A New Sparse Subspace Clustering Algorithm for Hyperspectral Remote Sensing Imagery

Han Zhai, *Student Member, IEEE*, Hongyan Zhang, *Member, IEEE*, Liangpei Zhang, *Senior Member, IEEE*, Pingxiang Li, *Member, IEEE*, and Antonio Plaza, *Fellow, IEEE*

***Abstract*— Robust techniques such as sparse subspace clustering (SSC) have been recently developed for hyperspectral images (HSIs) based on the assumption that pixels belonging to the same land-cover class approximately lie in the same subspace. In order to account for the spatial information contained in HSIs, SSC models incorporating spatial information have become very popular. However, such models are often based on a local averaging constraint, which does not allow for a detailed exploration of the spatial information, thus limiting their discriminative capability and preventing the spatial homogeneity of the clustering results. To address these relevant issues, in this letter, we develop a new and effective ℓ_2-norm regularized SSC algorithm which adds a four-neighborhood ℓ_2-norm regularizer into the classical SSC model, thus taking full advantage of the spatial-spectral information contained in HSIs. The experimental results confirm the potential of including the spatial information (through the newly added ℓ_2-norm regularization term) in the SSC framework, which leads to a significant improvement in the clustering accuracy of SSC when applied to HSIs.**

***Index Terms*— Hyperspectral images (HSIs), ℓ_2-norm regularization, sparse subspace clustering (SSC), spectral clustering.**

I. Introduction

HYPERSPECTRAL images (HSIs) provide a wealth of spectral information, resulting in spectroscopic diagnostic features that support fine land-cover classification and clustering [1], [2]. Due to the difficulty of labeling a large number of training samples for the supervised classification of high-dimensional HSI data, clustering has been widely used in various applications and offers an unsupervised alternative. However, clustering is usually a very challenging task because of the large spectral variability and complex spatial structures present in HSIs.

To date, many different clustering methods with various working mechanisms have been proposed for HSIs, such as k-means [3], the unsupervised artificial immune network for remote sensing image classification [4], spectral curvature clustering [5], clustering by fast search and find of density peaks (CFSFDP) [6], and so on. However, due to the fact that only spectral information is used to discriminate the different classes and the spatial contextual information present in the data is largely ignored, a large amount of salt-and-pepper noise generally appears in the final cluster map, and the discriminative ability is limited. Previous studies have demonstrated that spatial information should be simultaneously utilized together with spectral information to assist with the clustering of HSIs [7]. In recent years, many spectral-spatial clustering algorithms have been proposed to enhance the clustering accuracy by incorporating spatial information, such as fuzzy c-means clustering with spatial constraints (FCM_S1) [8], Markov random field clustering [9], and spatially constrained k-means [10]. Unfortunately, most of these methods still suffer from errors due to the uniform feature point distribution within the feature space, which is often caused by the large spectral variability observed in HSIs.

Recently, a robust technique known as the sparse subspace clustering (SSC) model has been successfully applied to perform HSI clustering and has shown great potential [11], [12]. However, the SSC model only focuses on analyzing the spectral features without considering the spatial information, which limits both its discriminative capability and the spatial homogeneity of the final clustering result. In fact, the sparse coefficient matrix should be piecewise smooth because of the spatial relationship and contextual dependence between the representation coefficient vector of the center pixel and its neighbors. In our previous work [11], the spectral-spatial SSC (SSC-S) algorithm was proposed to improve the clustering accuracy. SSC-S can help to guarantee spatial smoothness and can reduce the representation bias by enforcing a local (eight-connected neighborhood) averaging constraint on the coefficient matrix, based on the assumption that pixels in a local window often have similar spectra and belong to the same class. However, this local averaging strategy for incorporating the spatial information is very heuristic and can often be unrepresentative, especially in complex land-cover distribution areas.

To address this issue, this letter introduces a new ℓ_2-norm regularized SSC (L2-SSC) algorithm for HSIs that fully exploits the spatial-spectral information contained in the data. This method incorporates spatial contextual information into the SSC framework by means of the ℓ_2-norm regularizer. Compared with SSC-S, this approach utilizes the spatial information to constrain the spectral representation in a more elaborate way. Adjacent pixels in the four-connected neighborhood

Manuscript received September 22, 2016; revised November 1, 2016; accepted November 2, 2016. Date of publication November 18, 2016; date of current version December 26, 2016. This work was supported by the National Natural Science Foundation of China under Grant 41571362 and Grant 41431175. *(Corresponding author: Hongyan Zhang.)*

H. Zhai, H. Zhang, L. Zhang, and P. Li are with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan 430079, China (e-mail: zhanghongyan@whu.edu.cn).

