This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3132683

Superpixel-level Global and Local Similarity Graph-based Clustering for Large Hyperspectral Images

Haishi Zhao, *Student Member,* Fengfeng Zhou*, Senior Member, IEEE*, Lorenzo Bruzzone, *Fellow, IEEE*, Renchu Guan, *Member, IEEE* and Chen Yang, *Member, IEEE*

***Abstract*—Due to the scarcity of labeled samples, clustering in hyperspectral images (HSIs) has a great potential and application value. However, current clustering methods are mainly pixel-level techniques which neglect the large spectral variability of a scene, and suffer from massive time and memory consumption when dealing with large HSIs. In this paper, we propose a superpixel-level global and local similarity graph-based clustering (SGLSC) algorithm that can classify ground objects exploiting spectral and spatial dimensions with reasonable time and memory consumption on large HSIs. The proposed SGLSC, exploits the superpixel concept, which is treated as a homogeneous entity, into the clustering process. For modeling the essential structure of HSIs, a similarity graph combing the global and local information is constructed and inserted into the spectral clustering to partition the superpixel-level graph structure. The proposed method was tested on three benchmark HSIs data sets and compared with some advanced literature algorithms. Experiments demonstrate that it can obtain promising results.**

***Index Terms*—Clustering, superpixels, similarity graph, global and local structure information, large hyperspectral images, remote sensing.**

I. Introduction

HYPERSPECTRAL sensors integrate imaging and spectral technology for acquiring hyperspectral images (HSIs) in hundreds of narrow and contiguous spectral channels [1]. HSIs can capture discriminative features and fine differences among various land-cover materials. This is a very important information for many application fields such as environmental monitoring, precision agriculture, mineral exploration and urban planning [2-6].

If a sufficient number of labeled samples is provided, the problem of HSI classification can be addressed by supervised methods [7] (e.g., support vector machine (SVM) [8], extreme learning machine (ELM) [9] and deep learning-based approaches [10-12]) which achieved excellent performances in recent years. However, it is difficult to have a large amount labeled samples due to the high cost in label collection and practical limitations. On the contrary, HSI unsupervised classification or clustering, which does not employ any label samples, has a strong practical significance. Nevertheless, clustering is a very challenging task due to the high-dimensional features space, the large spectral variability, and the complex ground objects distribution presented in HSIs [13-16].

Recently, a wide variety of clustering algorithms used for HSIs have been proposed. According to clustering mechanism, they can be divided into five main categories: centroid-based, density-based, probability-based, bionics-based and graph-based methods.

As the most classical method, the centroid-based clustering algorithms [17-20] were earliest introduced in HSIs clustering. They usually start with a random initialization and iteratively update the cluster centers and the corresponding pixel partitions until a certain stop criterion is met. However, such algorithms often assume a spherical cluster structure in the feature space, which is not suitable for HSIs with large spectral variability. In the density-based methods, the samples in high-density regions partitioned by the sparse regions are regarded as different clusters [21-23]. However, due to the high dimensions of HSIs, pixels are usually intrinsically sparse in their feature space, which prevents the density-based methods from distinguishing high-density regions from sparse ones. Probability-based methods [24], [25] group pixels based on posterior probability maximization criterion. They assume that the pixels belonging to the same class in HSIs obey a certain probability distribution with specific parameters. Nonetheless, this assumption is often difficult to satisfy for the complex structure and large spectral variability in the HSIs. Bionics-based methods introduce the biological heuristic models, e.g., evolution algorithms, swarm intelligence algorithm and artificial neural networks, into HSIs clustering. Examples of these methods are the self-organizing

This work was supported in part by the National Natural Science Foundation of China under Grant 61572228, the Science-Technology Development Plan Project of Jilin Province of China under Grant 20190303006SF and 20190302107GX, the Industrial Innovation Special Funds Project of Jilin Province under Grant 2019C053-5 and 2019C053-7, and the Open Funds Project of Key Laboratory of Lunar and Deep Space Exploration LDSE201906 (*Corresponding authors: Chen Yang and Renchu Guan*)

H. S. Zhao, R. C. Guan, and F.F. Zhou are with the Key Laboratory for Symbol Computation and Knowledge Engineering of the Ministry of Education, College of Computer Science and Technology, Jilin University, Changchun, 130012, China (e-mail: guanrenchu@jlu.edu.cn).

L. Bruzzone is with the Department of Information Engineering and Computer Science, University of Trento, 38050 Trento, Italy (e-mail: lorenzo.bruzzone@unitn.it).

C. Yang is with the College of Earth Sciences, Jilin University, Changchun, 130061, China, and Key Laboratory of Lunar and Deep Space Exploration, National Astronomical Observatories, Chinese Academy of Sciences, Beijing, 100012, China (e-mail: yangc616@jlu.edu.cn).

map (SOM) [26] and the unsupervised extreme learning machine (US-ELM) [28], [29]. In general, such approaches also suffer from similar obstacles as mentioned above and fail to achieve satisfactory clustering accuracy for HSIs.

Graph learning can effectively express the essential structure and even the complex higher-order relationships of data [30, 31]. Thus graph-based methods have been widely applied to HSI processing. Similarly, for hyperspectral image clustering tasks, graph-based methods show impressive performance and great potential due to the flexible construction of a similarity graph. In general, the graph-based clustering methods [32-34] represent the collection of objects as a graph in which the (dis)similarity relationships among different objects is measured by the similarity graph. These methods typically employ a two-step strategy. First, a similarity graph (i.e., similarity matrix) is constructed, where the data points are the nodes and the relationships are the weights. The relationships between the data points can be computed by various similarity measures. Then, the clustering can be achieved based on the graph theory optimization process. It should be emphasized that the quality of similarity graph plays a decisive role in the final clustering performance.

In consequence, many improved graph-based clustering algorithms have been applied to HSIs, with the goal to construct more robust similarity graphs. Zhai *et. al.* [35, 36] introduced spatial information into the sparse self-representation learning framework to construct the similarity matrix. The nonlinear kernels was adopted to model the nonlinear structure of HSIs [37]. Zhang *et. al.* [38] proposed an unsupervised HSI clustering method based on the robust manifold matrix factorization technique, which can jointly perform the HSI dimension reduction and clustering tasks and achieve good results. In addition, Huang *et. al.* [39] proposed a bipartite graph partition based coclustering with joint sparsity for HSIs clustering task. This method incorporated the structured dictionary constraint into a joint sparsity constrained optimizing model to construct the similarity graph. Furthermore, the recent hot graph convolution network was utilized to generate the similarity graph to exploit the intrinsic structure information of HSIs [40, 41]. Unfortunately, these methods usually suffer from a large amount of computation and huge memory requirements when applied to large HSIs. The detailed reasons for which these methods have issues in handling large HSIs are described in Section II.

In recent years, several methods have been presented for large HSIs clustering. Wang *et al.* [42], [43] presented two clustering algorithms for large HSIs, i.e., the fast spectral clustering with anchor graph (FSCAG) [42] and the scalable graph-based with nonnegative relaxation (SGNCR) [43]. The two methods choose a small number of pixels (i.e., anchors) to build a small size similarity graph for reducing computational complexity. However, the anchors are selected in a random way that neglects the structure information in HSIs and affects the ability to explore the intrinsic similarity between pixels. Wei *et. al.* [44] proposed a fast spectral clustering method based on anchor graph and spatial information. Yang *et. al.* [45] proposed a novel method for large HSI clustering tasks inspired by the spectral embedding and clustering with adaptive neighbors. Huang *et. al.* [46] proposed a bipartite graph partition with graph nonnegative matrix factorization (NMF) for coclustering of large HSIs, which fully explores the spectral and spatial information in HSIs and relieves the computational consumption and memory space required with the orthonormal constrained NMF-based bipartite graph partition. However, these methods involve lots of complicated processing and the performance also need to be improved. Zhai *et al.* [47] converted the HSI clustering task into a pixel-wise sparse recovery problem and proposed joint spare coding-based clustering (JSCC) algorithms for large HSIs, which can efficiently reduce the computational complexity and memory consumption. However, a structured dictionary needs to be generated in advance. The performance of the JSCC depends on the quality of structure dictionary whose purity is difficult to be guaranteed.

Hyperspectral images have rich spectral and spatial context information. The high volume of spectral channels with a large number of pixels make the performance of the traditional clustering method degrades or even fail to work. For the spatial domain, the distribution of ground object is usually irregular, but the pixels in local areas are more likely to belong to the same class. It is necessary to design a clustering approach that can consider the characteristics of the spatial features and spectral information. Many papers [48-50] have exploited the superpixel concept [51], which represents a small homogeneous region in the images in which pixels have similar properties. For the HSI clustering task, superpixels are mostly used to provide spatial context information, such as in dictionary or representation learning constrained by superpixels [46, 47].

In this paper, we propose a superpixel-level global and local similarity graph-based clustering (SGLSC) algorithm which can accurately group ground objects integrating the spectral and spatial information with reasonable time and memory consumption on large HSIs. The main contributions of the proposed SGLSC method are highlighted as follows.

1) Inspired by the effectiveness of superpixel concept in the analysis of HSI, we introduce superpixel to address the challenges of high computational complexity and memory consumption for large HSIs clustering. The superpixels can significantly reduce the number of samples on large HSIs clustering tasks, and alleviate the spectral variability of pixels located in a local region.
2) A new global and local superpixel-level similarity graph is constructed, in which the relationships of each superpixel with its neighbors are measured to model the essential structure of HSIs. Due to the comprehensive superpixel-level similarity graph, the inherent structure of ground objects in HSI can be completely exploited to improve the clustering performance.
3) A lightweight large HSIs clustering algorithm, i.e., SGLSC is presented by inserting constructed superpixel-level similarity graph into the spectral clustering. Extensive experiments show that our proposed SGLSC achieves both high accuracy and high

efficiency on some benchmark datasets compared with advanced methods recently proposed in the literatures.

The rest of this paper is organized as follows. Section II describes the related work and motivation. Section III presents the details of the proposed SGLSC algorithm. Section IV describes the experimental results and the comparison between the proposed method and some state-of-the-art and latest methods on three HSI data sets. Finally, Section V concludes the paper and presents the future research.

II. Related Works and Motivation

Based on graph theory, graph-based clustering methods have recently shown advanced performance in HSI clustering task. These methods generally construct a similarity graph (i.e., similarity matrix) $W \subset \mathrm{R}^{N\times N}$, whose elements indicate the affinity between the corresponding two pixels, to model the similarities among different hyperspectral pixels. A similarity graph is commonly constructed by various similarity measures combining with the KNN strategy [52], the ε–ball strategy [53], and the full connection strategy [54]. After that, graph cut algorithms are adopted to segment the similarity graph into several disjoint subgraphs by minimizing the correlation between subgraphs and maximizing the correlation within subgraphs. The spectral clustering (SC) [55, 56] is a typical graph-based clustering algorithm, whose optimization problem can be formulated as follow:

$$\min_{F^TF=I} Tr(H^TLH) \tag{1}$$

where $H = \{h_1, h_2, \cdots, h_M\} \subset \mathbb{R}^{N\times M}$ is the indicator matrix, M indicates the number of clusters, $L=D - W$ is the Laplacian matrix, D denotes the degree matrix, which is a diagonal matrix and can be computed as $D_{ii} = \sum_j W_{ji}$. The optimization problem of SC is usually solved by singular value decomposition (SVD). The optimal H can be obtained by extracting the eigenvectors corresponding to the first M small eigenvalues of the Laplace matrix L. Then, k-means is commonly applied to the eigenvectors to assign a cluster index to each sample.

