A Graph-Regularized Non-local Hyperspectral Image Denoising Method

Ling Lei¹, Binqian Huang¹, Minchao Ye¹(✉), Hong Chen¹, and Yuntao Qian²

¹ Key Laboratory of Electromagnetic Wave Information Technology and Metrology of Zhejiang Province, College of Information Engineering, China Jiliang University, Hangzhou 310018, China
yeminchao@cjlu.edu.cn
² College of Computer Science, Zhejiang University, Hangzhou 310027, China

Abstract. A lot of hyperspectral images (HSIs) are corrupted by noises when they are captured. Noise removal is an essential pre-processing for the noisy HSIs. Though denoising algorithms for common (grayscale or RGB) images have been studied for decades, HSIs have their inherent characteristics, so denoising algorithms for HSIs need to be specially designed. In this work, we have developed a non-local denoising algorithm for HSIs based on multi-task graph-regularized sparse non-negative matrix factorization (MTGSNMF). MTGSNMF delivers noise removal in both spatial and spectral views. In spatial view, patch-based sparse recovery is performed by sparse non-negative matrix factorization (SNMF), which conducts noise suppression and local pattern preservation. Graph regularization is imposed on the SNMF model for maintaining the non-local similarities between patches. In spectral view, spectral structure is extracted by multi-task learning, i.e., denoising tasks of different bands are bound by sharing the same coefficient matrix. By exploiting the non-local similarity in spatial view and spectral structure in spectral view, MTGSNMF achieves superior denoising performance on HSI datasets.

Keywords: Hyperspectral image denoising · Non-local denoising · Graph regularization · Nonnegative matrix factorization · Multi-task learning

1 Introduction

Hyperspectral images (HSIs), more or less, are corrupted by noises. The noises in HSIs can greatly affect the applications on HSIs, e.g., pixel classification, target detection, etc. As a necessary pre-processing, noise reduction plays an important role in applications of HSIs. In earlier researches, heavily noisy bands were

Supported partly by the National Natural Science Foundation of China (grant numbers 61701468 and 62071421), and partly by the National Key Research and Development Program of China (grant number 2018YFB0505000).

© Springer Nature Switzerland AG 2021
M. Nguyen et al. (Eds.): ISGV 2021, CCIS 1386, pp. 327–340, 2021.
https://doi.org/10.1007/978-3-030-72073-5_25
directly removed. For example, most noisy bands in Indian Pines data were typically dropped before classification [5]. However, some researches suggested that noisy bands still contain useful information that may contribute to the subsequent applications and thus should not be removed [11]. Therefore, HSI denoising has become a hot topic in recent years. Different from a common color image, an HSI contains hundreds of spectral bands, forming a data cube that spanned by two spatial dimensions and one spectral dimension. The spectral-spatial structure is desired to be preserved when denoising algorithm is performed on an HSI. Various algorithms have been proposed for HSI denoising in recent years. For examples, structured sparse representation on three-dimensional blocks [16], spatio-spectral total variation [1], tensor factorization [17], block-matching and 4D filtering (BM4D) [18], etc. Aforementioned methods have achieved inspiring results. Nevertheless, we believe there is still space for improvements.

In our previous research work, we proposed a nonnegative matrix factorization (NMF)-based algorithm HSI denoising, namely multi-task sparse NMF (MTSNMF) [19]. MTSNMF is a patch-based denoising algorithm. Each band image is divided into overlapping patches. The patches are denoised by sparse recovery using sparse NMF (SNMF). By incorporating multi-task learning, MTSNMF binds together denoising tasks of all bands. The multi-task learning scheme contributes to full utilization of cross-band correlations embedded in HSIs. It showed great success and surpassed most of existing HSI denoising algorithms. However, MTSNMF can still be improved by taking the intrinsic properties of HSIs into consideration.

Non-local similarity is a commonly used property in image denoising. It implies that image patches in different positions of an image may share similar patterns. The related works include 2D/3D non-local means filtering (NLMF) [3,15], block-matching and 3D/4D filtering (BM3D/BM4D) [8,13], non-local sparse recovery [14,16], etc. A latest research combined global low-rank denoising and non-local denoising [10].

