

An Improved Camouflage Target Detection Using Hyperspectral Image Based on Block-Diagonal and Low-Rank Representation

Fei Li^{1,2}, Xiuwei Zhang^{1,2}(\boxtimes), Lei Zhang^{1,2}, Yanning Zhang^{1,2} Dongmei Jiang^{1,2}, and Genping Zhao³

¹ School of Computing Science, Northwestern Polytechnical University, Xi'an 710071, Shaanxi, China xwzhang@nwpu.edu.cn

Shaanxi Provincial Key Laboratory of Speech and Image Information Processing, Xi'an 710071, Shaanxi, China

³ School of Computers, Guangdong University and Technology, Guangzhou 510006, China

Abstract. Accurate camouflage target distinction is often resorted to hyperspectral spectral imaging technique as for the rich spectral information contained in hyperspectral images. In this paper, a novel blockdiagonal representation based camouflage target detection method is proposed for hyperspectral imagery. To better represent the multi-mode cluster background, an hyperspectral image is first clustered into different background clusters according to their spectral features. Then, an orthogonal background dictionary is learned for each cluster via a principle component analysis (PCA) learning scheme. The background and camouflage target often show different structures when projected onto those dictionaries. The former exhibits block-diagonal structure while the latter shows sparse structure. Inspired by this fact, we cast the block-diagonal structure into a low-rank representation model. With proper optimization of such model, the sparse camouflage targets can be accurately separated from the block-diagonal background. Experimental results on the real-world camouflage target datasets demonstrate that the proposed method outperforms the state-in-art hyperspectral camouflage target detection methods.

Keywords: Hyperspectral image \cdot Camouflage target detection Block-diagonal structure \cdot Sparse representation \cdot Dictionary learning

© Springer Nature Switzerland AG 2018

J.-H. Lai et al. (Eds.): PRCV 2018, LNCS 11259, pp. 384–395, 2018. https://doi.org/10.1007/978-3-030-03341-5_32

X. Zhang—Supported by the National Natural Science Foundation of China (No. 61231016, No. 61303123, No. 61273265, No. 61701123), the Fundamental Research Funds for the Central Universities (No. 3102015JSJ0008).

1 Introduction

Camouflage target plays an important role in modern warfare. Therefore, various strategic and tactical targets rely on the camouflage or covert technology to avoid being discovered. How to quickly and accurately detect and identify camouflage targets has become an important research topic for the military target detection. To conceal target or reduce its detectability, camouflaged and stealth materials are widely used to reduce the differences of scattering and radiation intensity between target and background in optical bands. With the rapid development of spectral imaging technology, an hyperspectral image (HSI) [16] contains not only the two-dimensional geometric spatial information of the target, but also the one-dimensional spectral information of the target. Since, most of the camouflage targets are designed to hide mainly at certain bands, which makes it hard to realize all-wave stealth [11], hyperspectral imaging technology [13] shows great potential for the camouflage target detection by utilizing the abundant spectral information.

To give a better reconnaissance result of camouflage target, researchers investigated the hyperspectral camouflage target detection methods on early days [3,5,6], such as, HUA *et al.* [5] employed the constrained energy minimization (CEM) hyperspectral target detection method to extract the camouflage targets, YANG *et al.* [11] utilized the spectral angle distance and mathematical morphology to detect the camouflage target in hyperspectral image. The common strategies of these methods is to use the prior spectrum information of the camouflage target to complete the detection task, but it is not practical in real applications. The reason is that the prior spectral information of camouflage target is hardly available as it is often effected by some uncertain environmental factors, such as the absorption and scattering of the atmosphere, change of illumination, spectral response of sensor, etc. Additionally, the military secrets also interferes with the acquisition of spectra information of camouflage target detection method.

With the ideas discussed above, this paper proposes a novel camouflage target detection method via hyperspectral image. Since there is no prior spectral information for the camouflage target, we first cluster the HSI into different background clusters according to their spectral features in order to describe the background accurately. Then, a spectral-based background dictionary is learned for each cluster through the principle component analysis (PCA) learning method. Based on the representation theory, the background can exhibit block-diagonal structure while camouflage target has sparse property when the HSI is represented onto those background sub-dictionaries. Following this investigation, the block-diagonal-based low-rank and sparse representation model has been built. When solving this model, the HSI can be decomposed into the background part and sparsity part. Since the camouflage target has a sparse property, we can extract it from the sparsity part.