However, a huge challenge in the computational requirements of graph-based clustering methods for large and high-dimensional HSIs is encountered in the solution of the optimization problem with SVD. In addition, the memory requirements for storing the entire similarity matrix of large HSIs are also very expensive. Given a HSI data set $X \subset \mathbb{R}^{L\times N}$, where N is the number of pixels and L is the number of bands, which contains M classes, the memory space and time complexity required to compute the similarity matrix are $O(N^2)$ and $O(N^2L)$, respectively. Moreover, at least a $O(N^2M)$ complexity is required to obtain the M eigenvectors of the Laplace matrix derived from the similarity matrix. For instance, the popular Salinas Valley hyperspectral benchmark image contains 111,104 pixels (i.e., the spatial size is 512×217), which makes the similarity matrix consumes more than 90GB memory space, resulting in little applicability of these methods to large HSI data sets in real-world scenarios.

From the aforementioned analysis, it can be clearly observed that the number of data points has a critical impact on computational complexity and memory consumption when using the conventional graph-based clustering methods on large HSIs. However, it is well known that the pixels of HSI in local areas are more likely to belong to the same category, and that one HSI can be over-segmented into different disjoint homogeneous regions (i.e., superpixels) in which pixels have similar properties. The number of superpixels is far smaller than the number of pixels in the original HSI. Therefore, superpixel oriented graph-based clustering can significantly mitigate the computation and memory requirements for large HSI clustering tasks at a large extent.

From the clustering performance perspective, the quality of similarity graph directly affects final clustering accuracy. Therefore, the similarity graph should be able to model the relationships among ground objects in HSI as realistically as possible. Due to the complexity and inhomogeneity of ground objects, objects belonging to the same category are usually distributed in different positions in the HSI. Meanwhile, because of the over-segmentation characteristics of superpixels, a region having pixels associated with the same class label may be divided into different adjacent superpixels. Consequently, the similarity graph should be able to both model the global information (i.e., a superpixel should be able to connect with the non-adjacent superpixels belonging to the same category), and the local information (i.e., the adjacent superpixels should have a high probability of belonging to the same category).

As a result, we firstly construct the superpixel-level global similarity graph and the local similarity graph, which can complement each other. Then a new superpixel-level similarity graph integrating the global and local information is generated to model the essential structure of large HSIs and thus improve the clustering accuracy. Section III describes the proposed algorithm in detail.

III. Proposed SGLSC Technique

In this section, the proposed SGLSC technique is described in detail. First, the given hyperspectral data is oversegmented into a certain number of superpixels for getting homogeneous regions. Second, we build a superpixel-level global similarity graph by mean of self-expression among different superpixels. Then a superpixel-level local similarity graph is constructed based on the spatial structure of the superpixel map. Finally, the clustering results are obtained by spectral clustering with combination of the global and local similarity graphs. The flowchart of the proposed SGLSC algorithm is illustrated in Fig. 1.

A. Superpixel Segmentation

The large number of pixels contained in HSIs dataset brings high computational complexity for the calculation of similarity in the clustering process. The entropy rate superpixel segmentation (ERS) [57] method is adopted to oversegment the hyperspectral imagery into a small number of superpixels with high homogeneity and compactness. The ERS is one of the representative superpixel segmentation algorithms based on the graph theory. It takes into account both the entropy rate of a random walk term ensuring the compactness and homogeneity

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3132683

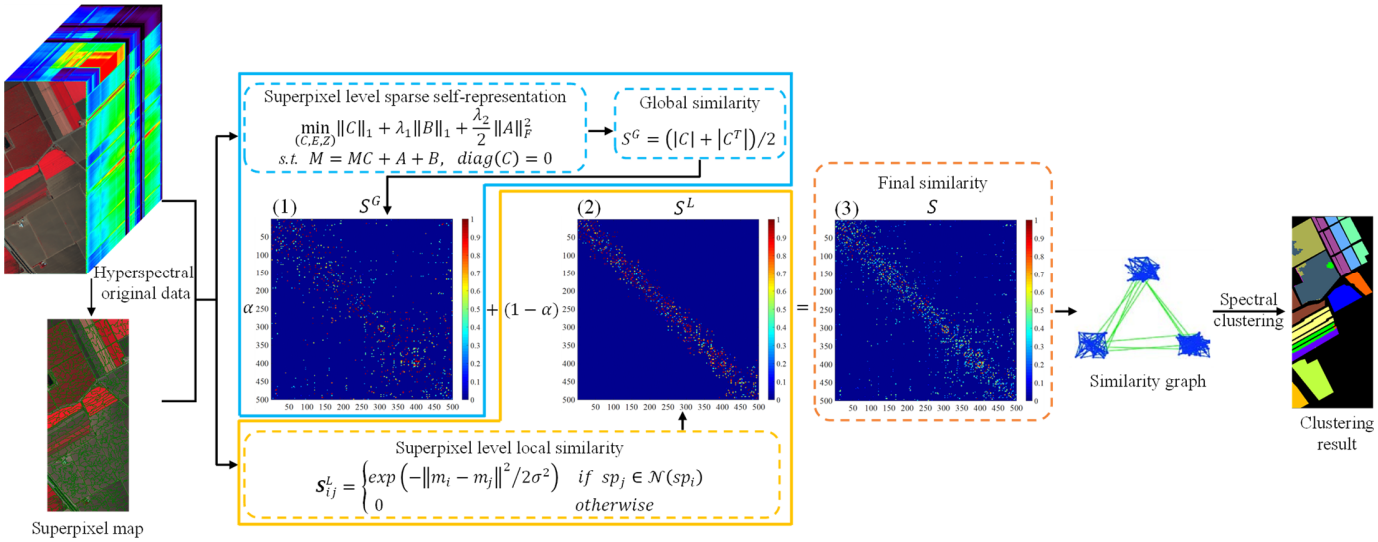

Fig. 1. Flowchart of the proposed SGLSC approach. (1) superpixel-level global similarity graph S^G, (2) superpixel-level local similarity graph S^L and (3) superpixel-level local similarity graph S on the Salinas Valley data set.

of superpixels and the balance term controlling the number of superpixels to segment the given image.

Given a HSI image $X = \{x_1, x_2, \cdots, x_N\} \subset \mathbb{R}^{L\times N}$, where $x_i = \{x_{i1}, x_{i2}, \cdots, x_{iL}\}$, N is the number of pixels in the HSI, and L is the number of spectral bands. When the ERS is applied to the HSI data set X, the superpixels set $SP = \{sp_1, sp_2, \cdots, sp_K\}$ is generated, where $sp_k = \{x_1^k, x_2^k, \cdots, x_{n_k}^k\}$ ($k = 1, 2, \cdots, K$) is the kth superpixel, K denotes the number of superpixels, x_i^k is the ith pixel vector in sp_k and n_k denotes the number of pixel vectors in the sp_k. Each superpixel is treated as a homogeneous entity. The mean spectrum $M = \{m_1, m_2, \cdots, m_K\} \subset \mathbb{R}^{L\times K}$ corresponding to the superpixel set SP is used as superpixel feature, in which m_k indicates the mean spectrum of the kth superpixel computed as follows:

$$m_k = \frac{1}{n_k}\sum_{i=1}^{n_k} x_i^k \tag{2}$$

The superpixel feature is used for the similarity graph construction instead of pixels. It is worth noting that the number of superpixels is far less than that of the pixels in the HSIs (i.e. $K<<N$), thus the computing time and memory can be tremendously reduced. More importantly, the superpixel can retain the local structure information of the ground objects as adjacent pixels with similar spectral features usually belong to the same category. Therefore, the superpixel features can alleviate the influence of spectral variability in the HSIs.

B. Superpixel-level Global Similarity Graph

Considering the characteristics of high dimensionality and large spectral variability of HSIs, the similarity graph based on traditional distance measure cannot adequately reflect relationships between data points. Recently, the graph-based subspace clustering method has achieved accurate results in HSI clustering tasks. It takes the whole HSI data points as an over-complete dictionary and models the similarity between data points based on a self-representation learning pattern with a certain prior constraint (such as sparsity or low-rank [58, 59]).

In this sub-section, a superpixel-level global similarity graph is built by the self-representation learning to measure the relationships between superpixels. According to the definition of superpixel feature matrix M in Section III-*A*, we assume that C is the representation coefficient matrix of superpixel feature matrix M under the self-representation learning framework, and $m_i \subset \mathbb{R}^L$ and $c_{:,i} \subset \mathbb{R}^K$ denote the columns of M and C in the vector form, respectively. The i-th superpixel m_i is represented by the other superpixels with sparsity constraint defined as follows:

$$m_i = \sum_{j=1, j\neq i}^{K} c_{ji} m_j, \quad \|c_{:,j}\|_0 \rightarrow min \tag{3}$$

where $\|c_{:,j}\|_0$ denotes the ℓ_0-norm which means the number of nonzero elements reflecting the sparsity of the solution, and c_{ji} is the value of contribution of the j-th superpixel (i.e., m_j) to the process of reconstruction of the i-th superpixel (i.e., m_i). It is natural to understand that larger c_{ji} values indicate that a higher similarity between m_j and m_i.

It should be noted that the ℓ_0-norm optimization problem corresponds to a NP-hard problem with no guarantee of the uniqueness and stability of the solution. Fortunately, according to the literature [60], ℓ_0-norm can be relaxed to ℓ_1-norm, which can be solved efficiently by multiple convex optimization algorithms [61-63]:

$$m_i = \sum_{j=1, j\neq i}^{K} c_{ji} m_j, \quad \|c_{:,j}\|_1 \rightarrow min \tag{4}$$

Equation (4) can be rewritten for superpixels by using the matrix formulation:

$$\min_{C} \|C\|_1 \quad s.t.\ M = MC,\ diag(C) = 0 \tag{5}$$

Furthermore, HSI data are inevitably corrupted by a variety of noises and even outliers in the process of data acquisition and transmission. Therefore, for better modeling the affinity relationship of objects collected in the complex HSI scenes, the self-representation problem can be ultimately reformulated as:

$$\min_{(C,E,Z)} \left\{ \|C\|_1 + \lambda_1 \|B\|_1 + \frac{\lambda_2}{2} \|A\|_F^2 \right\}$$
$$s.t.\ M = MC + A + B,\ diag(C) = 0 \quad (6)$$

where $C = [c_1, c_2, \ldots, c_K] \subset \mathbb{R}^{K\times K}$ indicates the sparse representation coefficient matrix in which the i-th column refers to the sparse representation coefficient of the i-th superpixel m_i, and the ℓ_1-norm of C (i.e., $\|C\|_1$) ensures the sparsity of the representation coefficient matrix C; $A \subset \mathbb{R}^{K\times K}$ represents the noise, and the Frobenius norm facilitates that noise matrix has small values; $B \subset \mathbb{R}^{K\times K}$ represents the sparse outliers, and the ℓ_1-norm is adopted to guarantee sparsity. Two parameters $\lambda_1 > 0$ and $\lambda_2 > 0$ are used to balance the contributions of the three terms (i.e., the sparsity of the representation coefficient, the degree of noise and the magnitude of outliers). The constraint item $diag(C) = 0$ is used to exclude the trivial solutions, i.e., a data point is expressed by itself.

The self-representation learning optimization problem in (6) can be efficiently solved by the alternating direction method of multipliers (ADMM) [63]. The nonzero elements in the representation coefficient of each superpixel demonstrate that the corresponding superpixels belong to the same subspace (i.e., the same class). Considering that the other superpixels are taken into account in the optimization process of sparse representation coefficient of a certain superpixel, we define the superpixel-level global similarity graph S^G. It models the weight on the edge between the superpixels and can be constructed by the sparse representation coefficient matrix C as follows:

$$S^G = \frac{1}{2}(|C| + |C|^T) \quad (7)$$

Equation (7) expresses that the weight on the edge between the superpixel m_i and m_j is equal to $\frac{1}{2}(|C_{ij}| + |C_{ji}|)$. According to the assumption that nonzero elements of the sparse representation coefficient of m_i demonstrate that the corresponding superpixels belong to the same category including m_i, it is intuitive that m_j is chosen to represent m_i if $|C_{ji}| > 0$. However, m_j may not select m_i in its sparse representation. Therefore, (7) is adopted to obtain a symmetric similarity matrix to guarantee that m_i and m_j are connected to each other regardless of whether m_i and m_j have chosen each other in the self-representation process. The process of construction of the superpixel-level global similarity graph S^G is summarized as Algorithm 1.

