A Multiple Criteria-based Spectral Partitioning Method for Remotely Sensed Hyperspectral Image Classification

Yi Liu[a], Jun Li[b], Antonio Plaza[a], and Yanli Sun[a,c]

[a]Hyperspectral Computing Laboratory (Hypercomp), Department of Technology of Computers and Communications, Escuela Politécnica, University of Extremadura, Cáceres, E-10071, Spain.
[b]School of Geography and Planning and Guangdong Key Laboratory for Urbanization and Geo-Simulation, Sun Yat-sen University, Guangzhou, P. R. China.
[c]State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China.

ABSTRACT

Hyperspectral remote sensing offers a powerful tool in many different application contexts. The imbalance between the high dimensionality of the data and the limited availability of training samples calls for the need to perform dimensionality reduction in practice. Among traditional dimensionality reduction techniques, feature extraction is one of the most widely used approaches due to its flexibility to transform the original spectral information into a subspace. In turn, band selection is important when the application requires preserving the original spectral information (especially the physically meaningful information) for the interpretation of the hyperspectral scene. In the case of hyperspectral image classification, both techniques need to discard most of the original features/bands in order to perform the classification using a feature set with much lower dimensionality. However, the discriminative information that allows a classifier to provide good performance is usually class-dependent and the relevant information may live in weak features/bands that are usually discarded or lost through subspace transformation or band selection. As a result, in practice, it is challenging to use either feature extraction or band selection for classification purposes. Relevant lines of attack to address this problem have focused on multiple feature selection aiming at a suitable fusion of diverse features in order to provide relevant information to the classifier. In this paper, we present a new dimensionality reduction technique, called multiple criteria-based spectral partitioning, which is embedded in an ensemble learning framework to perform advanced hyperspectral image classification. Driven by the use of a multiple band priority criteria that is derived from classic band selection techniques, we obtain multiple spectral partitions from the original hyperspectral data that correspond to several band subgroups with much lower spectral dimensionality as compared with the original band set. An ensemble learning technique is then used to fuse the information from multiple features, taking advantage of the relevant information provided by each classifier. Our experimental results with two real hyperspectral images, collected by the reflective optics system imaging spectrometer (ROSIS) over the University of Pavia in Italy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) over the Salinas scene, reveal that our presented method, driven by multiple band priority criteria, is able to obtain better classification results compared with classic band selection techniques. This paper also discusses several possibilities for computationally efficient implementation of the proposed technique using various high-performance computing architectures.

Keywords: Hyperspectral remote sensing, image classification, spectral partitioning, multiple band selecting criteria, high-performance computing.

Further author information: (Send correspondence to Yi Liu)
Yi Liu: E-mail: yiliu@unex.es
Jun Li: E-mail: lijun48@mail.sysu.edu.cn
Antonio Plaza: E-mail: aplaza@unex.es
Yanli Sun: E-mail: sunyl@radi.ac.cn
This work is supported by the CSC (China Scholarship Council) program.

High-Performance Computing in Geoscience and Remote Sensing VI, edited by Bormin Huang, Sebastián López, Zhensen Wu, Jose M. Nascimento, Jun Li, Valeriy V. Strotov, Proc. of SPIE Vol. 10007, 100070C
 · doi: 10.1117/12.2241299

1. INTRODUCTION

Hyperspectral image analysis has developed significantly during the past two decades.[1,2] However, the collection of labeled data has been quite time-comsuming and expensive, leading to the imbalance between high spectral dimensionality and limited labeled samples.[3,4] In order to deal with the issue, dimensionality reduction technique has been widely developed and used.[5,6] Classic dimensionality reduction techniques can be separated into two categories: 1) feature extraction, and 2) band selection. Considering the diversity of available algorithms and techniques, feature extraction turns out to be more flexible and widely used, including well-known approaches in the literature such as principal component analysis (PCA),[7] independent component analysis (ICA),[8] manifold learning (ML)[9,10] and subspace-based approaches.[11,12] However, feature extraction generally transforms the original information after projecting the data into a certain feature space,[13] which may be a challenge for certain applications that require meaningful spectral signatures according to their physical interpretation.[14,15] In turn, band selection algorithms are more effective in preserving the original information due to their capacity for selecting the most informative spectral bands among hundreds or even thousands of bands with great correlation and redundancy with supervised or unsupervised ways[16–18] while, at the same time, retaining their original physical meaning. As a matter of fact, the discriminative information that allows a classifier to provide good performance may often live in weak features/bands that are usually discarded or lost after feature transformation or band selection. As a result, in practice, it is challenging to use either feature extraction or band selection for classification purposes. And thus meanwhile, it is usually a hard task to select the best performing dimensionality reduction technique in practice.

