Spatial-Spectral Clustering with Anchor Graph for Hyperspectral Image

Qi Wang, *Senior Member, IEEE,* Yanling Miao, Mulin Chen, and Yuan Yuan, *Senior Member, IEEE*

Abstract—**Hyperspectral image (HSI) clustering, which aims at dividing hyperspectral pixels into clusters without labeled training data, has drawn significant attention in practical applications. Recently, many graph-based clustering methods, which construct an adjacent graph to model the data relationship, have shown dominant performance. However, the high dimensionality of HSI data makes it hard to construct the pairwise adjacent graph. Besides, abundant spatial structures are often overlooked during the clustering procedure. In order to better handle the high dimensionality problem and preserve the spatial structures, this paper proposes a novel unsupervised approach called spatial-spectral clustering with anchor graph (SSCAG) for HSI data clustering. The SSCAG has the following contributions: 1) the multiscale filtering module is utilized to smooth the homogeneous regions, so that it can increase the similarity and consistency of neighboring pixels and capture the multiple views of a local region with different scales; 2) a new similarity metric is proposed to embed the spatial-spectral features into the combined adjacent graph, which can mine the intrinsic property structure of HSI data; 3) the AG-based strategy is adopted to construct the adjacent graph by a neighbor assignment scheme without hyperparameters, and its optimization employs SVD to replace eigenvalue decomposition to reduce the computational complexity. Extensive experiments on three public HSI datasets show that the proposed SSCAG is competitive against the state-of-the-art approaches.**

Index Terms—**Hyperspectral image, graph-based clustering, anchor graph, spatial-spectral information.**

I. Introduction

HYPERSPECTRAL image (HSI) data obtained by hyperspectral imaging spectrometer provides abundant spatial structure and spectral information of the observed objects [1]–[3]. With very narrow diagnostic spectral bands (wavelength interval is generally 10 nm), HSI usually has high spectral resolution [4], and can effectively distinguish subtle objects and materials between land cover classes [5], [6]. Therefore, HSI has been applied to the real-world applications, such as vegetation investigation [7], resource exploration [8], environmental monitoring [9] and target identification [10]. Among current tasks, the clustering is a commonly used method for processing hyperspectral image. The purpose of HSI clustering is to partition all pixels into several clusters according to

Yanling Miao is with the School of Computer Science and the School of Artificial Intelligence, Optics and Electronics (iOPEN), Northwestern Polytechnical University, Xi'an 710072, China. (e-mail: skilamiaomyl@gmail.com).

Qi Wang, Mulin Chen, and Yuan Yuan are with the School of Artificial Intelligence, Optics and Electronics (iOPEN), Northwestern Polytechnical University, Xi'an 710072, China. (e-mail: crabwq@gmail.com, chenmulin001@gmail.com, y.yuan1.ieee@gmail.com).

their intrinsic properties in an unsupervised way. By assigning groups, the points in one group have high similarity, while those in different groups show great differences. Different from supervised classification [11]–[14], clustering does not require labeled training samples, which has a wide range of applications in practice.

Though numerous techniques have been developed, HSI clustering remains to be a challenging issue. According to the working mechanism, existing clustering approaches mainly include these categories: centroid-based approaches [15]–[17], density-based approaches [18], [19], graph-based approaches [20], [21] and subspace approaches [22]–[24]. In particular, graph-based methods have obtained widespread attention due to their excellent performance [25], which use the eigenvalues of similarity matrix of data to reveal the cluster structure, where graph construction in these methods plays an essential role in the final clustering accuracy [20], [26], [27]. In the past two decades, the graph-based models have obtained widespread attention [28]–[30], which explore the pairwise adjacent graph between pixels and capture the data structure. Hang et al. [29] designed a regularization method to incorporate the spatial structure into the local graph discriminant embedding model (LGDE) naturally. The proposed model adopted the superpixel-level intraclass graphs as a regularization term for LGDE, where the intraclass graph was built to describe the spatial structure information for each superpixel. Liu et al. [30] proposed a spatial-spectral method with low-rank representation (LRR) and hypergraph learning, which preserves the spatial information and local manifold structure by fusing the locality constraint into LRR. In addition, the high-order relationships among pixels can be captured effectively by using hypergraph model. These methods usually have high time complexity which comes from two major parts: (1) the construction of $n \times n$ adjacent graph takes $\mathcal{O}(n^2 d)$, where n and d are the number of pixels and spectral bands respectively; (2) the eigenvalue decomposition on the graph Laplacian matrix costs $\mathcal{O}(n^2 c)$ or $\mathcal{O}(n^3)$, where c is the number of land cover types. To address the above problems, some anchor graph-based approaches have been proposed in the field of data mining. The Anchor Graph (AG) algorithm [31] utilizes a small number of anchors enough to cover the whole points to construct the large-scale adjacent graph. It reduces the computational complexity to $\mathcal{O}(ndm)$, where m is the number of anchors. Motivated by this, some AG-based variants have been developed over the past years in the HSI tasks [32]–[34].

However, the AG-based methods rarely utilize the spatial information within pixels, which limits their discriminant ca-
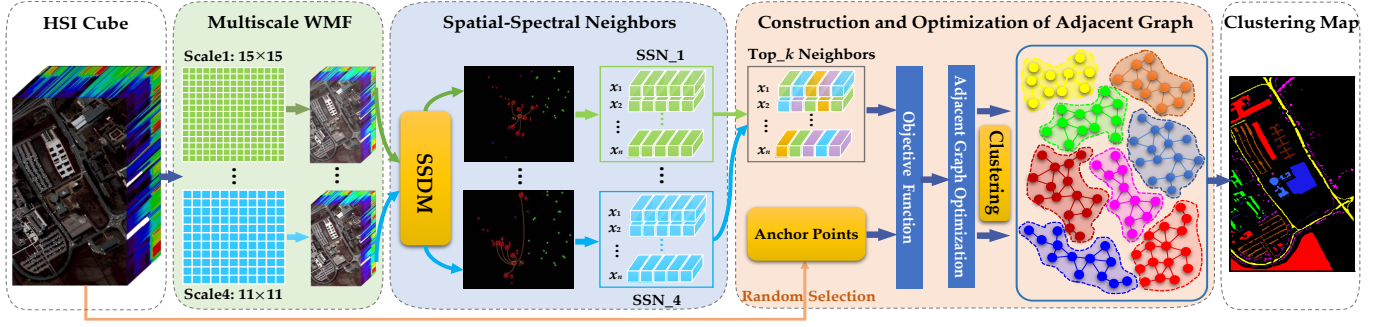
Fig. 1. Overview architecture of the proposed SSCAG. Each scale WMF is separately used to preprocess the HSI, then we get the corresponding processed HSIs. For each processed HSI, the defined spatial-spectral distance metric (SSDM) is utilized to obtain spatial-spectral nearest neighbors of each pixel, respectively. Reorder the obtained neighbors of each pixel from each scale WMF, we get the final top-k neighbors of each pixel. Next, we utilize anchor points and k nearest spatial-spectral neighbors to construct objective function. By adopting an effective neighbor assignment strategy with spatial information, the final clustering is obtained. **SSDM**: spatial-spectral distance metric. **SSN_***: the spatial-spectral neighbors set obtained from the previous WMF with scale $*$ and SSDM module.

pability for real-world applications [35]. Some spatial-spectral combined attempts have been made in graph-based approaches through spatial correlations and spectral information. This category of methods mainly utilize the spatial correlations in the filtering preprocessing, which filter local homogeneous regions with single scale [36], [37]. Nevertheless, they cannot fully capture the spatial structures for three reasons: 1) the HSI data may include small and large homogeneous regions simultaneously, but the local homogeneous regions can not be covered accurately with the single-scale filter. This phenomenon makes spatial-spectral algorithms fail to describe different spatial structures of HSI; 2) most of them just consider the spatial structures in the HSI data preprocessing process, and fail to incorporate these inherent structures into the clustering process; 3) the graph is commonly constructed by k nearest neighbors method with Euclidean distance in the original data space, which is sensitive to noise and often fails to capture the intrinsic structure, especially for the pixels distributed near the boundary of classes.

To overcome the aforesaid shortcomings, this paper proposes a novel spatial-spectral clustering with anchor graph (SSCAG) for HSI data. The main contributions can be summarized as follows.

• We design a multiscale spatial filtering module to preprocess HSI data according to the distribution properties of different datasets, which considers small and large homogeneous regions simultaneously. The multiscale filtering increases the local neighboring pixel consistency of HSI, and provides rich complementary information with different scales.

• We introduce a distance metric that skillfully combines spatial structure and spectral features simultaneously, which is used to find effective neighbors and construct the adjacent graph. In this way, the spatial-spectral collaboration information is integrated into the objective function.

• We adopt the AG-based strategy to construct the adjacent graph, which models the relationships between the points and anchors, instead of the complete-graph. Different from previous works, the adjacent graph is constructed by a neighbor assignment scheme without hyperparameters, and its optimiza-

tion employs SVD to replace eigenvalue decomposition, to reduce the computational complexity.

The reminder of this paper is organized as follows. Section II introduces an overview of traditional and AG-based HSI clustering methods. Section III describes the proposed SSCAG in detail. Section IV presents the extensive experimental results and corresponding analysis. Finally, the conclusion is summarized in Section V.

II. RELATED WORK

In this section, we first briefly review the related works on traditional clustering methods, and then some existing AG-based methods are introduced.

A. Hyperspectral Clustering Methods

Many fundamental researches on HSI data clustering have been proposed. The traditional methods contains centroid-based approaches [15]–[17], density-based approaches [18], [19], graph-based approaches [20], [21], [38] and subspace approaches [22]–[24].