In this work, we add the non-local similarity information into our previous work of MTSNMF [19]. During denoising, the non-local similarities between patches are represented by a graph. Then a graph regularization is imposed on MTSNMF, forming a new denoising model named multi-task graph-regularized sparse NMF (MTGSNMF). The proposed MTGSNMF has following advantages: 1) local structures and patterns are maintained by the sparse approximation of SNMF; 2) non-local similarities are kept by the graph regularization between patches; 3) spectral correlation is represented by sharing a common coefficient matrix among all bands. Shown by the experimental results, MTGSNMF achieves superior results when compared to other noise removal algorithms.

2 Related Work

The proposed MTGSNMF is a combination of MTSNMF [19] and graph-regularized NMF (GNMF) [4]. The basics MTSNMF and GNMF are introduced in this section.
2.1 MTSNMF

MTSNMF is a multi-task denoising algorithm, where noise elimination in each band image is seen as a single task. In each 2D band image, overlapping image patches are extracted via a sliding window. Each patch sized $\sqrt{N} \times \sqrt{N}$ is reshaped to a column vector $x \in \mathbb{R}^{N \times 1}$ and thus can be recovered by a low-dimensional and sparse approximation:

$$
\min_v \|x - Uv\|_2^2 + \lambda \|v\|_1 \quad \text{s.t.} \quad U \geq 0, v \geq 0,
$$

where $U \in \mathbb{R}^{N \times R}$ is a trained basis matrix, or alternatively called a dictionary, $R$ is the dictionary size (number of basis atoms in $U$), $v \in \mathbb{R}^{R \times 1}$ is the sparse representation (coefficient vector) of $x$ based on the dictionary $U$, and $\|v\|_1 = \sum_{i=1}^{R} |v_i|$ is the $\ell_1$-norm regularization for sparsity, with $\lambda$ controlling the sparseness. It is worth noting that if nonnegative constant is imposed on $v$, $\ell_1$-norm can be simplified to $\|v\|_1 = \sum_{i=1}^{R} v_i$. Assume that there are $M$ patches within each band image, we can merge their denoising models into a unified one:

$$
\min_{U, V} \|X - UV\|_F^2 + \lambda \|V\|_1 \quad \text{s.t.} \quad U \geq 0, V \geq 0,
$$

where $X = [x_1, x_2, \ldots, x_M]$ is the input data matrix containing all patches within a band image, $V = [v_1, v_2, \ldots, v_M]$ is the coefficient matrix, and $\| \cdot \|_F$ is the Frobenius norm. This comes to a SNMF model on each band.

Considering the spectral correlation among different bands, it is proposed in MTSNMF that different band-wise noise removal tasks can be bound by sharing a common coefficient matrix $V$ across all band images. Thus the MTSNMF model can be mathematically represented by

$$
\min_{U_1, \ldots, U_K, V} \sum_{k=1}^{K} (\|X_k - U_k V\|_F^2 + \lambda_k \|V\|_1) \quad \text{s.t.} \quad U_k \geq 0 (k = 1, \ldots, K), V \geq 0,
$$

where $X_k$ is the patch sample matrix of $k$th band image, $U_k$ is the basis matrix of the $k$th band image, $V$ is the common coefficient matrix shared across all the band images, which maintains the spectral correlation. MTSNMF has shown inspiring signal-to-noise ratio (SNR) in our previous work [19].

2.2 GNMF

GNMF model was originally proposed by Cai et al. in [4]. The motivation of GNMF is to maintain the relationship between samples (or so called the structure of data manifold) during the factorization. Suppose we have a nonnegative input matrix $X = [x_1, x_2, \ldots, x_M] \in \mathbb{R}^{N \times M}$ containing $M$ samples with $N$ dimensions. An adjacent matrix $W \in \mathbb{R}^{M \times M}$ is defined to represent the
similarity graph, where \( w_{ij} \) stands for the similarity between \( x_i \) and \( x_j \). Then the objective function of GNMF can be defined as

\[
\min_{U, V} \|X - UV\|^2_F + \frac{\mu}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} w_{ij} \|v_i - v_j\|^2_2
\]

s.t. \( U \geq 0, V \geq 0 \),

where \( U \in \mathbb{R}^{N \times R}_+ \) is the basis matrix, \( V = [v_{i1}, v_{i2}, \ldots, v_{iM}] \in \mathbb{R}^{R \times M}_+ \) is the coefficient matrix, \( v_{i \cdot} \) and \( v_{j \cdot} \) stand for the \( i \)th and the \( j \)th columns of matrix \( V \). In GNMF, if \( x_i \) and \( x_j \) are similar, \( w_{ij} \) and \( w_{ji} \) will be large, and then \( v_{i \cdot} \) and \( v_{j \cdot} \) are forced to be close to each other under the graph regularization. \( \mu \) is the regularization parameter controlling the strength of graph regularization. GNMF and its variants have been successfully applied to HSIs for preserving the similarities between pixels [6, 21].