As an alternative strategy, spectral partitioning that aims mainly at rearranging the original spectral bands in the hyperspectral image has been used in recent developments. However, as opposed to band selection, spectral partitioning does not necessarily discard most of the original spectral bands to achieve lower dimensionality.[19] Instead, spectral partitioning generates several groups of band subsets from the original spectral bands, so that each band subset is a so-called spectral partition containing a much lower number of spectral bands as compared with the original hyperspectral image, meanwhile, the union of multiple subsets can make up to the full original image.[20] Therefore, spectral partitioning may effectively provide multiple views of the original hyperspectral image by obtaining several subgroups that can simultaneously exploit most of the original spectral information in the hyperspectral scene without discarding a large proportion. In other words, different subgroups of spectral bands can be used to provide different classification results.[20] The diversity of classifiers constructed with the subgroups of bands provides the possibility to obtain a very robust classifier ensemble, which can be achieved by combining the classification results obtained from each of the subgroups using different ensemble learning strategies, such as bagging and boosting,[21,22] decision combination via majority voting,[23] and multiple classifier systems (MCS).[24,25] The diversity obtained in the generation of the multiple views is one of the keys for successful spectral partitioning prior to classification.[24,26,27]

Bearing in mind the aforementioned issues, our motivations to introduce a new spectral partitioning technique in this work can be summarized as follows: 1) to exploit multiple band priority criteria in spectral partitioning instead of using only single one; 2) to facilitate multiple-classifier feature learning by considering different views (spectral partitions) of the original data; 3) to explore the relevance of multiple criteria based band features, which are effectively aligned into multi-subgroups by the employed criteria. Motivated by the issues in mind, in this paper our main work is to develop a new strategy for spectral partitioning that can effectively utilize the spectral bands that are more relevant to different band priority criteria. Here we exploit multiple band priority criteria to perform spectral partitioning, where each spectral partition corresponds to one of the criteria. As for the partitions, we are trying to generate high diversities, considering that the different criteria can provide strong and different perspectives into the original hyperspectral data. The diversities are then used to motivate the classification ensemble.

The remainder of this paper is organized as follows. Section 2 describes our proposed approach. In Section 3, we discuss experimental results obtained using three well-known hyperspectral data sets: the Reflective Optics System Imaging Spectrometer (ROSIS) Pavia University scene and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Salinas scene. The experiments suggest that our presented method leads to competitive results when compared to other state-of-the-art approaches in the field. Conclusions and hints at plausible future research lines are given in section 4.

2. CLASS-ORIENTED SPECTRAL PARTITIONING METHOD

2.1 Proposed spectral partitioning strategy

Let us denote by $\mathbf{x} \equiv \{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \in \mathbb{R}^{d \times n}$ the original spectral signatures of the hyperspectral image $\mathbf{B}^{1 \times d}$ with n pixels indexed by $\mathcal{S} : \{1, 2, \cdots, n\}$ and d wavebands Ω ($\Omega = \{\mathbf{B}^1, \cdots, \mathbf{B}^d\}$ and $|\Omega| = d$). Classification is the process that assigns each pixel $\mathbf{x}_i = \{x_{i1}, x_{i2}, \cdots, x_{id}\}, i \in \{1, \cdots, n\}$ with a label $y_i \in \mathcal{L} : \{1, 2, \cdots, N\}$, where N is the number of classes of interest in the scene. Spectral partitioning aims at separating and reassigning the d spectral bands Ω into a group of band subsets, namely spectral partitions $\{SP_i\}$, $i \in \{1, 2, \cdots, M\}$, so that:

$$SP_i \bigcap_{\forall\ i,j \in \{1,\cdots,d\}, i \neq j} SP_j \subseteq \emptyset \quad \text{AND} \quad \bigcup_{i \in \{1,\cdots,d\}} \{SP_i\} \subseteq \Omega. \tag{1}$$

which means that usually the union of the spectral partitions gives the whole set of original spectral bands of the hyperspectral image. This brings different perspectives of the original hyperspectral image.[19,20] Given $|SP_i|$ as the number of bands in partition SP_i, we can infer from Eqs. (1) that $\sum_{i=1}^{N} |SP_i| = d$.

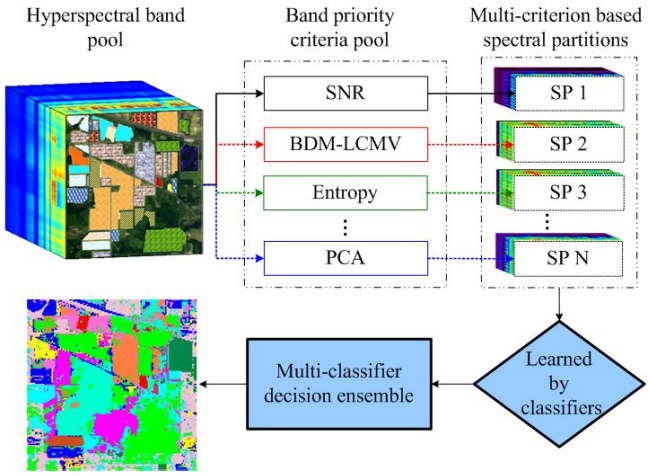


Figure 1. Flowchart of the proposed class-oriented spectral partitioning prior to classification approach.

