World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:10, No:10, 2016

A Spatial Hypergraph Based Semi-Supervised Band Selection Method for Hyperspectral Imagery Semantic Interpretation

Akrem Sellami, Imed Riadh Farah

Abstract—Hyperspectral imagery (HSI) typically provides a wealth of information captured in a wide range of the electromagnetic spectrum for each pixel in the image. Hence, a pixel in HSI is a high-dimensional vector of intensities with a large spectral range and a high spectral resolution. Therefore, the semantic interpretation is a challenging task of HSI analysis. We focused in this paper on object classification as HSI semantic interpretation. However, HSI classification still faces some issues, among which are the following: The spatial variability of spectral signatures, the high number of spectral bands, and the high cost of true sample labeling. Therefore, the high number of spectral bands and the low number of training samples pose the problem of the curse of dimensionality. In order to resolve this problem, we propose to introduce the process of dimensionality reduction trying to improve the classification of HSI. The presented approach is a semi-supervised band selection method based on spatial hypergraph embedding model to represent higher order relationships with different weights of the spatial neighbors corresponding to the centroid of pixel. This semi-supervised band selection has been developed to select useful bands for object classification. The presented approach is evaluated on AVIRIS and ROSIS HSIs and compared to other dimensionality reduction methods. The experimental results demonstrate the efficacy of our approach compared to many existing dimensionality reduction methods for HSI classification.

Keywords—Hyperspectral image, spatial hypergraph, dimensionality reduction, semantic interpretation, band selection, feature extraction.

I. Introduction

IN the last decades, the technological evolution of optical sensors has provided remote sensing analysts with rich spatial, spectral, and temporal information. Nowadays, with the development of hyperspectral remote sensing imaging technology, we can capture HSI with hundreds of contiguous bands across the electromagnetic spectrum. The HSI, referred to as "data cube", is a kind of 3-D datum with two spatial dimensions and one spectral dimension. The use of HSI is becoming more and more widespread, such as target detection, changes detection, and object classification. A pixel in HSI is a high-dimensional vector of intensities with a large spectral range and a high spectral resolution. Therefore, the semantic interpretation is an essential task for HSI analysis extracting the semantics of sensor data. It means the understanding and the semantic interpretation of image contents just like humans do. Semantic image interpretation is a problem of visual perception, i.e. the perception of our environment by visual sensors. Visual perception is the act of sensing a scene (its visible objects, structures and events), of recognizing it and of describing it with symbols [1]-[5]. The results of semantic interpretation can be object categorization but also event, situation or scenario recognition. Semantic image interpretation results can be used for different purposes like making decision (diagnosis problem), like monitoring issues (visual surveillance, health care monitoring), and so on. We focused in this paper on object classification as HSI semantic interpretation. However, HSI classification still faces some issues, among which are the following [6], [7]: the spatial variability of spectral signatures, the high number of spectral bands, and the high cost of true sample labeling. Therefore, the high number of spectral bands and the low number of training samples pose the problem of the curse of dimensionality [8]. Hence, the Dimensionality Reduction is necessary to overcome the different issues to improve the semantic interpretation of HSI.

The remainder of this paper is organized as follows. Section II introduces the related work about dimensionality reduction of HSI. In Section III, we present the proposed approach. We first introduce the hypergraph embedding model as well as spatial hypergraph structure. Then, we go into details about semi-supervised band selection based on the proposed spatial hypergraph. In Section IV, we show experimental results of HSI classification using our approach with a comparative study with respect to well established approaches of the field. We conclude and suggest future work in the last section.

A. Sellami and I. R. Farah are with RIADI Laboratory, National School of Computer Science, Manouba, Tunisia 2010 and the ITI Department, Telecom Bretagne, Brest-Iroise 29238 France (e-mail: akrem.sellami@telecom-bretagne.eu, imed.farah@telecom-bretagne.eu).

II. Related Work

The high dimensionality of HSI not only increases the computational complexity but also may degrade classification accuracy [9]. Hence, the dimensionality reduction seeks to decrease computational complexity of input data while some desired intrinsic information of the

data is preserved to improve the classification accuracy. The problem of dimensionality reduction [10] can be defined as follows: Let us consider a data set represented by $n \times D$ matrix S consisting of n data vectors $i \in \{1, 2, ..., n\}$ with dimensionality D, and d be its intrinsic dimensionality d (where $d < D$, and often $d << D$). Intrinsic dimensionality means that the points in data set S are lying on or near a manifold with dimensionality d that is embedded in the D-dimensional space. Dimensionality reduction methods transform data set S with dimensionality D into a new data set Z with dimensionality d, while retaining as much as possible the geometry of the data. In general, neither the geometry of the data manifold, nor the intrinsic dimensionality d of the data set S are known. Therefore, dimensionality reduction is an ill-posed problem that can only be solved by assuming certain properties of the data (such as its intrinsic dimensionality). In the HSI context, the projection methods extract the d components, with $d < D$, such that

$$Z = T \times S \tag{1}$$

where Z is the reduced matrix, T is a linear transformation matrix and S is the different spectral signatures.