The centroid-based approaches usually cluster HSI data based on the similarity measure. Li *et al.* [39] proposed an improved k-means clustering to identify the mineral types from HSI of mining area, which uses dimensionless similarity measurement methods to obtain the mapping results, enhancing spectral absorption features. Chen *et al.* [17] proposed a fuzzy c-means (FCM) clustering with spatial constraints, which contributes to the introduction of fuzziness for belongingness of each pixel and exploits the spatial contextual information. Salem *et al.* [40] developed hyperspectral feature selection for HSI clustering, which utilizes revisited FCM with spatial and spectral features to enhance the clustering. The density-based approaches form clusters by dense regions in the feature space. Tu *et al.* [41] applied density peak clustering for noisy label detection of HSI, which considers the spatial correlations of adjacent pixels to define the local densities of the training set, improving the performance of classifiers. The graph-based methods find clusters by learning the adjacent graph of data. Zhang *et al.* [20] took the spectral correlations

and spatial information into sparse subspace clustering (SSC), which obtains a more accurate coefficient matrix for building the adjacent graph and promotes the clustering performance. To overcome the single sparse representation in SSC model, Yan *et al.* [21] proposed two adjacent graph based on overall sparse representation vector and dynamic weights selection method, which better uses the spectral relationship and spatial structure during adjacent graph construction. Subspace clustering models the internal complex structures of HSI according to the union of subspaces, and explores the underlying adjacency between points through self-representation learning. Zhai *et al.* [22] proposed a scalable and sketched subspace clustering method, which is based on nonlocal means regularization and reweighted sparse constraint, to mine structural information and learn more favorable representations. Huang *et al.* [23] designed a unified model based on scalable subspace clustering, which integrates the learning of a concise dictionary and robust subspace representation, to improve the robustness of the model to noise. Afterwards, they developed a multiview subspace clustering model [24], which combines local and nonlocal spatial features from each view in a unified framework, learning the view-specific structure.

As mentioned above, there are some drawbacks about them. The centroid-based approaches are sensitive to initialization and noise. The density-based approaches are not suitable for HSI data, because it is hard to find density peaks in high-dimensional sparse feature space. The traditional graph-based is inefficient in large scale data, which has high computational complexity.

B. AG-based Methods

To address the large computational complexity of graph-based methods based on the adjacent matrix that consists of all pixels, AG-based methods have been developed, which select only a few important and representative points to construct adjacent graph. The main step of AG-based methods is relation matrix $\mathbf{Z}$ construction that models the relationships among anchors and pixels. The matrix $\mathbf{Z}$ is usually constructed based on common similarity measurements, such as the Gaussian kernel distance. To obtain a more accurate $\mathbf{Z}$, it can be learned by optimizing the function $\sum_{j=1}^{m} \|x_i - u_j\|_2^2 z_{ij} + \gamma \|\mathbf{Z}\|_F^2$, where x_i and u_j represent arbitrary pixels and anchor points from HSI matrix, respectively. Motivated by recent development in the AG construction, He *et al.* [32] proposed fast semi-supervised learning with AG for HSI data, which constructs a naturally sparse and scale invariant AG, alleviating the computation burden. Wang *et al.* [33] developed a scalable AG-based clustering method, which adds the non-negative relaxation to AG model. With this, the clustering results are directly obtained without adopting k-means. To handle the defect that AG-based methods usually ignore the spatial information, Wang *et al.* [42] presented fast spectral clustering with AG for HSI data, which fuses the spatial information by using the mean of neighboring pixels to reconstruct center pixel. Wei *et al.* [43] utilized the spatial correlation by adopting weighted mean filtering (WMF) to filter hyperspectral pixels, which considers the local neighborhood relationship within a window. Zhou *et*

al. [36] proposed a regularized local discriminant embedding model with spatial-spectral information, which incorporates spatial feature into the dimensionality reduction procedure, to achieve optimal discriminative matrix by the minimization of local spatialspectral scatter. Feng *et al.* [37] designed the discriminate margins with spatial-spectral neighborhood pixels, which effectively captures the discriminative features and learns the structures of HSI data. Luo *et al.* [44] constructed the intraclass spatial-spectral hypergraph by considering the coordinate relationship and similarity between adjacent samples, which can better make the high-dimensional spatial features embedded in low-dimensional space.

III. SPATIAL-SPECTRAL CLUSTERING WITH ANCHOR GRAPH

In this section, we detail the proposed SSCAG algorithm from four aspects: weighted mean filtering (WMF), spatial-spectral neighbors, anchor-based adjacent graph construction and optimization. The overview architecture of SSCAG is shown in Fig. 1.

A. WMF

To smooth the homogeneous regions and reduce the interference of noisy points in the HSI, the WMF [36] is employed to preprocess the pixels by utilizing spatial information. Suppose that each pixel of HSI can be denoted as a vector $x_i \in \mathbb{R}^d$ ($i = 1, 2, ..., n$), where d is the number of spectral bands and n refers to the number of HSI pixels. The HSI data is denoted as $\mathbf{X} = [x_1, x_2, ..., x_n]^T \in \mathbb{R}^{n \times d}$, and s_{ij} represents the similarity between x_i and x_j, $\mathbf{S} \in \mathbb{R}^{n \times n}$. Let $\ell(x_i) \in \{1, 2, ..., c\}$ be the class label of the pixels $\{x_1, x_2, ..., x_n\}$ respectively, where c is the number of classes. Assuming that the coordinate of pixel x_i is denoted as (p_i, q_i), the adjacent pixels centered at x_i can be defined as

$$\Omega(x_i) = \left\{ x_i(p, q) \mid \begin{array}{l} p \in [p_i - t, p_i + t] \\ q \in [q_i - t, q_i + t] \end{array} \right\}, \qquad (1)$$

where $t = (w - 1)/2$, w indicates the size of neighborhood window and is a positive odd number. The pixels in the neighborhood space $\Omega(x_i)$ are also represented as $\{x_{i1}, x_{i2}, ..., x_{i(w^2-1)}\}$, where $w^2 - 1$ is the number of neighbors of x_i.

The reconstructed pixel $\hat{x}_i$ by WMF is defined with a weighted summation, i.e.,

$$\hat{x}_i = \frac{x_i + \sum_{k=1}^{w^2-1} v_k x_{ik}}{1 + \sum_{k=1}^{w^2-1} v_k}, \quad x_{ik} \in \Omega(x_i), \qquad (2)$$

where $v_k = exp\{-\gamma_0 \|x_i - x_{ik}\|_2^2\}$ is the weight that represents the spectral similarity between x_{ik} and x_i. The parameter γ_0 is empirically set to be 0.2 in the experiments, which reflects the degree of filtering. After filtering, the consistency of pixels in the same homogeneous regions is guaranteed. Since HSI may include homogeneous regions of different sizes simultaneously, we use multiscale WMF to obtain potential spatial structures of HSI. The abundant information with multiscale complementarity is beneficial to enhance the description for local homogeneous regions, thereby improving the performance of clustering.

B. Spatial-Spectral Neighbors

To better explore the comprehensive characteristics of HSI data, we introduce a new distance metric to seek effective neighbors by incorporating the spatial structures and spectral features.

The pixels in the HSI are spatially related [45]. Especially, the adjacent pixels in a local homogeneous area have the spatial distribution consistency of land objects, which consist of the same materials and belong to the same category [20]. Therefore, neighboring pixels are used to measure the spatial and spectral similarity. Suppose that a pixel x_i and its neighbors in $\Omega(x_i)$ form a local pixel patch $P(x_i) = \{x_{i1}, x_{i2}, ..., x_{iw^2}\}$. Let the spatial feature matrix be $\hat{\mathbf{L}} = [l_1, l_2, ..., l_n] \in \mathbb{R}^{2 \times n}$, where l_i denotes the coordinate of x_i. The distance d_{ss} of x_i and x_j can be defined as

$$d_{ss}(x_i, x_j) = \frac{\sum_{h=1}^{w^2} v_{ih} \|x_{ih} - \hat{x}_j\|_2}{\sum_{h=1}^{w^2} v_{ih}}, \quad x_{ih} \in P(x_i), \quad (3)$$

where v_{ih} is the weight that indicates the spatial similarity between pixel x_{ih} and $\hat{x}_j$. The weight can be obtained by a kernel function, which is denoted as

$$v_{ih} = exp\left\{ -\|l_{ih} - l_j\|_2^2 / \sigma_j^2 \right\}, \quad (4)$$

where l_{ih} and l_j are the coordinates of x_{ih} and x_j, respectively. The spatial distance $\|l_{ih} - l_j\|_2$ is defined by the Euclidean distance between their coordinates, and σ_j is set as the average of $\sum_{h=1}^{w^2} \|l_{ih} - l_j\|_2$, i.e.,

$$\sigma_j = \frac{1}{w^2} \sum_{h=1}^{w^2} \|l_{ih} - l_j\|_2. \quad (5)$$

Note that the values of $\mathbf{X}$ and $\hat{\mathbf{L}}$ are normalized within [0,1]. Eq. (4) and Eq. (5) enforce the pixels with larger spatial distances to have a smaller similarity.

For HSI data, the spectral neighbors may contain the pixels with similar spectrum, but they are in different classes. The spatial neighbors only consider the coordinate distance of pixels, which may include pixels placed in different classes, especially at the boundary. According to Eq. (3), the proposed $d_{ss}(x_i, x_j)$ combines the spatial and spectral features simultaneously, where $\|x_{ih} - \hat{x}_j\|_2$ denotes the spectral similarity of two pixels, and v_{ih} is the corresponding spatial similarity. Furthermore, it not only considers the neighboring pixels in the patch $P(x_i)$, but also employs the adjacent pixels in the $\Omega(x_j)$. Hence, d_{ss} chooses the effective neighbors by collaborating spatial and spectral distance. After obtaining the spatial-spectral neighbors set of each pixel with different scales of WMF, we choose k nearest neighbors for constructing the objective function in the next part.

C. Anchor-based Adjacent Graph Construction

To reduce the computational complexity of constructing adjacent graph, we exploit an anchor-based strategy to learn the adjacent graph, and design an efficient strategy to accelerate the optimization. Similar to the previous works [46]–[51], the anchor-based strategy mainly consists of two steps: 1) anchors generation; 2) adjacent graph construction.

Anchors Generation: in large-scale clustering, the anchors are usually generated by k-means method or random sampling. We employ the random selection method to generate m anchors ($m \ll n$), since its computational complexity is $\mathcal{O}(1)$. Let $\mathbf{U} = [u_1, u_2, ...u_m]^T \in \mathbb{R}^{m \times d}$ denotes the anchor set, and u_{ϕ_i} represents the set of k-nearest anchors for x_i.