A. Plaza is with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, Escuela Politecnica, University of Extremadura, E-10071 Cáceres, Spain.

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2016.2625200

model are thus incorporated into the analysis process, which can effectively promote piecewise smoothness of the sparse coefficient matrix. The ℓ_2-norm spatial regularizer can better account for the spatial homogeneity of the final clustering result, as neighboring pixels generally belong to the same thematic class and are likely to have similar representation coefficients with respect to the same sparse representation basis, according to the scheme of sparse representation.

II. L2-Norm Regularized Sparse Subspace Clustering Algorithm

In this section, the SSC model is first briefly reviewed. We then show how we integrate the ℓ_2-norm regularizer into the SSC model to effectively incorporate the spatial information, which can significantly enhance the clustering results for HSIs by exploiting the piecewise smoothness of the sparse coefficient matrix.

A. Clustering HSIs With Sparse Subspace Clustering

In the classical SSC model for HSIs, all the pixels are assumed to be sampled from a union of subspaces $\bigcup_{i=1}^{l} S_i$, with each subspace corresponding to a certain land-cover type, where l denotes the number of subspaces or classes. It then exploits the self-representation property of the data to represent each pixel as a linear or affine combination of the others [13]. In addition, it depends on the assumption that the nonzero coefficients lie in the corresponding class-dependent low-dimensional subspace [11]. With the 2-D reordered HSI matrix being used as the dictionary, the SSC optimization problem can be modeled as follows:

$$\begin{aligned} &\min_{\mathbf{C},\mathbf{N}} \ \|\mathbf{C}\|_1 + \lambda \|\mathbf{N}\|_F^2 \\ &\text{s.t. } \mathbf{Y} = \mathbf{YC} + \mathbf{N}, \quad \text{diag}(\mathbf{C}) = 0, \ \mathbf{C}^T 1 = 1 \end{aligned} \tag{1}$$

where $\mathbf{Y} \in \mathbb{R}^{D\times \text{MN}}$ is the 2-D HSI matrix, D represents the number of bands, M refers to the height of the data, N stands for the width of the data, $\mathbf{C} \in \mathbb{R}^{\text{MN}\times\text{MN}}$is the sparse coefficient matrix, $\mathbf{N} \in \mathbb{R}^{D\times \text{MN}}$ is the representation error matrix, and λ is the sparsity/noise tradeoff parameter. The constraint $\text{diag}(\mathbf{C}) = 0$ is used to eliminate the trivial solution of representing a pixel as itself [13]. In addition, the constraint $\mathbf{C}^T 1 = 1$ shows that it is a case of an affine subspace, with 1 denoting a vector consisting of only ones. The ℓ_1-norm regularization can promote the sparseness of $\mathbf{C}$ and is guaranteed to be subspace-preserving [13].

This model can be easily solved using the alternating direction method of multipliers (ADMM) [14]. We then use the obtained sparse coefficient matrix $\mathbf{C}$ to construct the adjacent matrix $\mathbf{W} \in \mathbb{R}^{\text{MN}\times\text{MN}}$, which is called the similarity graph [11], [13], with each element standing for the similarity between two pixels

$$\mathbf{W} = |\mathbf{C}| + |\mathbf{C}|^T. \tag{2}$$

The "symmetrization" form in (2) is adopted to strengthen the connection of the graph, and the final clustering result is achieved by applying the spectral clustering algorithm to the similarity graph [15], [16].

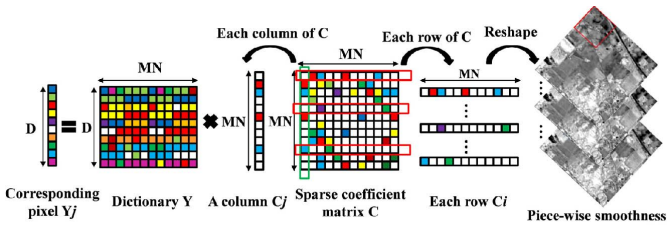

Fig. 1. Graphical interpretation of each column and each row of the sparse coefficient matrix $\mathbf{C}$.