3 Non-local Denoising for HSIs with MTGSNMF

3.1 Non-local Denoising

The idea of non-local denoising was firstly proposed in non-local means filtering (NLMF) [3]. In NLMF, non-local similarity of two pixels is measured by Gaussian weighted Euclidean distance:

\[
w(i, j) = \frac{1}{Z(i)} \exp\left( -\frac{\|N_i - N_j\|^2_2}{h^2} \right),
\]

where \( \|N_i - N_j\|^2_2 \) is Gaussian weighted distance, \( h \) controls the degree of filtering, and \( Z(i) = \sum_j \exp\left( -\frac{\|N_i - N_j\|^2_2}{h^2} \right) \) is a normalization factor. NLMF calculates the denoised pixel by weighted average of all pixels using the weights calculated by Eq. (5).

After NMLF, Block Matching 3D (BM3D) algorithm was proposed for non-local denoising, where similar patches are stacked into a 3D signal block, and 3D filtering is performed for collaborative denoising [7]. Then non-local similarity was adopted in non-local sparse denoising, where similar patches are put together and a joint sparse representation is adopted to achieve similar sparse pattern between them [14].

Following aforementioned researches, we propose a new non-local denoising approach based on graph modeling in this work.

3.2 Modeling the Non-local Similarity by Graph

Non-local similarity is an inherent property of natural images. It has been widely utilized in the applications of HSIs, e.g., denoising [12, 15], classification [22], super-resolution [9], etc. In this work, we adopt the patch-based modeling for
non-local similarity in a remotely sensed HSI, which is illustrated by Fig. 1. Similar patches may cover the same type of land cover object or contain the same material. It is worth noting that patches in a HSI are 3D volumetric patches (i.e., 3rd-order tensors) rather than 2D patches in grayscale images. The similarity measurement between volumetric patches takes both spatial and spectral similarities into account.

![Similar 3D volumetric patches](image)

**Fig. 1.** The non-local similarity in a HSI.

Graph is a powerful tool to represent pairwise similarities between samples (i.e., volumetric patches here). Moreover, graph is easy to be embedded into recovery/denoising models. Supposed that we have $M$ volumetric patches inside a HSI, each sized $\sqrt{N} \times \sqrt{N} \times K$, where $\sqrt{N} \times \sqrt{N}$ is the spatial size (number of pixels), while $K$ is the spectral size (number of bands). Then an adjacent matrix $W \in \mathbb{R}^{M \times M}$ is defined on the non-local similarity graph, where $w_{ij}$ represents the similarity between the $i$th volumetric patch $X_i$ and the $j$th volumetric patch $X_j$. To calculate $w_{ij}$, a spectral-spatial Euclidean distance is adopted

$$d(X_i, X_j) = \|X_i - X_j\|_F,$$

(6)

where $\| \cdot \|_F$ is the Frobenius norm. With the distance defined, we propose a fast method to build a sparse graph. First, a k-means clustering is performed on all patches using the spectral-spatial Euclidean distance defined in Eq. (6). Then a 0-1 weighting is performed to build the graph:

$$w_{ij} = \begin{cases} 
1, & X_i \text{ and } X_j \text{ are in the same cluster,} \\
0, & \text{otherwise.}
\end{cases}$$

(7)
It needs to be mentioned that a direct clustering on a noisy HSI may lead to an incorrect non-local similarity graph: two originally dissimilar patches may look similar after being corrupted by noises. So it is recommended to conduct a coarse signal-noise separation first, and then build the non-local similarity graph on the “clean” signal component. In this work, the coarse signal-noise separation is accomplished with the algorithm proposed in [20].