The rest of this paper is organized as follows. In Sect. 2, we give a detailed description of the proposed model, the spectral-based background dictionary

construction method and the optimization procedure for the proposed model. The experiments and results analyses are provided in Sect. 3, and Sect. 4 concludes the paper.

2 The Proposed Method

2.1 Background Block-Diagonal Structure for the Hyperspectral Image

It is important to describe the background accurately without prior spectral information of camouflage target. However, due to the cluttered imaging scene, an HSI often contains different categories of materials. And the corresponding background is inhomogeneous but multi-mode. Therefore, to guarantee the camouflage target detection accuracy, it is crucial to consider the multi-mode structure in background modeling. A promising way for the multi-mode structure capturing is to apply the clustering method. As it specializes in collecting similar pixels into a homogeneous cluster and dispersing different pixels into various clusters. In such way, the multi-mode structure is represented with different clusters. In this study, we propose to incorporate the clustering method and dictionary learning scheme to depict the multi-mode structure of background in the representation based detection framework. Through clustering the background, we obtain several homogeneous clusters and different clusters exhibit obvious discrepancy. When being represented on the concatenation of all dictionaries learned from each cluster, the representation matrix of the HSI exhibits obvious block-diagonal structure.

To clarify this point, we first decompose the input HSI X into a background part as well as camouflage target part, which can be formulated as:

$$\boldsymbol{X} = \boldsymbol{B}_{bg} + \boldsymbol{C},\tag{1}$$

where B_{bg} is the background part and C is the camouflage target part. As discussed above, the background can be represented by a reasonable dictionary while the camouflage target C not. Thus, we can represent $B_{bg} = DZ$ and reformulated Eq. (1) as:

$$\boldsymbol{X} = \boldsymbol{D}\boldsymbol{Z} + \boldsymbol{C},\tag{2}$$

where $D = [D_1, D_2, \dots, D_k]$ contains k background sub-dictionaries which is learned from each cluster independently, D_i corresponds to *i*-th sub-dictionaries, Z is the background representation matrix of the HSI. Suppose the HSI X can be divided into k cluster, it is easy to permute the columns in X according to the cluster result as $X = [X_1, X_2, \dots, X_k]$, X_i represent *i*-th cluster and each column in X_i denotes the spectra of a specific pixel. According to [8,12] when D_i and X_i are exactly sampled from independent subspaces, Eq. (2) can reveal the subspace membership among the samples. Therefore, with clustering and permuting the original HSI, the background representation matrix Z will exhibit a block-diagonal structure which means that the background also has this structure characteristics.

2.2 Block-Diagonal Structure Based Low-Rank and Sparse Representation Model

Based on above discussion, the proposed Block-Diagonal Structure Based Low-Rank and Sparse Representation (BDSLRSR) is implemented by integrating the multi-mode structure background and sparse camouflage as follows:

$$\min_{\boldsymbol{Z},\boldsymbol{C}} \quad \operatorname{rank}(\boldsymbol{Z}) + \lambda \|\boldsymbol{C}\|_{2,1},$$
s.t $\boldsymbol{X} = \boldsymbol{D}\boldsymbol{Z} + \boldsymbol{C},$

$$(3)$$

where rank(·) denotes the rank function, parameter $\lambda > 0$ is used to balance the effects of the two parts, and $\|\cdot\|_{2,1}$ is the $\ell_{2,1}$ norm defined as the sum of ℓ_2 norm of the column of a matrix. $X = [x_{11}, \ldots, x_{1n_1}, \ldots, x_{k1}, \ldots, x_{kn_k}] \in \mathbb{R}^{b \times n}$ is a sorted 2-D HSI matrix according to the cluster processing (suppose that there are k clusters for the HSI, x_{ij} (i = 1, ..., k; j = 1, ..., n) is j-th pixel of the i-th cluster, $n_1 + \cdots + n_k = n$ is the total number of samples, b is the number of hyperspectral bands), DZ denotes the background part, D is the background dictionary learned by each cluster, Z denotes the background block-diagonal representation coefficients, and C denotes the remaining part corresponding to the camouflage target. The reason for the sparsity of camouflage target in Eq. (3) is that the dictionary D stands for background characteristics only and can not be utilized to represent camouflage target reasonably. Moreover, there are very low amounts of camouflage target in the data X compared with the background pixels, thus the camouflage target may have sparsity property rather than lowrank property [14]. Consequently, it is reasonable to add the sparse constraint into camouflage target as shown in Eq. (3).

