}_2) = \\ &\|\mathcal{Y} - \mathcal{X}\times_1 \mathrm{P}_1 \times_2 \mathrm{P}_2\|_F^2 + \|\mathcal{Z} - \mathcal{S}\times_3 \mathrm{P}_3\|_F^2 + \\ &\mu\left\|\mathcal{X} - \mathcal{U} + \frac{\mathcal{V}_1}{2\mu}\right\|_F^2 + \mu\left\|\mathcal{X} - \mathcal{S} + \frac{\mathcal{V}_2}{2\mu}\right\|_F^2 + \\ &\lambda\sum\nolimits_i \|\mathcal{R}_i\mathcal{U} - \mathcal{G}_i\times_1 \mathrm{U}_{\mathrm{i},1}\times_2 \mathrm{U}_{\mathrm{i},2}\times_3 \mathrm{U}_{\mathrm{i},3}\|_F^2 + \beta\|\mathcal{G}_i\|_1,\end{aligned} \tag{6}$$

where $\mathcal{V}_1$ and $\mathcal{V}_2$ are the Lagrange multipliers, μ is the positive scalars. We shall break Eq. (6) into five sub-problems and iteratively update each variable via fixing the other ones. (I) $\langle \mathcal{G}_i, \mathrm{U}_{\mathrm{i},1}, \mathrm{U}_{\mathrm{i},2}, \mathrm{U}_{\mathrm{i},3}\rangle$ sub-problem:

$$\begin{aligned}\min_{\mathcal{G}_i,\mathrm{U}_{\mathrm{i},1},\mathrm{U}_{\mathrm{i},2},\mathrm{U}_{\mathrm{i},3}} & \lambda\sum\nolimits_i \|\mathcal{R}_i\mathcal{U} - \mathcal{G}_i\times_1 \mathrm{U}_{\mathrm{i},1}\times_2 \mathrm{U}_{\mathrm{i},2}\times_3 \mathrm{U}_{\mathrm{i},3}\|_F^2 \\ & + \beta\|\mathcal{G}_i\|_1,\end{aligned} \tag{7}$$

According to [13, 14], we can solve the orthogonal dictionaries $\mathrm{U_{i,1}}, \mathrm{U_{i,2}}, \mathrm{U_{i,3}}$ by the SVD-based method and the sparse core tensor $\mathcal{G}_i$ by tensor-based iterative shrinkage thresholding algorithm.

(II) $\mathcal{U}$ sub-problem:

$$\begin{aligned}\min_{\mathcal{U}} \mu \left\| \mathcal{X} - \mathcal{U} + \frac{\mathcal{V}_1}{2\mu} \right\|_F^2 + \\ \lambda \sum\nolimits_i \|\mathcal{R}_i \mathcal{U} - \mathcal{G}_i \times_1 \mathrm{U_{i,1}} \times_2 \mathrm{U_{i,2}} \times_3 \mathrm{U_{i,3}}\|_F^2,\end{aligned} \tag{8}$$

By solving the Eq. (7), we can form a new tensor $\mathcal{L}$ as an intermediate variable through aggregating all patches to the corresponding positions in the HSI $\mathcal{X}$ and averaging on overlapped areas. Then, its equivalent form is

$$\min_{\mathcal{U}} \mu \left\| \mathcal{X} - \mathcal{U} + \frac{\mathcal{V}_1}{2\mu} \right\|_F^2 + \lambda \left\| \mathcal{U} - \mathcal{L} \right\|_F^2, \tag{9}$$

The objective function can be solve by ridge regression strategy.

$$\mathcal{U} = (\mu + \lambda)^{-1} (\mu \mathcal{X} + \frac{\mathcal{V}_1}{2} + \mathcal{L}), \tag{10}$$

(III) $\mathcal{S}$ sub-problem:

$$\min_{\mathcal{S}} \|\mathcal{Z} - \mathcal{S} \times_3 \mathrm{P_3}\|_F^2 + \mu \left\| \mathcal{X} - \mathcal{S} + \frac{\mathcal{V}_2}{2\mu} \right\|_F^2, \tag{11}$$

Along the third-mode unfolding, we have

$$\min_{S_{(3)}} \left\| \mathrm{Y_{(3)}} - \mathrm{P_3} S_{(3)} \right\|_F^2 + \mu \left\| \mathrm{X_{(3)}} - S_{(3)} + \frac{\mathcal{V}_{2(3)}}{2\mu} \right\|_F^2, \tag{12}$$