Algorithm 1: Superpixel-level Global Similarity Graph

Input:

1) Dataset: superpixel feature matrix $M = \{m_1, m_2, \cdots, m_K\} \subset \mathbb{R}^{L\times K}$;

2) Parameters: the regularization parameters λ_1 and λ_2.

Procedure:

1. Construct the sparse self-representation learning model with (6) and obtain the sparse representation coefficient matrix C by the ADMM;
2. Normalize C by column with $c_i \leftarrow c_i / \|c_i\|_\infty$;
3. Establish the superpixel-level global similarity graph S^G according to (7).

Output:

The superpixel-level global similarity graph S^G.

C. Superpixel-level Local Similarity Graph

A superpixel as homogeneous entity also has spatial adjacency, i.e. the spatial neighbors may belong to the same category. In this sub-section, a superpixel-level local similarity graph S^L is constructed for modeling the spatial adjacency between a superpixel and its neighbors.

1) Neighborhood of superpixels

In a hyperspectral image, the regular 4 or 8-adjacent areas are used to describe the neighborhoods of pixels. Unlike the pixel-level case, the spatial distribution of superpixels is irregular, and the neighborhoods of superpixels should be defined differently. For two different superpixels $sp_a = \{x_1^a, x_2^a, \cdots, x_{n_a}^a\}$ and $sp_b = \{x_1^b, x_2^b, \cdots, x_{n_b}^b\}$ ($a, b = 1, 2, \ldots, K$, and $a \neq b$), the distance between sp_a and sp_b is defined as the minimum Manhattan distance between the pixels x_i^a in sp_a and x_j^b in sp_b. It can be formulated as follows:

$$Dist(sp_a, sp_b) = \min_{\substack{\forall x_i^a \in sp_a \\ \forall x_j^b \in sp_b}} \{\mathcal{M}(x_i^a, x_j^b)\} \quad (8)$$

where $Dist(sp_a, sp_b)$ indicates the distance between sp_a and sp_b, $\mathcal{M}(x_i^a, x_j^b)$ is the Manhattan distance between x_i^a and x_j^b, it is defined as follows:

$$\mathcal{M}(x_i^a, x_j^b) = |p_i^a - p_j^b| + |q_i^a - q_j^b| \quad (9)$$

where p_i^a and q_i^a indicate the location (i.e., the row and column number) of pixel x_i^a in the image, and p_j^b and q_j^b indicate the location of pixel x_j^b. The neighborhoods of superpixel sp_a can be defined based on the (8) as follows:

$$sp_b \in \mathcal{N}(sp_a) \quad if\ Dist(sp_a, sp_b) = 1,\ \ b = 1,2,\cdots,K \quad (10)$$

where $\mathcal{N}(sp_a)$ denotes the set of neighborhoods of superpixel sp_a; $Dist(sp_a, sp_b) = 1$ means that at least one pair x_i^a and x_j^b of superpixels are adjacent.

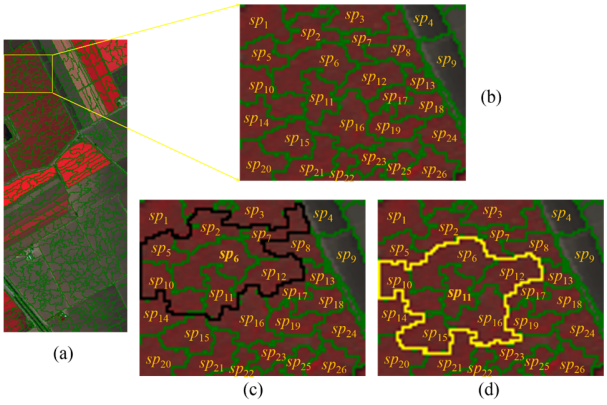

Fig. 2. Example of superpixel neighborhoods. (a) Superpixel map of the Salinas Valley dataset. (b) Enlarged view of the local area marked by the yellow box in (a). (c) Shape of neighborhoods of superpixel sp_6, i.e, the region surrounded by black lines. (d) Shape of neighborhoods of superpixel sp_{11}, i.e, the region surrounded by yellow lines.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3132683

An illustrative example of superpixel neighborhoods is showed in Fig. 2. From Fig. 2. (c), one can observe that the neighborhoods of superpixel sp_6 are superpixels sp_2, sp_5, sp_7, sp_{10}, sp_{11}, and sp_{12}, i.e., $\mathcal{N}(sp_6) = \{sp_2, sp_5, sp_7, sp_{10}, sp_{11}, sp_{12}\}$. From the Fig. 2. (c) and (d), we can observe that the shape of neighborhoods of superpixel is irregular, and that the shapes of neighborhoods of different superpixels are also different. This indicates that the superpixel neighborhoods are more flexible and more consistent with the distribution of ground objects in comparison with the conventional pixel-level neighborhood (i.e., 4-neighborhood, 8-neighborhood and so on).

2) Construction of local similarity graph based on superpixel neighborhoods

In order to measure the spatial adjacency of superpixels, the Gaussian kernel function, $G(x,y) = exp(-\|x-y\|^2/2\sigma^2)$ where σ is a hyperparameter, is adopted to calculate the similarity between superpixel sp_i and its neighbor sp_j. The superpixel-level local similarity graph S^L is expressed as:

$$\boldsymbol{S}_{ij}^{L} = \begin{cases} exp\left(-\|m_i - m_j\|^2/2\sigma^2\right) & if\ \ sp_j \in \mathcal{N}(sp_i) \\ 0 & otherwise \end{cases} \quad (11)$$

where m_i indicates the superpixel feature of sp_i defined in (2), $\|\cdot\|$ denotes the length of a vector, and the parameter σ controls the decay rate and is empirically set to 1 in the following experiments.

The construction process of superpixel-level local similarity graph S^L is given in Algorithm 2.

Algorithm 2: Superpixel-level Local Similarity Graph

Input:

1) Dataset: superpixel feature matrix $M = \{m_1, m_2, \cdots, m_K\} \subset \mathbb{R}^{L\times K}$;
2) Parameters: the parameter σ.

Procedure:

1) Obtain neighborhoods set of superpixel sp_a using (8) and (10);
2) Construct the superpixel-level local similarity graph S^L with (11).

Output:

The superpixel-level local similarity graph S^L.

D. Superpixel-level Global and Local Similarity-based Clustering

In this paper, a graph-based clustering method, i.e., spectral clustering [35-40], is adopted to achieve final clustering results. The basic idea of spectral clustering is to segment the graph based on the criterion that the weight sum of edges between different subgraphs should be as small as possible, whereas the weight sum of edges within the same subgraph should be as large as possible. The final clustering accuracy depends on the quality of the similarity graph, and the computational complexity is proportional to the square of the dimension of the similarity graph.

To consider both the superpixel-level global similarity graph and the superpixel-level local similarity graph, an integrated similarity graph S is defined as follows:

$$S = \alpha S^G + (1-\alpha)S^L,\ \ 0 \le \alpha \le 1 \quad (12)$$

where $\alpha \in [0, 1]$ indicates a trade-off parameter tuning the significance of global and local similarity graphs.

The procedure of the proposed SGLSC is illustrated in detail in Algorithm 3.

Compared with traditional pixel-level clustering methods, the proposed superpixel-level approach has two advantages. On the one hand, for a given HSI dataset, the number of superpixels K is far less than the number of the pixels. This reduces the dimension of the similarity graph considerably and thus guarantees the efficiency of clustering process. On the other hand, in the superpixel pattern, the comprehensive consideration of global and local similarity can improve clustering performance.

Algorithm 3: SGLSC algorithm

Input:

1) Dataset: hyperspectral image dataset $X = \{x_1, x_2, \cdots, x_N\} \subset \mathbb{R}^{L\times N}$;
2) Parameters: the number of superpixel K, the regularization parameters λ_1 and λ_2, the parameter σ, and the trade-off parameter α.

Procedure:

1) Apply ERS to HSI dataset X, and obtain the superpixels set $SP = \{sp_1, sp_2, \ldots, sp_K\}$ ($k = 1, 2, \ldots, K$);
2) Compute the superpixel feature $M = \{m_1, m_2, \cdots, m_K\} \subset \mathbb{R}^{L\times K}$ corresponding to SP according to (2);
3) Construct the superpixel-level global similarity graph S^G using Algorithm 1;
4) Construct the superpixel-level local similarity graph S^L with Algorithm 2;
5) Compute the superpixel-level global-local similarity graph S according to Equation (12);
6) Apply spectral clustering to the superpixel-level global-local similarity graph S.

Output:

The clustering results.

IV. Experimental Results and Analysis

In this section, extensive experiments on three widely used hyperspectral benchmarks are carried out to verify the effectiveness of the proposed SGLSC technique. All the experiments are conducted on a Windows computer with the Intel(R) Core(TM) i7-4790 3.60 GHz CPU and 32-Gb RAM. And the two comparison methods EKGCSC [39] and 3DCAE [66] are implemented with Python 3.6, the other methods with Matlab R2016b.

A. Description of Data Sets

The performance of the proposed SGLSC is assessed on three HSI datasets, i.e., the Indian Pines, the Salinas Valley, and the Pavia Center benchmarks, which are briefly described in the following.

1) Indian Pines: This historical data set was collected by National Aeronautics and Space Administration using the Airborne Visible/Infrared Imaging Spectrometer sensor (AVIRIS) over northwest Indiana in June 12, 1992. It contains 145×145 pixels and has a spatial resolution of 20m. After discarding 20 water absorption bands, the remaining 200

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3132683

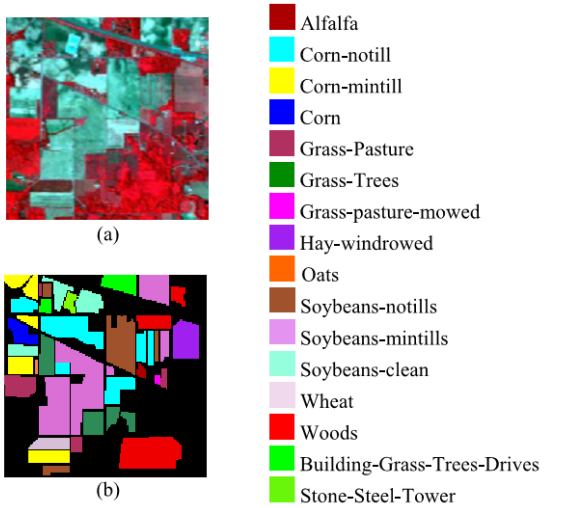

Fig. 3. (a) False color composite of images of bands 57, 27, and 17; (b) Available ground truth map (Indian Pines dataset).

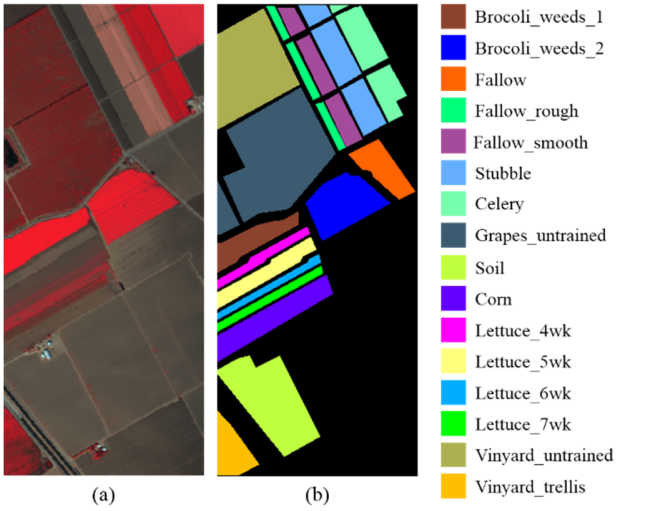

Fig. 4. (a) False color composite images of bands 50, 27, and 17; (b) Available ground truth map (Salinas Valley dataset).