In HSI context, the different dimensionality reduction can be roughly categorized into Features Extraction (FE) and Band Selection (BS). The feature extraction method is based on data transformation [9], [11]. It reduces the dimensionality by transforming the original spectral bands from HSI into a new low-dimensional space through projection. However, the band selection method (primitive selection) reduces the dimensionality by selecting a subset with the most of characteristics of original HSI [12]. In HSI, every spectral band corresponds to a 2-dimensional image, which can be considered as a feature. According to the availability of class labels, band selection can be further divided into two types: Supervised and unsupervised. Supervised band selection methods as Mutual Information (MI) [13], [14] select the most discriminative features by measuring the correlation with the class labels. Without any priori information of class labels, unsupervised feature selection methods select the most informative features to preserve the information of original features as much as possible [15]-[17]. In practice, the collection of class labels in HSI needs the field exploration and verification by experts, which is expensive and difficult due to excessive labor cost. Thus, unsupervised methods are more practical for HSI processing. Unsupervised band selection can be implemented by bands ranking [16], [17] or bands clustering [18], [19]. For bands ranking, every band is ranked based on certain criteria firstly. Then, top bands are selected by a given dimensionality or threshold. Chang and Wang [16] proposed a constrained band selection (CBS) method. In CBS, when the band has large information divergence with other bands, it is higher priority. In [17], a weighted principal component (WPC) is devised as the criterion, and an adaptive thresholding algorithm based on moving control chart is used to determine the number of selected bands. Although these methods adopt different criteria, all of them aim to select individually informative bands. However, the combination of individually informative bands may not be the most informative bands set. The reason is that these selected bands may have large redundancy and provide little extra information. Also, most traditional methods belong to the feature extraction category, such as Fisher's Linear Discriminant Analysis (FLDA) [20], Principal Component Analysis (PCA) [21], Locality Preserving Projections (LPP) [9], and Isometric Feature Mapping (ISOMAP) [22]. However, these methods usually change the physical characteristics of each original spectral band. Despite the different motivations of these aforementioned methods, Yan et al. [23] have proposed a general graph embedding framework for dimensionality reduction, and many popular methods, e.g., PCA, FLDA, and LPP, could be considered as special cases within this framework. This graph embedding model aims to represent each vertex of graph as a low-dimensional vector that embodies some statistical and geometric properties of a data set. However, it considers only the pairwise relationship between two data set, which fail to capture the complex relationships of the data set. In order to represent the higher order relationships of data set, a first model has been proposed by Berge et al. [24]. This model is called hyperedge, which consists to connect more than two vertices in the hypergraph. Fig. 1 shows the difference between the edge and the hyperedge. Huang et al. [25] proposed a learning framework based on weighted hypergraph for image retrieval, which images are taken as vertices. Ji et al. [26] developed a method for HSI classification using the spatial and spectral information, in which the relationship between the different pixels is represented in a hypergraph structure. Yuan et al. [27] proposed a spatial hypergraph embedding model for feature extraction. This method has been applied to HSI classification. In order to improve the semantic interpretation especially in the object classification of HSI, we present in this paper a semi-supervised method based on spatial hypergraph embedding model to represent higher order relationships and to embody the different weights of the spatial neighbors corresponding to the centroid of pixel. Then, the spatial hypergraph based semi-supervised band selection has been developed to select useful bands for object classification.

III. Proposed Method

After the extensive study on main techniques for dimensionality reduction in Section II, we present in this section our method. Fig. 2 is divided into three phases; the first phase aims to construct a spatial hypergraph to represent higher and complex relationships between the different

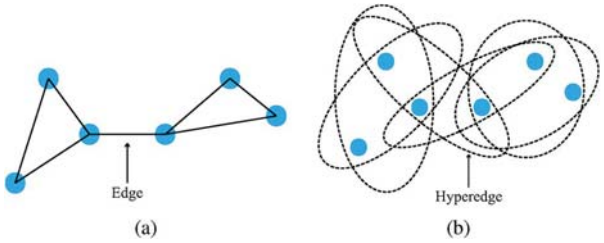

Fig. 1 (a) Each edge is built by the 2-nearest-neighbor method (b) Each hyperedge is built by each vertex and its two nearest neighbors

pixels. In fact, in order to jointly take into account the spatial and spectral information, we have presented a band selection method based on spatial hypergraph to select the useful and relevant bands to improve the semantic interpretation of HSI especially in object classification accuracy. Finally, we have applied the SVM classifier to classify the obtained bands with our presented band selection method.