Adjacent Graph Learning: let $\mathbf{Z} \in \mathbb{R}^{n \times m}$ be the adjacent graph, where z_{ij} denotes the similarity between x_i and u_j. $\mathbf{Z}$ is constructed by k-nearest neighbors method. Traditional approaches usually adopt kernel-based strategy to assign neighbors [52], which always bring extra hyperparameters. Inspired by [17] and [53], we design an effective neighbor assignment strategy. The nearest anchors assignment of x_i can be seen as solving the objective function:

$$\min_{z_i^T \mathbf{1} = 1, z_{ij} \geq 0} \sum_{j=1}^{m} \underbrace{\|x_i - u_j\|_2^2 z_{ij}}_{\mathcal{J}_1} + \underbrace{\alpha \|\widetilde{x}_i - u_j\|_2^2 z_{ij}}_{\mathcal{J}_2} + \underbrace{\gamma z_{ij}^2}_{\mathcal{J}_3}, \quad (6)$$

where z_i^T is the i-th row of $\mathbf{Z}$, and $\|x_i - u_j\|_2^2$ is the square of Euclidean distance between pixel x_i and anchor u_j. $\widetilde{x}_i$ denotes the average of k spatial-spectral neighbors about x_i, which can be calculated in section III-B. Similarly, $\|\widetilde{x}_i - u_j\|_2^2$ is the distance between $\widetilde{x}_i$ and u_j. It can be seen that the objective function (6) contains three terms: $\mathcal{J}_1$, $\mathcal{J}_2$ and $\mathcal{J}_3$. $\mathcal{J}_1$ utilizes the spectral features to learn $\mathbf{Z}$; $\mathcal{J}_2$ incorporates the local spatial-spectral feature to learn $\mathbf{Z}$, where α controls the tradeoff between $\mathcal{J}_1$ and $\mathcal{J}_2$; $\mathcal{J}_3$ is the regularization term, which is to prevent the trivial solution of Eq. (6), where γ is the regularization parameter.

Let us define

$$e_{ij} = \|x_i - u_j\|_2^2, \quad (7)$$

$$\widetilde{e}_{ij} = \|\widetilde{x}_i - u_j\|_2^2, \quad (8)$$

where $e_i \in \mathbb{R}^{m \times 1}$ denotes a vector with the j-th element as e_{ij}. Similarly, $\widetilde{e}_i \in \mathbb{R}^{m \times 1}$ is a vector with the j-th element as $\widetilde{e}_{ij}$. Combining Eq. (7) and Eq. (8), we obtain

$$E_{ij} = e_{ij} + \alpha \widetilde{e}_{ij}, \quad (9)$$

and denote $E_i \in \mathbb{R}^{m \times 1}$ as a vector with the j-th element as E_{ij}, then the objective function (6) can be expressed as

$$\min_{z_i^T \mathbf{1} = 1, z_{ij} \geq 0} \frac{1}{2} \left\| z_i + \frac{1}{2\gamma} E_i \right\|_2^2. \quad (10)$$

The Lagrangian function of Eq. (10) is

$$\mathcal{L}(z_i, \eta, \beta_i) = \frac{1}{2} \left\| z_i + \frac{E_i}{2\gamma} \right\|_2^2 - \eta(z_i^T \mathbf{1} - 1) - \beta_i^T z_i, \quad (11)$$

where η and $\beta_i^T \geq 0$ are the Lagrangian multipliers. To achieve the optimal z_i^*, it should satisfy that the derivative of Eq. (11) with respect to z_i^* is equal to zero, i.e.,

$$z_i^* + \frac{E_i}{2\gamma} - \eta \mathbf{1} - \beta_i = \mathbf{0}. \quad (12)$$

Then the j-th element of z_i^* is

$$z_{ij}^* + \frac{E_{ij}}{2\gamma} - \eta \mathbf{1} - \beta_{ij} = 0. \quad (13)$$

By the KKT condition and constraints, to achieve the optimal solution z_i^* to the function (6) that has exactly k nonzero values, the η and γ are

$$\eta = \frac{1}{k} + \frac{1}{2k\gamma} \sum_{j=1}^{k} E_{ij}, \qquad (14)$$

$$\gamma = \frac{k}{2} E_{i,k+1} - \frac{1}{2} \sum_{j=1}^{k} E_{ij}. \qquad (15)$$

Therefore, the optimal solution z_{ij}^* is as follows:

$$z_{ij}^* = \frac{E_{i,k+1} - E_{ij}}{kE_{i,k+1} - \sum_{j'=1}^{k} E_{ij'}}. \qquad (16)$$

For the detailed deviation, please refer to [53]. The calculation of z_{ij}^* only involves the basic operations: addition, subtraction, multiplication and division, which ensures the efficiency of the proposed method. In addition, there are no hyperparameters which may affect the stability of model in Eq. (16). Then, we propose to normalize $\mathbf{Z}$ to be a doubly stochastic matrix.

Doubly Stochastic Graph Construction: after getting the adjacent graph $\mathbf{Z}$, the normalized adjacent graph $\mathbf{S}$ can be computed as:

$$\mathbf{S} = \mathbf{Z}\boldsymbol{\Lambda}^{-1}\mathbf{Z}^T, \qquad (17)$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix whose j-th element is represented as $\Lambda_{jj} = \sum_{i=1}^{n} z_{ij}$, and $\boldsymbol{\Lambda} \in \mathbb{R}^{m \times m}$. Intuitively, the element s_{ij} of matrix $\mathbf{S}$ is expressed as $s_{ij} = z_i^T \Lambda^{-1} z_j$ that satisfies $s_{ij} = s_{ji}$. Moreover, with $z_{ij} \geq 0$ and $z_i^T \mathbf{1} = 1$, it can be proved that matrix $\mathbf{S}$ is positive semidefinite and doubly stochastic [31]. The above property of $\mathbf{S}$ is crucial for reducing the computation cost of optimization, and the details will be given later.

D. Optimization of Anchor Graph

The complete graph-based clustering methods model the similarity between any pair of pixels with an adjacent matrix $\tilde{\mathbf{Z}}$, which group HSIs with a graph cut algorithm. Spectral clustering (SC) [3] is a typical example, which usually utilizes the normalized cut algorithm to do graph cutting, and its optimization problem can be expressed as

$$\min_{\mathbf{F}^T\mathbf{F}=\mathbf{I}} \mathrm{Tr}(\mathbf{F}^T\mathbf{L}\mathbf{F}), \qquad (18)$$

where $\mathbf{F} \in \mathbb{R}^{n \times c}$ denotes clustering indicators matrix. The Laplacian matrix is $\mathbf{L} = \mathbf{D} - \tilde{\mathbf{Z}}$, where the i-th element of the degree matrix $\mathbf{D}$ is denoted as $d_{ii} = \sum_{j=1}^{n} \tilde{z}_{ij}$. The optimal solution of $\mathbf{L}$ is obtained by enforcing eigenvalues decomposition on $\mathbf{L}$, which is comprised of eigenvectors corresponding to the smallest c eigenvalues of $\mathbf{L}$. Inspired by SC, we group HSI by using SC to the adjacent matrix $\mathbf{S}$ that encodes the relationships among anchors and pixels. Therefore, the Laplacian matrix is $\mathbf{L} = \mathbf{D} - \mathbf{S}$, where the i-th element of $\mathbf{D}$ is denoted as $d_{ii} = \sum_{j=1}^{n} s_{ij}$. However, this solution process requires $\mathcal{O}(n^2c)$, which is not applicable for

large HSI clustering. According to the natural characteristics of $\mathbf{S}$, we have

$$d_{ii} = \sum_{j=1}^{n} s_{ij} = \sum_{j=1}^{n} z_i^T \Lambda^{-1} z_j = z_i^T \sum_{j=1}^{n} \Lambda^{-1} z_j = z_i^T \mathbf{1} = 1. \qquad (19)$$

Therefore, the degree matrix $\mathbf{D}$ equals to the identity matrix $\mathbf{I}$ exactly. Moreover, it can be seen that the matrix $\mathbf{S}$ is automatically normalized. Considering that $\mathbf{L} = \mathbf{D} - \mathbf{S} = \mathbf{I} - \mathbf{S}$, Eq. (18) is equivalent to solving this problem:

$$\max_{\mathbf{F}^T\mathbf{F}=\mathbf{I}} \mathrm{Tr}(\mathbf{F}^T\mathbf{S}\mathbf{F}). \qquad (20)$$

Note that $\mathbf{S}$ can also be denoted as

$$\mathbf{S} = \mathbf{A}\mathbf{A}^T, \qquad (21)$$

where $\mathbf{A} = \mathbf{Z}\Lambda^{-1/2}$, then we apply SVD on $\mathbf{A}$ as follows:

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T, \qquad (22)$$

where $\mathbf{U} \in \mathbb{R}^{n \times n}$, $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times m}$ and $\mathbf{V} \in \mathbb{R}^{m \times m}$. With respect to the property of SVD, there are $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ and $\mathbf{V}^T\mathbf{V} = \mathbf{I}$. So $\mathbf{S} = \mathbf{A}\mathbf{A}^T = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T\mathbf{V}\boldsymbol{\Sigma}^T\mathbf{U}^T = \mathbf{U}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T\mathbf{U}^T = \mathbf{U}\boldsymbol{\Sigma}^2\mathbf{U}^T$. It indicates that the column vectors of $\mathbf{U}$ are the eigenvectors of $\mathbf{S}$. Instead of using eigenvalue decomposition on $\mathbf{S}$, we adopt SVD on $\mathbf{A}$ to get the relaxed continuous solution of $\mathbf{F}$, which only costs $\mathcal{O}(nmc + m^2c)$. Then, we perform k-means method on $\mathbf{F}$ to get the discrete solution.

Algorithm 1 SSCAG

Input: HSI data $\mathbf{X} \in \mathbb{R}^{n \times d}$, class number c, window size w, anchor number m, neighbor number k, the balance parameter α and $\gamma_0 = 0.2$.