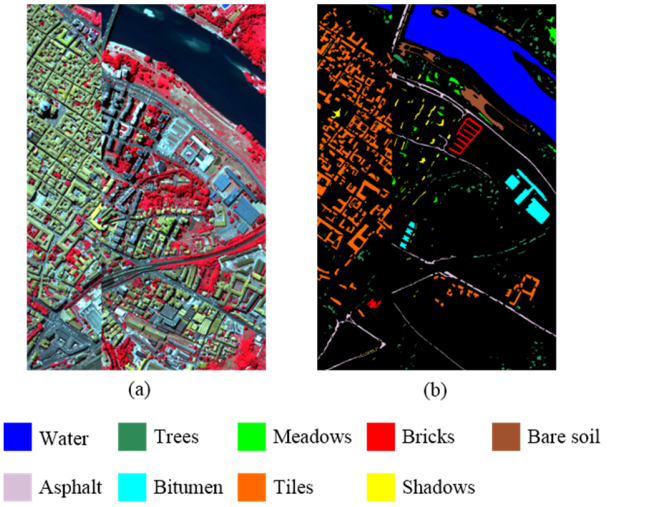

Fig. 5. (a) False color composite images of bands 102, 56, and 31; (b) Available ground truth map (Pavia Center dataset).

spectral bands are used in the experiments. There are 16 land-cover classes in total in this scene, which are mainly different categories of crops and vegetation. Fig.3 shows the false color composition of the data set and the available ground truth map.

2) Salinas Valley: The second data set was collected by AVIRIS over Salinas Valley, Southern California, in 1998. It contains 224 spectral bands from 400 to 2500 nm and a spatial size of 512×217 pixels with a high spatial resolution of 3.7 m. Twenty water absorption bands (i.e., in this case bands: 108-112, 154-167 and 224) were discarded. This area is mainly composed of some subtypes of vegetation, bare soil, and vineyard fields. The ground truth contains 16 classes with relatively compact land-cover distributions. The false-color image composition and reference land-cover map are shown in Fig. 4.

3) Pavia Center: This data set was obtained by the Reflective Optics System Imaging Spectrometer System (ROSIS) sensor with 115 spectral channels over the center of Pavia, Italy. After removing the low SNR bands, the dataset used for experiments has 102 bands. The image scene has a size of 1096×715 pixels with a spatial resolution of 1.3m. There are nine classes of interest contained in the image. The false-color image of Pavia Center dataset and the corresponding ground reference map are shown in Fig. 5.

B. Experimental Settings

1) Compared methods

To comprehensively assess the validity of the proposed SGLSC technique, three common methods (i.e., FCM [18] US-ELM [28] and Scalable-SC [64]), and seven recently proposed methods for HSI clustering (i.e., EKGCSC [39], SRDL [65], FSCAG [41], SGCNR [42], FSCS [44], JSCC [43] and 3DCAE [66]) are selected as benchmarks for comparison. The properties of the seven recently proposed HSI clustering methods are summarized in Table I in terms of suitability for large HSIs, exploitation of spatial-spectral information, use of deep learning techniques, and exploitation of graph-based methods. All the comparative clustering algorithms are briefly described below.

1) FCM [18]: This is a *C*-means based on fuzzy theory HSIs clustering.
2) US-ELM [28]: This is an unsupervised extreme learning machine with manifold regularization used for HSIs clustering.
3) Scalable-SC [64]: This is a scalable spectral clustering with cosine similarity used for HSIs clustering.
4) EKGCSC [39]: The HSI is clustered with the similarity graph constructed by graph convolution network and nonlinear kernel technology.
5) SRDL [65]: Spectral-spatial diffusion geometry is incorporated into the diffusion learning algorithm for HSIs clustering.
6) FSCAG [41]: Anchor graph is designed to reduce the computational complexity. The FSCAG can efficiently perform the large HSIs clustering task.

7) SGCNR [42]: The anchor graph and nonnegative relaxation is used to cluster large HSIs.
8) FSCS [44]: The spectral–spatial information and anchor graph are combined to fast cluster large HSIs.
9) JSCC [43]: The sparse representation reconstruction error is used to cluster large HSIs.
10) 3DCAE [66]: The deep architecture (i.e., 3-D convolutional autoencoder) is introduced to cluster large HSIs.

TABLE I
PROPERTIES OF SEVEN RECENTLY PROPOSED METHODS FOR HSIS CLUSTERING TASK USED IN OUR EXPERIMENTS

Methods	Large HSI	Spatial-Spectral based	Deep- Learning based	Graph based
EKGCSC			√	√
SRDL		√		
FSCAG	√			√
SGCNR	√			√
FSCS	√	√		√
JSCC	√	√		
3DCAE	√	√	√	

For the three HSI data sets, the parameters adopted in each compared clustering method have been tuned to be optimal. For the FSCAG, SGCNR and JSSC, the parameters are set according to the respective literature.

2) Quality indices

In order to assess the performance of the proposed approach and the seven compared methods, four quantitative evaluation indicators including overall accuracy (OA), kappa coefficient (Kappa), normalized mutual information (NMI) and producer's accuracy (PA) are adopted. All these quantitative indices range from 0 to 1. Meanwhile, the execution time of all clustering methods is reported to assess the computing complexity, and the visual clustering results of different methods are given for more intuitive comparative analysis.

C. Parameter Sensitivity Analysis

In the proposed SGLSC, there are three parameters, i.e., the number of superpixels K, the regularization parameter λ and the trade-off coefficient α. The sensitivity of each parameter is firstly evaluated on all three HSI data sets. The influence of the three parameters of the proposed SGLSC technique on the three considered HSIs is shown in Figs. 6-8. In addition, the best OAs of the first three clustering methods (i.e., those that achieved the best performance among the seven compared methods) are given to quantitatively show the performance of the proposed SGLSC algorithm.

1) Number of superpixels

The number of superpixels K is an important parameter for the proposed SGLSC. It controls the homogeneity in each superpixel and the samples size in the clustering. In the experiments, the parameter K is set in the range [1000, 2000] for the Indian Pines and the Pavia Center data sets, and [500, 1500] for the Salinas Valley data set, with a step of 250.

The clustering accuracy and execution time of the proposed SGLSC technique are shown in Fig. 6. It can be seen that the proposed SGLSC approach is robust to the different number of superpixels. The OA of SGLSC is relatively steady over a wide range of K values for the three HSI data sets. Furthermore, the performance of the proposed SGLSC is far better than other three best compared methods on the three hyperspectral images. Given the complex spatial structure and irregular distribution of land-covers of Indian Pines and Pavia Center data sets, the optimal values of parameter K are relatively large numbers, i.e., 1500 and 1750, respectively. On the contrary, the OA curve of Salinas Valley image describes declining trend with the increase of the number of superpixels [Fig. 6 (b)]. The best performance is obtained when the parameter K is 500. For the execution time, the time cost of the proposed SGLSC algorithm is proportional to the number of superpixels.

2) Regularization parameter

The proposed SGLSC requires to tune the two regularization parameters (i.e., λ_1 and λ_2) in (5), which adjust the weights of the noise constraint term and the outlier constraint term. In practice, for reducing the number of hyperparameters, the regularization parameters λ_1 and λ_2 are set to be equal, i.e., $\lambda_1 = \lambda_2 = \lambda$. In the experiments, we set λ vary from 5 to 50 with a step of 5 for the Indian Pines data set, from 500 to 1500 with the interval of 100 for the Salinas Valley data set, and from 10 to 90 with steps of 10 for the Pavia Center data set. The achieved OAs are shown in Fig. 7.

Fig. 7 shows that the parameter λ exerts a large influence on the clustering accuracy of the proposed SGLSC approach. Nevertheless, the SGLSC mostly achieves higher OAs than the three best compared algorithms on the considered HSI data sets. For the three HSI data sets, the optimal parameter λ is set as 15, 1000 and 40 in the following experiments, respectively. In practical applications, the suitable setting of parameter λ can be easily fine-tuned for a specific hyperspectral image.

3) Trade-off coefficient

The trade-off coefficient α balances the ratio of superpixel-level global and local information. In the experiments, the parameter α gets its value in the range of [0, 1] with a step of 0.1.

Fig. 8 illustrates the behavious of the OAs of the proposed SGLSC algorithm versus the parameter α. For the Indian Pines and Salinas Valley data sets, one can observe that the proposed SGLSC with integrated superpixel-level global and local similarity graph ($0< \alpha < 1$) always achieve better performance than when only either the superpixel-level global similarity graph ($\alpha = 1$) or the superpixel-level local similarity graph ($\alpha = 0$) are used. This indicates that the global and local similarity information are complementary and both beneficial for the improvement of HSI clustering accuracy. In details, from Fig. 8 (a)-(c), one can see that the best OA is obtained when α is set to 0.7, 0.5, 0.9 with different proportions of the superpixel-level global similarity graph and the superpixel-level local similarity graph. Therefore, the optimal α used are 0.7, 0.5, 0.9 for the Indian Pines, the Salinas Valley and the Pavia Center data sets, respectively.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3132683

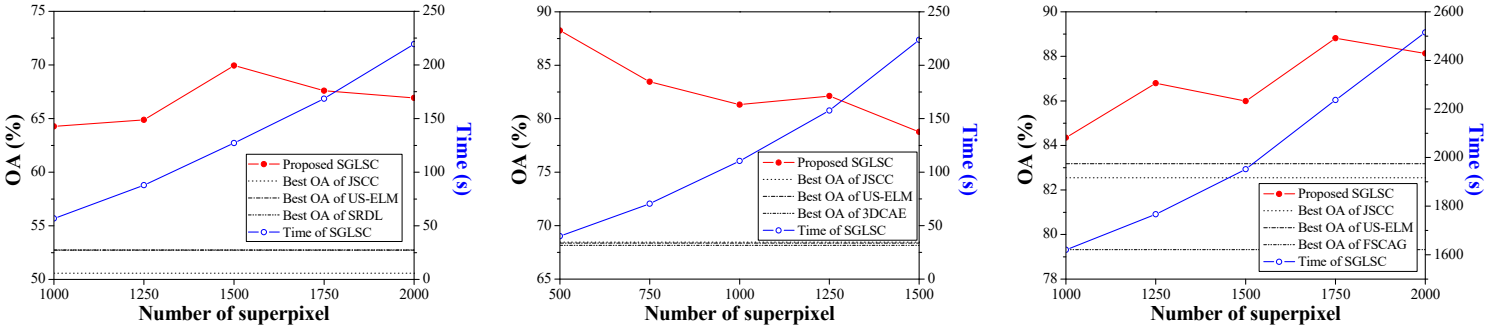

Fig. 6. OA and execution time versus the number of superpixels. (a) Indian Pines data set, (b) Salinas Valley data set, (c) Pavia Center data set.