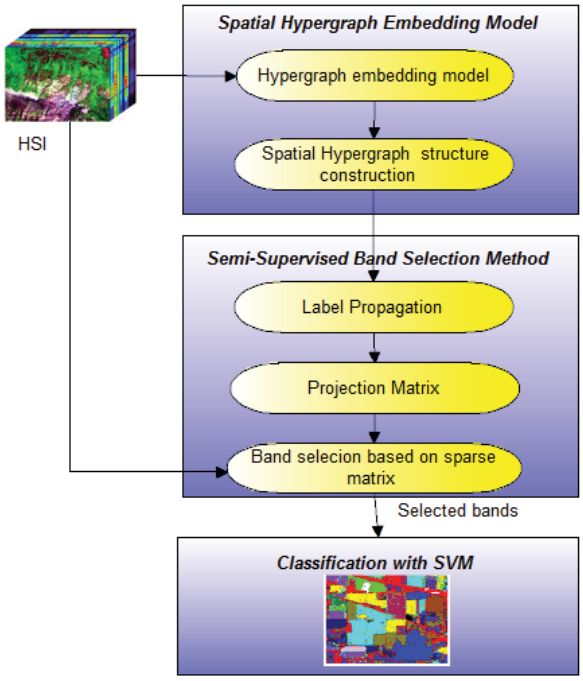

Fig. 2 Flowchart of presented method

A. Dimensionality Reduction and Classification Performance

In our experiment, we used both AVIRIS and ROSIS HSI. Fig. 6 shows the spectral signatures for eight classes extracted from Indian Pines HSI and nine spectral signatures from University of Pavia.

B. Spatial Hypergraph Embedding Model

Given a data set of HSI denoted by $V = [v_1, ..., v_N] \in R^{d\times N}$, where d is the dimensionality of the HSI and N is the number of pixels, it can be used to construct a graph $G = (V, \varepsilon, w)$ is an hypergraph with the vertex set V, the hyperedge set $\varepsilon = \{E_1, ..., E_M\}$, and each hyperedge $E \in \varepsilon$ is a positive weight $w(E)$. A hyperedge E_j is constructed by a centroid vertex v_j and its K-nearest neighbors. The incidence matrix $H \in R^{|v|\times\varepsilon}$ of the hypergraph is represented as

$$H_{ij} = h(v_i, E_j) = \begin{cases} 1, & \text{if } v_i \in E_j. \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

The hyperedge weight w_i is computed as:

$$w_i = w(E_i) = \sum_{v_j \in E_i} exp\left(-\frac{||v_j - v_i||_2^2}{h}\right) \tag{3}$$

Based on H and w, the vertex degree of each vertex $v_i \in V$ is

$$d_i = d(v_i) = \sum_{j=1}^{M} w_j Hij \tag{4}$$

and the edge degree of hyperedge $E_i \in \varepsilon$ is

$$\delta_i = \delta(E_i) = \sum_{j=1}^{N} H_{ji} \tag{5}$$

Let D_v, D_e, and W denote the diagonal matrices containing the vertex degree d, the hyperedge degree δ, and the weight of hyperedge w, respectively. As follows, the objective function of binary hyperedge constraint is formulated as

$$\begin{aligned} &\frac{1}{2}\sum_{E\in\varepsilon}\sum_{u,v\in E}\frac{w(E)h(u,E)h(v,E)}{\delta(E)}||P^T u - P^T v||_2^2 \\ &= \frac{1}{2}\sum_{k=1}^{M}\sum_{i,j=1}^{N}\frac{w_k H_{ik} H_{jk}}{\delta k}||P^T v_i - P^T v_j||_2^2 \\ &= trace(P^T V L V^T P) \end{aligned} \tag{6}$$

where $L = D_v - HWD_e^{-1}H^T$. Finally, the binary hyperedge embedding model is formulated as

$$P^* = \underset{P}{\text{argmax}} = \frac{trace(P^T V D_v V^T P)}{trace(P^T V L V^T P)} \tag{7}$$

In this hypergraph, each hyperedge represents a set of the vertices, including affinity information. Then, the pixels inside a small spatial neighborhood are often made up of the same materials [28]; hence the spatial neighborhood is very appropriate to construct the hypergraph. However, some neighbors around the edges may belong to different classes comparing the centoid one as showing Fig. 3. Therefore, a spatial hypergraph structure is carried out to embody the different weights of the spatial neighbors corresponding to the centroid one. Hence, we use the spatial neighborhood to construct a hypergraph $G' = (V', \varepsilon', w')$ where each pixel v_j is taken as centroid vertex, and the hyperedge $E'_j \in \varepsilon'$ is constructed by this centroid pixel and its T spatial neighbors

