Parallel Multi-view Low-rank and Sparse Subspace Clustering for Unsupervised Hyperspectral Image Classification

Long Tian and Qian Du*
*Mississippi State University, Mississippi, USA
E-mail: du@ece.msstate.edu Tel: +1-662-3252035

Abstract — **In this paper, parallel multi-view low-rank sparse subspace clustering (MLRSSC) is investigated for unsupervised classification of remotely sensed hyperspectral imagery. A 3-dimensional (3D) hyperspectral image contains abundant spectral and spatial information. Such diverse information can be considered as multiple views during the clustering process. In this paper, multiple spectral views are generated from correlated spectral band groups, a decorrelated and denoised view from principal components, and spatial views from morphological features. To make such a computational expensive clustering technique applicable to large-scale remote sensing images, parallel MLRSSC is implemented to non-overlapping 3D blocks. Experimental results demonstrate that the performance of the MLRSSC is better than other subspace clustering based algorithms.**

I. INTRODUCTION

A hyperspectral remote sensing image contains hundreds of spectral bands for the same image scene on the Earth. Due to very high spectral resolution, hyperspectral imaging can be used to distinguish objects with subtle spectral discrepancies. Because of this special advantage, hyperspectral imaging has been widely applied in remote sensing monitoring [1-6]. Clustering is one of popular techniques in image processing, and also for hyperspectral image processing [7-9]. By considering both spectral and spatial information, a clustering algorithm separates spectral pixels into different clusters. As an unsupervised technique, clustering is more challenging than supervised classification using labeled samples. On the other hand, an unsupervised algorithm requires no label information that is often difficult or too expensive to obtain. Thus, unsupervised classification through clustering is of great interest to remote sensing applications.

The *k*-means clustering is a classical method, which is sensitive to initial conditions and prone to be stuck in local optima. Furthermore, the clustering results are centroid-based, but a hyperspectral image may not have this nature due to its very high data dimension. Subspace clustering has been developed for high-dimensional dataset, where the data is clustered into multi-subspace and a low-dimensional subspace is achieved to fit each group of pixels. Recently, sparse subspace clustering (SSC) and low-rank subspace clustering (LSC) [10-12] are proposed to find affinity matrices for effective clustering, where an affinity matrix defines the similarity between pixels. The SSC algorithm uses the sparsest representation for each pixel with pixels in its group, and the local structure of data can be maintained. The LSC algorithm introduces low-rank constraint into self-representation matrix, and the global structure of data is preserved. In order to contain both local and global information in dataset, the low-rank sparse subspace clustering (LRSSC) algorithm is proposed which combines the low-rank and sparsity constraints [13].

In machine learning area, a dataset usually is acquired from multiple sources or contain different features, where multi-view learning technique has been deployed [14, 15], since traditional single-view learning could not represent all the features or sources properly in dataset. For a hyperspectral dataset, it can be treated as image with varied sources (e.g., spectral bands covering visible to shortwave infrared channels) or features (e.g., spatial and contextual information), where a multi-view learning algorithm could be applied. In this case, hyperspectral imagery is a perfect dataset for multi-view learning. For instance, Li *et al*. utilized multiple morphological features for hyperspectral image classification [16].

In particular, multi-view learning and LRSSC are incorporated in Ref. [17] as multi-view low-rank sparse subspace clustering (MLRSSC) to deal with multi-features or multi-sources in a dataset. We also apply the MLRSSC for hyperspectral image clustering where both spectral and spatial views are constructed [18]. Due to the very large spatial size of remote sensing data, subspace clustering type of algorithms are not directly applicable. In this paper, to maintain local spatial information, the parallel version of MLRSSC is applied to non-overlapping 3D blocks, and the final clustering result is produced by merging those of individual ones. Such a technique can be simply implemented in parallel machines to significantly reduce the overall computing time.