B. L2-Norm Regularized Sparse Subspace Clustering Model

By reordering the 2-D matrix $\mathbf{C} \in \mathbb{R}^{\text{MN}\times\text{MN}}$ into a 3-D cube $\hat{\mathbf{C}} \in \mathbb{R}^{M\times N\times\text{MN}}$ along the rows, as shown in Fig. 1, each band of $\hat{\mathbf{C}}$ denotes the representation coefficient distribution of the whole image with respect to a single atom, and can be seen as a specific "fractional abundance" with respect to an "endmember" in the unmixing domain (using spectral unmixing jargon). Considering that a specific land-cover material should be regionally distributed in the image, i.e., two spatially neighboring pixels in an HSI usually have a high probability of belonging to the same thematic class, and then, their representation coefficients should also be very close, with respect to the same sparse basis, according to the working mechanism of the SSC framework. Therefore, each band of $\hat{\mathbf{C}}$ should be piecewise smooth, which means that it is reasonable to apply a spatial constraint to it to improve the clustering performance.

Based on this fact, it is natural for us to utilize the neighboring coefficient vectors to constrain the target coefficient vector to reduce the representation bias. In the SSC-S model [11], the mean of the local window is utilized to regularize the center representation coefficient vector, assuming that pixels in a small window should share the same dominant subspace. However, this local averaging constraint is heuristic. That is to say, it fails to exploit the accurate relationship between the target coefficient vector and its neighbors, especially for the case in which the difference between the mean coefficient in the local window and the center coefficient is large. Hence, SSC-S obtains a relatively limited improvement over the performance of SSC.

At this point, considering the ability of the total variation (TV) regularizer to promote piecewise smoothness and preserve edges [17], it seems natural to add the TV spatial regularization term into the SSC model framework to model the sparse optimization problem in the following formulation, to better account for the spatial neighborhood information and promote the piecewise smoothness of $\mathbf{C}$:

$$\begin{aligned} &\min_{\mathbf{C},\mathbf{N}} \|\mathbf{C}\|_1 + \frac{\lambda}{2}\|\mathbf{N}\|_F^2 + \frac{\alpha}{2}\text{TV}(\mathbf{C}) \\ &\text{s.t. } \mathbf{Y} = \mathbf{YC} + \mathbf{N}, \quad \text{diag}(\mathbf{C}) = 0, \ \mathbf{C}^T 1 = 1 \end{aligned} \tag{3}$$

where α trades off the spectral and spatial terms, with the TV constraint defined as

$$\text{TV}(\mathbf{C}) = \sum_{\{i,j\}\in\gamma} \|\mathbf{C}_i - \mathbf{C}_j\|_1. \tag{4}$$

Algorithm 1 L2-SSC for HSIs

Input:

1) A 2-D matrix of the HSI containing a set of points $\{y_i\}_{i=1}^{MN}$, in a union of l affine subspaces $\{S_i\}_{i=1}^{l}$;
2) Parameters, including the cluster number l and the regularization parameters λ and α.

Main algorithm:

1) Construct the L2-SSC optimization model (5) and solve it to obtain $\mathbf{C}$ using ADMM;
2) Normalize the columns of $\mathbf{C}$ as $\mathbf{C}_i \leftarrow \frac{\mathbf{C}_i}{\|\mathbf{C}_i\|_\infty}$;
3) Construct the similarity graph with (2);
4) Apply spectral clustering to the similarity graph to obtain the final clustering results.

Output :

A 2-D matrix which records the class labels of the HSI clustering result.

This is a vector extension of anisotropic TV [17], where γ denotes the set of horizontal and vertical neighbors. However, it should be noted that the optimization problem in (3), although convex, is very hard to solve owing to the nonsmooth term (the TV spatial regularizer [17]). In addition, unlike the unmixing problem, where the endmember dictionary is usually very small, the dictionary $\mathbf{Y}$ in the SSC framework is much larger, which leads to a much higher dimensional sparse coefficient matrix $\mathbf{C}$ compared with the abundance matrix. As a result, it would take a very long time to achieve convergence for the optimization problem in (3).