1: **for** $i = 1$ **to** n **do**
2:    seek the neighbors of x_i according to Eq. (1);
3:    **for** $k = 1$ **to** $w^2 - 1$ **do**
4:       calculate the weights: $v_k = exp\{-\gamma_0\|x_i - x_{ik}\|^2\}$;
5:    **end for**
6:    receive the filtered point $\hat{x}_i$ by Eq. (2);
7: **end for**
8: Obtain $\hat{\mathbf{X}} = [\hat{x}_1, \hat{x}_2, ..., \hat{x}_i, ..., \hat{x}_n]$;
9: **for** $i = 1$ **to** n **do**
10:   calculate the spatial-spectral distance d_{ss} by Eq. (3);
11:   find k spatial-spectral nearest neighbors for x_i, computing the average of neighbors and saving as $\tilde{x}_i$;
12: **end for**
13: Generate m anchors by random sampling, obtaining anchor matrix $\mathbf{U} \in \mathbb{R}^{d \times m}$;
14: Find k nearest neighbors of x_i in $\mathbf{U}$, saving the index set ϕ_i;
15: Define function $\mathcal{L}(z_i) = \|x_i - u_{\phi_i}\|_2^2 z_i + \alpha\|\tilde{x}_i - u_{\phi_i}\|_2^2 z_i + \gamma\|z_i\|_2^2$;
16: Obtain γ and the adjacent graph $\mathbf{Z}$ by Eq. (16) and Eq. (17);
17: Design the normalized adjacent graph $\mathbf{S} = \mathbf{Z}\boldsymbol{\Lambda}^{-1}\mathbf{Z}^T$, $\Lambda_{jj} = \sum_{i=1}^{n} z_{ij}$;
18: Obtain $\mathbf{F}$ by applying SVD on matrix $\mathbf{A}$, and $\mathbf{A} = \mathbf{Z}\Lambda^{-1/2}$;
19: Utilizing k-means on $\mathbf{F}$ for final clustering results;

Output: Label matrix about c classes

The detail steps of SSCAG are described in **Algorithm 1**. The computational complexity of **Algorithm 1** mainly comes from four parts: 1) the WMF-based spatial preprocessing takes $\mathcal{O}(ndw^2)$; 2) obtaining m anchors by random selection costs $\mathcal{O}(1)$, and the spatial-spectral combined distance d_{ss} is calculated with $\mathcal{O}(n^2w^2)$; 3) the cost of constructing $\mathbf{Z}$ is

$\mathcal{O}(ndm)$, and achieving the relaxed continuous eigenvalue of $\mathbf{F}$ needs $\mathcal{O}(nmc + m^2c)$; 4) the k-means which is used to get the final clustering results takes $\mathcal{O}(nmc\tau)$, where τ is the number of iterations. Note that $m \ll n$, $c \ll d$, w and τ are small. Therefore, the final computational complexity of SSCAG is $\mathcal{O}(n^2w^2 + ndm)$.

IV. EXPERIMENTS

In this section, to demonstrate the effectiveness of the proposed SSCAG method, the verification experiments are conducted on public hyperspectral datasets. Some excellent algorithms are adopted as competitors. Finally, the experimental results and corresponding analysis are provided.
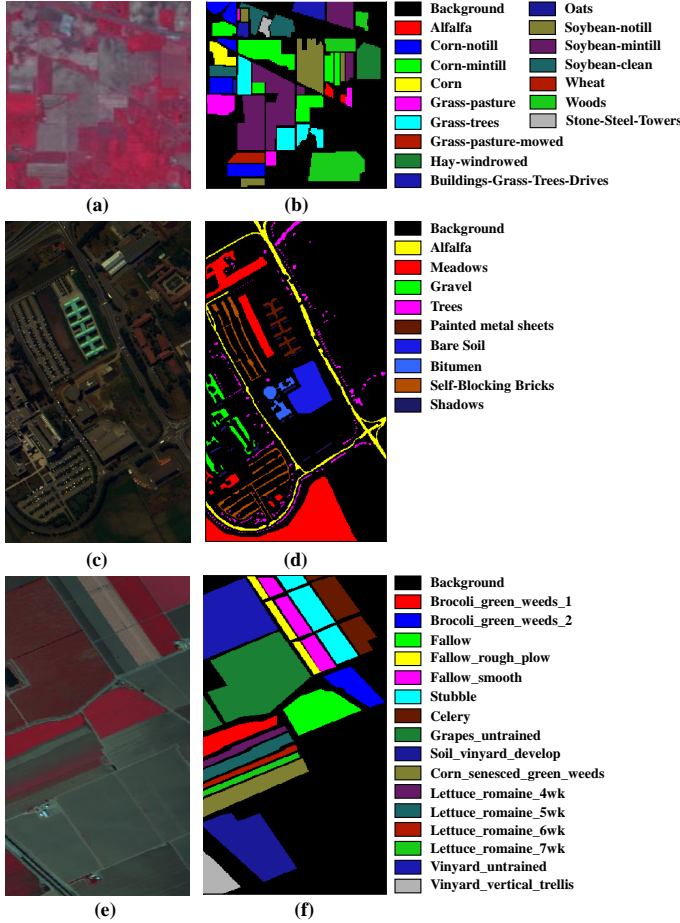
Fig. 2. The RGB false color images and ground truth maps of three HSI datasets. (a) Indian Pines image of bands 50, 27 and 17. (c) Pavia University image of bands 60, 30 and 2. (e) Salinas image of bands 70, 27 and 17. (b), (d) and (f) are corresponding ground truth maps respectively.

A. Dataset Description

The detailed information of three hyperspectral datasets are introduced as follows. Their false color images and ground truths are displayed in Fig. 2.

1) Indian Pines: The hyperspectral image was acquired via AVIRIS device (spatial resolution is $20\ m$) over the northwestern Indiana in 1992, as shown in Fig. 2 (a). The spatial size of this area is 145×145 pixels. Each pixel contains 220 spectral bands ranging from 0.4 to 2.5 μm. Due to noise and water absorption phenomena, 20 channels are removed, leaving 200 bands to be used for experiments. The dataset contains 16 land cover types, and the distribution of samples is clear in Fig. 2 (b).

2) Pavia University: The scene was captured by the German ROSIS sensor (spatial resolution is $1.3\ m$) over the Pavia in 2002, as displayed in Fig. 2 (c). The HSI data contains 610×340 pixels and 115 spectral bands, and the spectral range is 0.43 to 0.86 μm. With discarding 12 noise and water absorption bands, 103 bands are left for classification task. This dataset includes 9 ground-truth classes, and the distribution of them is shown in Fig. 2 (d).

3) Salinas: Like the Indian Pines image, the Salinas data was also taken by AVIRIS sensor over the Salinas Valley in 1998, as exhibited in Fig. 2 (e). Unlike Indian Pines, its spatial resolution is $3.7\ m$. Similarly, the 204 bands are left after excluding 20 water absorption bands. The size of Salinas image is 512×217. All pixels in HSI data are divided into 16 categories whose distribution is illustrated in Fig. 2 (f).

B. Experimental Setup

1) Comparison Algorithms: To validate the superiority of the proposed SSCAG, several HSI clustering methods are considered as follows. *Basic clustering methods*: FCM_S1 [17] and NL-SSLR [22]. FCM commonly use Euclidean distance as similarity measure to achieve the clusters. It belongs to soft clustering, which uses the membership to identify the relationship between samples and each cluster. FCM_S1 has enhanced robustness of the original clustering algorithms by exploiting spatial contextual information. NL-SSLR is a scalable subspace clustering method based on nonlocal means regularization and reweighted sparse constraint, which explores the intrinsic adjacency between pixels and mines the complex structural information in a computationally affordable level, to make it feasible for large HSIs. *Graph-based clustering methods*: SCC [38], JSCC [38], and USPEC [34]. SCC utilizes the structured dictionary and the recovery residual minimization criterion to build the sparse coding-based clustering model, to relieve computational obstacles. Based on SCC, JSCC considers the spatial information and homogeneity of regular neighborhood in the model, which uses super-pixel neighborhood and enhances the smoothness of clustering. USPEC is a scalable and robust clustering model, which contains two stages: constructing a set of base clusters by ultra-scalable spectral clustering and combining multiple clusters into the consensus clustering with more robustness. *AG-based clustering methods*: SGCNR [33], FSCAG [42] and FSCS [43]. The first method SGCNR only adopts the spectral feature, while other methods FSCAG and FSCS consider spatial and spectral information. SGCNR builds an adjacent graph based on AG and the nonnegative relaxation conditions, which directly obtains the clustering indicators. FSCAG considers spectral correlation and spatial neighborhood properties of the HSI data to construct anchor graph. FSCS uses the spatial nearest pixels to reconstruct the center point within a window, which combines local spatial structure with spectral information.
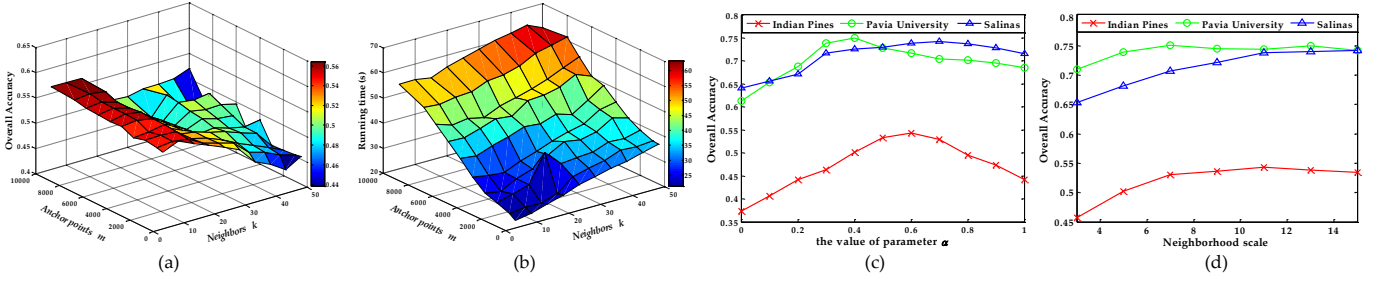
Fig. 3. Parameter analysis in different situations. (a) OAs of SSCAG on Indian Pines dataset under different m and k. (b) Running time of SSCAG on Indian Pines dataset under different m and k. (c) OAs of SSCAG on three datasets with different values of α. (d) Effects of the scale w on datasets.

TABLE I
THE EFFECTS OF THE MULTISCALE SPATIAL WMF. THE OPTIMAL VALUE OF EACH ROW IS HIGHLIGHTED IN BOLD.

w	Indian				PaviaU				Salinas			
	OA	AA	Kappa	Times	OA	AA	Kappa	Times	OA	AA	Kappa	Times
1*1	0.4708	0.4515	0.4027	4.9	0.7011	0.6893	0.6154	34.7	0.6769	0.7496	0.6415	23.7
3*3	0.4956	0.4904	0.4311	6.5	0.7147	0.6913	0.6301	35.5	0.7050	0.7893	0.6746	26.9
5*5	0.5004	0.4821	0.4361	7.2	0.7225	0.6954	0.6387	36.2	0.7159	0.7787	0.6841	30.4
7*7	0.5198	0.5586	0.4571	13.5	**0.7484**	**0.7443**	**0.6539**	61.7	0.7171	0.7798	0.6855	33.8
9*9	0.5327	0.5425	0.4676	13.6	0.7396	0.7201	0.6532	83.6	0.7180	0.7827	0.6865	37.1
11*11	**0.5421**	**0.5681**	**0.4826**	19.3	0.7340	0.7187	0.6527	100.4	0.7198	0.7808	0.6888	38.3
13*13	0.5405	0.5515	0.4717	19.7	0.7323	0.7012	0.6488	122.8	0.7243	0.7888	0.6944	39.6
15*15	0.5371	0.5359	0.4762	19.8	0.7305	0.7002	0.6452	129.1	**0.7351**	**0.7987**	**0.6968**	40.3
11*11, 7*7,15*15	**0.5427**	**0.5684**	**0.4818**	21.3	**0.7504**	**0.7493**	**0.6747**	130.8	**0.7419**	**0.8067**	**0.6993**	41.8

2) Evaluation Indices: Three quantitative metrics are used to evaluate clustering performance, including overall accuracy (OA), average accuracy (AA) and Kappa coefficient. OA represents the proportion of correctly classified pixels in HSI. AA denotes the average of the classification accuracy of each category. The value of OA and AA ranges from 0 to 1. Here, the higher accuracy value is, the better clustering performance is. Kappa coefficient is a metric combined commission error and omission error, which can evaluate the overall consistency. Its value range is [0,1]. The larger Kappa values indicate better consistency.