After getting the sparsity matrix C, the role of *i*-th pixel can be determined as follows:

$$r(\boldsymbol{x}_i) = \|[\boldsymbol{C}]_{:,i}\|_2 = \sqrt{\sum_j ([\boldsymbol{C}]_{j,i})^2} \gtrsim \delta,$$
(4)

where $\|[C]_{:,i}\|_2$ denotes the ℓ_2 norm of the *i*-th column of C, δ is the segmentation threshold, and if $r(\mathbf{x}_i) > \delta$, \mathbf{x}_i is determined as the camouflage target; otherwise, \mathbf{x}_i is labeled as the background.

The main advantages in our model are as follows:

- 1. We adopt the cluster method to describe background which can exploit background information and characteristics more accurately. This kind of detailed feature has not been considered in the former low-rank-based methods which regarded the background as a whole.
- 2. The block-diagonal structure is utilized to represent the multi-mode structure information of background based on the cluster result and it is more robust than the low-rank structure. Because, low-rank structure depends on the feature consistency of pixels, and a slight variation may cause the background to be full-rank. While the block-diagonal structure depends on the feature dissimilarity of pixels, which is more robust to feature variation.

3. We employ dictionary leaning method to obtain background dictionary which can extract background feature efficiently. The later section will give a detailed explanation.

2.3 Spectral Feature Based Dictionary Learning

Generally, the background dictionary has a great impact on the representationbased hyperspectral unsupervised target detection methods [7, 15]. To construct a robust background dictionary, we utilize the k-means [1] method to divide the hyperspectral data into k clusters and each cluster can represent one background material roughly. By this way, the multi-mode characteristics of background can be well exhibited through selecting a reasonable k (k should be larger than the true number of ground material clusters in order to make sure that the k cluster represent all the ground materials).

Through clustering background, the camouflage target will be assigned to one of cluster. Then, we adopt the PCA technique for dictionary learning. It has been shown that the significant components in PCA deliver the major information of the data. In a given cluster, the major information comes from the background pixels. Thus, we remove the less significant components after PCA to eliminate the negative effect of anomalies on the learned dictionary. Finally, we obtain the background dictionary D after using the PCA learning algorithm for each cluster.

The advantages of spectral feature based dictionary learning technique are as follows:

- 1. By using the cluster way to represent background, both the diversity and multi-mode structure information of background can be well described explicitly. Moreover, the low-rank property of background is enhanced, which is helpful to increase the separability of camouflage targets and background.
- 2. The PCA learning scheme enables us to learn clean background dictionary by neglecting those less significant principle component.

The entire flow of the proposed method can be shown in Fig. 1. It can be seen clearly that the proposed method mainly contains two modules, dictionary learning and block-diagonal structure based low-rank and sparse representation. Given an HSI, we first divide it into different clusters by utilizing the k-means method. Then, a robust background dictionary D can be learned through PCA method for each cluster. With the background dictionary D and re-ordered HSI X corresponding the dictionary D, the block-diagonal structure based low-rank and sparse representation model in Eq. (3) can be built. Through solving this model, we can get the sparse matrix C containing the camouflage target. As a result, the targets are extracted from this sparse matrix by Eq. (4).

2.4 Optimization Procedure

This section will show the detailed procedure of how to solve the BDSLRSR model. The model in Eq. (3) is non-convex and NP-hard. An effective way to mitigate this problem is to relax Eq. (3) into the following convex problem:

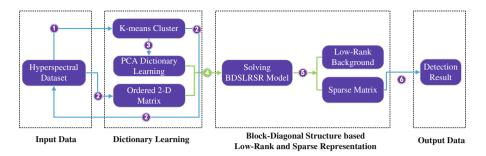


Fig. 1. Framework of the proposed method.

$$\min_{\boldsymbol{Z},\boldsymbol{C}} \quad \|\boldsymbol{Z}\|_* + \lambda \|\boldsymbol{C}\|_{2,1},$$
s.t $\boldsymbol{X} = \boldsymbol{D}\boldsymbol{Z} + \boldsymbol{C}.$

$$(5)$$

where the nuclear norm $\|\cdot\|_*$ is utilized to replace the original rank regularization. It has been shown that the solution of Eq. (3) is equal to that of Eq. (5) when some mild conditions hold [8].

