AN EFFICIENT ALGORITHM FOR HYPERSPECTRAL IMAGE CLUSTERING

Yushu Pan [1], *Yuchen Jiao* [2], *Tiejian Li* [3], *and Yuantao Gu* [2]

[1]School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China
[2]Department of Electronic Engineering, Tsinghua University, Beijing 100084, China
[3]State Key Laboratory of Hydroscience and Engineering, Tsinghua University, Beijing 100084, China

ABSTRACT

Hyperspectral images (HSIs) clustering problem is a challenge and valuable task due to its inherent complexity and abundant spectral information. Sparse subspace clustering (SSC) and SSC-based methods are widely used in this problem and demonstrate excellent performance. However, considering that HSIs are usually of high dimension, these methods have expensive computing complexity because of the usage of SSC. To solve this problem, we propose a novel approach called SuperPixel and Angle-based HyperSpectral Image Clustering (SPAHSIC). It first extracts the local spectral and spatial information between pixels by superpixel segmentation, and then applies spectral clustering on the similarity matrix built based on subspace principal angles. We implement experiments on real datasets and get a high accuracy, which indicates the effectiveness of our algorithm.

***Index Terms*—** Hyperspectral images (HSIs), superpixel, subspace, principal angle, spectral clustering

1. INTRODUCTION

Hyperspectral images (HSIs) refer to the high-dimensional images acquired by sampling the solar reflection of earth surface. It collects a large amount of very narrow spectral spectrum, including the entire visible, near-IR, mid-IR, and thermal-IR, and thus contains much more information than traditional images [1]. HSIs have important applications in many fields, including mineral exploration [2], precision agriculture [3], and many others.

A fundamental problem in these applications, called HSI classification, is to separate pixels in HSI into different groups according to different land covers. According to whether examples are labeled, existing algorithms are generally divided into two categories, i.e., supervised algorithms and unsupervised algorithms. In real-world scenarios, label information required by supervised algorithms is unlikely or very expensive to be available. Therefore, applying HSI classification without label information, known as HSI clustering, is a well-motivated problem and has received much attention.

Some traditional clustering algorithms are directly applied to HSI clustering, including K-means [4], fuzzy c-means (FCM) [5], the density-based spatial clustering of applications with noise (DBSCAN) [6]. Due to limited discriminative information in spectral domain, most of these algorithms behave not so-well, especially when dealing with complex ground objects with a large diversity. In recent years, the subspace model for HSI is proposed, which assumes that pixels of the same land-cover class lie in the same subspace. Based on this model, SSC and SSC-based algorithms are introduced and demonstrate the current state-of-art performance [7, 8]. The high accuracy of many methods, such as spectral-spatial SSC (SSC-S) [9], joint sparsity based SSC (JSSC) [10] and spectral-spatial low-rank subspace clustering (SS-LRSC) [11], is achieved by utilizing spatial information under SSC model. However, because SSC requires optimizing a self-representation matrix, it brings high computing complexity and leads to expensive time cost.

In order to reduce the complexity of HSI clustering, we propose a new algorithm called SuperPixel and Angle-based HyperSpectral Image Clustering (SPAHSIC for short). SPAHSIC includes two steps. In the first step, raw pixels in the original HSI are divided in to a set of multiple superpixels based on local spatial and spectral information. In the second step, each superpixel is treated as a subspace and a similarity matrix is built based on the principal angles. Then spectral clustering is adopted to incorporate the global information and obtain a clustering result. The proposed algorithm does not need to calculate self-representation matrix and has much lower computing complexity. Besides, experiments implemented on two real datasets indicate that SPAHSIC achieves higher accuracy compared with the existing methods [1].

2. BACKGROUND

2.1. Superpixel Segmentation and SLIC Algorithm

Superpixel segmentation is the process of partitioning an image into non-overlapping multiple sub-regions, $\{P_i\}_{i=1,2,\cdots,K}$. Each superpixel contains the pixels with similar spectral information and spatial information. Many algorithms were proposed to solve this problem for traditional RGB images, such as watershed superpixel [12], entropy rate superpixel segmentation [13], and simple linear iterative clustering algorithms (SLIC) [14]. Among them, SLIC is popular for its segmentation performance and efficiency in computation and storage. Besides, it is straightforward to extend to supervoxel segmentation. We refer SLIC to design the superpixel segmentation part in our algorithm.