C. Experimental Results and Analyses

1) Parameter Settings: The whole pixels of the Indian Pines, Pavia University and Salinas were used as testing data. Set the number of clusters equal to the number of the ground truth labels in each dataset. For basic clustering methods and graph-based methods, all the parameters are manually adjusted to the optimum. Every algorithm is repeated 20 times to avoid the bias. For AG-based methods, the optimal parameter settings of SGCNR is followed as [33]. For FSCAG, the anchors number m, neighbors number k and balance parameter α are $\{m = 500, k = 5, \alpha = 0.6\}$ for Indian Pines, $\{m = 1000, k = 5, \alpha = 0.3\}$ for Pavia University, $\{m = 1000, k = 5, \alpha = 0.8\}$ for Salinas. For FSCS, the window size w, the anchors number m, neighbors number k and balance parameter α for three datasets are $\{w = 11, m = 500, k = 5, \alpha = 0.2\}$, $\{w = 7, m = 1000, k = 5, \alpha = 0.1\}$ and $\{w = 15, m = 1000, k = 5, \alpha = 0.2\}$. For the proposed SSCAG, we choose three different window scales, i.e., $7 \times 7, 11 \times 11$ and 15×15, to preprocess every dataset. The corresponding parameters are set as $\{m = 500, k = 5, \alpha = 0.6\}$, $\{m = 1000, k = 5, \alpha = 0.4\}$, $\{m = 1000, k = 5, \alpha = 0.7\}$ for three HSI datasets, respectively. The specific parameters analysis of proposed SSCAG is in the next part.

2) Parameter Analysis: The proposed SSCAG contains four parameters that need to be set and tuned, i.e. m, w, k and α. As mentioned in Section III, the computational complexity mainly depends on the size of parameters m and w. The parameter k mainly controls the sparsity of matrix $\mathbf{Z}$, which affects the computational complexity little. Concretely, we implement the sensitivity experiments of m and k on Indian Pines dataset.

According to Fig. 3 (a), the fluctuation range of OAs is small. We can see that the OAs of the proposed SSCAG are relatively stable with changing the parameters m and k. As shown in Fig. 3 (b), the running time is mainly related to m, while the influence of k is little. Hence, to speed up SSCAG algorithm and ensure good accuracy, we choose $\{m = 500, k = 5\}$ for the Indians Pines dataset. Similarly, we select $\{m = 1000, k = 5\}$ for the other large datasets.

Fig. 3 (c) illustrates the OAs of three HSI datasets with varying α. α is an important parameter in SSCAG, which is used to adjust the tradeoff between spectral feature and spatial-spectral feature during the adjacent graph learning. When α
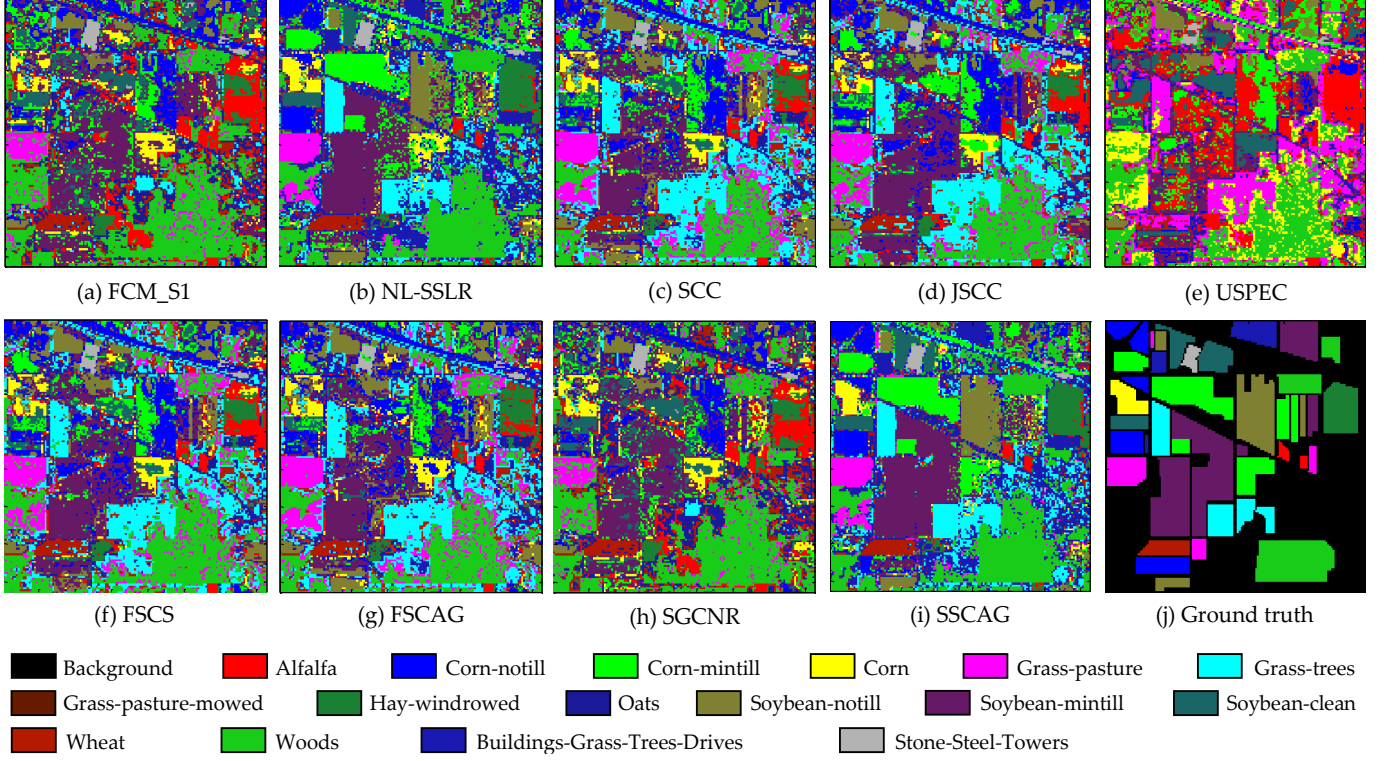
Fig. 4. The clustering maps obtained by different methods and the ground truth of Indian Pines.

TABLE II

QUANTITATIVE METRICS OF DIFFERENT COMPARISON METHODS AND THE PROPOSED SSCAG ON INDIAN PINES DATASET. THE OPTIMAL VALUE OF EACH ROW IS HIGHLIGHTED IN BOLD.