To this end, an effective ℓ_2-norm spatial regularizer is utilized instead of the TV regularizer to incorporate the spatial information, in order to further assist with the spectral analysis of SSC. This is also a four-connected neighborhood spatial constraint, like the TV regularizer, but it is much easier to solve than TV because it is both a convex and smooth term. Compared with the local averaging constraint, it incorporates the spatial information in a much more elaborate way and can reduce the artificial error caused by the deviation between the target coefficient vector and the neighboring coefficient vectors to a minimum. By incorporating the ℓ_2-norm spatial regularizer into the SSC model, the L2-SSC problem can be modeled as follows:

$$\min_{\mathbf{C},\mathbf{N}} \|\mathbf{C}\|_1 + \frac{\lambda}{2}\|\mathbf{N}\|_F^2 + \frac{\alpha}{2}\sum_{\{i,j\}\in\gamma}\|\mathbf{C}_i - \mathbf{C}_j\|_2^2$$
$$\text{s.t.} \quad \mathbf{Y} = \mathbf{YC} + \mathbf{N}, \quad \text{diag}(\mathbf{C}) = 0, \mathbf{C}^T 1 = 1. \tag{5}$$

Equation (5) can be efficiently solved with the ADMM [14]. In the same way as SSC, $\mathbf{C}$ is used to build the similarity graph. The final result can then be achieved by applying the spectral clustering algorithm to the similarity graph.

C. L2-SSC Algorithm Flowchart

The proposed L2-SSC is summarized in Algorithm 1.

III. Experimental Results and Discussion

In order to evaluate the performance of the proposed L2-SSC algorithm, the following spectral and spectral-spatial clustering methods were selected as benchmarks: k-means [3], FCM [18], CFSFDP [6], SSC [13], FCM_S1 [8], and SSC-S [11]. A common strategy adopted in clustering is to treat the number of clusters as a prior, and then manually determines the thematic information for each cluster group through cross-referencing between the clustering result and the image. In our experiments, all the parameters of each clustering method were manually tuned to the optimum. In order to thoroughly evaluate the clustering performance of each method, both visual cluster maps and quantitative evaluations of the precision (producer's accuracy, overall accuracy (OA), and kappa coefficient) are given.