Similar to the optimization of Eq. (9), we have the following closed-from solution

$$\mathrm{S_{(3)}} = (\mathrm{P_3}^T \mathrm{P_3} + \mu \mathrm{I})^{-1} \mathrm{P_3}^T \mathrm{Y_{(3)}} + \mu \mathrm{Z_{(3)}} + \frac{\mathrm{V_{2(3)}}}{2}, \tag{13}$$

(IV) $\mathcal{X}$ sub-problem:

$$\begin{aligned}\min_{\mathcal{X}} \|\mathcal{Y} - \mathcal{X} \times_1 \mathrm{P_1} \times_2 \mathrm{P_2}\|_F^2 + \mu \left\| \mathcal{X} - \mathcal{U} + \frac{\mathcal{V}_1}{2\mu} \right\|_F^2 + \\ \mu \left\| \mathcal{X} - \mathcal{S} + \frac{\mathcal{V}_2}{2\mu} \right\|_F^2,\end{aligned} \tag{14}$$

Let $\mathrm{M} = (\mathrm{P_1} \otimes \mathrm{P_2})^T$, then we have $\mathrm{Y_{(3)}}$=$\mathrm{X_{(3)}}\mathrm{M}$. When unfolding the Eq. (14) along third-mode, its equivalent form is

$$\begin{aligned}\min_{\mathrm{X_{(3)}}} \left\| \mathrm{Y_{(3)}} - \mathrm{X_{(3)}M} \right\|_F^2 + \mu \left\| \mathrm{X_{(3)}} - \mathrm{U_{(3)}} + \frac{\mathcal{V}_{1(3)}}{2\mu} \right\|_F^2 + \\ \mu \left\| \mathrm{X_{(3)}} - \mathrm{S_{(3)}} + \frac{\mathcal{V}_{2(3)}}{2\mu} \right\|_F^2,\end{aligned} \tag{15}$$

Its closed-from solution is as follows

$$\begin{aligned}\mathrm{Z_{(3)}} = (\mathrm{X_{(3)}M}^T + \mu \mathrm{U_{(3)}} - \frac{\mathrm{V_{1(3)}}}{2} + \mu \mathrm{S_{(3)}} - \frac{\mathrm{V_{2(3)}}}{2}) \\ (\mathrm{M_{(3)}^T M_{(3)}} + 2\mu)^{-1},\end{aligned} \tag{16}$$

(V) Multipliers updating

$$\begin{cases} \mathcal{V}_1 \leftarrow \mathcal{V}_1 + \rho\mu(\mathcal{X} - \mathcal{U}) \\ \mathcal{V}_2 \leftarrow \mathcal{V}_2 + \rho\mu(\mathcal{X} - \mathcal{S}) \end{cases} \tag{17}$$

where ρ is a parameter associated with convergence rate with fixed value, i.e., 1.05. The above five steps will stop until convergence, and we abbreviate the proposed method as TSR.

3. EXPERIMENT RESULTS AND ANALYSIS

We now test the proposed method on University of Pavia data, which is the size of $610 \times 340 \times 115$. We reserve 93 bands by removing the water vapor absorption bands, and only choose a sub-region of 256×256 for each band in the experiment. We choose three popular methods for comparisons, namely coupled spectral unmixing (CSU) [8], sparse spatio-spectral representation (SSSR) [9], and Bayesian sparse representation (BSR) [10]. Two objective metrics for evaluation including peak signal-to-noise ratio (PSNR) and root-mean-square error (RMSE) are employed to quantitatively evaluate the performance of all methods. Three different scaling factors (SF), i.e., s = 8, 16, 32 are used to test the effectiveness of the proposed method. The HR MSI and LR HSI are generated by same down-sampling methods with [2, 5, 7].

Table I presents the quantitative results, we can see that under different SFs the TSR method is superior to all comparison methods in PSNR and RMSE. For SF 8, TSR improves the PSNR about 2.0db than state-of-the-art CSU method, and gain of PSNR value of TSR is more amplified compared to CSU, up to 2.2db under SF 32. These observations provide evidence that the HSI-MSI fusion results obtained by TSR are more close to the ground truth in the PSNR and RMSE. In addition, we also give the visual comparison results on band 20 from all methods under SF 16 and their reconstruction error maps to the ground truth in Fig. 1. We can find that the TSR method can reconstruct more edges and details than others, and its reconstruction error is more close to zero. The improvement mainly comes from the tensor sparsity regularization which can well characterize the more refined spatial-spectral structure of the latent HSI.