In our study, we employ the standard alternative direction method of multipliers (ADMM) to solve the problem in Eq. (5). Specifically, we first reformulate Eq. (5) as follows

$$\min_{\substack{Z,E,J}} \|J\|_* + \lambda \|C\|_{2,1},$$
s.t $X = DZ + C, Z = J.$
(6)

Then, we can obtain the following Lagrangian function:

$$L = \|\boldsymbol{J}\|_{*} + \lambda \|\boldsymbol{C}\|_{2,1} + \operatorname{tr}(\boldsymbol{Y}_{1}^{T}(\boldsymbol{X} - \boldsymbol{D}\boldsymbol{Z} - \boldsymbol{C})) + \operatorname{tr}(\boldsymbol{Y}_{2}^{T}(\boldsymbol{Z} - \boldsymbol{J})) + \frac{\mu}{2}(\|\boldsymbol{X} - \boldsymbol{D}\boldsymbol{Z} - \boldsymbol{C}\|_{F}^{2} + \|\boldsymbol{Z} - \boldsymbol{J}\|_{F}^{2})$$
(7)

where Y_1 and Y_2 are Lagrange multipliers and $\mu > 0$ is the penalty coefficient. Similar as [8], given the Lagrangian function the detailed steps for solving Eq. (7) can be summarized into Algorithm 1. The detailed derivation for each step can be found in [8].

3 Experiments and Discussion

3.1 Comparison Methods and Evaluation Index

To give a well and objective evaluation for the proposed method, we employ 4 state-of-the-art hyperspectral unsupervised target detection methods for comparison. They are RX (RX) [9], Sparse Representation-based (SR) Unsupervised Target Detector, Cluster-Based Detector (CBAD) [2] and Low-Rank and Sparse

Algorithm 1. ADMM for BDSLRSR

Input: Sorted 2-D Hyperspectral image X, Sorted cluster-based background dictionary D. Initialize: $\boldsymbol{Z} = \boldsymbol{J} = \boldsymbol{C} = \boldsymbol{0},$ $Y_1 = Y_2 = 0, \ \lambda = 0.002$ $\mu = 10^{-6}, \, \mu_{\rm max} = 10^{6},$ $\rho = 1.1, \varepsilon = 10^{-8}.$ while Stopping criteria is not satisfied do 1. Fix the others and update J by $J = \arg\min \frac{1}{\mu} \|J\|_{*} + \frac{1}{2} \|J - (Z + Y_{2}/\mu)\|_{F}^{2};$ 2. Fix the others and update Z by $Z = (I + D^T D)^{-1} (\hat{D}^T (X - C))$ $+J + (D^T Y_1 - Y_2)/\mu);$ 3. Fix the others and update C by $C = \arg\min\frac{\lambda}{\mu} \|C\|_{2,1}$ $+\frac{1}{2}\|C-(X-DZ+Y_1)/\mu\|_F^2;$ 4. Update the multipliers: $Y_1 = Y_1 + \mu (X - DZ - C);$ $Y_2 = Y_2 + \mu(Z - J);$ 5. Update the parameter μ : $\mu = \min(\rho\mu, \mu_{\max});$ 6. Check the convergence conditions: $\|X - DZ - C\|_{\infty} < \varepsilon$ and $\|Z - J\|_{\infty} < \varepsilon$. **Output:** Sparsity matrix *C*.

Representation (LRASR) based Detection Method [10]. The proposed method is named Block-Diagonal Structure Based Low-Rank and Sparse Representation (BDSLRSR) method. The results are evaluated using receiver operating characteristic (ROC) curves and the area under such curves (AUC) [4].

3.2 Hyperspectral Datasets

In this study, we employed our Hyperspectral Imager shown in Fig. 2(a) bought from Zolix company¹ in Beijing to collect the hyperspectral datasets. The specific parameters of this imager are shown in Table 1. Three kinds of camouflage nets shown in Fig. 2(b)–(d) are utilized as the camouflage targets by placing them in the same color surroundings.

The first dataset was collected on the fourth floor of the School of Computer Science, Northwestern Polytechnical University. The camouflage target were woodland camouflage net (left) and desert camouflage net (right). There are 174×411 pixels in the whole data with 160 spectral bands as shown in Fig. 3(a). The ground truth of the target is shown in Fig. 3(b). The parameters

¹ http://www.zolix.com.cn/index.html.