II. Proposed Method

A. Low-rank sparse subspace clustering

Let a hyperspectral image data with D spectral bands and N pixels be denoted as $\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_i, \cdots, \mathbf{x}_N]$, where $\mathbf{x}_i$ is the i-th pixel vector. The low-rank and sparse representation of $\mathbf{X}$ can be formulated as the following minimization problem:

$$\min_{\mathbf{Z},\mathbf{E}} rank(\mathbf{Z}) + \gamma \|\mathbf{E}\|_1, \quad s.t. \;\; \mathbf{X} = \mathbf{AZ} + \mathbf{E} \tag{1}$$

where $\mathbf{A}$ is a dictionary, $\mathbf{Z}$ is the low-rank matrix, and $\mathbf{E}$ is a sparse error matrix.

If $\mathbf{A} = \mathbf{X}$ as in the low-rank subspace clustering problem for self representation, its optimization problem becomes:

$$\min_{\mathbf{Z}} \|\mathbf{Z}\|_*, \quad s.t. \;\; \mathbf{X} = \mathbf{XZ} \tag{2}$$

where the rank of $\mathbf{Z}$ is approximated by its nuclear norm.

Rather than using low-rank representation where the whole data space is applied as representation space, the sparse subspace clustering searches a small number of atoms for representation, which may contain more local information in the dataset. Then the minimization problem can be expressed as:

$$\min_{\mathbf{Z}} \|\mathbf{Z}\|_1, \quad s.t. \;\; diag(\mathbf{Z}) = 0 \tag{4}$$

where $diag(\mathbf{Z}) = 0$ is imposed to remove trivial solutions.

According to [19-22], sparse representation contains major local structure information of dataset, while low-rank representation focuses on global structure information of the dataset. Thus, the low-rank sparse subspace clustering (LRSSC) proposed by [13] is to combine sparse and low-rank representations so as to handle both global and local structure information. The LRSSC can be formulated as

$$\min_{\mathbf{Z}} \alpha_1 \|\mathbf{Z}\|_* + \alpha_2 \|\mathbf{Z}\|_1, \quad s.t. \;\; \mathrm{diag}(\mathbf{Z}) = 0 \tag{5}$$

where α_1 is low-rank and α_2 is sparsity constraints. Then a symmetric affinity matrix $\mathbf{W}$ can be calculated as:

$$\mathbf{W} = |\mathbf{Z}| + |\mathbf{Z}|^T \tag{6}$$

Finally, the spectral clustering [23] can be applied to achieve clustering.

B. Multi-view low-rank sparse subspace clustering

Intuitively, multi-view learning could improve the performance of unsupervised classification or clustering, because multi-view learning could provide more information than single-view learning. For the MLRSSC, let a new dataset with t views be denoted as $\tilde{\mathbf{X}} = [\mathbf{X}^1, \cdots, \mathbf{X}^i, \cdots, \mathbf{X}^t]$, where the i-th view $\mathbf{X}^i = \{\mathbf{x}_j^i\}_{j=1}^N \in R^{Di}$ containing D^i dimension features are extracted from the original data $\mathbf{X}$. Now the joint optimization problem with t views can be formulated as

$$\min_{\mathbf{Z}^1, \mathbf{Z}^2, \cdots, \mathbf{Z}^t} \sum_{i=1}^{t} \left(\alpha_1 \|\mathbf{Z}^i\|_* + \alpha_2 \|\mathbf{Z}^i\|_1 \right) + \lambda \sum_{1 \le i, j \le t, t \ne j} \|\mathbf{Z}^i - \mathbf{Z}^j\|_F^2,$$
$$s.t. \;\; \mathbf{X}^i = \mathbf{X}^i \mathbf{Z}^i, \; diag(\mathbf{Z}^i) = 0 \tag{7}$$

where the regularization parameter of each view λ should be different. However, for simplicity and practical usage, they are simply assumed identical. In addition to the low-rank and sparsity constraints, the third term in Eq. (7) encourages the representations from different views to be consistent if possible.