An algorithm parameter of SLIC is the desired number of superpixels, denoted as $\hat{K}$. SLIC works in the following way. It first initializes superpixel centers $\{C_i\}_{i=1,2,\cdots,\hat{K}}$ as grid points with uniform interval $S=(MN/\hat{K})^{1/2}$, where M and N denote the length and width of the image, respectively. To avoid choosing edge pixels or noisy pixels as the center, SLIC corrects C_i by replacing it with a neighbor pixel of the lowest gradient. The gradient is defined as

$$G(p,q)=\|\mathbf{I}(p+1,q)-\mathbf{I}(p-1,q)\|+\|\mathbf{I}(p,q+1)-\mathbf{I}(p,q-1)\|,$$

This work was partly funded by the National Natural Science Foundation of China (NSFC 61571263, 61531166005) and the National Key Research and Development Program of China (Project No. 2016YFE0201900, 2017YFC0403600). (Corresponding author: Yuantao Gu.)

[1]The supplementary downloadable material, including MATLAB codesfor all experiments, is available at http://gu.ee.tsinghua.edu.cn/publications/.

where $\mathbf{I}(p,q)$ denotes the (p,q)-th pixel of the processed image. Next for each pixel, we calcualte the distance between it and any cluster center whose $2S \times 2S$ neighbor area overlaps this pixel. The distance is defined as the sum of the color distance and the spatial distance normalized by the grid interval S.

$$\hat{D}(p_i, C_j) = d_{\mathrm{O}}(p_i, C_j) + (m/S) d_{\mathrm{A}}(p_i, C_j),$$

where

$$d_{\mathrm{A}}(p_i, C_j) = \sqrt{(x_i - \bar{x}_j)^2 + (y_i - \bar{y}_j)^2},$$
$$d_{\mathrm{O}}(p_i, C_j) = \sqrt{(l_i - \bar{l}_j)^2 + (a_i - \bar{a}_j)^2 + (b_i - \bar{b}_j)^2},$$

and (x_i, y_i), $(\bar{x}_j, \bar{y}_j)$ denotes the location of the i-th pixel and j-th cluster center, respectively. (l_i, a_i, b_i), $(\bar{l}_j, \bar{a}_j, \bar{b}_j)$ denotes the color represented in the CIELAB color space of the i-th pixel and j-th center, respectively. Then pixel i is associated to the nearest cluster center. After all pixels are associated, the centers are adjusted to be the mean vector of all the pixels belonging to the cluster. The l_1 norm is used to compute a residual error E between the new cluster center locations and previous cluster center locations. The assignment and adjustment steps can be repeated iteratively until the error E is smaller than some threshold τ. Pixels associated to the same center belong to the same superpixel.

2.2. Principal Angles

The principal angles (or canonical angles) between two subspaces provide the best way to characterize the relative subspace relationship [15]. Its definition is as below. Compared with Euclidean distance, principal angles ignore the influence of amplitude and reflect the similarity between two subspaces from a more accurate angle. We use it to measure the spectral distance between two pixels and build the similarity matrix of superpixels.

Definition 1. The principle angles $\{\theta_k\}_{k=1,2,\cdots,r}$ between two subspaces $\mathcal{S}_1$ and $\mathcal{S}_2$ of dimensions r, are recursively defined as

$$\cos\theta_k = \max_{\mathbf{x}_1 \in \mathcal{S}_1} \max_{\mathbf{x}_2 \in \mathcal{S}_2} \frac{\mathbf{x}_1^{\mathrm{T}} \mathbf{x}}{\|\mathbf{x}_1\| \|\mathbf{x}_2\|} = \frac{\mathbf{x}_{1k}^{\mathrm{T}} \mathbf{x}_{2k}}{\|\mathbf{x}_{1k}\| \|\mathbf{x}_{2k}\|},$$

with the orthogonality constraints $\mathbf{x}_i^{\mathrm{T}} \mathbf{x}_{ij} = 0$, $j = 1, \cdots, k-1$, $i = 1, 2$.

An alternative way of calculating principal angles is to use the singular value decomposition.