Spectral range	Spectral resolution	Detection element	Number of pixels	Frame rate	Interface
400 - 1000 nm	$2-3\mathrm{nm}$	Cooled CCD	1392×1040	$1162\mathrm{fps}$	USB

 Table 1. Parameters of hyperspectral imager



Fig. 2. Data acquisition equipments. (a) Hyperspectral Imager, (b) Woodland camouflage net, (c) Desert camouflage net, (d) Digital camouflage net.



Fig. 3. First dataset. (a) Pseudo-RGB of the scene, (b) Ground truth of the camouflage targets.

on this dataset are as follows: the number of clusters is 8 and the number of first principle components is 20.

The second dataset was collected at the BaiLu Tableland, Xian, Shaanxi Province. The camouflage targets in the data were woodland camouflage net (left) and digital camouflage net (right). The image size is 174×682 pixels with 160 spectral bands as shown in Fig. 4(a). The ground truth of the target is shown in Fig. 4(b). The parameters on this dataset are as follows: the number of clusters is 8 and the number of first principle components is 20.

The third data was also collected at the BaiLu Tableland, Xian. By using a woodland camouflage net, the roadside car was covered as a camouflage target. Its size is 87×181 pixels with 160 spectral bands as shown in Fig. 5(a). Figure 5(b) is the ground truth of the target. The parameters on this dataset are as follows: the number of clusters is 8 and the number of first principle components is 20.



Fig. 4. Second dataset. (a) Pseudo-RGB of the scene, (b) Ground truth of the camou-flage targets.



Fig. 5. Third dataset. (a) Pseudo-RGB of the scene, (b) Ground truth of the camou-flage targets.

3.3 Result Analysis

All the two-dimensional plots of detection results for these three datasets are shown in Figs. 6, 7 and 8. For giving a good visualization, we adopt the colormap image to show these results. Background pixels are represented by the blue pixels. The groundtruth location of the targets are shown in Figs. 3(b), 4(b) and 5(b). Compared with other methods, the proposed method has achieved better results both in background representation and background suppression. Although a small amount of man-made objects, such as buildings in Fig. 8, tele-

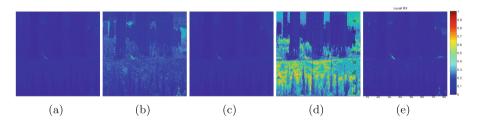


Fig. 6. Two-dimensional plots of the detection results obtained by different methods for the first dataset. (a) RX, (b) SR, (c) CBAD, (d) LRASR, (e) BDSLRSR. (Color figure online)

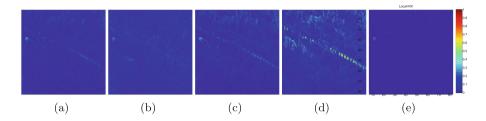


Fig. 7. Two-dimensional plots of the detection results obtained by different methods for the second dataset. (a) RX, (b) SR, (c) CBAD, (d) LRASR, (e) BDSLRSR. (Color figure online)

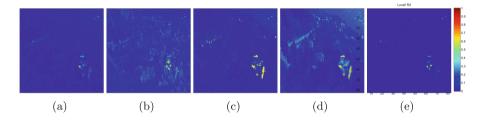


Fig. 8. Two-dimensional plots of the detection results obtained by different methods for the third dataset. (a) RX, (b) SR, (c) CBAD, (d) LRASR, (e) BDSLRSR. (Color figure online)

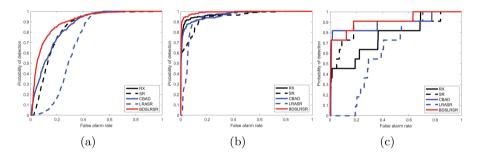


Fig. 9. Detection accuracy evaluation for the first dataset. (a) ROC curves. (b) AUC values.

graph poles in Fig. 7, etc., can be easily judged as a target of interest when the target is detected using spectral information, the proposed method can effectively suppress such objects and reduce interference by considering the overall block-diagonal structure information of the background. Finally, the background and the camouflage targets are effectively separated by solving a block-diagonal structure based low-rank and sparse representation model. Further more, the ROC curves and AUC values in Figs. 9 and 10 show that the proposed method achieved better detection results.

394 F. Li et al.

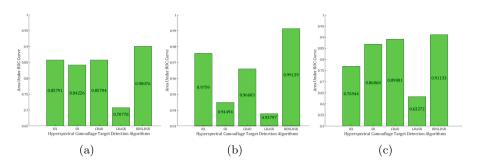


Fig. 10. Detection accuracy evaluation for the second dataset. (a) ROC curves. (b) AUC values.