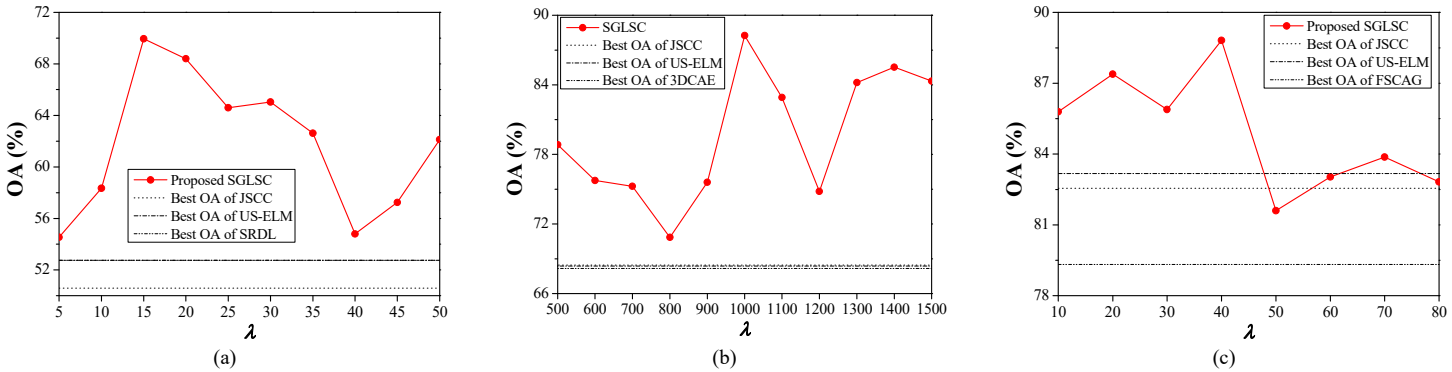

Fig. 7. OA versus the parameter λ. (a) Indian Pines data set, (b) Salinas Valley data set, (c) Pavia Center data set.

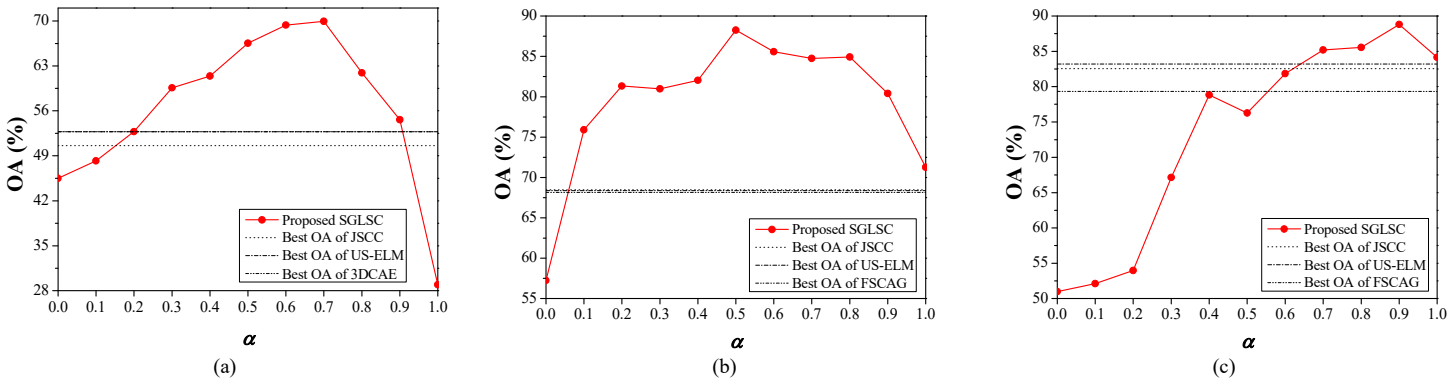

Fig. 8. OA versus the parameter α. (a) Indian Pines data set, (b) Salinas Valley data set, (c) Pavia Center data set.

D. Results on the Indian Pine Data Set

The first experiment was conducted on the Indian Pines image to verify the validity of the proposed SGLSC algorithm. The visual clustering results of compared methods are shown in Fig. 9 (a)-(l), and the corresponding quantitative assessments are listed in Table II. In the table, the optimal values of clustering accuracies are marked in bold, and the suboptimal values are underlined. From Fig. 9 and Table II, one can observe that the proposed SGLSC algorithm performs far better than the other methods in terms of OA, Kappa, and NMI.

The comparison among the three common methods shows that the FCM obtains the worst clustering result in this data set with significant amounts of misclassifications. In details, the FCM achieves a relatively good clustering results for the Hay-windrowed, Wheat and Oats categories, with the PA of 98.95%, 96.59% and 65% respectively. However, the accuracies of many other classes are all at a low level. The US-ELM method achieves third-best result in this experiment, achieving a good accuracy on most classes. However, due to the pixel-level based distance metric adopted in US-ELM, this method fails to model the local spatial information. The improved graph-based clustering method (i.e., Scalable-SC), is the most efficient and take very little time. However, it achieves an unsatisfactory clustering accuracy, with an OA of 46.43%.

For the recently proposed method for HSI clustering task, the EKFCSC obtained the medium level performance compared with other clustering methods on this data set. However, it should be noted that the execution time of the EKGCSC is extremely high, which may greatly reduce its application prospect in practical applications. The SRDL method achieves suboptimal results with an OA of 52.75%. Due to its exploitation of spatial information, it can be observed that there are little salt and pepper noise in the related cluster map (i.e., Fig. 9 (f)).

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3132683

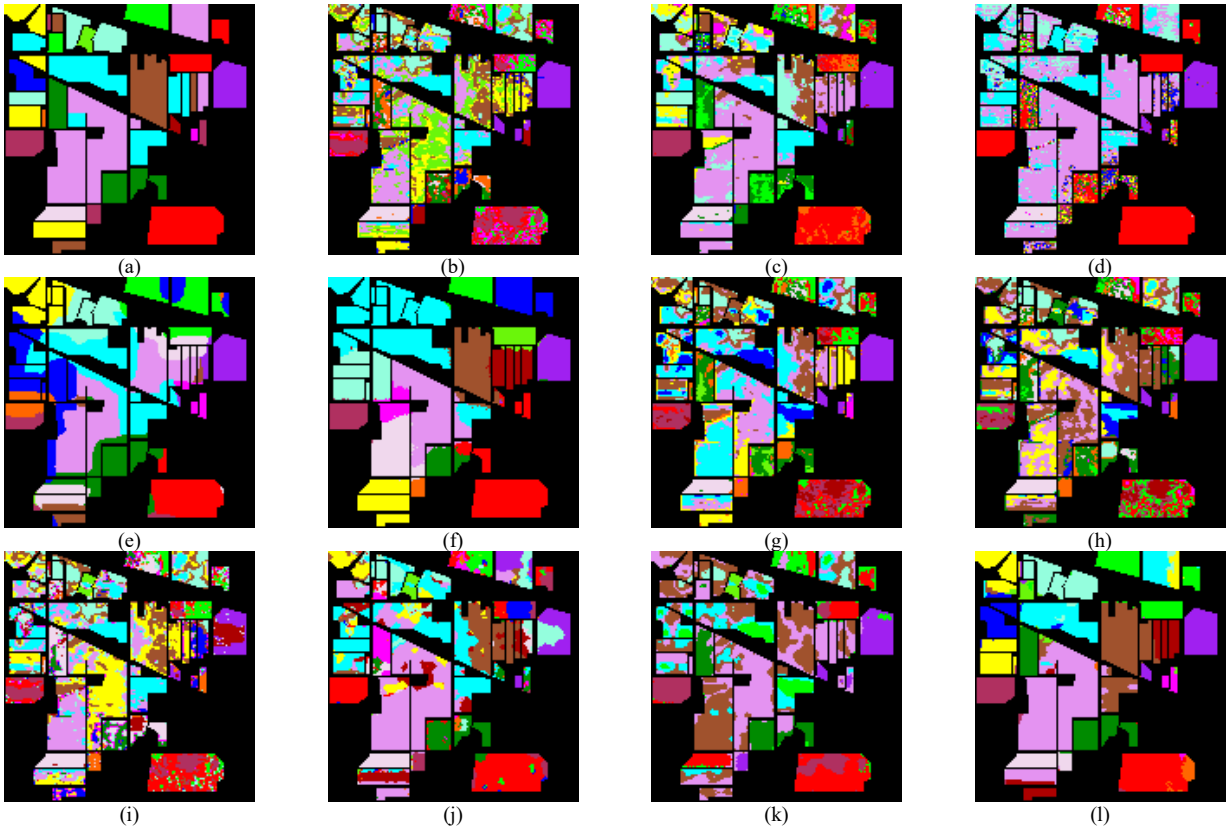

Fig. 9. Clustering maps of different methods on the Indian Pines data set. (a) Ground truth. (b) FCM. (c) US-ELM. (d) Scalable-SC. (e) EKGCSC. (f) SRDL. (g) FSCAG. (h) SGCNR. (i) FSCS. (j) JSSC. (k) 3DCAE. (l) Proposed SGLSC.

TABLE II
QUANTITATIVE ASSESSMENT OF CLUSTERING RESULTS BY USING SEVEN COMPARED METHODS ON THE INDIAN PINES DATA SET

Metric	Classes	FCM	US-ELM	Scalable-SC	EKGCSC	SRDL	FSCAG	SGCNR	FSCS	JSCC	3DCAE	Proposed SGLSC
PA (%)	Alfalfa	0.00	2.17	8.70	0.00	6.52	0.00	0.00	**10.87**	6.52	0.00	0.00
	Corn-notill	31.16	30.11	30.25	51.82	**67.09**	32.07	19.19	34.45	35.01	21.08	42.44
	Corn-mintill	23.86	34.58	0.00	32.53	32.53	16.27	11.45	34.82	27.47	12.29	**57.35**
	Corn	8.86	0.42	5.91	64.14	0.00	27.00	33.76	11.39	12.66	14.77	**100.00**
	Grass/Pasture	40.79	65.63	7.87	32.51	**65.84**	49.28	47.62	36.44	15.53	4.14	**65.84**
	Grass/Trees	35.62	53.97	18.08	55.75	39.04	56.85	55.62	25.07	56.71	90.55	**99.73**
	Grass/Pasture-mowed	0.00	0.00	3.57	**100.00**	0.00	3.57	0.00	53.57	0.00	0.00	**100.00**
	Hay-windrowed	98.95	98.54	98.12	**100.00**	98.12	99.16	99.37	55.86	65.69	97.91	**100.00**
	Oats	65.00	0.00	5.00	40.00	0.00	**75.00**	50.00	50.00	0.00	**75.00**	0.00
	Soybean-notill	30.04	42.18	5.76	10.08	74.79	33.44	51.03	31.38	45.99	43.21	**75.31**
	Soybean-mintill	30.79	69.86	**77.39**	47.78	41.83	32.91	32.99	32.02	68.43	60.73	74.26
	Soybean-clean	25.13	29.01	35.92	54.47	24.79	18.38	24.96	25.97	27.15	26.98	**74.20**
	Wheat	96.59	95.12	98.05	64.88	0.49	96.59	98.54	97.56	**100.00**	24.88	**100.00**
	Woods	26.40	69.09	**99.21**	65.77	71.46	33.36	29.09	42.37	70.20	71.46	62.37
	Bldg-Grass-Tree-Driver	16.32	34.97	10.62	61.66	76.94	16.84	12.69	16.84	40.93	42.49	**79.79**
	Stone-Steel-Tower	6.45	3.23	2.15	0.00	0.00	2.15	0.00	68.82	**88.17**	3.23	1.08
OA (%)		33.22	52.74	46.43	49.15	52.75	36.38	35.54	34.87	50.59	46.79	**69.95**
Kappa		0.2685	0.4617	0.3702	0.4353	0.4772	0.2922	0.2907	0.2808	0.4402	0.4001	**0.6638**
NMI		0.4238	0.5223	0.4632	0.4770	0.6418	0.4504	0.4223	0.3719	0.4892	0.4450	**0.6917**
Time (s)		60.27	24.47	**1.62**	11193.08	150.20	2.95	6.52	6.37	16.43	627.42	127.22

By comparison, the three graph-based clustering methods for large HSIs, i.e., the FSCAG, the SGNCR, and the FSCS, all take a few seconds. Nevertheless, it should be noted that the three methods are not able to obtain satisfactory clustering accuracies, and they achieve OAs of 36.38%, 35.54%, and 34.87%, respectively. The sparse representation recovery residual-based method for large HSIs (i.e., JSCC) achieves the third-best clustering accuracy with a small execution time equal to only 16.43s. However, the clustering accuracy of this method also cannot meet the needs of practical applications. The

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3132683

3DCAE method, which adopts the 3D convolutional neural network (CNN) to extract the features for the clustering task and integrates the spatial and spectral information, is selected for comprehensive comparison. However, it failed to achieve satisfactory result on this scene, with an OA of only 46.79%. This may be due to the fact that 3DCAE uses regular spatial information, which does not accurately match the ground object distribution.