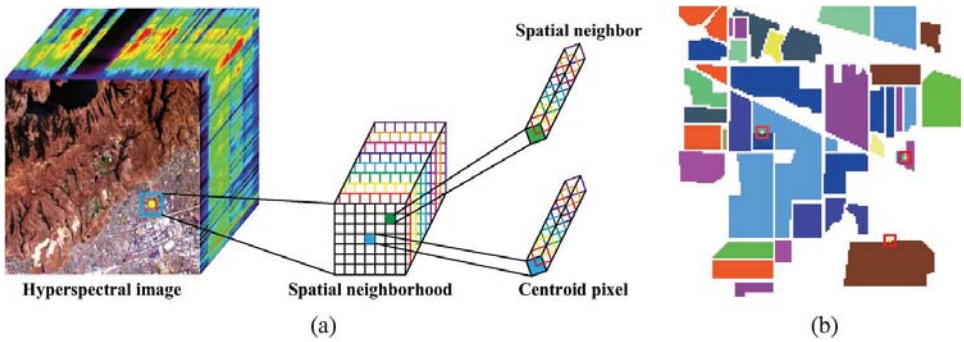

Fig. 3 (a) Spatial neighborhood in the HSI (b) Some spatial neighborhood

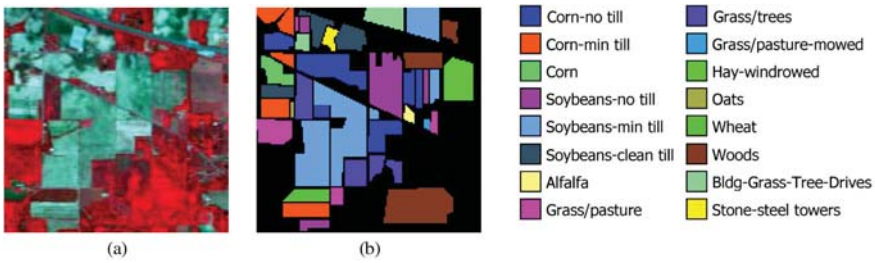

Fig. 4 (a) Color composite image, (b) Ground truth image of Indian Pines

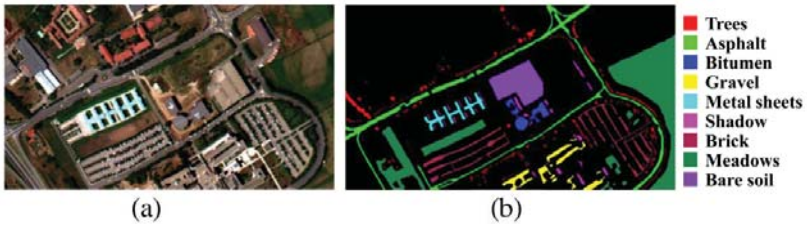

Fig. 5 (a) Color composite image, (b) Ground truth image of University of Pavia

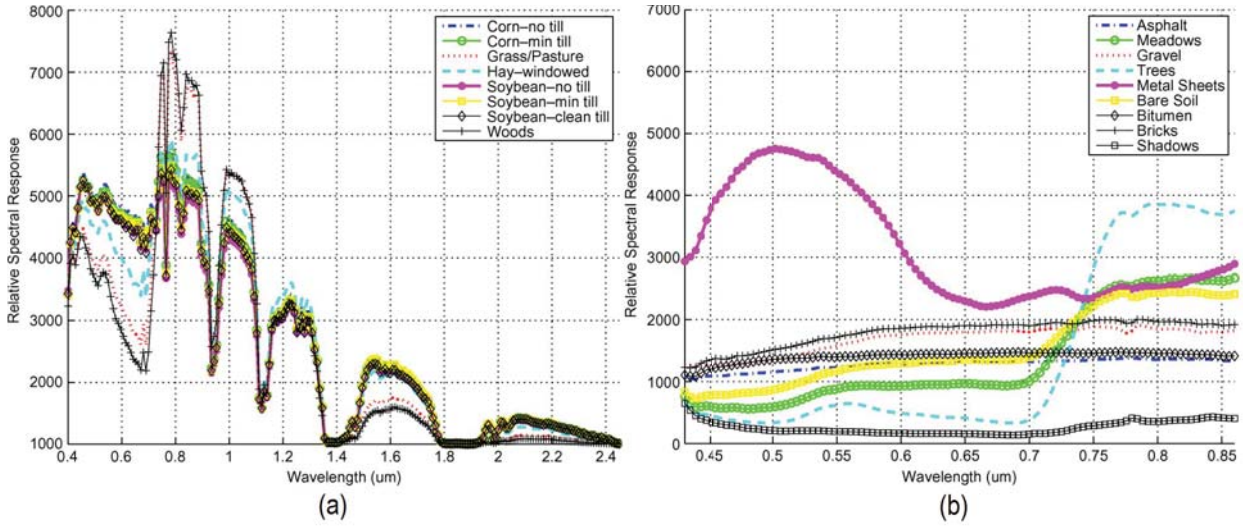

Fig. 6 (a) Spectral signatures of eight classes from Indian Pines (b) Spectral signatures of nine classes from University of Pavia

(this hyperedge connects $T+1$). Hence, the incidence matrix $H' \in R^{|v'|\times\varepsilon'}$ of the spatial hypergraph is represented as

$$H'_{ij} = \begin{cases} exp(-\frac{||v_i - v_j||_2^2}{h}), & \text{if } v_i \in E'_j \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

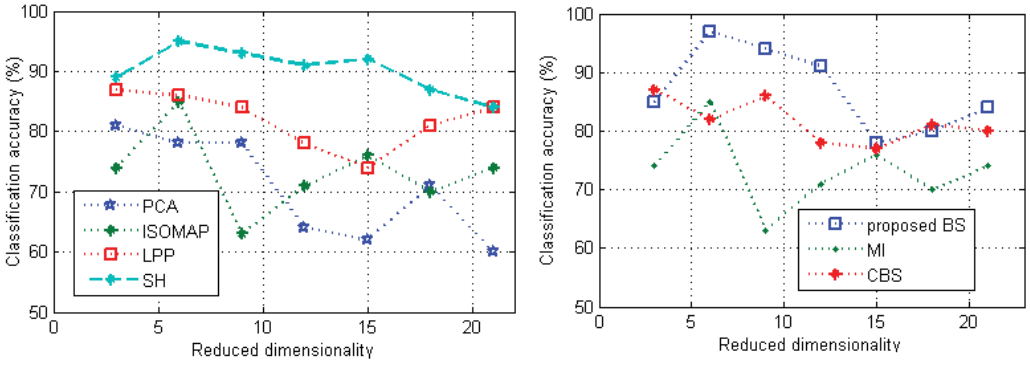

Fig. 7 (a) Classification accuracy of Indian Pines using different feature extraction methods (b) Classification accuracy of Indian Pines using different band selection methods