The alternating direction method of multipliers (ADMM) can be applied to solve this convex optimization problem in Eq. (7). Finally, spectral clustering is applied to the affinity matrix $\mathbf{W}$ which is also generated according to Eq. (6) using the $\mathbf{Z}$ from Eq. (7).

C. Parallel Multi-view low-rank sparse subspace clustering for Hyperspectral Images

Due to high spatial and spectral correlations in hyperspectral images, both spatial and spectral information should be considered. Accordingly, spatial and spectral views can be constructed. Specifically, in this research, spectral partitioning based on correlation coefficients is applied to generate multiple spectral views, and highly correlated spectral groups are considered as a type of views. Principal component analysis (PCA) is deployed to remove spectral correlation and eliminate noise, and the principal components are considered as a view. As for spatial views, we used morphological features which are extracted from the first principal component (PC), and the coarse, fine, high-contrast, low-contrast, horizontal and vertical features are used to form a matrix as another view. Then the MLRSSC algorithm can be applied to the constructed multi-view dataset $\tilde{\mathbf{X}}$ [18].

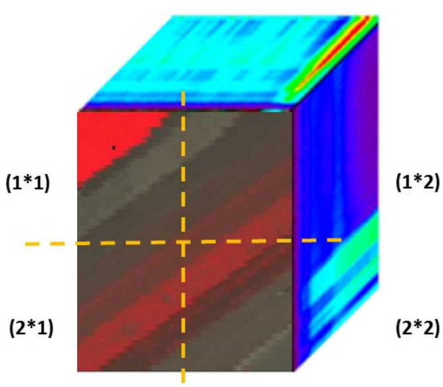
Due to very high computational cost, the subspace learning type of clustering cannot process large-scale images. Thus, we propose to partition an original image into non-overlapping blocks, and apply the MLRSSC algorithm to each block. The final clustering result can be generated by merging the results from each block through comparing spectral similarity of mean vectors. It is worth mentioning that to solve the similar computational problem of image clustering, a uniform data sampling technique is used in [12], where a pixel is chosen from an $M \times M$ local window, resulting in M^2 subimages. However, this technique is inapplicable here, because local spatial information needs to be preserved for multi-view clustering.

III. EXPERIMENT

The SalinasA dataset is acquired by the Airborne Visible and Infrared Imaging Spectrometer (AVIRIS) sensor over the Valley of Salinas, Central Coast of California, in 1998, where spatial resolution is 3.7m with 204 spectral bands. The false-color image of SalinasA and block partitioning are shown in Fig. 1. This image scene is about agricultural fields with 16 classes in total.

For spectral partitioning, the threshold of correlation coefficient is chosen as 0.8 according to Fig. 2, resulting in 29 spectral groups. This means there are 29 views based on the spectral features in the following processing.

Fig. 3 shows the clustering results with different number of PCs. When only PCs are used (as shown in blue bar in Fig. 3), the accuracy is improved with the number of PCs being increased to 10, and then the accuracy becomes almost stable. The accuracy reaches the maximum around using 50 PCs, which is about 85%. This means the major information in the original data are contained in the first 10 PCs. When spectral partitions are jointly used with PCs (orange bar in Fig.3), there is significant improvement in classification accuracy, where the highest accuracy can be as high as 94% with 50 PCs views and 29 spectral partition views. When spatial views are added, the accuracy can be further improved. Note that using 100 PCs cannot improve the accuracy, because minor PCs may contain noise only without discriminant information for class separation.



Blocks	Number of Classes	Number of Pixels
(1*1)	3	1265
(1*2)	4	1367
(2*1)	4	1344
(2*2)	3	1446

Fig. 1. The block partitioning of SalinasA cube, and classes and labeled pixels in each block.

Table I compares several subspace clustering methods with the MLRSSC, which include SSC, LRSSC, MLRSSC with PCs only, MLRSSC with spectral partitions only, and MLRSSC with morphological features only. According to the results, MLRSSC with all views can provide the best performance, yielding the accuracy as high as 88%. The classical SSC and LRSSC produce less accurate results, which are 82% and 84% respectively. The MLRSSC with spectral partition views and the MLRSSC with PC views offer lower accuracy than the MLRSSC with all views. In particular, the performance of MLRSSC with morphological features only is the worst, which may indicate that spectral information is more important than spatial information in this hyperspectral image classification problem.