Lemma 1. [16] Assume $\mathbf{U}_1$, $\mathbf{U}_2$ is orthonormal basis of r-dimensional subspace $\mathcal{S}_1$, $\mathcal{S}_2$, respectively. Denote λ_k as the k-th singular value of $\mathbf{U}_1^{\mathrm{T}} \mathbf{U}_2$ and θ_k as the k-th principal angle between $\mathcal{S}_1$ and $\mathcal{S}_2$. Then $\cos\theta_k = \lambda_k$ for $k = 1, \cdots, r$.

3. PROPOSED ALGORITHM

Denote $\mathbf{H} \in \mathbb{R}^{M \times N \times L}$ a hyperspectral image, where M, N represents the height and width of the data respectively and L denotes the number of bands. Denote $\mathbf{h}_{m,n} \in \mathbb{R}^L$ as the spectral signature of the (m,n)-th pixel with length L. The proposed SPAHSIC includes two steps. Step 1 is the superpixel segmentation with an angle-based distance, which uses the high spatial correlation between pixel values, which partitions adjacent pixels lying in the same subspace into a superpixel. Step 2 consists of computing the affinity matrix of superpixels and clustering them, which merges superpixels lying in the same subspace into one class.

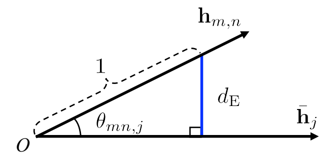

Fig. 1. Principal angle between the column spaces spanned by $\mathbf{h}_{m,n}$ and $\bar{\mathbf{h}}_j$.

3.1. Angle-based Superpixel Segmentation

In this step, we want to partition HSI into many superpixels, where pixels in the same superpixel lie in the same subspace. To this end, we improve SLIC from two aspects. In order to take advantage of the subspace property of hyperspectral images, our first improvement is to introduce a new measure d_{E} to evaluate the spectral distance between two pixels. d_{E} is defined as the principal angle between the column spaces spanned by $\mathbf{h}_{m,n}$ and $\bar{\mathbf{h}}_j$, as shown in Fig. 1.

$$d_{\mathrm{E}}(\mathbf{h}_{m,n}, \bar{\mathbf{h}}_j) = \sin\theta_{mn,j} = \sqrt{1 - (\mathbf{h}_{m,n}^{\mathrm{T}} \bar{\mathbf{h}}_j)^2 / (\|\mathbf{h}_{m,n}\|^2 \|\bar{\mathbf{h}}_j\|^2)},$$

where $\bar{\mathbf{h}}_j$ denotes the spectral signature of the j-th cluster center. The motivation of the above definition is that the intensity of the spectral information mainly depends on the transmission energy loss and is thus meaningless for superpixel segmentation. The distance between the (m,n)-th superpixel and the cluster center is defined as $D(\mathbf{h}_{m,n}, \bar{\mathbf{h}}_j) = d_{\mathrm{E}}(\mathbf{h}_{m,n}, \bar{\mathbf{h}}_j) + (m/S) d_{\mathrm{A}}(\mathbf{h}_{m,n}, \bar{\mathbf{h}}_j)$, where $d_{\mathrm{A}}(\cdot,\cdot)$ and $d_{\mathrm{E}}(\cdot,\cdot)$ denotes, respectively, the measure on spAtial distance and that on spEctral distance.

Another contribution is the calculation of gradient. To further reduce the effects of noise, we define the gradient of pixels as follows:

$$\begin{aligned} G(p,q) = &\|\mathbf{h}(p,q) - \mathbf{h}(p+1,q)\| + \|\mathbf{h}(p,q) - \mathbf{h}(p-1,q)\| \\ &+ \|\mathbf{h}(p,q) - \mathbf{h}(p,q+1)\| + \|\mathbf{h}(p,q) - \mathbf{h}(p,q-1)\|. \end{aligned}$$

Besides, we count the number E of pixels that their label change in successive iterations and set threshold τ as 5. If E is smaller than τ, we jump out of the loop. Finally, to avoid the number of pixels of some superpixel is too small, we merge such superpixels to their neighborhoods by changing their label to the nearby superpixels' labels. Then, we calculate the current superpixel number K.

The influence of desired number of superpixels $\hat{K}$ and scaling factor m is discussed in Section 4.