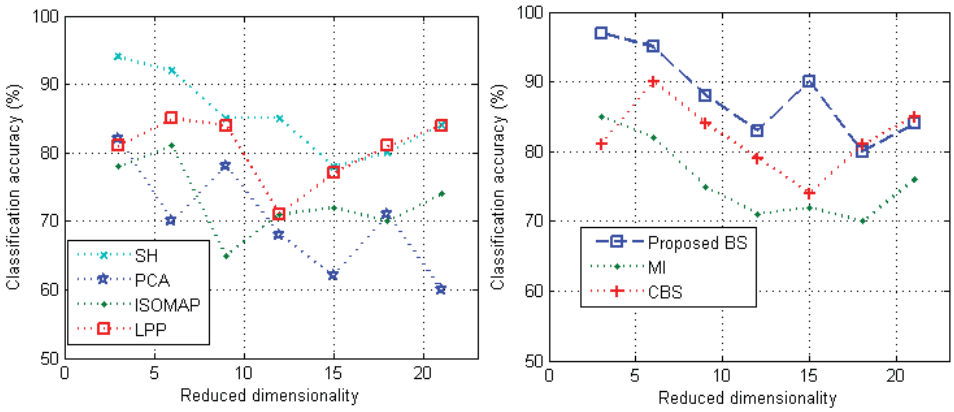

Fig. 8 (a) Classification accuracy of University of Pavia using different feature extraction methods (b) Classification accuracy of University of Pavia using different band selection methods

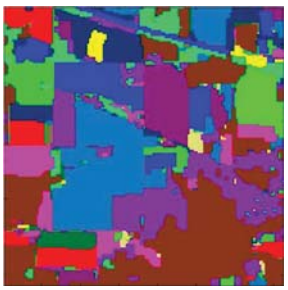

Fig. 9 Classification of Indian Pines HSI with SVM classifier

Fig. 10 Classification of University of Pavia HSI with SVM classifier

Then, the spatial hypergraph model is formulated as

$$P^* = \underset{P}{\operatorname{argmax}} = \frac{trace(P^T V D'_v V^T P)}{trace(P^T V L' V^T P)} \quad (9)$$

where $L' = D'_v - H'W'(D'_e)^{-1}(H')^T$.

C. Semi-Supervised Band Selection Method

The available data set of HSI is denoted by $H = [h_1, ..., h_N]$, where the first N_1 pixels belong to the labeled data set $H_1 = [h_1, ..., h_{N1}]$, where the corresponding label vector is denoted by $z = [z_1, ..., z_{N1}]$ and the rest is the unlabeled data set $H_2 = [h_{N_1+1}, ..., h_N]$. Since the available data are used to construct a spatial hypergraph. After obtaining the optimal projection matrix P^* via the hypergraph embedding model, the labeled data set H_1 is projected into its reduced feature set $X_1 = [(P^*)^T h_1, ..., (P^*)^T h_{N1}]$. In fact, the semi-supervised band method can be used to select the relevant spectral bands to perform HSI classification. Hence, a projection matrix Q is used to classify feature set X by $Y = XQ$. Then, we sort all bands based on sparse matrix S. This phase consists of three

steps, which include label propagation, projection matrix Q, and band sort based on sparse matrix S.

1) Label Propagation: In this step, we use a semi-supervised label propagation method [29]. Let us consider a feature set $X = [x_1, ..., x_l, ..., x_n] \in R^{n \times m}$, there l labeled pixels and $u = n - l$ unlabeled pixels. Here we set a class $C + 1$ to record the outlier data. We define an initial label matrix:

$$L = [(l_1)^T, (l_2)^T, ..., (l_N)^T] \in R^{N \times C+1} \tag{10}$$

$$L_{ij} = \begin{cases} 1, & \text{if } x_i \text{ is labeled as } l_i = j \text{ or } j = C + 1 \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

Given a following spreading function:

$$F(t + 1) = \lambda PF(t) + (1 - \lambda)L \tag{12}$$

where λ is a parameter in $(0, 1)$. The above step should be iteratively repeated until convergence. if F^* denotes the limit of the sequence $F(t)$, the final estimation function F^* can be computed as follows:

$$F^* = \lim_{t \to +\infty} F(t) = (1 - \lambda)(I - \lambda P)^{-1} L \tag{13}$$

2) Projection Matrix Q: In this step, we aim to get an affine matrix Q from $Y = XQ$. Hence, we build a linear classifier $y = Q^T x + b$ where x is a sample in feature set and b is a bias term. If y is close to t_j where $t_j = [0, ..., 0, 1, 0, ..., 0]^T$, x will be classified into class j. We can assume that the result of classification from the linear classifier is equal to the above semi-supervised learning matrix F, and we can define a regression function as:

$$\underset{Q,b}{\operatorname{argmin}} \alpha ||Q||^2 \sum_{i=1}^{N} \sum_{j=1}^{C} F_{i,j} ||Q^T x_i + b - t_j||^2 \tag{14}$$

where $F_{i,j}$ is the label propagation matrix from (13).