The proposed SGLSC approach achieves the best clustering performance, including the OA, Kappa and NMI of 69.95%, 0.6638 and 0.6917, respectively. Furthermore, the accuracy of SGLSC is far better than the second best result on this data set, with 17.20%, 0.1866 and 0.0499 improvement in term of OA, Kappa and NMI, respectively. In particular, the proposed SGLSC achieves the best PA for ten land-cover types, with the classes of Corn, Grass/Pasture-mowed, Hay-windrowed, and Wheat having accuracies of 100%. Overall, the proposed SGLSC algorithm exhibits superiority to the other compared clustering methods in this experiment.

E. Results on the Salinas Valley Data Set

The second experiment was conducted on the Salinas Valley data set. Fig. 10 shows the clustering maps obtained by the proposed SGLSC approach and the compared methods. Table III reports the OA, Kappa, NMI, PA and execution time for different methods. Similarly to the results of the Indian Pines data set, the proposed SGLSC approach is far superior to other considered methods.

For the common clustering methods (i.e., FCM, Scalable-SC, and US-ELM), the FCM obtains the worst clustering accuracy on this data set, with the lowest OA of 58.20%, and its clustering map displays a large number of misclassifications and noise. The Scalable-SC can efficiently obtain the clustering result for this large hyperspectral image, with a running time of only 14.11s. However, the clustering accuracy of this method is not satisfactory, with an OA of only 64.55%. The US-ELM achieves the third-best accuracy among these clustering methods in this experiment, but at the cost of requiring the second-largest running time. In details, from Fig. 10(c), one can observe that there are a lot of misclassification between the classes of Grapes_unstrained and Vinyard_unstrained, and approximately half of the land-cover of Fallow is misclassified as Lettuce_5wk.

It is unfortunate that the EKGCSC and SRDL methods cannot deal with this data set because of the large memory consumption and the high computing complexity. By comparison, the FSCAG, the SGNCR, and the FSCS can also efficiently cluster this large hyperspectral image, with execution times of 34.02s, 41.75s, and 34.19s, respectively. However, they cannot achieve satisfactory clustering accuracies, with OAs of only 65.89%, 60.24%, and 59.97%, respectively.

In contrast, the JSCC can categorize different land cover types with the second best clustering OA of 68.46%. It should be noted that the execution time of the JSCC is relatively high (i.e., 451.25s). The 3DCAE obtains acceptable clustering accuracies, and a suboptimal NMI (i.e., 0.8271), which is far better than other considered methods. It should also be noted that the 3DCAE consumes the most running time on this scene.

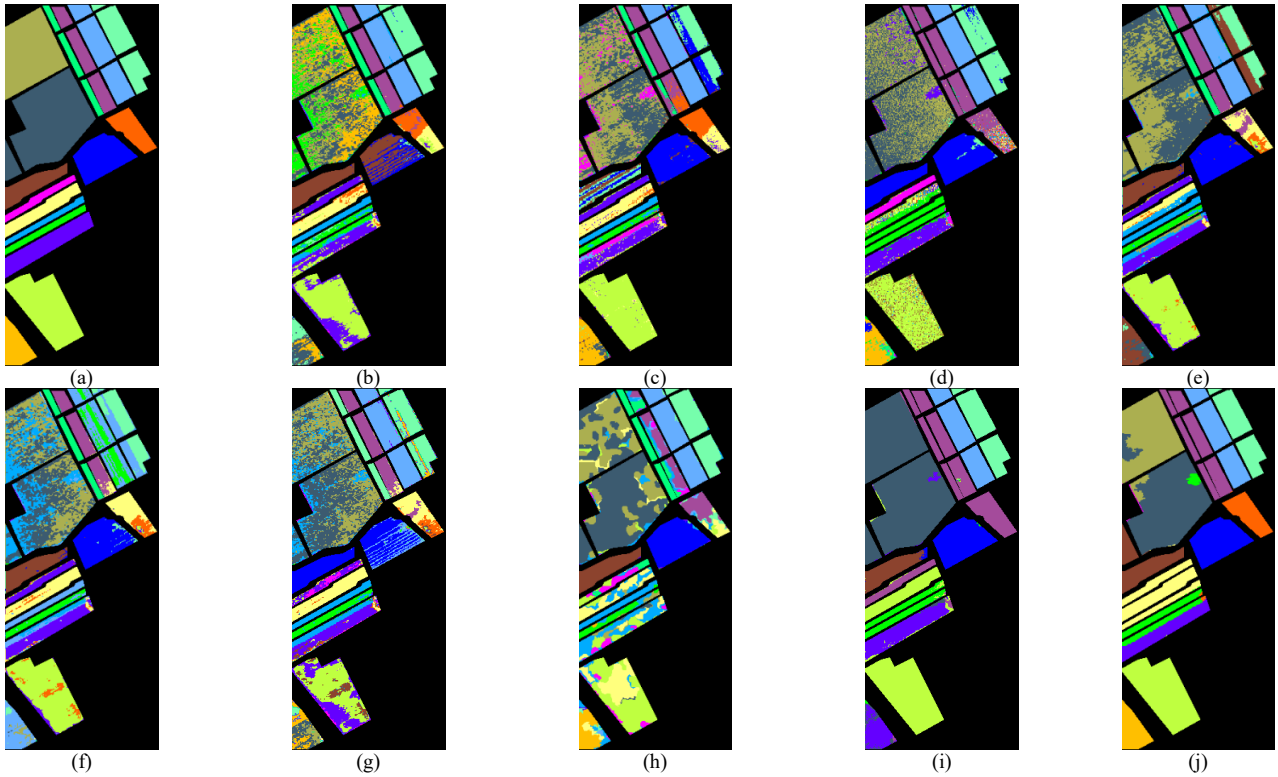

Fig. 10. Clustering maps of different methods on the Salinas Valley data set. (a) Ground truth. (b) FCM. (c) US-ELM. (d) Scalable-SC. (e) FSCAG. (f) SGCNR. (g) FSCS. (h) JSSC. (i) 3DCAE. (j) Proposed SGLSC.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3132683

TABLE III
QUANTITATIVE ASSESSMENT OF CLUSTERING RESULTS BY USING SIX COMPARED METHODS ON THE SALINAS VALLEY DATA SET

Metric	Classes	FCM	US-ELM	Scalable-SC	FSCAG	SGCNR	FSCS	JSCC	3DCAE	Proposed SGLSC
PA (%)	Brocoli_weeds_1	99.75	31.66	0.00	96.71	95.87	0.00	99.65	94.82	**100.00**
	Brocoli_weeds_2	35.29	96.22	95.57	98.55	95.38	68.65	97.72	99.70	**100.00**
	Fallow	51.82	57.19	13.36	16.85	24.44	24.34	0.00	0.00	**100.00**
	Fallow_rough	98.64	69.01	0.50	97.92	97.56	0.00	**100.00**	0.00	0.00
	Fallow_smooth	97.65	83.68	97.61	97.80	89.13	90.93	47.80	99.10	**99.81**
	Stubble	96.36	98.99	99.77	95.91	59.86	95.58	98.53	99.82	**99.90**
	Celery	98.85	61.83	94.72	54.26	63.51	80.78	98.46	99.86	**100.00**
	Grapes_unstrained	33.50	45.45	52.80	52.23	46.59	58.83	76.27	**97.98**	87.12
	Soil	73.14	97.03	79.53	91.20	87.57	56.62	57.36	**100.00**	**100.00**
	Corn	48.87	56.62	82.40	60.31	61.53	55.46	51.68	**86.49**	70.99
	Lettuce_4wk	0.00	0.00	**94.85**	0.09	0.75	34.27	42.70	0.00	0.00
	Lettuce_5wk	81.94	89.15	12.45	49.71	92.79	98.13	58.33	0.00	**100.00**
	Lettuce_6wk	99.24	98.14	0.00	**99.34**	0.00	98.80	0.00	0.00	0.00
	Lettuce_7wk	61.50	81.68	**98.60**	86.45	87.66	88.22	63.18	94.67	98.41
	Vinyard_unstrained	31.48	59.89	53.40	50.61	38.79	50.17	55.88	0.00	**92.38**
	Vinyard_trellis	25.84	81.96	76.43	0.00	0.00	33.09	63.70	0.00	**100.00**
OA (%)		58.20	68.37	64.55	65.89	60.24	59.97	68.46	68.16	**88.25**
Kappa		0.5470	0.6504	0.6074	0.6212	0.5652	0.5578	0.6496	0.6389	**0.8695**
NMI		0.6837	0.7314	0.7479	0.7433	0.7042	0.6665	0.7211	0.8271	**0.9062**
Time (s)		321.32	932.76	**14.11**	34.02	41.75	34.19	451.25	2596.44	40.19

The proposed SGLSC approach obtains the best clustering performance compared with the other methods (see Table III). In details, the proposed SGLSC has 19.79%, 0.2191 and 0.0791 improvements of OA, Kappa and NMI compared with the second best result, respectively. Furthermore, the proposed SGLSC achieves the best PA in most land-cover types, and there are seven classes whose accuracies are 100%. The running time of the SGLSC is 40.19s, which is less than the time required by the JSCC. This demonstrates the excellent capabilities of the proposed SGLSC on this data set.

F. Results on the Pavia Center Data Set

A large-size hyperspectral image, i.e., Pavia Center data set, is used in the third experiment to evaluate the performance of the proposed SGLSC approach. From the clustering maps shown in Fig. 11 and the quantitative assessments listed in Table IV, one can see that the proposed SGLSC achieves the best clustering performance.

The FCM obtains the poorest clustering accuracy, with a small OA of 67.93% and a clustering map including a large number of misclassifications. The Scalable-SC is still able to quickly cluster large HSIs, with a running time of only 303.32s. However, its accuracy is low, with an OA of only 73.25%. The US-ELM achieves the second best accuracy, i.e., 83.18%, among all these methods on this data set. However, it is not suitable to be extended to practical applications due to the huge running time, i.e., 39675.63s.

For the three graph-based clustering methods for large HSIs, the SGCNR also achieves poor clustering performances, with an OA of 67.58%. The FSCAG performs better than the SGCNR in terms of both clustering accuracy and running time on this date set. The FSCS achieves an unsatisfactory clustering performance, with an OA of 75.58%. However, it is most efficient and takes only a small amount of time (i.e., 67.32s).

The JSCC obtains the third-best clustering accuracy, i.e., 82.55%, with a relatively higher running time, i.e., 8733.11s. However, it achieves the best result in term of NMI. The 3DCAE only achieves a moderate clustering performance and requires a relatively long running time with respect to other methods on this scene.

In comparison, the proposed SGLSC approach achieves the best clustering performance and yields smooth clustering map. It performs best for three land cover types, i.e., Bricks, Baresoil and Tiles, among all other methods. Moreover, the proposed SGLSC increases of 5.63% and 0.0833 the OA and the Kappa over the second best method (i.e., the US-ELM), and obtains the second best NMI on this data set. In addition, the running time of the proposed SGLSC approach is 2236.53s, which is close to the running time of the SGCNR. However, the OA of the proposed SGLSC is 21.23% higher than that of the SGCNR. This experiment from both the visual results and quantitative results further confirmed the effectiveness of the proposed SGLSC approach.