3.2. Superpixel Clustering

Assume we get K superpixels $P_1, \cdots, P_K$ in step 1 and each superpixel lies in a subspace. We first calculate the first r principal components $\{\mathbf{U}_j \in \mathbb{R}^{L \times r}\}_{j=1,2,\cdots,K}$ via PCA for each superpixel. Then the distance between two superpixels can be defined as

$$d^2(P_j, P_k) = \sum_{i=1}^{r} \sin^2 \phi_i,$$

where ϕ_i denotes the i-th principal angle between subspaces spanned by the columns of $\mathbf{U}_j$ and $\mathbf{U}_k$, as shown in Fig. 2. The calculation of principal angles can be completed according to Lemma 1. The similarity matrix $\mathbf{A} \in \mathbb{R}^{K \times K} = (a_{jk})_{jk}$, where $a_{jk} = \exp\left(-D^2(P_j, P_k) / \left(2\sigma^2\right)\right)$, and σ is the variance of the Gaussian kernel function. In this work, $2\sigma^2$ is set as 7. Finally, we apply the spectral clustering to the similarity matrix $\mathbf{A}$, and assign each pixel the label the same as that of the superpixel it belongs to. The SPAHSIC algorithm is summarized in Algorithm 1.

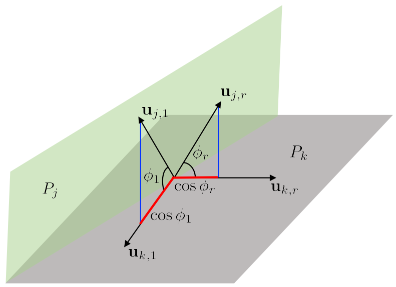

Fig. 2. Principal angles between two superpixels P_j and P_k, where only the first and the last angles are plotted.

3.3. Analysis of computational complexity

In this section, we analyze the computational complexity of two steps separately. For step 1, the computational complexity of SLIC for normal images is $O(MN)$ in each iteration [14]. Unlike normal SLIC for RGB images with just three spectral channels, SPAHSIC is required to compute $D(\mathbf{h}_{m,n}, \bar{\mathbf{h}}_j)$ with all spectral. Thus, the computational complexity of step 1 is $O(LMN)$ in each iteration.

In step 2, the computational complexity of applying PCA to the i-th superpixel is $O(rLn_i)$, where n_i is the number of pixels in superpixel P_i. As a result, the computational complexity for all superpixels is $\sum_{i=1}^{t} O(rLn_i) = O(rLMN)$. Based on $\mathbf{U}_i$, the computational complexity for calculating the affinity matrix $\mathbf{A}$ and spectral clustering is $O(rLK^2)$ and $O(K^3)$, respectively, where K denotes the number of superpixels. Then the overall computational complexity is $O(rLMN) + O(rLK^2) + O(K^3)$. Generally, the intrinsic dimension of each superpixel is low, which indicates that r is small. Considering that the distribution of the same land cover is usually concentrated, superpixel number K is also small compared with MN. As a result, the overall computational complexity of step 2 can be written as $O(rLMN)$.

4. PARAMETER SELECTION

Parameter $\hat{K}$ represents the desired number of superpixels. Considering that pixels in the same superpixel should belong to the same class, K should be larger than the number of classes. If K is small, pixels from different classes will be classified to the same superpixel, which leads to incorrect orthonormal bases $\mathbf{U}_i$. On the other side, if $\hat{K}$ is too large, it is of high probability that some superpixels only contain few pixels, which increases the estimation error of $\mathbf{U}_i$. Generally, we take $\hat{K}$ slightly larger than the three times of the class numbers.

Scaling factor m is the weight of d_{A} in the definition of D. It controls the compactness of a superpixel. The greater the value of m, the more spatial proximity is emphasized. Its selection is related to the geometric resolution of the images. When the geometric resolution is large, the real distance of each pixel center is far, spatial information should be used less, i.e., m is smaller, and vice versa. For example, if the resolution is 20 meters, m can be taken as 0.08, while if resolution is 1.3 meters, m is 0.06.

Subspace dimension r is determined by the energy distribution on the principal components of the superpixel. For each superpixel, we can calculate the ratio of the energy of the first r principal components with respect to the whole energy, and get its distribution denoted as $h_r(\mu)$, $\mu \in [0,1]$.As r increases, the energy of the first r singular values increases and $h_r(\mu)$ changes. r is taken as the minimum value when the correlation coefficient between $h_r(\mu)$ and $h_{r+1}(\mu)$ almost does not change.