3) Band Sort Based on Sparse Matrix: Through the above steps, the HSI can be expressed as $Y = XQ$ where Q is a projection matrix from (14). In fact, we apply an band selection method to find an optimal subset of bands. This method constrains the projection matrix $Q = [Q_1, ..., Q_c]$ via a sparse selection matrix $S = [S_1, ..., S_c]$. This can be represented by the following minimization problem

$$\underset{S}{\operatorname{argmin}} \sum_{k=1}^{c} ||Y_k - XS_k||^2 \tag{15}$$

where S_k is an m-dimensional vector. Hence, we aim to select t bands from a total of m bands. In fact, we use score band as criterion to evaluate every band from the selection matrix S, which is defined as

$$bandscore(j) = \max_k ||S_{j,k}|| \tag{16}$$

where j is the band index, $S_{j,k}$ is the jth line of S_k.

D. Classification with SVM

This phase aims to classify HSI using a linear Support Vector Machines (SVM) classifier. The main goal of SVM classifier is to characterize classes using geometrical criteria instead of statistical ones.

Let us consider a training set consisting of N vectors from the $d-$dimensional features space $(x_i \in \Re^d, i = 1, 2, ..., N)$, with class label $y_i \in \{+1, -1\}$. SVM [30] is represented by function $f(x, \alpha) \to y$, where α denotes the parameters of the classifier. The SVM approach consists in finding the optimal hyperplane that maximizes the distance between the closest training sample to the separating hyperplane, so that:

- Samples with different labels are located on either side of the hyperplane;
- The distance of the closest vectors to the hyperplane on either side is maximum. These vectors are called *support vectors* and the distance separating them to the hyperplane is the optimal margin.

The hyperplane is defined by $w.x + b = 0$, where w and b are parameters of the hyperplane. Vectors that are not on this hyperplane verify: $w.x + b \gtrless 0$. Thereby, the SVM can be defined as $f(x, \alpha) = sgn(w.x + b)$. In order to find such a hyperplane, one should estimate the parameters (b and w) so that:

$$y_i(w.x_i + b) > 0, \quad \text{with} \quad i = 1, 2, ..., N \tag{17}$$

The distance between the closest training sample and the separating hyperplane can be set to $1/||w||$ with a simple rescaling of the hyperplane parameters w and b such that:

$$\min_{i=1,2,...,N} = \{y_i(w.x_i + b)\} \geq 1 \tag{18}$$

Therefore, the geometric margin between the two classes equals $2/||w||$. Hence, the maximization of the margin leads to the following constrained optimization problem:

$$\begin{cases} \text{minimize: } \frac{1}{2}||w||^2 \\ \text{subject to: } y_i(w.x_i + b) \geq 1, \quad i = 1, 2, ..., N \end{cases} \tag{19}$$

The SVM classifier is applied to each feature points with selected bands from the HSI.

IV. Experiments and Discussion

Here, we validate our presented method with several HSI data set and present experimental results demonstrating the benefits of spatial hypergraph embedding model and semi-supervised band selection for HSI classification.

A. HSI Data Sets

Two hyperspectral data sets collected by different hyperspectral sensors are used in our experiments.

1) The first HSI used in experiments represents the Indiana Pines region in northwest Indiana by the Airborne Visible/ Infrared Imaging Spectrometer

(AVIRIS) sensor in 1992. It contains 145×145 pixels, with each pixel having 220 spectral bands covering the range of $375-2500$ nm with a spatial resolution of 20 m/pixel and consists of 16 ground-truth classes. The advantage of this HSI is the availability of a reference ground truth image, which can be used for training and testing. Fig. 4 shows a color composite image of the Indian Pines data set along with the ground-truth image.

2) The second HSI was collected by the Reflective Optics System Imaging Spectrometer (ROSIS) optical sensor over the urban area of Pavia University, Italy. It consists of $610 \times 340 \times 115$ datacube. The spectral range is from 430 to 860 nm with a spatial resolution of 1.3 m per pixel. The University of Pavia HSI is collected from an urban area and consequently consists of small buildings, materials, and trees. Fig. 5 (a) shows a color composite image of University of Pavia, whereas Fig. 5 (b) shows the nine ground-truth classes.

The number of reduced dimensionality has been fixed $n = 3, 6, 9, 12, 18, 21$ features. The number of the nearest neighbors $K = 5$ for both HSIs. In classification phase, In Indian Pines HSI, approximately 8600 labeled pixels are employed to train and test the efficacy of the presented system. This data set is partitioned into approximately 1496 training pixels and 7102 test pixels. Also, the number of training and testing samples used for the University of Pavia data set are 1476 and 7380, respectively.