G. Discussion

According to the experimental results presented in the previous section, one can see that the proposed SGLSC almost always outperforms the other clustering algorithms in terms of OAs while requires less execution time. In this section, the effectiveness and computational complexity of the SGLSC are analyzed and discussed in detail.

1) Analysis of the effectiveness of proposed SGLSC

In the proposed SGLSC, superpixel-level global and local similarity graph S (consisting of global similarity graph S^G and local similarity graph S^L) is presented for modelling the distribution of ground objects more realistically. From the

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3132683

experimental results in Fig. 8 (Section IV *C.*), one can see that leveraging on both S^G and S^L provides the best accuracy on all three HSI data sets. This indicates that the superpixel-level global and local similarity graph S can effectively model the essential structure of the HSIs. A further analysis of Fig. 1 (1) - (3) (the visualization results of S^G, S^L and S on the Salinas Valley data set) indicates that the global similarity graph S^G and local similarity graph S^L are complementary to each other. The superpixel-level global similarity graph S^G enables the connection between superpixels belonging to the same category that are not adjacent to each other, which is consistent with the real-world situation in which objects belonging to the same category may be distributed in different locations in the HSI. The superpixel-level local similarity graph S^L models the relationship between neighboring superpixels, which corresponds to the fact that superpixels in close proximity are more likely to belong to the same category. As a result, the final similarity graph S integrating the S^G and S^L ensures that the proposed SGLSC method can achieve efficient performance.

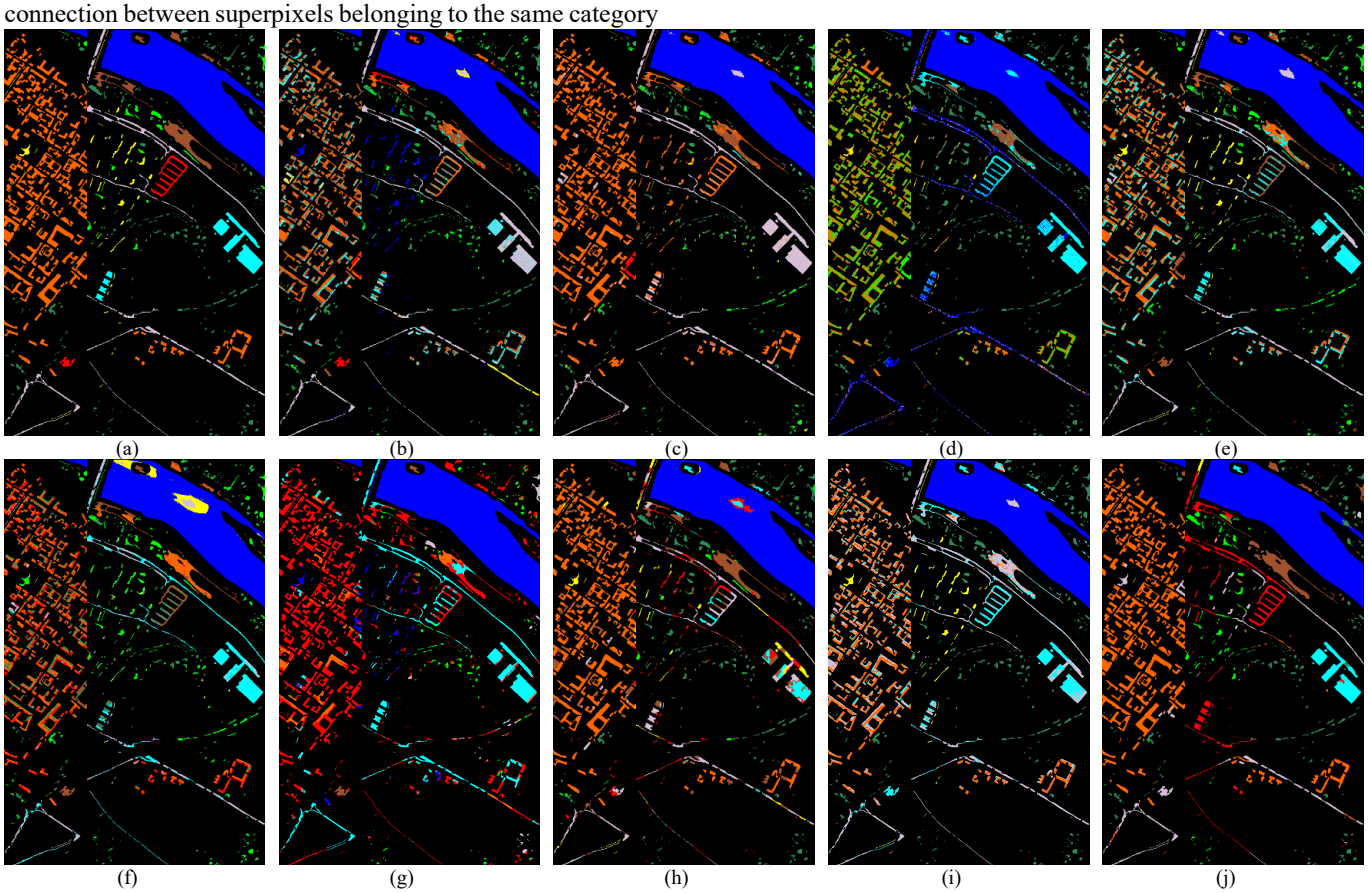

Fig. 11. Clustering maps of different methods on the Pavia Center data set. (a) Ground truth. (b) FCM. (c) US-ELM. (d) Scalable-SC. (e) FSCAG. (f) SGCNR. (g) FSCS. (h) JSSC. (i) 3DCAE. (j) Proposed SGLSC.

TABLE IV
QUANTITATIVE ASSESSMENT OF CLUSTERING RESULTS BY USING SIX COMPARED METHODS ON THE PAVIA CENTER DATA SET

Metric	Classes	FCM	US-ELM	Scalable-SC	FSCAG	SGCNR	FSCS	JSCC	3DCAE	Proposed SGLSC
PA (%)	Water	99.11	99.36	99.29	98.76	93.40	**100.00**	98.03	99.35	99.61
	Trees	63.56	64.73	**100.00**	46.14	39.01	29.38	85.89	99.95	68.83
	Meadows	17.54	21.75	0.00	61.68	**84.53**	42.52	28.58	0.00	31.52
	Bricks	19.85	0.00	0.11	0.00	0.52	15.23	19.33	0.00	**79.70**
	Baresoil	36.53	50.29	46.54	38.81	31.39	15.78	77.28	0.00	**77.32**
	Asphalt	87.49	**99.34**	14.40	90.59	49.22	76.25	53.72	83.43	48.77
	Bitumen	24.76	0.00	84.23	80.32	**84.29**	0.00	31.56	65.38	83.71
	Tiles	39.81	92.01	55.65	63.75	47.04	79.30	84.63	63.62	**97.66**
	Shadows	0.10	6.64	36.85	**99.83**	0.10	0.00	38.53	98.22	0.00
OA (%)		67.93	83.18	73.25	79.32	67.58	75.58	82.55	78.08	**88.81**
Kappa		0.5633	0.7586	0.6272	0.7161	0.5680	0.6456	0.7579	0.6987	**0.8419**
NMI		0.6727	0.7728	0.8008	0.7366	0.6812	0.5971	**0.8358**	0.7136	0.8181
Time (s)		692.22	39675.63	303.32	1111.65	2232.45	**67.32**	8733.11	7617.43	2236.53

2) Computational Complexity

The clustering process for the classical spectral clustering (SC) algorithm (described in Section II) can be divided into three main steps: similarity graph construction, singular value decomposition (SVD) to determine the indicator matrix, and k-means clustering to obtain the optimal cluster index. Therefore, the time complexity of SC for a given hyperspectral image X with N pixels is $O(N^2) + O(N^3) + O(N)$, corresponding to the above three steps, while the space complexity is $O(N^2) + O(N^2) + O(N)$.

Assuming that the hyperspectral image X is over-segmented into K superpixels ($K \ll N$), the proposed SGLSC needs to construct the superpixel-level global and local similarity graphs. The time and space complexities for constructing the global similarity graph according to Eq. (6) are $O(K^3)$ and $O(K^2)$, and the time and space complexities for constructing the local similarity graph according to Eq. (11) are $O(K^2)$ and $O(K^2)$, respectively. Thus, the overall time and space complexities of the proposed SGLSC are $O(K^2) + O(K^3) + O(K^3) + O(K)$ and $O(K^2) + O(K^2) + O(K^2) + O(K)$, respectively. The SGLSC can greatly reduce the time and space complexity as K is much smaller than N. The reduction in space complexity, in particular, makes SGLSC suitable for large size HSIs. For example, the Salinas Valley benchmark image contains 111,104 pixels (i.e., the spatial size is 512×217), requiring 91.97GB memory space for similarity graph constructed at pixel level. However, we can see from the experimental results in Section IV *C. 1)* that SGLSC achieved the best performance with 500 superpixels. The superpixel-level similarity graph took only 1.86MB memory space, which is significantly less than the pixel-level similarity graph.

V. Conclusion

This paper proposed a large HSIs clustering approach, i.e., the SGLSC, in which the superpixel concept is introduced to greatly decrease both the computation cost and the memory consumption on large size hyperspectral images. Moreover, a new similarity graph is constructed to fully exploit both the global and the local information for modeling the essential structure of HSIs. The strategy of combining superpixel-level global and local similarity graph significantly improves the clustering accuracy of the complex ground objects. In the experiments, we demonstrated that the SGLSC is significantly and consistently superior to some recently presented methods, including the graph-based, spatial spectral-based and CNN-based methods on three widely used benchmark data sets. In particular, the proposed SGLSC exhibits efficient computational properties when clustering large HSIs. In addition, the proposed superpixel-level global and local similarity graph enables the revisit of conventional graph-based methods for large HSIs clustering.

In future developments of this work, we will further improve two aspects of the proposed SGLSC. On the one hand, the number of superpixels for different data sets are determined by repeated experiments. An adaptive automatic strategy for estimating the optimal number of superpixels for different data sets will be investigated. On the other hand, from the methodology perspective, despite the proposed method has shown accurate clustering results and low computational complexity for large HSIs, it still requires efficiency optimization. To this end, we will design a one-step similarity graph construction strategy that can simultaneously consider global and local similarities and adaptively adjust the weights between them to further improve the clustering accuracy.

References

[1] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. M. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.

[2] Heiden, U., Heldens, W., Roessner, S., Segl, K., Esch, T. and Mueller, A. "Urban structure type characterization using hyperspectral remote sensing and height information," *Landscape Urban Plan.*, vol. 102, no. 4, pp. 361-375, Apr. 2012.

[3] I. Herrmann, U. Shapira, S. Kinast, A. Karnieli, and D. J. Bonfil, "Ground-level hyperspectral imagery for detecting weeds in wheat fields," *Precis. Agric.*, vol. 14, no. 6, pp. 637–659, Dec. 2013.

[4] C. McCann, K. S. Repasky, R. Lawrence, and S. Powell, "Multi-temporal mesoscale hyperspectral data of mixed agricultural and grassland regions for anomaly detection," *ISPRS J. Photogramm. Remote Sens.*, vol. 131, pp. 121–133, Sep. 2017.

[5] G. Camps-Valls, D. Tuia, L. Bruzzone, and J. Atli Benediktsson, "Advances in hyperspectral image classification: Earth monitoring withstatistical learning methods," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 45–54, Jan. 2014.

[6] R. D. P. M. Scafutto, C. R. de Souza Filho, and W. J. de Oliveira, "Hyperspectral remote sensing detection of petroleum hydrocarbons in mixtures with mineral substrates: Implications for onshore exploration and monitoring," *ISPRS J. Photogramm. Remote Sens.*, vol. 128, pp. 146–157, Jun. 2017.