Algorithm 1 SuperPixel and Angle-based HyperSpectral Image Clustering (SPAHSIC)

Input: HSI $\mathbf{H} \in \mathbb{R}^{M \times N \times L}$, desired number of superpixels $\hat{K}$, scaling factor m, subspace dimension r.
Output: Label matrix of HSI clustering $\mathbf{H}_{\mathrm{label}}$.

Initialize centers of clusters $\{C_i\}_{i=1,2,\cdots,\hat{K}}$ by sampling the data at regular step $S = \sqrt{MN/\hat{K}}$;
Compute the gradient of neighborhoods of the centers C_i and move the centers to the lowest gradient points;
repeat
  for each pixel **do**
    According to the angle-based distance, assign each pixel to the best matched center;
  end for
  Update the cluster centers as the mean of the pixels of the same label;
  Calculate residual error E;
until $E \le \tau$;
Search for small superpixels with fewer than r pixels and merge them into their nearby big superpixels.
Calculate the current number K of superpixels and the orthonormal bases $\mathbf{U}$ of each superpixel with Subspace dimension r;
Compute the affinity matrix of all superpixels $\mathbf{A}$ by principal angles;
Applying spectral clustering to the affinity matrix $\mathbf{A}$ to get the final result $\mathbf{H}_{\mathrm{label}}$;

5. EXPERIMENTAL RESULTS

Like previous work [7], [10] and [11], we apply SPAHSIC to two real datasets: 1) Indian Pines dataset and 2) the University of Pavia dataset.

We choose two widely used clustering methods K-means [4], FCM [5], and the state-of-the-art methods vanilla SSC [8], SSC-S [9] and JSSC [10] for comparison. Among them, K-means and FCM measure the distance between pixels by Euclidean distances without utilizing subspace property. Vanilla SSC produces a representation coefficient matrix by ADMM and applies it to spectral clustering to get the result. It focuses on subspace property but ignores the spatial information. To add the spatial information in clustering, SSC-S introduces the mean constraint of the representation coefficients and JSSC enforces pixels within a local region having similar representation coefficients. However, these three methods have high computing complexity compared to SPAHSIC.

Two common performance measurements, overall accuracy (OA) and Kappa coefficient κ, which measures the agreement between the predicted clustering results and the ground truths, are used for quantitative assessment of the clustering performances. We run all methods on the same computer and record the time they need to complete clustering.

5.1. AVIRIS DataSet: Indian Pines Image

The first experiment is conducted on part of Indian Pines dataset, which is acquired by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensors from the Northwestern Indiana Indian Pines test site in June 1992 [17]. The size is 145×145 pixels with a 20m geometric resolution. We remove 20 water absorption and noisy bands (104-108, 150-163, 220) from the original 220 bands, and use the leaving 200 bands for experiment. A typical subimage

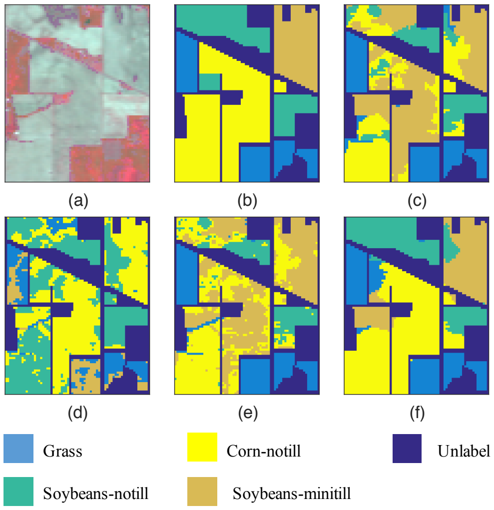

Fig. 3. Cluster maps of different methods with the Indian Pines image: (a) False-color image (RGB 45, 30, 20), (b) Ground truth, (c) K-means, (d) FCM, (e) SSC, (f) SPAHSIC.