We run SVM on the extracted features and selected bands using our presented method and the state of the art methods, i.e., PCA, LPP, ISOMAP, MI, and CBS. Fig. 7 shows the accuracy of each features extraction method with various numbers of features. Also, Fig. 8 shows the accuracy of each band selection method with various numbers of bands. It is obvious that the presented method gives better classification rates compared to the other projection methods.

In Indian Pines HSI, the better classification is given by the presented method with the number of dimensions $n = 6$ and classification accuracy is equal to 95%. Whereas with presented spatial hypergaph, we have obtained 93.25%. Also, we have obtained a better classification of University of Pavia HSI with our method. In fact, the number of dimensions $n = 3$, the classification accuracy using spatial hypergraph is equal to 94% and the classification accuracy using our method is equal to 97%. It appears from the results that the pixels are correctly classified of both HSIs, which is qualified as almost perfect.

B. Experimental Time

We compare the running times (in seconds) spent by each method. The experiment results are tested on a PC with I5-42100U CPU, 8-GB memory, 64-bits Windows 7 OS using MATLAB 2014b. Table I shows the running times for dimensionality reduction. Among these dimensionality reduction methods, PCA is the fastest one. Spatial hypergraph spends the longest time since it needs to compute the nearest neighbors.

TABLE I
RUNNING TIMES (IN SECONDS) OF DIMENSIONALITY REDUCTION FOR THE INDIAN PINES AND THE UNIVERSITY OF PAVIA HSIS USING DIFFERENT METHODS

Method	PCA	LPP	ISOMAP	SH	CBS	MI	Proposed BS
Indian Pines	0.149	3.022	20.15	22.80	0.325	5.124	10.235
University of Pavia	0.162	5.022	21.95	24.35	0.214	4.951	7.854

V. CONCLUSION

In this paper, we have presented a spatial hypergraph embedding model for feature extraction. It can represent higher order relationships than the graph embedding model. To make the hypergraph embedding model fit for HSI, we presented an spatial hypergraph structure, which incorporates the spatial information. In fact, a semi-supervised band selection method based on spatial hypergraph has been presented to improve the classification of HSI.

REFERENCES

[1] A. Radoi, R. Tanase, and M. Datcu, "Semantic interpretation of multi-level change detection in multi-temporal satellite images," in *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International*. IEEE, 2015, pp. 4157–4160.

[2] M. Ivasic-Kos, M. Pavlic, and P. Poscic, "The analysis and overview of semantic image interpretation," in *Information Technology Interfaces, 2009. ITI'09. Proceedings of the ITI 2009 31st International Conference on*. IEEE, 2009, pp. 181–186.

[3] G. T. Papadopoulos, C. Saathoff, M. Grzegorzek, V. Mezaris, I. Kompatsiaris, S. Staab, and M. G. Strintzis, "Comparative evaluation of spatial context techniques for semantic image analysis," in *Image Analysis for Multimedia Interactive Services, 2009. WIAMIS'09. 10th Workshop on*. IEEE, 2009, pp. 161–164.

[4] C. Hudelot, N. Maillot, and M. Thonnat, "Symbol grounding for semantic image interpretation: from image data to semantics," in *Computer Vision Workshops, 2005. ICCVW'05. Tenth IEEE International Conference on*. IEEE, 2005, pp. 1875–1875.

[5] M. Ivašić-Kos, M. Pavlić, and M. Matetić, "Data preparation for semantic image interpretation," in *Information Technology Interfaces (ITI), 2010 32nd International Conference on*. IEEE, 2010, pp. 181–186.

[6] S. Chen and D. Zhang, "Semisupervised dimensionality reduction with pairwise constraints for hyperspectral image classification," *Geoscience and Remote Sensing Letters, IEEE*, vol. 8, no. 2, pp. 369–373, 2011.

[7] H. Huang and M. Yang, "Dimensionality reduction of hyperspectral images with sparse discriminant embedding," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 53, no. 9, pp. 5160–5169, 2015.

[8] H. Huang, J. Li, and J. Liu, "Enhanced semi-supervised local fisher discriminant analysis for face recognition," *Future Generation Computer Systems*, vol. 28, no. 1, pp. 244–253, 2012.

[9] W. Li, S. Prasad, J. E. Fowler, and L. M. Bruce, "Locality-preserving dimensionality reduction and classification for hyperspectral image analysis," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 50, no. 4, pp. 1185–1198, 2012.

[10] J. Khoder, R. Younes, and F. B. Ouezdou, "Stability of dimensionality reduction methods applied on artificial hyperspectral images," in *Computer Vision and Graphics*. Springer, 2012, pp. 465–474.

[11] B.-C. Kuo, C.-H. Li, and J.-M. Yang, "Kernel nonparametric weighted feature extraction for hyperspectral image classification," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 47, no. 4, pp. 1139–1155, 2009.

[12] J. Feng, L. Jiao, F. Liu, T. Sun, and X. Zhang, "Unsupervised feature selection based on maximum information and minimum redundancy for hyperspectral images," *Pattern Recognition*, vol. 51, pp. 295–309, 2016.