4. CONCLUSION

In this paper, we propose a novel method for HSI-MSI fusion by tensor sparsity regularization. The proposed method considers intrinsic sparsity, where the nonlocal similarity between spatial cubes and the global correlation across all bands are considered fully. Each cube group contains similar

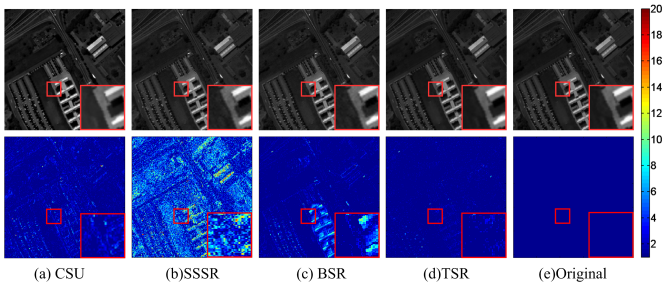

Fig. 1. Reconstruction results of different methods for band 20 from the Pavia University when scaling factor s = 16. The first row is the fusion results and the second row is the corresponding fusion error maps for the different methods.

Table 1. PSNR and RMSE Comparison for Different Fusion Methods.

SFs	Metrics	CSU	SSSR	BSR	TSR
8	PSNR	40.87	32.62	38.92	**42.76**
	RMSE	1.115	2.902	1.028	**0.828**
16	PSNR	39.73	31.35	38.30	**41.30**
	RMSE	1.341	2.999	1.382	**1.232**
32	PSNR	38.49	30.13	38.17	**40.63**
	RMSE	1.999	3.635	1.863	**1.678**

structures, its tensor-based sparsity property can be regarded as very valuable priors. Experimental results reveal that the proposed methods outperform the state-of-the-art methods in term of visual inspection and quantitative assessment.

5. REFERENCES

[1] J. Yang, Y.-Q. Zhao, and J. C.-W. Chan, "Learning and transferring deep joint spectral–spatial features for hyperspectral classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4729–4742, 2017.

[2] W. Liao, F. V. Coillie, L. Gao, L. Li, B. Zhang, and C. Jocelyn, "Deep learning for fusion of apex hyperspectral and full-waveform lidar remote sensing data for tree species mapping," *IEEE Access*, vol. 6, pp. 68 716–68 729, 2018.

[3] J. Yang, Y.-Q. Zhao, and J. Chan, "Hyperspectral and multispectral image fusion via deep two-branches convolutional neural network," *Remote Sens.*, vol. 10, no. 5, p. 800, 2018.

[4] L. Zhang, W. Wei, C. Bai, Y. Gao, and Y. Zhang, "Exploiting clustering manifold structure for hyperspectral imagery super-resolution," *IEEE Trans. Image Process.*, vol. 27, no. 12, pp. 5969–5982, 2018.

[5] 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, 2018.

[6] R. Kawakami, J. Wright, Y. W. Tai, Y. Matsushita, and K. Ikeuchi, "High-resolution hyperspectral imaging via matrix factorization," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*

[7] 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, 2012.

[8] C. Lanaras, E. Baltsavias, and K. Schindler, "Hyperspectral super-resolution by coupled spectral unmixing," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 3586–3594.

[9] N. Akhtar, F. Shafait, and A. Mian, "Sparse spatio-spectral representation for hyperspectral image super-resolution," in *Computer Vision – European Con. Comput. Vis.* Springer, 2014, pp. 63–78.

[10] N. Akhtar, F. Shafait, and A. Mian, "Bayesian sparse representation for hyperspectral image super resolution," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 3631–3640.

[11] W. Dong, F. Fu, G. Shi, X. Cao, J. Wu, G. Li, and X. Li, "Hyperspectral image super-resolution via non-negative structured sparse representation," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 2337–2352, 2016.

[12] J. Xue, Y. Zhao, W. Liao, and S. G. Kong, "Joint spatial and spectral low-rank regularization for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 1940–1958, 2018.

[13] Q. Xie, Q. Zhao, D. Meng, Z. Xu, S. Gu, W. Zuo, and Z. Lei, "Multispectral images denoising by intrinsic tensor sparsity regularization," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 1692–1700.

[14] J. Xue, Y. Zhao, W. Liao, and J. Chan, "Nonlocal tensor sparse representation and low-rank regularization for hyperspectral image compressive sensing reconstruction," *Remote Sens.*, 2019.

[15] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.