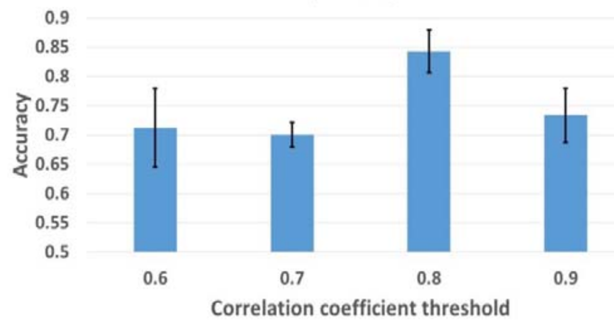


Fig. 2 clustering accuracy with different correlation coefficient threshold.

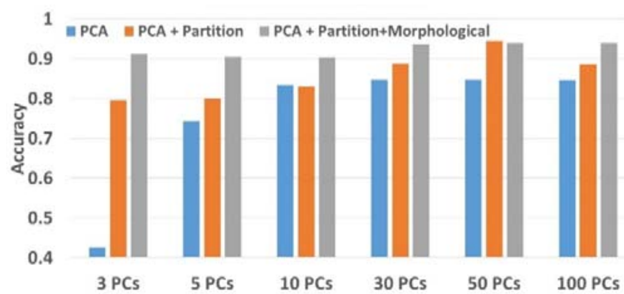


Fig. 3 The clustering accuracy with varied views in the SalinasA experiment.

TABLE I
ACCURACY OF CLUSTERING IN EACH BLOCK OF SALINASA DATASET WITH VARIED ALGORITHMS

BLOCK	SSC	LRSSC	MLRSSC (PCA)	MLRSSC (SPECTRAL PARTITION)	MLRSSC (MORPHOLOGY)	MLRSSC (ALL)
(1*1)	0.934	0.960	**0.976**	0.960	0.666	0.950
(1*2)	0.929	0.914	0.831	0.894	0.655	**0.941**
(2*1)	0.694	0.731	0.851	0.713	0.742	**0.757**
(2*2)	0.719	0.768	0.565	0.739	0.631	**0.874**
OA	0.819	0.843	0.806	0.826	0.673	**0.881**

IV. Conclusions

In this paper, parallel MLRSSC is applied to hyperspectral image clustering, which makes the computational expensive MLRSSC feasible to large-scale remote sensing images. The multiple views are extracted by spectral partitioning, morphological filtering, and PCA. The MLRSSC can outperform other single-view subspace clustering methods, such as SSC, LRSSC.

Acknowledgment

This work is partially supported by the National Natural Science Foundation of China (Grant No. 61420106007).

References

[1] J. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, “Hyperspectral remote sensing data analysis and future challenges,” *IEEE Geoscience and remote sensing magazine* vol. 1, pp. 6-36, Jun. 2013.

[2] M. Farid, and L. Bruzzone, “Classification of hyperspectral remote sensing images with support vector machines,” *IEEE Transactions on geoscience and remote sensing,* vol. 42, pp. 1778-1790, Aug. 2004.

[3] A. Sumarsono and Q. Du, “Low-rank subspace representation for estimating the number of signal subspaces in hyperspectral imagery,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 11, pp. 6286-6292, Nov. 2015.

[4] W. Sun, A. Halevy, J. Benedetto, W. Czaja, C. Liu, H. Wu, B. Shi, and W. Li, “UL-Isomap based nonlinear dimensionality reduction for hyperspectral imagery classification,” *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 89, pp. 25-36, 2014.

[5] X. Bian, C. Chen, Y. Xu, and Q. Du, “Robust hyperspectral image classification by multi-layer spatial-spectral sparse representations,” *Remote Sensing,* vol. 8, pp. 985, 2016.