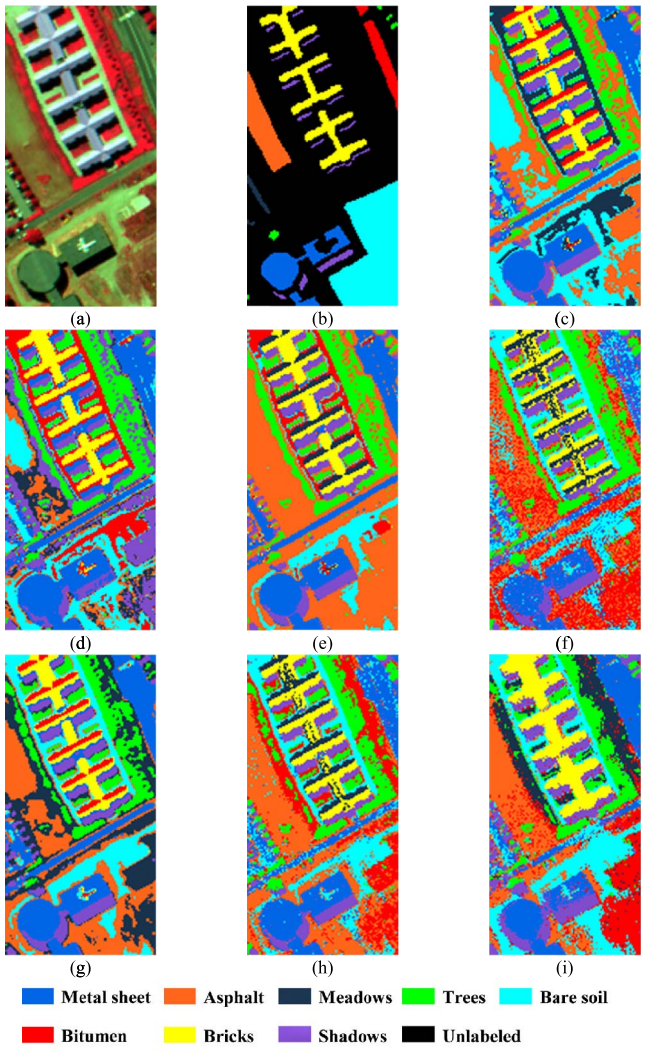

Fig. 2. Cluster maps of the different methods with the ROSIS Pavia University image. (a) False-color image (R: 102, G: 56, B: 31). (b) Ground truth. (c) k-means. (d) FCM. (e) CFSFDP. (f) SSC. (g) FCM_S1. (h) SSC-S. (i) L2-SSC.

A. Experimental Results and Analysis

The proposed method was evaluated on two widely used hyperspectral data sets: the Pavia University image and the Washington DC Mall image. The first scene was collected by the Reflective Optics System Imaging Spectrometer (ROSIS) sensor, at a size of 610 × 340 × 103 with a 1.3-m geometric resolution and nine main classes. A typical subset of

TABLE I
QUANTITATIVE EVALUATION OF THE DIFFERENT CLUSTERING ALGORITHMS FOR THE ROSIS PAVIA UNIVERSITY IMAGE

Method	Class	k-means	FCM	CFSFDP	SSC	FCM_S1	SSC-S	L2-SSC
Producer's accuracy (%)	Metal sheet	**100**	82.59	**100**	85.58	99.72	98.60	97.34
	Asphalt	21.09	25.36	74.60	52.60	60.32	**95.83**	60.32
	Meadows	**0**	**0**	**0**	**0**	**0**	**0**	**0**
	Trees	66.67	99.88	**100**	**100**	87.24	**100**	87.67
	Bare soil	35.05	27.86	20.59	22.04	24.89	25.79	**37.20**
	Bitumen	0	0	0	**0.71**	0	0	**0.71**
	Bricks	58.63	76.19	60.61	53.84	60.99	52.32	**93.91**
	Shadows	**100**	57.62	99.87	98.61	**100**	99.45	92.97
OA (%)		47.79	47.54	52.12	43.97	52.24	52.04	**63.30**
Kappa		0.3629	0.3737	0.4421	0.3495	0.4481	0.4419	**0.5681**

TABLE II
QUANTITATIVE EVALUATION OF THE DIFFERENT CLUSTERING ALGORITHMS FOR THE WASHINGTON DC MALL IMAGE

Method	Class	k-means	FCM	CFSFDP	SSC	FCM_S1	SSC-S	L2-SSC
Producer's accuracy (%)	Water	**100**	**100**	**100**	99.49	**100**	98.41	98.84
	Grass	84.46	87.40	97.32	97.39	87.41	95.69	**98.66**
	Tree	97.36	97.09	99.17	6.30	**100**	21.18	79.53
	Square	30.27	32.23	23.24	0	92.23	94.90	**95.44**
	Roof	97.35	89.93	96.76	48.68	97.74	94.06	**98.21**
	Road	41.56	45.25	78.46	**97.04**	46.32	79.53	79.22
OA (%)		82.42	82.80	86.16	76.53	90.92	88.73	**94.55**
Kappa		0.7558	0.7610	0.8066	0.6569	0.8726	0.8421	**0.9229**

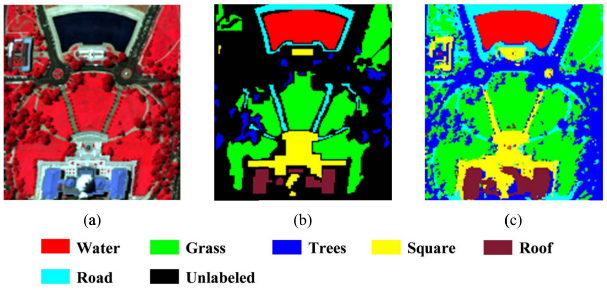

Fig. 3. Cluster map of the L2-SSC method with the Washington DC Mall image. (a) False-color image (R: 63, G: 52, B: 36). (b) Ground truth. (c) L2-SSC.