| Class	FCM_S1	NL-SSLR	SCC	JSCC	USPEC	FSCS	FSCAG	SGCNR	SSCAG
Alfalfa	0.1252	0.1646	0.0923	0.1324	0	0.0187	0.1143	0.1547	**0.1749**
Corn-notill	0.2334	0.4222	0.2779	0.5366	**0.5627**	0.3245	0.3397	0.2765	0.4097
Corn-mintill	0.3404	0.0050	0.4153	0.3041	0.1858	0.1467	0.4190	0.4131	**0.4431**
Corn	0.1466	0.2732	0.2615	0.3604	0.1160	0.2236	0.3051	0.2079	**0.5402**
Grass-pasture	**0.6099**	0.5582	0.4626	0.5777	0.3009	0.5983	0.5373	0.5928	0.5039
Grass-trees	0.6988	0.7132	0.5253	0.6585	0.8225	0.7837	0.7987	0.7124	**0.8395**
Grass-pasture-mowed	0.7985	0.2766	0.1274	0.5783	0	**0.8004**	0.7665	0.7548	0.7829
Hay-windrowed	0.8727	0.8787	0.8858	0.7227	0.8872	0.8282	0.8573	0.8825	**0.9497**
Oats	**0.4500**	0.1410	0.1556	0.2573	0.0286	0.4315	0.3535	0.4459	0.3462
Soybean-notill	0.2920	0.1699	0.2769	**0.4857**	0.3544	0.2925	0.3501	0.2735	0.4543
Soybean-mintill	0.4792	**0.7755**	0.3937	0.5880	0.4890	0.5605	0.4038	0.3934	0.5033
Soybean-clean	0.2124	0.2869	0.1276	0.1912	0.2561	0.1192	0.2590	0.2753	**0.3531**
Wheat	0.8244	0.8094	0.8600	0.6141	0.3350	0.8356	0.8438	0.8261	**0.8617**
Woods	0.4495	**0.9268**	0.3882	0.7012	0.8501	0.4949	0.7863	0.4877	0.8503
Buildings-Grass-Trees-Drives	0.3054	0.3476	0.2897	**0.3558**	0.3405	0.0736	0.1716	0.2881	0.2877
Stone-Steel-Towers	0.7845	0.8919	0.6913	0.7773	**0.9121**	0.7957	0.8042	0.8182	0.8867
OA	0.4128	0.5112	0.3864	0.4873	0.3778	0.4409	0.4708	0.4531	**0.5427**
AA	0.4469	0.4857	0.3894	0.4901	0.4027	0.4642	0.5069	0.4877	**0.5684**
Kappa	0.3512	0.4443	0.3094	0.4257	0.3141	0.3816	0.4019	0.3931	**0.4818**
Time(s)	163.7	12906	45.9	43.6	133.8	23.6	43.1	42.3	**21.3**
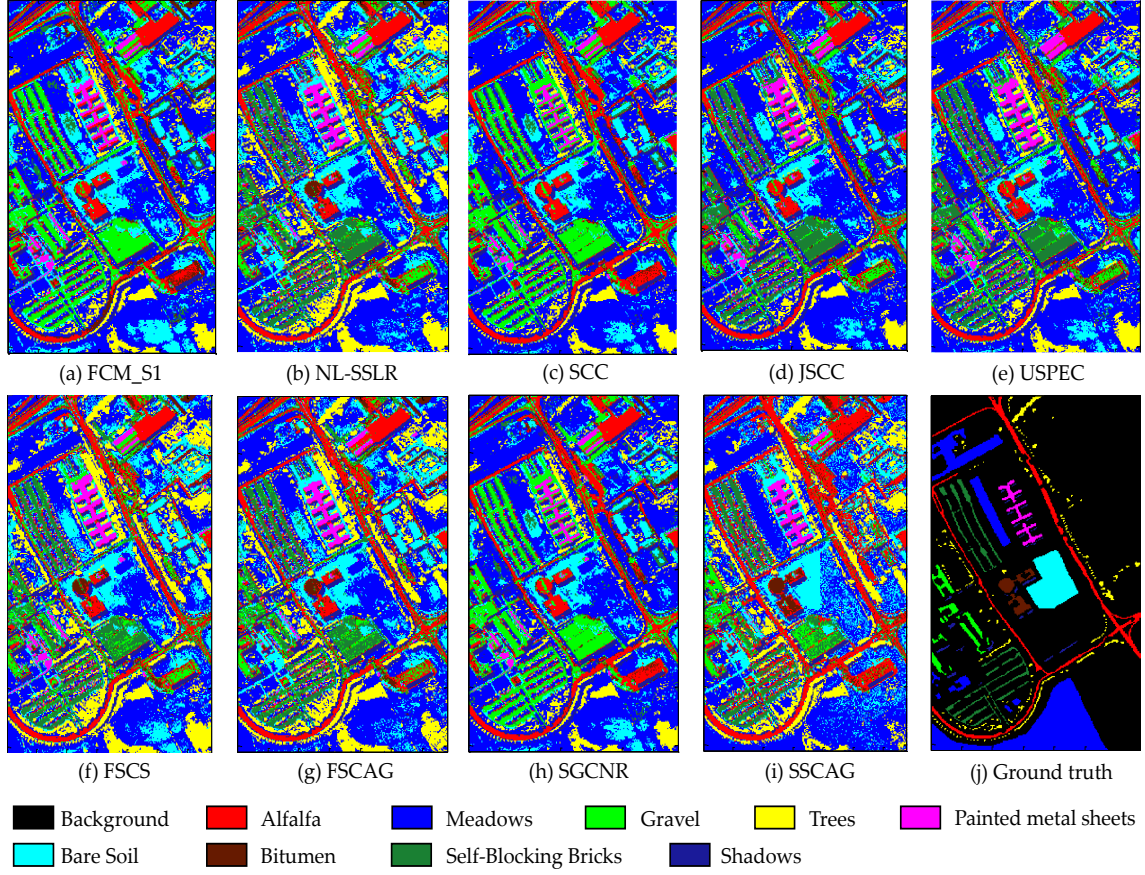
Fig. 5. Clustering maps and ground truth of Pavia University.

is set to 0, it reveals the classification accuracy without the second term in the objective function. We can see that the OA of the proposed SSCAG is relatively the worst value with $\alpha = 0$, while the results are all better when α takes other values. From Fig. 3 (c), the result demonstrates that the optimal value of α is 0.6, 0.4 and 0.7 for Indian Pines, Pavia University and Salinas datasets, respectively.

Fig. 3 (d) and Table I exhibit the OAs of three datasets with seven different neighborhood scales. The parameter w is the window scale of WMF that affects the result of preprocessing. Here, we adopt eight different scales (i.e., 1×1, 3×3, 5×5, ..., 15×15) to investigate their impact. As shown in Fig. 3 (d) and Table I, the best OAs for Indian Pines, Pavia University and Salinas are obtained, when the window scale is 11, 7 and 15 respectively. This indicates that different HSI datasets involve different spatial structures, even the same dataset may include small and large homogeneous regions simultaneously. So it is difficult to determine a best scale in advance. Therefore, we adopt three scales (7×7, 11×11 and 15×15) in the multiscale WMF framework, which not only exempts from the scale selection but also provides multiple views of a local homogeneous region. By doing this, it obtains the complementary information and improves the final clustering performance.

3) Clustering Performance Comparison: In this section, experiments are conducted on three HSI datasets. The parameter

analysis of the proposed approach is also discussed.

a) Performance on Indian Pines dataset: The quantitative results of the methods are given in Table II. It can be revealed that SSCAG obtains better results than competitors from the perspective of OA, AA and Kappa, and gains the higher accuracy in most classes. Moreover, the AG-based methods (SGCNR, FSCAG, FSCS and SSCAG) outperform the centroid-based clustering method (FCM_S1) and some graph-based clustering methods (SCC and USPEC) as shown Table II. Among these competitors, SCC and USPEC only focus on the spectral information. The other algorithms (FCM_S1, NL-SSLR, JSCC, FSCAG, FSCS and SSCAG) combine spectral feature and spatial structure to enforce HSI clustering. Corresponding to the Table II, the quantitative accuracy of SCC and USPEC is lower (i.e., OAs reduce by at least 5%) than the accuracy of other algorithms. This reveals that incorporating spatial information is beneficial to HSI clustering. Furthermore, compared with FSCS, the proposed SSCAG improves over 10% in terms of OA. Because the WMF with multiscale windows outperforms the optimal results in single scale, which reaches higher OAs by complementary information in different scales.

Then the final clustering maps obtained with each algorithm are also visualized in Fig. 4. The proposed SSCAG obtains much smoother clustering maps than other methods. Especially, SSCAG has a superior performance than SGCNR

TABLE III
QUANTITATIVE METRICS OF DIFFERENT COMPARISON METHODS AND THE PROPOSED SSCAG ON PAVIA UNIVERSITY DATASET. THE OPTIMAL VALUE
OF EACH ROW IS HIGHLIGHTED IN BOLD.

Class	FCM_S1	NL-SSLR	SCC	JSCC	USPEC	FSCS	FSCAG	SGCNR	SSCAG
Asphalt	0.7525	0.8315	0.7064	0.6169	0.6890	0.8298	0.8311	0.6289	**0.8378**
Meadows	0.6395	0.8104	0.6233	0.5954	0.8197	0.8090	0.8141	0.7521	**0.8245**
Gravel	**0.9877**	0.5410	0.8553	0.8774	0.6883	0.5389	0.5331	0.2857	0.5440
Trees	0.6220	0.7730	0.8956	**0.9467**	0.7856	0.7313	0.7116	0.7892	0.7908
Painted metal sheets	0.5438	0.5222	0.8163	0.9095	**0.9987**	0.9548	0.9956	0.9728	0.9738
Bare Soil	0.2755	0.8616	0.6473	0.5510	0.4461	0.4007	0.3570	0.3897	**0.8957**
Bitumen	0	0.5783	0.5263	0.5068	0.5215	0.5774	0.5708	0.2915	**0.5813**
Self-Blocking Bricks	0.0039	0.6876	0.5078	0.6028	0.3002	0.6648	0.6996	**0.8621**	0.7451
Shadows	0.9975	0.9989	0.2063	0.3887	0.2897	**1.0000**	**1.0000**	0.9980	**1.0000**
OA	0.5341	0.7340	0.6358	0.6579	0.6431	0.7089	0.7306	0.6275	**0.7504**
AA	0.5358	0.7379	0.6427	0.6661	0.6154	0.7230	0.7248	0.6633	**0.7992**
Kappa	0.4334	0.6539	0.5987	0.6005	0.6121	0.6324	0.6497	0.5751	**0.6747**
Time(s)	979.0	68326	2945.7	1350.4	184.03	191	137.5	495.4	**130.8**

(a) FCM_S1   (b) NL-SSLR   (c) SCC   (d) JSCC   (e) USPEC

(f) FSCS   (g) FSCAG   (h) SGCNR   (i) SSCAG   (j) Ground truth

Background   Brocoli_green_weeds_1   Brocoli_green_weeds_2   Fallow   Fallow_rough_plow   Fallow_smooth

Stubble   Celery   Grapes_untrained   Soil_vinyard_develop   Corn_senesced_green_weeds   Lettuce_romaine_4wk

Lettuce_romaine_5wk   Lettuce_romaine_6wk   Lettuce_romaine_7wk   Vinyard_untrained   Vinyard_vertical_trellis
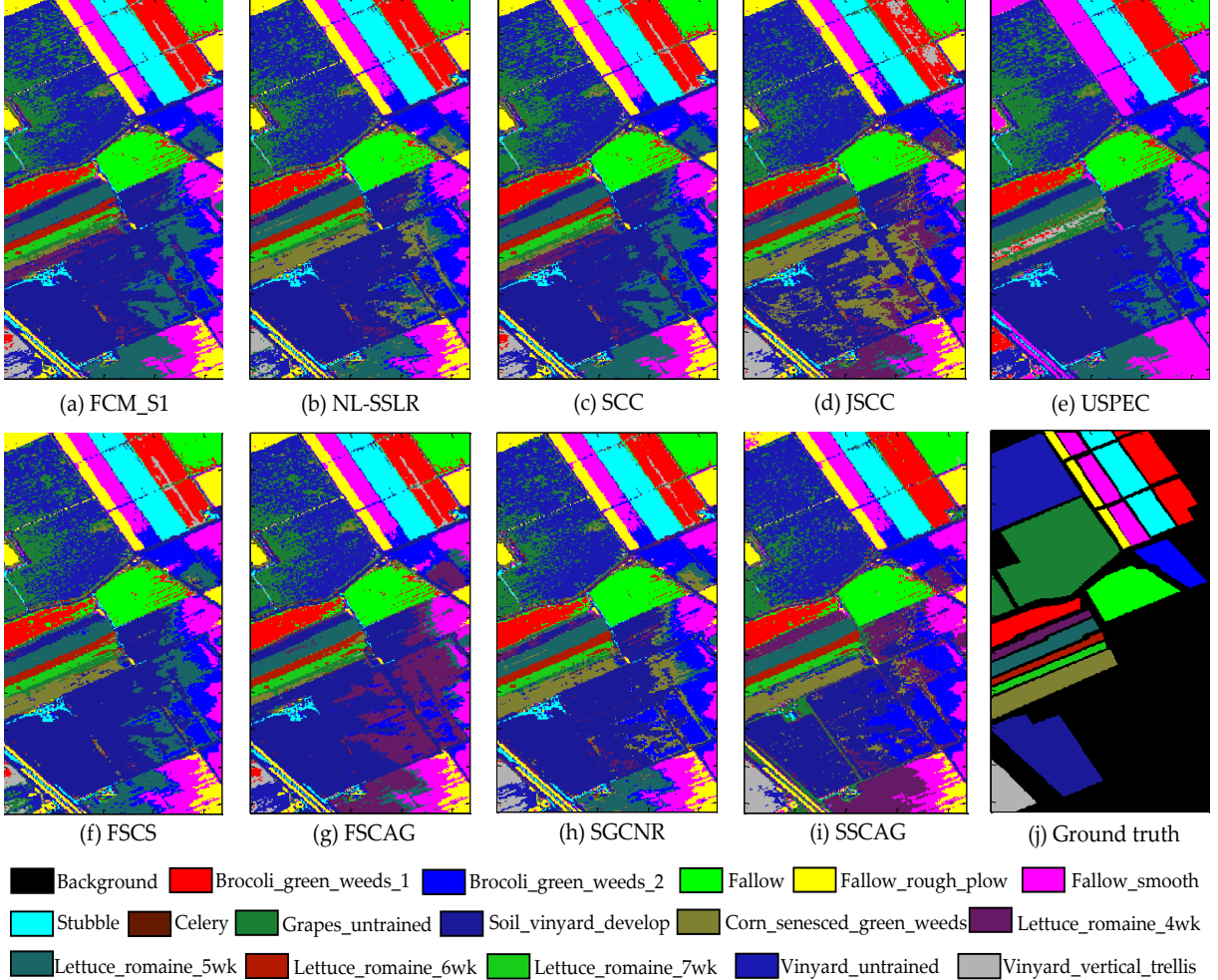
Fig. 6. The clustering maps obtained by different methods and the ground truth of Salinas.