[6] J. Peng, and Q. Du, “Robust joint sparse representation based on maximum correntropy criterion for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing,* vol. 55, no. 12, pp. 7152-7164, 2017.

[7] Y. Tarabalka, J. A. Benediktsson and J. Chanussot, “Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques,” *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2973–2987, Aug. 2009.

[8] H. Xie, A. Zhao, S. Huang, J.hang, S. Liu, X. Xue, X. Luo, H. Pan, Q. Du, and X. Tong, “Unsupervised hyperspectral remote sensing image clustering based on adaptive density,” *IEEE Geoscience and Remote Sensing Letters*, vol. 15, pp. 632-636, Apr. 2018.

[9] A. Paoli, F. Melgani, and E. Pasolli, “Clustering of hyperspectral images based on multiobjective particle swarm optimization,” *IEEE transactions on geoscience and remote sensing,* vol. 47, pp. 4175-4188, Dec. 2009.

[10] E. Elhamifar, and R. Vidal, “Sparse subspace clustering: Algorithm, theory, and applications,” *IEEE transactions on pattern analysis and machine intelligence,* vol. 35, pp. 2765-2781, 2013.

[11] P. Favaro, R. Vidal, A. Ravichandran, “A closed form solution to robust subspace estimation and clustering,” *CVPR*, pp. 1801–1807, 2011.

[12] A. Sumarsono and Q. Du, “Low-rank subspace representation for supervised and unsupervised classification of hyperspectral image classification,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, pp. 4188-4195, Sep. 2016.

[13] Y. Wang, H. Xu, and C. Leng, “Provable subspace clustering: When LRR meets SSC,” *Advances in Neural Information Processing Systems*, pp. 64-72. 2013.

[14] C. Xu, D. Tao, and C. Xu “A survey on multi-view learning,” *arXiv preprint arXiv:1304.5634,* 2013.

[15] J. Zhao, X. Xie, X. Xu, and S. Sun, “Multi-view learning overview: Recent progress and new challenges,” *Information Fusion,* vol. 38, pp. 43-54, Nov. 2017.

[16] J. Li, X. Huang, P. Gamba, J. Bioucas-Dias, L. Zhang, J. Benediktsson, and A. Plaza, “Multiple feature learning for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing* vol, 53, pp. 1592-1606, Mar. 2015.

[17] M. Brbić and I. Kopriva, “Multi-view low-rank sparse subspace clustering,” *Pattern Recognition,* vol. 73, pp. 247-258, Jan. 2018.

[18] L. Tian, Q. Du, I. Kopriva, and N. Younan, “Spatial-spectral based multi-view low-rank sparse subspace clustering for hyperspectral imagery,” *Proceedings of IEEE Geoscience and Remote Sensing Symposium*, Valencia, Spain, July 2018.

[19] E. Elhamifar, G. Sapiro, and R. Vidal, “See all by looking at a few: Sparse modeling for finding representative objects,” *in Proc. IEEE CVPR*, pp. 1600-1607, 2012.

[20] H. Shen and J. Z. Huang, “Sparse principal component analysis via regularized low rank matrix approximation,” *J. Multivariate Anal*., vol. 99, pp. 1015–1034, Jul. 2008.

[21] R. Rubinstein, R. Zibulevsky, and M. Elad, “Double sparsity: Learning sparse dictionaries for sparse signal approximation,” *IEEE Trans. Signal Process*., vol. 58, pp. 1553–1564, Mar. 2010.

[22] G. Liu, Z.Lin, Y. Yu, “Robust subspace segmentation by low-rank representation,” *in: Proceedingsof the 27th International Conferenceon Machine Learning (ICML-10),* pp.663–670, 2010.

[23] A. Y. Ng, M. I. Jordan, Y. Weiss, “On spectral clustering: Analysis and an algorithm,” *Advances in Neural Information Processing Systems*, pp. 849–856, 2001.