Table 1. Quantitative evaluation of the different clustering algorithms for the Indian Pines image

Class	K-means	FCM	SSC	SSC-S	JSSC	SPAHSIC
Grass	**100.0**	54.5	**100.0**	**100.0**	**100.0**	99.9
Corn-notill	50.5	70.7	47.8	58.1	74.0	**83.6**
Soybeans-notill	65.4	0.0	75.3	68.1	86.2	**88.3**
Soybeans-mintill	41.9	55.7	54.9	64.8	**87.8**	84.6
OA	57.4	49.7	64.2	68.1	86.4	**87.5**
κ	0.43	0.28	0.51	0.55	0.81	**0.82**
Time (s)	**2.18**	6.12	363.48	-	-	5.11

of the size 85 $\times$ 70 is used, which contains four main classes. The false color and ground truth are showed in Fig. 3.

Based on our discussion in Section 4, we choose $\hat{K} = 20$, $m = 0.06$, $r = 3$. the quantitative evaluations of the clustering results including OA and κ are provided in Table 1. Particularly, for methods SSC-S and JSSC, we copy the result of [9] and [10], to ensure the best parameters being chosen. The results demonstrate that SPAHSIC achieves the best OA and κ among all methods. The cluster maps of different methods are shown in Fig. 3. It can be observed that SPAHSIC provides a more smoother classification. This is because the superpixel segmentation step enforces the spatial correlation between pixels. Running time is listed in the last line of the table. Time for SSC-S and JSSC is not recorded. However, considering that they are modified under SSC model, it has similar computational complexity as SSC-S, which is very high. Only K-means runs faster than SPAHSIC on this dataset, but it has much lower accuracy. Moreover, the computational complexity of K-means is $O(LMNK)$ in each iteration. As the number of pixels increases, K-means takes longer time than SPAHSIC, which is verified in the next experiment.

5.2. ROSIS Urban DataSet: University of Pavia, Italy

We use part of University of Pavia dataset from the Reflective Optics System Imaging Spectrometer (ROSIS) sensor in this experiment. The size of original dataset is 610$\times$340 with a 1.3m geometric resolution and the data has 103 spectral channels [9]. A typical subimage of size 200 $\times$ 100 is chosen, which contains eight main classes. The false color and ground truth are showed in Fig. 4.

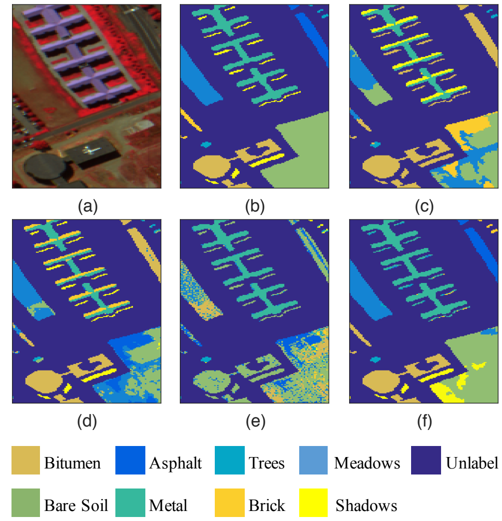

Fig. 4. Cluster maps of the different methods with the University of Pavia image: (a) False-color image (RGB 99, 67, 30), (b) Ground truth, (c) K-means, (d) FCM, (e)SSC, (f) SPAHSIC.

Table 2. Quantitative evaluation of the different clustering algorithms for the University of Pavia Image

Class	K-means	FCM	SSC	SSC-S	JSSC	SPAHSIC
Bitumen	0.0	0.2	21.7	0.0	99.1	**100.0**
Asphalt	59.9	86.6	38.2	95.9	11.8	**98.6**
Trees	81.0	60.3	69.8	**100.0**	98.4	49.2
Meadows	63.0	61.4	96.2	0.0	99.4	**99.9**
Bare soil	43.8	40.6	49.6	25.8	68.5	**82.5**
Metal	**100.0**	**100.0**	0.0	98.6	99.8	99.8
Brick	0.0	0.0	0.0	**52.3**	0.0	0.0
Shadows	0.0	**100.0**	0.0	98.6	89.0	39.6
OA	51.5	58.5	46.0	52.0	79.35	**87.5**
κ	0.41	0.50	0.27	0.44	0.74	**0.84**
Time (s)	37.31	19.11	1.22E4	-	-	**15.16**

Based on our discussion, the geometric resolution of Pavia University is smaller than the Indian Pines's, so m is chosen as 0.08. $\hat{K} = 25$, $r = 4$. Similarly, we copy the result of SSC-S and JSSC in [9] and [10], to ensure the best parameters being chosen. The running time for these two methods is not recorded. The results of this experiment are shown in Table 2 and the clustering maps are shown in Fig. 4. According to the results, SPAHSIC has the best performance and consumes the least time. K-means is not as efficient as SPAHSIC as the dimension of HSI increases.