4 Conclusion

With the hyperspectral image, this paper describes a new camouflage target detection method. To represent the background more accurate, spectral-based cluster strategy is employed to exhibit multi-mode structure of the background. Then, the dictionary of each cluster is obtained by utilizing the PCA method and the whole background dictionary consists of the learned sub-dictionary of each cluster. Next, we cast the block-diagonal structure and background dictionary into a low-rank and sparse representation model. After solving this model, camouflage targets are extracted from the sparsity part. Compared with the traditional cluster-based and low-rank methods, this proposed method can achieve a better detection result, since it simultaneously consider the low-rank, multimodal, and block diagonal structure properties of the background.

References

- 1. Arthur, D., Vassilvitskii, S.: k-means++: the advantages of careful seeding. In: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 1027–1035. Society for Industrial and Applied Mathematics (2007)
- Carlotto, M.J.: A cluster-based approach for detecting man-made objects and changes in imagery. IEEE Trans. Geosci. Remote Sensing 43(2), 374–387 (2005)
- Chen, Y., Chen, X., Zhou, J., Ji, Y., Shen, W.: Camouflage target detection via hyperspectral imaging plus information divergence measurement. In: International Conference on Optoelectronics and Microelectronics Technology and Application, p. 102440F (2017)
- Fawcett, T.: An introduction to ROC analysis. Pattern Recognit. Lett. 27(8), 861– 874 (2006)
- Hua, W., Guo, T., Liu, X.: Camouflage target reconnaissance based on hyperspectral imaging technology. In: International Conference on Optical Instruments and Technology: Optoelectronic Imaging and Processing Technology (2015)
- Kim, S., Shim, M.S.: Cooperative spectral and spatial feature fusion for camouflaged target detection. In: SPIE Defense + Security, p. 94721M (2015)

- Li, F., Zhang, Y., Zhang, L., Zhang, X., Jiang, D.: Hyperspectral anomaly detection using background learning and structured sparse representation. In: 2016 IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2016, Beijing, China, 10–15 July 2016, pp. 1618–1621 (2016). https://doi.org/10.1109/IGARSS. 2016.7729413, https://doi.org/10.1109/IGARSS.2016.7729413
- Liu, G., Lin, Z., Yan, S., Sun, J., Yu, Y., Ma, Y.: Robust recovery of subspace structures by low-rank representation. IEEE Trans. Pattern Anal. Mach. Intell. 35(1), 171–184 (2013). https://doi.org/10.1109/TPAMI.2012.88
- Reed, I.S., Yu, X.: Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. IEEE Trans. Acoust. Speech Signal Process. 38(10), 1760–1770 (1990)
- Xu, Y., Wu, Z., Li, J., Plaza, A., Wei, Z.: Anomaly detection in hyperspectral images based on low-rank and sparse representation. IEEE Trans. Geosci. Remote Sens. 54(4), 1990–2000 (2016)
- Yang, J., Hua, W., Ma, Z., Zhang, Y.: Detection of camouflaged targets using hyperspectral imaging technology. In: International Symposium on Photoelectronic Detection and Imaging 2013: Imaging Spectrometer Technologies and Applications, vol. 8910, p. 891006. International Society for Optics and Photonics (2013)
- 12. Zhang, L., et al.: Adaptive importance learning for improving lightweight image super-resolution network. arXiv preprint arXiv:1806.01576 (2018)
- Zhang, L., Wei, W., Bai, C., Gao, Y., Zhang, Y.: Exploiting clustering manifold structure for hyperspectral imagery super-resolution. IEEE Trans. Image Process. 27, 5969 (2018)
- Zhang, L., Wei, W., Shi, Q., Shen, C., van den Hengel, A., Zhang, Y.: Beyond low rank: a data-adaptive tensor completion method. arXiv preprint arXiv:1708.01008 (2017)
- Zhang, L., Wei, W., Zhang, Y., Shen, C., Den Hengel, A.V., Shi, Q.: Dictionary learning for promoting structured sparsity in hyperspectral compressive sensing. IEEE Trans. Geosci. Remote Sens. 54(12), 7223–7235 (2016)
- Zhang, L., Wei, W., Zhang, Y., Shen, C., van den Hengel, A., Shi, Q.: Cluster sparsity field: an internal hyperspectral imagery prior for reconstruction. Int. J. Comput. Vis. 126, 1–25 (2018)