![Diagram of non-local similarity graph](image)

**Fig. 2.** The model framework of MTGSNMF.

### 3.3 Noise Removal Based on MTGSNMF

With the non-local similarity graph defined in the previous sub-section, we present our MTGSNMF model here. The model framework is displayed in Fig. 2. By adding the non-local similarity graph regularization to Eq. (3), we can get the proposed MTGSNMF model

\[
\min C(U_1, \ldots, U_K, V) = \\
\sum_{k=1}^{K} \|X_k - U_k V\|_F^2 + \left( \sum_{k=1}^{K} \lambda_k \right) \|V\|_1 + \frac{\mu}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} w_{ij} \|v_i - v_j\|_2^2 \\
\text{s.t. } U_k \geq 0 (k = 1, \ldots, K), V \geq 0.
\]

The symbols in Eq. (8) have the same meanings as that in Eqs. (3) and (4). There are three terms in the cost function $C$ of Eq. (8):
1. $\sum_{k=1}^{K} \|X_k - U_k V\|_F^2$ is the recovery error to ensure an approximation of the original noisy image;

2. $\left(\sum_{k=1}^{K} \lambda_k\right) \|V\|_1$ is the sparse regularization to accomplish noise reduction;

3. $\frac{\mu}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} w_{ij} \|v_{j} - v_{i}\|_2^2$ is the graph regularization to maintain non-local similarity during the recovery. If $\mu$ is set to zero, MTGSNMF degenerates to MTSNMF.

By combining these three terms, a non-local similarity preserving denoising model is developed.

### 3.4 Optimization Algorithm for MTGSNMF

To minimize Eq. (8), we first define the following matrices and variables:

\[ X = [X_1^T, X_2^T, \ldots, X_K^T]^T, \]  
\[ U = [U_1^T, U_2^T, \ldots, U_K^T]^T, \] 
\[ \lambda = \sum_{k=1}^{K} \lambda_k, \] 
\[ D = \text{diag}(\sum_{j=1}^{M} w_{1j}, \sum_{j=1}^{M} w_{2j}, \ldots, \sum_{j=1}^{M} w_{Mj}), \] 
\[ L = D - W. \]

Then Eq. (8) can be simplified to

\[
\min C(U, V) = \|X - UV\|_F^2 + \lambda \|V\|_1 + \mu \text{Tr}(VLV^T) \\
\text{s.t.} \quad U \geq 0, V \geq 0. 
\]

The partial derivatives are:

\[
\nabla_U C = -2XV^T + 2UVV^T, \\
\nabla_V C = -2U^TX + 2U^TUUV + \lambda + 2\mu VVL^T.
\]

Considering the KKT conditions $U \odot \nabla_U C = 0$ and $V \odot \nabla_V C = 0$ ($\odot$ is the element-wise multiplication of two matrices), we can obtain the updating rules

\[
U = U \odot \frac{XV^T}{UVV^T}, \\
V = V \odot \frac{U^TX + \mu WV^T}{U^TUUV + \frac{\lambda}{2} + \mu VDV^T}.
\]

By initializing $U$ and $V$ with nonnegative random values and applying Eqs. (17) and (18) iteratively, the minimization of Eq. (14) can be achieved.
4 Experiments

To evaluate the denoising performance of MTGSNMF, experiments are carried out on synthetic dataset and Indian Pines dataset.

4.1 Compared Algorithms and Parameter Settings

The following algorithms are compared:

1. **BM4D**: block-matching and 4D filtering (BM4D) is a denoising algorithm for volumetric data based on non-local similarity [13]. In BM4D, Gaussian noise model is assumed and noise standard deviation is automatically estimated.

2. **NLMF**: non-local means filtering was initially proposed for 2D image denoising [3]. It was then extended to a 3D version for HSI denoising [15], which is selected for comparison in this work. NLMF has a parameter \(h\) controlling the strength of filtering, which is set as \(h \in \{0.01, 0.02, \ldots, 0.05\}\) in the experiments.

3. **NGmeet**: NGmeet is a state-of-the-art algorithm for HSI denoising brought out in CVPR 2019 [10]. NGmeet needs the noise variance \(\sigma_0^2\) to be the input. Since the true noise variance is actually unknown, a noise estimation is performed with the algorithm in [2].

4. **MTGSNMF**: MTGSNMF is the algorithm proposed in this work. It is an extension of MTSNMF [19] with non-local similarity graph embedding. The parameters of MTGSNMF include

(a) patch size \(N = 7 \times 7 = 49\),

(b) dictionary size \(R \in \{100, 200\}\),

(c) sparse regularization parameter \(\lambda_k = \hat{\sigma}_k \sqrt{2 \log(R)}\), where \(\hat{\sigma}_k\) is the estimated noise standard deviation of the \(k\)th band image using the algorithm in [2],

(d) graph regularization parameter \(\mu \in \{0, 50, \ldots, 200\}\) for synthetic data and \(\mu \in \{0, 1, 10, 100\}\) for Indian Pines data.