TABLE IV
QUANTITATIVE METRICS OF DIFFERENT COMPARISON METHODS AND THE PROPOSED SSCAG ON SALINAS DATASET. THE OPTIMAL VALUE OF EACH
ROW IS HIGHLIGHTED IN BOLD.

Class	FCM_S1	NL-SSLR	SCC	JSCC	USPEC	FSCS	FSCAG	SGCNR	SSCAG
Brocoli green weeds 1	0.9826	**1.0000**	0.7979	0.3987	0	0.9300	0.9474	0.9501	0.9204
Brocoli green weeds 2	0.4367	0.9413	0.5694	**0.9459**	0.6465	0.4610	0.3695	0.5203	0.5678
Fallow	0.7071	0.6715	0.5434	0.3774	0.8725	0.7376	0.7125	0.7825	**0.8820**
Fallow rough plow	0.9819	0.7692	0.9589	**0.9988**	0	0.9507	0.9579	0.9241	0.9087
Fallow smooth	0.8972	0.8445	0.8445	0.6857	0.5988	0.8578	0.8496	0.7088	**0.8985**
Stubble	0.9545	0.9989	0.9609	0.9683	**0.9995**	0.9518	0.9587	0.9590	0.9790
Celery	0.9055	0.9484	0.9876	**0.9968**	0.8648	0.8860	0.8847	0.9388	0.9117
Grapes untrained	0.4133	0.5992	**0.7889**	0.7757	0.5852	0.3817	0.4486	0.3415	0.4626
Soil vinyard develop	**0.9164**	0.8320	0.8317	0.7787	0.8725	0.8538	0.8710	0.7917	0.8336
Corn senesced green weeds	0.2739	**0.8536**	0.4099	0.2286	0.7805	0.5946	0.3239	0.5023	0.6199
Lettuce romaine 4wk	0	0.6025	0.5478	0.7584	0.5090	0.2314	0.1083	0.6526	**0.8968**
Lettuce romaine 5wk	0.9889	0.9465	0.0955	0.5135	0.8461	0.8954	0.9805	**0.9953**	0.9525
Lettuce romaine 6wk	0.9782	0.1640	0.9624	0.9323	0.4848	**0.9847**	0.7888	0.9536	0.9480
Lettuce romaine 7wk	0.8457	0.3960	0.8837	0.8284	0	0.8944	0.8557	0.8501	**0.9042**
Vinyard untrained	0.6295	0.0029	0.4229	0.5164	0	0.5631	0.5506	**0.6982**	0.6385
Vinyard vertical trellis	0.2561	**0.6926**	0.2926	0.5759	0.5198	0.4802	0.4161	0.5058	0.5825
OA	0.6754	0.7243	0.6687	0.6899	0.6907	0.7053	0.6569	0.7208	**0.7419**
AA	0.6980	0.7015	0.6811	0.7049	0.7167	0.7284	0.6890	0.7547	**0.8067**
Kappa	0.5907	0.6944	0.6372	0.6581	0.6608	0.6676	0.6216	0.6771	**0.6993**
Time(s)	967.4	96207	176.7	299.4	91.9	**25.7**	28.9	61.7	41.8

TABLE V
RUNNING TIME OF DIFFERENT METHODS ON THREE HSI DATASETS

Dataset	FCM_S1	NL-SSLR	SCC	JSCC	USPEC	FSCS	FSCAG	SGCNR	SSCAG
Indian Pines	127.7s	12906.0s	45.9s	43.6s	133.8s	23.6s	43.1s	42.3s	**21.3s**
Pavia University	979.0s	68326.1s	2945.7s	1350.4s	184.0s	191s	137.5s	495.4s	**130.8s**
Salinas	990.8s	96207.2s	176.7s	299.4s	91.9s	**25.7s**	28.9s	61.7s	41.8s

and FSCAG, which indicates that WMF-based spatial preprocessing significantly provides an accurate description of HSI homogeneous areas. Moreover, the preprocessing increases OAs by enhancing the similarity and consistency of adjacent pixels. It reflects that multiscale WMF can provide more precise information for distance metric to construct an effective spatial-spectral adjacency graph in clustering procedure.

b) Performance on Pavia University dataset: To observe the experimental results visually and quantitatively, Fig. 5 and Table III show the visual clustering maps and the corresponding classification accuracy, respectively. As shown in Table III, the AG-based clustering methods give superior results than the classic centroid-based method, especially for the Bitumen class. After incorporating the spatial characteristics into the anchor graph model, FSCAG, FSCS and SSCAG yield better accuracy in almost all classes compared with SGCNR, particularly for the Shadows class. This indicates that the combination of spatial and spectral information is beneficial to improve HSI clustering performance. The proposed SSCAG obtains

the best results for 5 classes, except for the Gravel class, which also achieves better accuracy for other 3 classes. The OA and kappa coefficient of SSCAG achieves to 0.7504 and 0.6747, respectively. It illustrates that SSCAG is an effective and superior algorithm for HSI clustering. As is shown in Fig. 5, SSCAG gains more smoother clustering map than other methods, which is consistent with the results in Table III.

c) Performance on Salinas dataset: To visualize the experimental results, the clustering maps are illustrated in Fig. 6. The quantitative accuracy of comparison algorithms and SSCAG are exhibited in Table IV. According to the visual results, it reveals that the pixels in the Grapes_untrained class and Vinyard_untrained class regions are most wrongly assigned. That is because the Grapes_untrained and Vinyard_untrained have highly similar feature information, and their spectral curves are very close. From Fig. 6 (g), the Lettuce_romaine_7wk class is discriminated well, and its accuracy reaches to 0.9042 in Table IV, which is higher than that of other methods. As shown in Table IV, the results indicate that the proposed SSCAG yields

the best classification accuracies, especially the OA and Kappa coefficient are 0.7419 and 0.6993 respectively. Furthermore, it can be observed that the Lettuce_romaine_4wk class is effectively assigned by SSCAG, in contrast, the recognition accuracy of other algorithms except SGCNR is close to 0. This phenomenon also justifies the effectiveness of SSCAG.

4) Running Time Comparison: We display the running time to quantitatively compare the complexity of all algorithms, as shown in Table V. All of the experimental results are conducted in MATLAB R2014a on a PC of Intel Core i7-9700F 3.00GHz CPU with 16 GB RAM. As shown in Table V, FSCS consumes the least running time on Salinas dataset, but its accuracy is generally not high comparing with SGCNR and SSCAG (see Table IV). From Table V, the proposed SSCAG method is faster than the other algorithms on the HSI datasets except Salinas dataset. However, according to Table II, III and IV, SSCAG provides best results than other methods, especially for AAs, which are 6%, 6% and 5% higher than the second best results on Indian Pines, Pavia University and Salinas datasets respectively. Therefore, the slight increase in running time is acceptable for the improvement in clustering performance.

V. CONCLUSION

In this paper, a novel spatial-spectral clustering with anchor graph (SSCAG) is proposed to efficiently cluster HSI data. Firstly, the multiscale spatial WMF is utilized to enhance the local pixel consistency and distinguish the structures across different classes. To facilitate the graph construction, we select representative anchors and exploit their relationship with the points, which reduces the computational complexity of model. Secondly, a new spatial-spectral distance metric is proposed to find representative neighbors by skillfully using spatial structure and spectral features, to integrate the intrinsic adjacency between pixels into the proposed model. Finally, the neighbors assignment strategy is used to learn the optimal adjacent graph adaptively, and SVD is performed on $\mathbf{S}$ to get the final clustering results. Extensive experiments on three public HSI datasets verify the efficiency and effectiveness of the proposed SSCAG. In the future, we mainly focus on designing a fusion strategy with multi-graph construction to better handle the tasks of HSI clustering.

REFERENCES

[1] L. Xu, A. Wong, F. Li, and D. A. Clausi, "Intrinsic representation of hyperspectral imagery for unsupervised feature extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 2, pp. 1118–1130, 2016.

[2] Q. Wang, Q. Li, and X. Li, "A fast neighborhood grouping method for hyperspectral band selection," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 6, pp. 5028–5039, 2021.

[3] H. Zhai, H. Zhang, P. Li, and L. Zhang, "Hyperspectral image clustering: Current achievements and future lines," *IEEE Geosci. Remote Sens. Mag.*, vol. 9, no. 4, pp. 35–67, 2021.

[4] C. Deng, X. Liu, C. Li, and D. Tao, "Active multi-kernel domain adaptation for hyperspectral image classification," *Pattern Recognit.*, vol. 77, pp. 306–315, 2018.

[5] Q. Wang, Z. Meng, and X. Li, "Locality adaptive discriminant analysis for spectral-spatial classification of hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 11, pp. 2077–2081, 2017.