[13] J. M. Sotoca and F. Pla, "Supervised feature selection by clustering using conditional mutual information-based distances," *Pattern Recognition*, vol. 43, no. 6, pp. 2068–2081, 2010.

[14] B. Guo, R. I. Damper, S. R. Gunn, and J. D. Nelson, "A fast separability-based feature-selection method for high-dimensional remotely sensed image classification," *Pattern Recognition*, vol. 41, no. 5, pp. 1653–1662, 2008.

[15] L. Zhang, C. Chen, J. Bu, and X. He, "A unified feature and instance selection framework using optimum experimental design," *Image Processing, IEEE Transactions on*, vol. 21, no. 5, pp. 2379–2388, 2012.

[16] C.-I. Chang and S. Wang, "Constrained band selection for hyperspectral imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 44, no. 6, pp. 1575–1585, 2006.

[17] S. B. Kim and P. Rattakorn, "Unsupervised feature selection using weighted principal components," *Expert systems with applications*, vol. 38, no. 5, pp. 5704–5710, 2011.

[18] W. Jian, "Unsupervised intrusion feature selection based on genetic algorithm and fcm," in *Information Engineering and Applications*. Springer, 2012, pp. 1005–1012.

[19] M. Breaban and H. Luchian, "A unifying criterion for unsupervised clustering and feature selection," *Pattern Recognition*, vol. 44, no. 4, pp. 854–865, 2011.

[20] M. Sugiyama, "Dimensionality reduction of multimodal labeled data by local fisher discriminant analysis," *The Journal of Machine Learning Research*, vol. 8, pp. 1027–1061, 2007.

[21] P. Deepa and K. Thilagavathi, "Feature extraction of hyperspectral image using principal component analysis and folded-principal component analysis," in *Electronics and Communication Systems (ICECS), 2015 2nd International Conference on*. IEEE, 2015, pp. 656–660.

[22] L. Ding, P. Tang, and H. Li, "Isomap-based subspace analysis for the classification of hyperspectral data," in *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International*. IEEE, 2013, pp. 429–432.

[23] S. Yan, D. Xu, B. Zhang, H.-J. Zhang, Q. Yang, and S. Lin, "Graph embedding and extensions: a general framework for dimensionality reduction," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 1, pp. 40–51, 2007.

[24] C. Berge and E. Minieka, *Graphs and hypergraphs*. North-Holland publishing company Amsterdam, 1973, vol. 7.

[25] Y. Huang, Q. Liu, S. Zhang, and D. N. Metaxas, "Image retrieval via probabilistic hypergraph ranking," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 3376–3383.

[26] R. Ji, Y. Gao, R. Hong, Q. Liu, D. Tao, and X. Li, "Spectral-spatial constraint hyperspectral image classification," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 52, no. 3, pp. 1811–1824, 2014.

[27] H. Yuan and Y. Y. Tang, "Learning with hypergraph for hyperspectral image feature extraction," *Geoscience and Remote Sensing Letters, IEEE*, vol. 12, no. 8, pp. 1695–1699, 2015.

[28] A. Soltani-Farani, H. R. Rabiee, and S. A. Hosseini, "Spatial-aware dictionary learning for hyperspectral image classification," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 53, no. 1, pp. 527–541, 2015.

[29] F. Nie, S. Xiang, Y. Liu, and C. Zhang, "A general graph-based semi-supervised learning with novel class discovery," *Neural Computing and Applications*, vol. 19, no. 4, pp. 549–555, 2010.

[30] P. Mitra, C. Murthy, and S. K. Pal, "Unsupervised feature selection using feature similarity," *IEEE transactions on pattern analysis and machine intelligence*, vol. 24, no. 3, pp. 301–312, 2002.

Akrem Sellami received the MSc degrees in computer science from the University of Jendouba, Tunisia, in 2012. Currently, he is working toward the PhD degree with the National School of Computer Sciences Engineering, University of Manouba, Manouba, Tunisia. He is also a Permanent Researcher at RIADI Laboratory, University of Manouba, since 2009. His work is mainly related with dimensionality reduction, pattern recognition, signal processing, and machine learning applied to remote sensing hyperspectral images. He also enjoyed a research scholarship from the ITI Department in Telecom Bretagne where he is currently a Ph.D. student. Mr. Sellami is a member of Arts-Pi Tunisia.

Imed Riadh Farah is currently Professor at ISAMM Institute, University of Manouba, Tunisia. He received the Engineer degree in 1991 from ENIG School Engineering, MD degree in 1995 from ISG Institute, and the PhD degree in computer science from National School of Computer Science (ENSI), Tunis, Tunisia in 2003. Since November 2009, he has been Associate Professor and then Professor of computer science. He is an Associate Researcher in the Department ITI-Telecom Bretagne, Brest, France, since January 2009. Since 2011, he is head of the Higher Institute of Multimedia Art of Manouba (ISAMM). His research interests are artificial intelligence, data mining, image processing, recognition and interpretation, especially for remote sensing applications, and earth observation.