It should be clarified that some algorithms will be tested with a series of parameter settings as listed above. In the comparisons between different algorithms, we take the best result (with optimal parameter setting) of each algorithm.

---

1 The code of BM4D is available at https://www.cs.tut.fi/~foi/GCF-BM3D/BM4D_v3p2.zip.

2 The code of NLMF is available at https://www.mathworks.com/matlabcentral/fileexchange/27395-fast-non-local-means-1d-2d-color-and-3d.

3 The code of NGmeet is available at https://github.com/quanmingyao/NGMeet.

4 The code of the proposed MTGSNMF is available at https://github.com/yeminchao/MTGSNMF.
4.2 Experimental Results on Synthetic Data

Since the quantitative criteria of denoising need a clean image for reference, a noiseless HSI named Reno is adopted for generating the synthetic data. A subset sized $200 \times 200 \times 356$ is selected for experiments. The clean band images can be found in Figs. 4a and 5a. Gaussian noise is added to each band image, where noise standard deviation varies from band to band. We set the noise standard deviation of the $k$th band image as $\sigma_k = 0.2\bar{x}_k$, where $\bar{x}_k$ is the mean value of the $k$th band image. Noisy band images can be found in Figs. 4b and 5b.

![Fig. 3. The PSNR of MTGSNMF with respect to parameter $\mu$ in synthetic data.](image)

Peak signal-to-noise ratio (PSNR) is the most commonly adopted quantitative criterion of denoising and thus is adopted in this work for performance evaluation. To show the influence of graph regularization, the PSNR of MTGSNMF with respect to graph regularization parameter $\mu$ is plotted in Fig. 3. It can be found in Fig. 3 that 1) the optimal setting of $\mu$ is related to dictionary size $R$; 2) a more redundant dictionary with $R = 200$ leads to better results. The PSNR achieved by different algorithms are listed in Table 1. It can be recognized that MTGSNMF wins the comparison by about 1dB higher PSNR than the state-of-the-art algorithm NGmeet. The visual denoising results on band 2 and band 340 are shown in Figs. 4 and 5. It can be seen that BM4D and NLMF cause over-blurring in a heavily noisy band (e.g. band 2), and NGmeet and MTGSNMF generate comparable visual results.

|                  | BM4D  | NLMF  | NGmeet | MTGSNMF |
|------------------|-------|-------|--------|---------|
| PSNR             | 39.9571 | 37.3884 | 42.6819 | 43.6181 |

**Table 1.** PSNR achieved by different algorithms
Fig. 4. Band 2 of synthetic data before and after denoising.

Fig. 5. Band 340 of synthetic data before and after denoising.
4.3 Experimental Results on Indian Pines Data

We select Indian Pines image as the real-world data, since it is a severely noisy HSI. Since Indian Pines image does not have a clean version for reference, PSNR is no longer available. Hence the visual results and classification accuracies are provided. Visual results of the compared algorithms on two representative bands are displayed in Figs. 6 and 7. In Indian Pines image, BM4D produces noticeable artifacts, especially on the heavily noisy band 220. NLMF gets blurred results, making small objects unrecognizable. Both NGmeet and MTGSNMF produce clean and clear results, but less stripes remain in the results of MTGSNMF (comparing Figs. 7d and 7e). Classification experiments with SVM are conducted on the denoised data cubes. In each land cover class, 50 samples are randomly selected as training samples, except for small classes, where 15 samples are selected for training. Remaining labeled samples are all regarded as test samples. The overall accuracy (OA), average accuracy (AA) and kappa coefficient ($\kappa$) are reported in Table 2. The results show that classification accuracy is not
always positively related to denoising performance, i.e., the best accuracies are obtained with the over-blurring denoising result with NLMF (see Figs. 6c and 7c). The proposed MTGSNMF yields a compromise between visual effect and classification accuracy.

5 Conclusion

In this research, we have proposed a multi-task graph-regularized sparse non-negative matrix factorization (MTGSNMF) model for noise reduction of HSIs. The non-local similarity is modeled with a graph through a clustering on volumetric patches. Then a graph regularization is imposed on the MTSNMF-based denoising model, forming the proposed MTGSNMF model. The advantages of MTGSNMF include: 1) patch-based modeling preserves local patterns; 2) sparse recovery delivers noise removal; 3) binding denoising tasks on different bands can keep the spectral structures; 4) graph regulation maintains non-local similarity of the HSI. Experimental results on both synthetic and real-world datasets prove the effectiveness of MTGSNMF.