[6] S. Mei, J. Hou, J. Chen, L.-P. Chau, and Q. Du, "Simultaneous spatial and spectral low-rank representation of hyperspectral images for classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 5, pp. 2872–2886, 2018.

[7] S. Susan, S. Shinde, and S. Batra, "Vegetation-specific hyperspectral band selection for binary-to-multiclass classification," in *Proc. IEEE 16th India Council Int. Conf*, 2019, pp. 1–4.

[8] F. Gan, S. Liang, P. Du, F. Dang, K. Tan, H. Su, and Z. Xue, "Chesre: A comprehensive public hyperspectral experimental site and data set for resources exploration," in *Proc. 7th Workshop Hyperspectral Image Signal Process., Evol. Remote Sens.*, 2015, pp. 1–4.

[9] Y. Wan, X. Hu, Y. Zhong, A. Ma, L. Wei, and L. Zhang, "Tailings reservoir disaster and environmental monitoring using the uav-ground hyperspectral joint observation and processing: A case of study in xinjiang, the belt and road," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 9713–9716.

[10] Q. Wang, Z. Yuan, Q. Du, and X. Li, "Getnet: A general end-to-end 2-d cnn framework for hyperspectral image change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 3–13, 2019.

[11] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Spectralspatial hyperspectral image segmentation using subspace multinomial logistic regression and markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 809–823, 2012.

[12] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Hyperspectral image classification via kernel sparse representation," in *Proc. IEEE Int. Conf. Image Process.*, 2011, pp. 1233–1236.

[13] X. Kang, S. Li, and J. A. Benediktsson, "Spectralspatial hyperspectral image classification with edge-preserving filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2666–2677, 2014.

[14] J. Li, P. R. Marpu, A. Plaza, J. M. Bioucas-Dias, and J. A. Benediktsson, "Generalized composite kernel framework for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 9, pp. 4816–4829, 2013.

[15] J. A. Hartigan and M. A. Wong, "Algorithm as 136: A k-means clustering algorithm," *J. R. Stat. Soc. Ser. C-Appl. Stat.*, vol. 28, no. 1, pp. 100–108, 1979.

[16] J. C. Bezdek, R. Ehrlich, and W. Full, "Fcm: The fuzzy c-means clustering algorithm," *Comput. Geosci.*, vol. 10, no. 2-3, pp. 191–203, 1984.

[17] 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. Part B-Cybern.*, vol. 34, no. 4, pp. 1907–1916, 2004.

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

[19] S. Vijendra, "Efficient clustering for high dimensional data: Subspace based clustering and density based clustering," *Information Technology Journal*, vol. 10, no. 6, 2011.

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

[21] Q. Yan, Y. Ding, J. J. Zhang, Y. Xia, and C. H. Zheng, "A discriminated similarity matrix construction based on sparse subspace clustering algorithm for hyperspectral imagery," *Cogn. Syst. Res.*, vol. 53, pp. 98–110, 2019.

[22] H. Zhai, H. Zhang, L. Zhang, and P. Li, "Nonlocal means regularized sketched reweighted sparse and low-rank subspace clustering for large hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 4164–4178, 2021.

[23] S. Huang, H. Zhang, and A. Piurica, "Subspace clustering for hyperspectral images via dictionary learning with adaptive regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–17, 2022.

[24] S. Huang, H. Zhang, and A. Piurica, "Hybrid-hypergraph regularized multiview subspace clustering for hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–16, 2022.

[25] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," in *Advances in neural information processing systems.* Cambridge, MA, USA: MIT Press, 2001, pp. 849–856.

[26] X. Li, M. Chen, F. Nie, and Q. Wang, "Locality adaptive discriminant analysis," in *Proc. 36th Int. Joint Conf. Artif. Intell.*, 2017, pp. 2201–2207.

[27] X. Li, M. Chen, F. Nie, and Q. Wang, "A multiview-based parameter free framework for group detection," in *Proc. 31st AAAI Conf. Artif. Intell.*, 2017, pp. 4147–4153.

[28] B. Rasti, D. Hong, R. Hang, P. Ghamisi, X. Kang, J. Chanussot, and J. A. Benediktsson, "Feature extraction for hyperspectral imagery: The

evolution from shallow to deep: Overview and toolbox," *IEEE Geosci. Remote Sens. Mag.*, vol. 8, no. 4, pp. 60–88, 2020.

[29] R. Hang and Q. Liu, "Dimensionality reduction of hyperspectral image using spatial regularized local graph discriminant embedding," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 11, no. 9, pp. 3262–3271, 2018.

[30] Q. Liu, Y. Sun, R. Hang, and H. Song, "Spatialspectral locality-constrained low-rank representation with semi-supervised hypergraph learning for hyperspectral image classification," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 10, no. 9, pp. 4171–4182, 2017.

[31] W. Liu, J. He, and S.-F. Chang, "Large graph construction for scalable semi-supervised learning," in *Proc. 27th Int. Conf. Mach. Learn.*, 2010, pp. 679–686.

[32] F. He, R. Wang, and W. Jia, "Fast semi-supervised learning with anchor graph for large hyperspectral images," *Pattern Recognit. Lett.*, vol. 130, pp. 319–326, 2020.

[33] 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, 2019.

[34] D. Huang, C.-D. Wang, J.-S. Wu, J.-H. Lai, and C.-K. Kwoh, "Ultra-scalable spectral clustering and ensemble clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 6, pp. 1212–1226, 2020.

[35] H. Huang, G. Shi, H. He, Y. Duan, and F. Luo, "Dimensionality reduction of hyperspectral imagery based on spatialspectral manifold learning," *IEEE Trans. Cybern.*, vol. 50, no. 6, pp. 2604–2616, 2020.

[36] Y. Zhou, J. Peng, and C. P. Chen, "Dimension reduction using spatial and spectral regularized local discriminant embedding for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 1082–1095, 2015.

[37] Z. Feng, S. Yang, S. Wang, and L. Jiao, "Discriminative spectralspatial margin-based semisupervised dimensionality reduction of hyperspectral data," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 2, pp. 224–228, 2015.

[38] H. Zhai, H. Zhang, L. Zhang, and P. Li, "Sparsity-based clustering for large hyperspectral remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 12, pp. 10 410–10 424, 2021.

[39] Z. Ren, L. Sun, Q. Zhai, and X. Liu, "Mineral mapping with hyperspectral image based on an improved k-means clustering algorithm," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 2989–2992.

[40] M. Ben Salem, K. Ettabaa, and M. S. Bouhlel, "Hyperspectral image feature selection for the fuzzy c-means spatial and spectral clustering," in *2016 International Image Processing, Applications and Systems*, 2016, pp. 1–5.

[41] B. Tu, X. Zhang, X. Kang, J. Wang, and J. A. Benediktsson, "Spatial density peak clustering for hyperspectral image classification with noisy labels," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 5085–5097, 2019.

[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, 2017.

[43] 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. 141 045–141 054, 2019.

[44] F. Luo, B. Du, L. Zhang, L. Zhang, and D. Tao, "Feature learning using spatial-spectral hypergraph discriminant analysis for hyperspectral image," *IEEE Trans. Cybern.*, vol. 49, no. 7, pp. 2406–2419, 2019.

[45] Q. Wang, X. He, and X. Li, "Locality and structure regularized low rank representation for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 911–923, 2019.

[46] C. Deng, R. Ji, W. Liu, D. Tao, and X. Gao, "Visual reranking through weakly supervised multi-graph learning," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2013, pp. 2600–2607.

[47] C. Deng, R. Ji, D. Tao, X. Gao, and X. Li, "Weakly supervised multi-graph learning for robust image reranking," *IEEE Trans. Multimedia*, vol. 16, no. 3, pp. 785–795, 2014.

[48] D. Cai and X. Chen, "Large scale spectral clustering via landmark-based sparse representation," *IEEE Trans. Cybern.*, vol. 45, no. 8, pp. 1669–1680, 2015.

[49] F. Nie, W. Zhu, and X. Li, "Unsupervised large graph embedding," in *Proc. 31st AAAI Conf. Artif. Intell.*, 2017, pp. 2422–2428.

[50] F. Nie, W. Zhu, and X. Li, "Unsupervised large graph embedding based on balanced and hierarchical k-means," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 4, pp. 2008–2019, 2022.

[51] X. Li, M. Chen, and Q. Wang, "Adaptive consistency propagation method for graph clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 4, pp. 797–802, 2020.

[52] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media, 2009.

[53] F. Nie, X. Wang, M. I. Jordan, and H. Huang, "The constrained laplacian rank algorithm for graph-based clustering," in *Proc. 30th AAAI Conf. Artif. Intell.*, 2016, pp. 1969–1976.

Qi Wang (M'15-SM'15) received the B.E. degree in automation and the Ph.D. degree in pattern recognition and intelligent systems from the University of Science and Technology of China, Hefei, China, in 2005 and 2010, respectively. He is currently a Professor with the School of Artificial Intelligence, Optics and Electronics (iOPEN), Northwestern Polytechnical University, Xi'an, China. His research interests include computer vision, machine learning, pattern recognition and remote sensing.
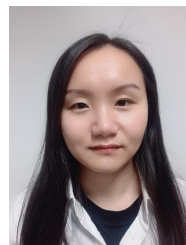
Yanling Miao received the B.E. degree in communication engineering and the M.S. degree in computer application technology from Henan Polytechnic University, Jiaozuo, China, in 2015 and 2019 respectively. She is currently pursuing the Ph.D. degree with the School of Computer Science and the School of Artificial Intelligence, Optics and Electronics (iOPEN), Northwestern Polytechnical University, Xi'an, China. Her research interests include hyperspectral image processing and computer vision.

Mulin Chen received the B.E. degree in software engineering and the Ph.D. degree in computer application technology from Northwestern Polytechnical University, Xi'an, China, in 2014 and 2019 respectively. He is currently a researcher with the School of Artificial Intelligence, Optics and Electronics (iOPEN), Northwestern Polytechnical University, Xi'an, China. His current research interests include computer vision and machine learning.

Yuan Yuan (M'05-SM'09) is currently a Full Professor with the School of Artificial Intelligence, Optics and Electronics (iOPEN), Northwestern Polytechnical University, Xi'an, China. She has authored or co-authored over 150 papers, including about 100 in reputable journals, such as the IEEE TRANSACTIONS AND PATTERN RECOGNITION, as well as the conference papers in CVPR, BMVC, ICIP, and ICASSP. Her current research interests include visual information processing and image/video content analysis.