6. CONCLUSION

We have proposed a novel algorithm for HSI clustering called SPAHSIC. It first divides HSI into superpixels that are treated as subspaces. Next, clustering is conducted on those subspaces based on principal angels. Superpixel segmentation provides local spatial information for clustering, and the latter clustering incorporates global information. Meanwhile, conducting clustering based on subspace also frees us from the high computational complexity of SSC. Two experiments on real datasets show the outstanding performance of SPAHSIC in both accuracy and computing efficiency.

7. REFERENCES

[1] T. M. Lillesand, *Remote Sensing and Image Interpretation*, JOHN WILEY & SONS, 1979.

[2] T. J. Cudahy, A. P. Rodger, P. S. Barry, P. Mason, M. Quigley, M. Folkman, and J. Pearlman, "Assessment of the stability of the hyperion swir module for hyperspectral mineral mapping using multi-date images from mount fitton, australia," in *IEEE International Geoscience and Remote Sensing Symposium*, June 2002, pp. 3504–3506.

[3] B. Hu, S. E. Qian, D. Haboudane, and J. Miller, "Impact of vector quantization compression on hyperspectral data in the retrieval accuracies of crop chlorophyll content for precision agriculture," in *IEEE International Geoscience and Remote Sensing Symposium*, June 2002, pp. 1655–1657.

[4] S. Lloyd, "Least squares quantization in pcm," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129–137, 1982.

[5] J. C. Bezdek, *Pattern recognition with fuzzy objective function algorithms*, Plenum Press, 1981.

[6] J. J. Buckley, "Fuzzy hierarchical analysis," *Fuzzy Sets & Systems*, vol. 17, no. 3, pp. 233–247, 1985.

[7] H. Zhai, H. Zhang, L. Zhang, P. Li, and A. Plaza, "A new sparse subspace clustering algorithm for hyperspectral remote sensing imagery," *IEEE Geoscience & Remote Sensing Letters*, vol. 14, no. 1, pp. 43–47, 2017.

[8] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 11, pp. 2765–2781, 2013.

[9] H. Zhang, H. Zhai, L. Zhang, and P. Li, "Spectralspatial sparse subspace clustering for hyperspectral remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 6, pp. 3672–3684, 2016.

[10] S. Huang, H. Zhang, and A. Pizurica, "Joint sparsity based sparse subspace clustering for hyperspectral images," in *IEEE International Conference on Image Processing*, Oct. 2018, pp. 3878–3882.

[11] J. Xu, N. Huang, and L. Xiao, "Spectral-spatial subspace clustering for hyperspectral images via modulated low-rank representation," in *IEEE International Geoscience and Remote Sensing Symposium*, July 2017, pp. 3202–3205.

[12] V. Machairas, E. Decencire, and T. Walter, "Waterpixels: Superpixels based on the watershed transformation," in *IEEE International Conference on Image Processing*, Oct. 2014, pp. 4343–4347.

[13] M. Liu, O. Tuzel, S. Ramalingam, and R. Chellappa, "Entropy rate superpixel segmentation," in *Computer Vision and Pattern Recognition*, June 2011, pp. 2097–2104.

[14] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Ssstrunk, "Slic superpixels compared to state-of-the-art superpixel methods.," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.

[15] A. Galntai, "Jordan's principal angles in complex vector spaces," *Numerical Linear Algebra with Applications*, vol. 13, no. 7, pp. 589–598, 2010.

[16] A. Bjrck and G. H. Golub, "Numerical methods for computing angles between linear subspaces," *Mathematics of Computation*, vol. 27, no. 123, pp. 579–594, 1973.

[17] H. Jiao, Y. Zhong, L. Zhang, and P. Li, "Unsupervised remote sensing image classification using an artificial dna computing," in *International Conference on Natural Computation*, July 2011, pp. 1341–1345.