References

1. Aggarwal, H.K., Majumdar, A.: Hyperspectral image denoising using spatio-spectral total variation. IEEE Geosci. Remote Sens. Lett. 13(3), 442–446 (2016)
2. Bioucas-Dias, J.M., Nascimento, J.M.P.: Hyperspectral subspace identification. IEEE Trans. Geosci. Remote Sens. 46(8), 2435–2445 (2008)
3. Buades, A., Coll, B., Morel, J.: A non-local algorithm for image denoising. In: Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 60–65 (2005)

4. Cai, D., He, X., Han, J., Huang, T.S.: Graph regularized nonnegative matrix factorization for data representation. IEEE Trans. Pattern Anal. Mach. Intell. 33(8), 1548–1560 (2011)

5. Camps-Valls, G., Bruzzone, L.: Kernel-based methods for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 43(6), 1351–1362 (2005)

6. Chen, H., Ye, M., Lu, H., Lei, L., Qian, Y.: Dual dictionary learning for mining a unified feature subspace between different hyperspectral image scenes. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium, pp. 1096–1099 (2019)

7. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Trans. Image Process. 16(8), 2080–2095 (2007)

8. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising with block-matching and 3D filtering. In: Proceedings of SPIE, Image Processing: Algorithms and Systems, Neural Networks, and Machine Learning, pp. 354–365 (2006)

9. Dian, R., Fang, L., Li, S.: Hyperspectral image super-resolution via non-local sparse tensor factorization. In: Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (2017)

10. He, W., Yao, Q., Li, C., Yokoya, N., Zhao, Q.: Non-local meets global: an integrated paradigm for hyperspectral denoising. In: Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 6861–6870 (2019)

11. Jia, S., Ji, Z., Qian, Y., Shen, L.: Unsupervised band selection for hyperspectral imagery classification without manual band removal. IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. 5(2), 531–543 (2012)

12. Li, J., Yuan, Q., Shen, H., Zhang, L.: Hyperspectral image recovery employing a multidimensional nonlocal total variation model. Signal Process 111, 230–248 (2015)

13. Maggioni, M., Katkovnik, V., Egiazarian, K., Foi, A.: Nonlocal transform-domain filter for volumetric data denoising and reconstruction. IEEE Trans. Image Process. 22(1), 119–133 (2013)

14. Mairal, J., Bach, F., Ponce, J., Sapiro, G., Zisserman, A.: Non-local sparse models for image restoration. In: Proceedings of IEEE International Conference On Computer Vision, pp. 2272–2279 (2009)

15. Qian, Y., Shen, Y., Ye, M., Wang, Q.: 3-D nonlocal means filter with noise estimation for hyperspectral imagery denoising. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium, pp. 1345–1348 (2012)

16. Qian, Y., Ye, M.: Hyperspectral imagery restoration using nonlocal spectral-spatial structured sparse representation with noise estimation. IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. 6(2), 499–515 (2013)

17. Xiong, F., Zhou, J., Qian, Y.: Hyperspectral restoration via $\ell_0$ gradient regularized low-rank tensor factorization. IEEE Trans. Geosci. Remote Sens. 57(12), 10410–10425 (2019)

18. Xu, P., Chen, B., Xue, L., Zhang, J., Zhu, L., Duan, H.: A new MNF-BM4D denoising algorithm based on guided filtering for hyperspectral images. ISA Trans. 92, 315–324 (2019)

19. Ye, M., Qian, Y., Zhou, J.: Multitask sparse nonnegative matrix factorization for joint spectral-spatial hyperspectral imagery denoising. IEEE Trans. Geosci. Remote Sens. 53(5), 2621–2639 (2015)
20. Ye, M., Chen, H., Ji, C., Lei, L., Qian, Y.: Spectral-spatial joint noise estimation for hyperspectral images. In: Proceedings of International Geoscience and Remote Sensing Symposium, pp. 230–233 (2019)

21. Ye, M., Zheng, W., Lu, H., Zeng, X., Qian, Y.: Cross-scene hyperspectral image classification based on DWT and manifold-constrained subspace learning. Int. J. Wavelets Multiresolut. Inf. Process. **15**(06), 1750062 (2017)

22. Zhang, H., Li, J., Huang, Y., Zhang, L.: A nonlocal weighted joint sparse representation classification method for hyperspectral imagery. IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. **7**(6), 2056–2065 (2014)