[7] X. Liu, R. Wang, Z. Cai, Y. Cai, and X. Yin, "Deep multigrained cascade forest for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 8169–8183, Oct. 2019.

[8] 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.

[9] Pal, Mahesh. "Extreme−learning−machine−based land cover classification." *Int. J. Remote Sens.*, vol. 30, no. 14, pp. 835-3841, Jul. 2009.

[10] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Deep learning for hyperspectral image classification: An overview," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6690–6709, Sep. 2019.

[11] L. He, J. Li, C. Liu, and S. Li, "Recent advances on spectral–spatial hyperspectral image classification: An overview and new guidelines," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1579–1597, Mar. 2018.

[12] N. Audebert, B. Le Saux and S. Lefevre, "Deep learning for classification of hyperspectral data: a comparative review," *IEEE Trans. Geosci. Remote Sens.*, vol. 7, no. 2, pp. 159-173, June 2019.

[13] U. Maulik and I. Saha, "Modified differential evolution based fuzzy clustering for pixel classification in remote sensing imagery," *Pattern Recognit.*, vol. 42, no. 9, pp. 2135–2149, Sep. 2009.

[14] H. Zhai, H. Zhang, L. Zhang, and P. Li, "Reweighted mass center based object-oriented sparse subspace clustering for hyperspectral images," *J. Appl. Remote Sens.*, vol. 10, no. 4, Nov. 2016, Art. no. 046014.

[15] Y. Wan, Y. Zhong, A. Ma and L. Zhang, "Multi-objective sparse subspace clustering for hyperspectral imagery," *IEEE Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2290-2307, April 2020.

[16] H. Zhai, H. Zhang, P. LI and L. Zhang, "Hyperspectral image clustering: current achievements and future lines," *IEEE Geosci. Remote Sens. Mag.*, doi: 10.1109/MGRS.2020.3032575.

[17] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A K-means clustering algorithm," *Appl. Stat.*, vol. 28, no. 1, pp. 100–108, 1979.

[18] J. C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms. New York, NY, USA: Plenum, 1981.

[19] 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.

[20] S. Niazmardi, S. Homayouni, and A. Safari, "An improved FCM algorithm based on the SVDD for unsupervised hyperspectral data classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 831–839, Apr. 2013.

[21] M. Ester, H. P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. Int. Conf. Knowl. Discov. Data Min*. (*KDD*), Aug. 1996, pp. 226–231.

[22] Ankerst, M., Breunig, M. M., Kriegel, H. P., & Sander, J., "OPTICS: Ordering Points to Identify the Clustering Structure," in *Proc. ACM SIGMOD Int. Conf. Manag. Data*. (*SIGMOD*), June, 1999, pp. 19-60.

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

[24] N. Acito, M. Diani, and G. Corsini, "Gaussian mixture model based approach to anomaly detection in multi/hyperspectral images," in *Proc. SPIE*, vol. 5982, Oct. 2005, Art. no. 59820O.

[25] C. Teodor, B. Alzenk, R. Constantinescu, and M. Datcu, "Unsupervised classification of EO-1 hyperion hyperspectral data using latent Dirichlet allocation," in *Proc. Int. Symp. Signals, Circuits Syst. (ISSCS)*, Jul. 2013, pp. 1–4.

[26] A. Baraldi and F. Parmiggiani, "A neural network for unsupervised categorization of multivalued input patterns: An application to satellite image clustering," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 2, pp. 305–316, Mar. 1995.

[27] Y. Zhong, L. Zhang, B. Huang, and P. Li, "An unsupervised artificial immune classifier for multi/hyperspectral remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp. 420–431, 2006.

[28] G. Huang, S. Song, J. N. D. Gupta and C. Wu, "Semi-Supervised and Unsupervised Extreme Learning Machines," *IEEE Trans. Cybern.*, vol. 44, no. 12, pp. 2405-2417, Dec. 2014.

[29] J. Xu, H. Li, P. Liu, and L. Xiao, "A novel hyperspectral image clustering method with context-aware unsupervised discriminative extreme learning machine," *IEEE Access*, vol. 6, pp. 16,176–16,188, Mar. 2018.

[30] F. Luo, H. Huang, Z. Ma and J. Liu, "Semisupervised Sparse Manifold Discriminative Analysis for Feature Extraction of Hyperspectral Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 10, pp. 6197-6211, Oct. 2016.

[31] F. Luo, L. Zhang, B. Du and L. Zhang, "Dimensionality Reduction With Enhanced Hybrid-Graph Discriminant Learning for Hyperspectral Image Classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 8, pp. 5336-5353, Aug. 2020.

[32] Y. Yan, G. Liu, S.Wang, J. Zhang, and K. Zheng, "Graph-based clustering and ranking for diversied image search," *Multimedia Syst.*, vol. 23, no. 1, pp. 41-52, 2017.

[33] P. Das and A. K. Das, "Graph-based clustering of extracted paraphrases for labelling crime reports," *Knowl.-Based Syst.*, vol. 179, pp. 55-76, Sep. 2019.

[34] H. Wang, Y. Yang, B. Liu, and H. Fujita, "A study of graph-based system for multi-view clustering," *Knowl.-Based Syst.*, vol. 163, pp. 1009-1019, Jan. 2019.

[35] 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.

[36] H. Zhai, H. Zhang, L. Zhang, P. Li, and A. Plaza, "A new sparse subspace clustering algorithm for hyperspectral remote sensing imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 1, pp. 43–47, Jan. 2017.

[37] H. Zhai, H. Zhang, X. Xu, L. Zhang, and P. Li, "Kernel sparse subspace clustering with a spatial max pooling operation for hyperspectral remote sensing data interpretation," *Remote Sens.*, vol. 9, no. 4, p. 335, Apr. 2017.

[38] L. Zhang, L. Zhang, B. Du, J. You and D. Tao, "Hyperspectral image unsupervised classification by robust manifold matrix factorization," *Inform Sciences*, vol. 485, pp.154-169, June 2019.

[39] N. Huang, L. Xiao and Y. Xu, "Bipartite Graph Partition Based Coclustering With Joint Sparsity for Hyperspectral Images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 12, pp. 4698-4711, Dec. 2019, doi: 10.1109/JSTARS.2019.2953378.

[40] Y. Cai, Z. Zhang, Z. Cai, X. Liu, X. Jiang and Q. Yan, "Graph convolutional subspace clustering: a robust subspace clustering framework for hyperspectral image," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 4191-4202, May 2021.

[41] N. Huang, L. Xiao, J. Liu and J. Chanussot, "Graph Convolutional Sparse Subspace Coclustering With Nonnegative Orthogonal Factorization for Large Hyperspectral Images," *IEEE Trans. Geosci. Remote Sens.*, doi: 10.1109/TGRS.2021.3096320.

[42] R. Wang, F. Nie, and W. Yu, "Fast spectral clustering with anchor graph for large hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 11, pp. 2003–2007, Nov. 2017.

[43] R. Wang, F. Nie, Z. Wang, F. He, and X. Li, "Scalable graph-based clustering with nonnegative relaxation for large hyperspectral image," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 7352–7364, Oct. 2019.

[44] Y. Wei, C. Niu, Y. Wang, H. Wang and D. Liu, "The Fast Spectral Clustering Based on Spatial Information for Large Scale Hyperspectral Image," *IEEE Access*, vol. 7, pp. 141045-141054, 2019, doi: 10.1109/ACCESS.2019.2942923.

[45] X. Yang, G. Lin, Y. Liu, F. Nie and L. Lin, "Fast Spectral Embedded Clustering Based on Structured Graph Learning for Large-Scale Hyperspectral Image," *IEEE Geosci. Remote Sens. Lett.*, doi: 10.1109/LGRS.2020.3035677.

[46] N. Huang, L. Xiao, Y. Xu and J. Chanussot, "A Bipartite Graph Partition-Based Coclustering Approach With Graph Nonnegative Matrix Factorization for Large Hyperspectral Images," *IEEE Trans. Geosci. Remote Sens.*, doi: 10.1109/TGRS.2021.3097358.

[47] H. Zhai, H. Zhang, L. Zhang and P. Li, "Sparsity-based clustering for large hyperspectral remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, doi: 10.1109/TGRS.2020.3032427.

[48] C. Yang, L. Bruzzone, H. Zhao, Y. Tan and R. Guan, "Superpixel-Based Unsupervised Band Selection for Classification of Hyperspectral Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 12, pp. 7230-7245, Dec. 2018.

[49] H. Zhao, L. Bruzzone, R. Guan, F. Zhou and C. Yang, "Spectral-Spatial Genetic Algorithm-Based Unsupervised Band Selection for Hyperspectral Image Classification," *IEEE Trans. Geosci. Remote Sens.*, doi: 10.1109/TGRS.2020.3047223.

[50] B. Tu, C. Zhou, D. He, S. Huang and A. Plaza, "Hyperspectral Classification With Noisy Label Detection via Superpixel-to-Pixel Weighting Distance," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 6, pp. 4116-4131, June 2020.

[51] Ren and Malik, "Learning a classification model for segmentation," in *Proc. Int. Conf. Comput. Vis. (ICCV)*, 2003, pp. 10-17.

[52] X. Zhu, C. Change Loy, and S. Gong, "Constructing robust affinity graphs for spectral clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2014, pp. 1450–1457.

[53] R. S. Zemel and M. Á. Carreira-Perpiñán, "Proximity graphs for clustering and manifold learning," in *Proc. Adv. Neural Inf. Process. Syst. (ANIPS)*, 2005, pp. 225–232.

[54] S. Liu, S. De Mello, J. Gu, G. Zhong, M. H. Yang, and J. Kautz, "Learning affinity via spatial propagation networks," in *Proc. Adv. Neural Inf. Process. Syst. (ANIPS)*, 2017, pp. 1520–1530.

[55] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," in *Proc. Conf. Workshop Neural Inf. Process. Syst*. (*NIPS*), 2002, pp. 849–856.

[56] U. Von Luxburg, "A tutorial on spectral clustering," *Statist. Comput.*, vol. 17, no. 4, pp. 395–416, 2007. doi: 10.1007/s11222-007-9033-z.

[57] M.-Y. Liu, O. Tuzel, S. Ramalingam, and R. Chellappa, "Entropy rate superpixel segmentation," in *Proc. Comput. Vis. Pattern Recognit.* (*CVPR*), 2011, pp. 2097–2104.

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

[59] R. Vidal and P. Favaro, "Low rank subspace clustering (LRSC)," *Pattern Recognit. Lett.*, vol. 43, pp. 47–61, Jul. 2014.

[60] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406-5425, Dec. 2006.

[61] S. Boyd and L. Vandenberghe, Convex Optimization. Cambridge Univ. Press, 2004.

[62] S.J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An interior-point method for large-scale l_1-regularized least squares," *IEEE J Sel Top Signal Process*, vol. 1, no. 4, pp. 606-617, Dec. 2007.

[63] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found.Trends Mach. Learn.*, vol. 3, no. 1, pp. 1-122, Jan. 2011.

[64] G. Chen, "Scalable spectral clustering with cosine similarity," in *Proc. Int. Conf. Pattern Recognit.* (*ICPR*), Aug. 2018, pp. 314–319.

[65] J. M. Murphy and M. Maggioni, "Spectral–Spatial Diffusion Geometry for Hyperspectral Image Clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 7, pp. 1243-1247, Jul. 2020.

[66] J. Nalepa, M. Myller, Y. Imai, K. -I. Honda, T. Takeda and M. Antoniak, "Unsupervised Segmentation of Hyperspectral Images Using 3-D Convolutional Autoencoders," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 11, pp. 1948-1952, Nov. 2020.