200 × 100 was selected as our test data, containing eight main classes [11]. The second data set was acquired by the hyperspectral digital imagery collection experiment sensor over the Washington DC Mall, at a size of 1208 × 307 × 109. Similarly, a typical subset of 160 × 145 was selected as the test data, with six main classes. The false-color composites and the cluster maps obtained with these two scenes are shown in Figs. 2 and 3, with the corresponding quantitative evaluations provided in Tables I and II, respectively. The best result of each row is shown in bold, with the second-best result underlined.

From Figs. 2 and 3 and Tables I and II, it can be clearly seen that the spectral-only clustering methods fail to achieve satisfactory results, but the spectral-spatial methods significantly improve the performance by considering the spatial information. Specifically, k-means and FCM obtain inferior clustering performances for both scenes, with large amounts of misclassification and salt-and-pepper noise due to their weak discriminative capability for HSIs. Although CFSFDP obtains a better performance with smoother cluster maps, the accuracy is still unsatisfactory, because the density model cannot deal well with the complex structure of HSIs. Compared with FCM, FCM_S1 improves the clustering performance to a large degree for both experiments to obtain much smoother cluster maps, which clearly reflects the importance of incorporating spatial information.

We now turn to the clustering results of the three subspace-based methods. The original SSC algorithm performs badly in both scenes, with a great deal of misclassification and salt-and-pepper noise, which leads to a low clustering accuracy. In comparison, SSC-S improves the performance of SSC to some degree by exploiting the inherent spatial-spectral duality of the HSIs with the local averaging constraint to obtain smoother cluster maps. However, these improvements are still limited and are far away from fully exploiting the potential of SSC. As the local averaging strategy is very heuristic and empirical, it fails to exploit the accurate spatial relationship and dependence between the target representation coefficient vector and the neighboring coefficient vectors. Compared with SSC-S, the proposed L2-SSC algorithm significantly improves the performance of SSC, decreasing the misclassification to a great degree and effectively promoting spatial smoothness by exploiting the spatial contextual information from a much more elaborate perspective. As a result, much smoother and more accurate clustering results are achieved. In summary, L2-SSC achieves the best performance in both the visual and quantitative evaluations. Improvements of more than 11% and 6% in OA are achieved compared with SSC-S for the Pavia University image and Washington dc image, respectively, which proves the effectiveness and superiority of the proposed L2-SSC method.

B. Parameter Analysis

There are two main parameters in the proposed L2-SSC algorithm: the regularization parameter λ and the spectral/spatial tradeoff parameter α. Parameter λ acts as a tradeoff

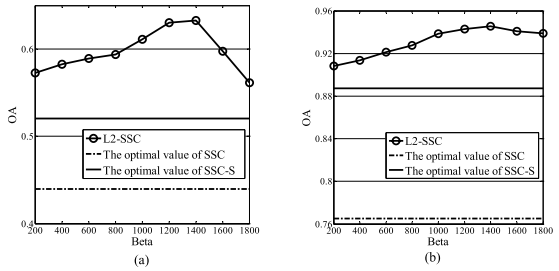

Fig. 4. Analysis of parameter β: change in the OA with various values of β. (a) Pavia University. (b) Washington DC Mall.

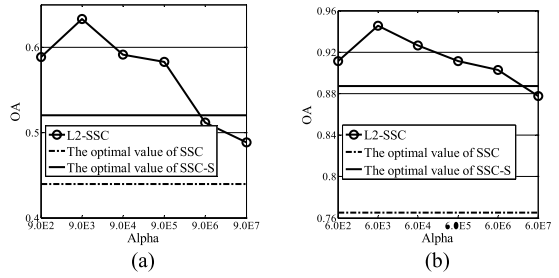

Fig. 5. Analysis of parameter α: change in the OA with various values of α. (a) Pavia University. (b) Washington DC Mall.

to balance the sparsity term and the data fidelity term. Indeed, it is decided by the following formulation [11], [13]:

$$\lambda = \beta/\mu \tag{6-a}$$

$$\mu \triangleq \min_{i} \max_{j \neq i} \left| y_i^T y_j \right|. \tag{6-b}$$

From (6-a) and (6-b), it can be clearly seen that λ is actually decided by β, as μ is fixed for a certain data set. Hence, in practice, we only need to fine-tune β. An accuracy change curve of the OA with various values of β is drawn for each data set to find the optimum value, as shown in Fig. 4. From Fig. 4, it can be observed that β is independent of the data set, to some degree, as the optimal value always lies in the range of [1000, 1600] and can be easily fine-tuned for each data set. In addition, the optimal clustering accuracy values for the SSC and SSC-S algorithms are also plotted on these figures. It can be seen that L2-SSC always achieves a better clustering accuracy than the optimal values of SSC and SSC-S, which further confirms its effectiveness.

Similarly, an OA versus α curve is drawn for the spectral/spatial tradeoff parameter α to explore its influence on the clustering performance, as shown in Fig. 5. As can be seen from Fig. 5, the spatial information plays a very important role in the clustering process. This parameter is also relatively stable, as the optimal value always falls in a narrow range of $(0, 10] \times 10^3$ and can be easily fine-tuned for a certain data set. Over a large range of α, L2-SSC can achieve a better clustering accuracy than the optimal values of SSC and SSC-S. Hence, the L2-SSC algorithm does make sense for real applications.

IV. Conclusion

In this letter, we have proposed a new L2-SSC algorithm for HSIs. By incorporating the spatial contextual information with the ℓ_2-norm regularizer, which more elaborately constrains the spectral representation from the four-connected neighborhood perspective, the proposed algorithm can accurately exploit the spatial contextual information and dependence between the target representation coefficient vector and the neighboring coefficient vectors. In this way, the piecewise smoothness of the sparse coefficient matrix and the homogeneity of the final clustering result can be effectively enhanced, leading to consistent results from both the spectral and spatial points of view. The experimental results clearly show that the proposed L2-SSC algorithm achieves a state-of-the-art clustering performance for HSIs. However, the proposed method could be further improved by the adaptive determination of the regularization parameters and efficient implementation in a high performance computing architecture, which will be in our future work.

References

[1] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.

[2] T. V. Bandos, L. Bruzzone, and G. Camps-Valls, "Classification of hyperspectral images with regularized linear discriminant analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 3, pp. 862–873, Mar. 2009.

[3] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982.

[4] Y. Zhong, L. Zhang, and W. Gong, "Unsupervised remote sensing image classification using an artificial immune network," *Int. J. Remote Sens.*, vol. 32, no. 19, pp. 5461–5483, 2011.

[5] G. Chen and G. Lerman, "Spectral curvature clustering (SCC)," *Int. J. Comput. Vis.*, vol. 81, no. 3, pp. 317–330, Mar. 2009.

[6] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science*, vol. 344, no. 6191, pp. 1492–1496, 2014.

[7] S. Li, B. Zhang, A. Li, X. Jia, L. Gao, and M. Peng, "Hyperspectral imagery clustering with neighborhood constraints," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 3, pp. 588–592, May 2013.

[8] S. Chen and D. Zhang, "Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 34, no. 4, pp. 1907–1916, Aug. 2004.

[9] Y. Tarabalka, M. Fauvel, J. Chanussot, and J. A. Benediktsson, "SVM- and MRF-based method for accurate classification of hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 736–740, Oct. 2010.

[10] M. Luo, Y. F. Ma, and H. J. Zhang, "A spatial constrained K-means approach to image segmentation," in *Proc. Joint Conf. 4th Int. Conf. Inf., Commun. Signal Process./4th Pacific Rim Conf. Multimedia*, Dec. 2003, pp. 738–742.

[11] H. Zhang, H. Zhai, L. Zhang, and P. Li, "Spectral–spatial sparse subspace clustering for hyperspectral remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3672–3684, Jun. 2016.

[12] W. Sun, L. Zhang, B. Du, W. Li, and Y. M. Lai, "Band selection using improved sparse subspace clustering for hyperspectral imagery classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2784–2797, Jun. 2015.

[13] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2765–2781, Nov. 2013.

[14] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, Mar. 2011.

[15] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient N-D image segmentation," *Int. J. Comput. Vis.*, vol. 70, no. 2, pp. 109–131, Nov. 2006.

[16] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.

[17] M. D. Iordache, J. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4484–4502, Nov. 2012.

[18] J. C. Bezdek, *Pattern Recognition With Fuzzy Objective Function Algorithms*. New York City, USA: Springer Science & Business Media, 2013.