With the aforementioned ideas in mind, Fig. 1 shows a general flowchart of our proposed multiple band priority criteria based spectral partitioning approach prior to classification. First, we build a pool of band priority criteria, which connect both the original hyperspectral bands and the given number N of spectral partitions. Each spectral partition associates with one of the criteria. Note that the multi-criteria pool is quite open to newly developed band selection techniques, which thus can be extended conveniently. In the second step, we start the focus on the first spectral partition and assign it with one spectral band that scores the most under its associated criterion. Afterward, we remove the assigned band from the original hyperspectral bands. Then the focus moves on to the second of the rest spectral partitions one by one until each partition obtains a new band. Thirdly, we repeat the operation of the second step until all the original hyperspectral spectral bands are assigned to multiple spectral partitions. The generated spectral partitions are thus multi-criterion based. In other words, the bands of each spectral partition correspond to their associated band priority criterion. Here we exploit the different used band priority criteria as the power of diversity generation for the following classification ensemble process.

2.2 Band selection criteria

Band selection techniques intend to select an appropriate band subset from the original data set to represent the data according to some optimality criterion.[28] Generally, band selection can be understood as an exhaustive searching process for all possible cases: $(L_{|\Omega_{BS}|}) = \frac{d!}{(d-|\Omega_{BS}|)!|\Omega_{BS}|!}$, with Ω_{BS} being the selected bands and L

being the number of all possible subset of selected bands, given a number of bands to be selected $|\Omega_{BS}|$.[28] A general way to perform the searching process is to solve the following optimization problem:

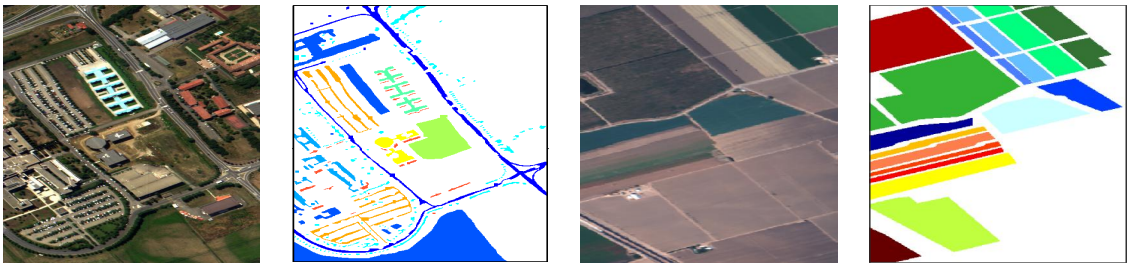
$$\Omega^*_{BS} = \arg \max_{\Omega_{BS} \subset \Omega, |\Omega_{BS}| = n_{BS}} J(\Omega_{BS}), \tag{2}$$

where $n_{BS} = |\Omega_{BS}|$ is the number of selected bands in subset Ω_{BS} and $J(\Omega_{BS})$ establishes the relative importance of a given spectral band in Ω_{BS}. Given the large number of spectral bands in a hyperspectral image, it is almost impossible to try all possible band combinations. It is also difficult to decide which spectral bands play a more relevant role, or to anticipate which combination is more useful. Consequentially, many available band selection algorithms are focused on defining target functions that calculate the priority score of a given spectral band.[29–31]

In this work, we build our multiple band priority criteria pool based on: signal-to-noise ratio-based criterion (SNR),[32] criteria of band dependence minimization-based linearly constrained minimum variance (BDM-LCMV) and band dependence constraint-based linearly constrained minimum variance (BDC-LCMV),[33] independent divergence-based criterion, higher order statistics-based kurtosis criterion (kurtosis) as well as the principle component analysis based criterion (PCA), five criteria in total as an instance for evaluating purposes. We have selected these five classic criteria as an instance because their strategies for band selection are widely used, which makes the algorithms quite representative of existing band selection techniques. Note that our presented framework is quite open and friendly to other existing developed band selection techniques. Besides, the presented framework can be straightforwardly parallelized since each spectral partition is actually learned independently by classifiers. Since each partition is trained by one classifier, we need a decision rule to fuse the individual classifiers. Let $\mathbf{p}_m(i)$ be the probability obtained by an SVM classifier for a given pixel i and partition m. $\mathbf{p}_m(i)$ provides the degree of membership of a pixel to different classes of interest. In this work, we use a soft ensemble strategy to combine the results obtained from all the partitions. Specifically, the probabilities resulting from all the different partitions in a given pixel are modeled by $\widehat{\mathbf{p}}(i) = \frac{1}{N}\sum_{m=1}^{N} \mathbf{p}_m(i)$, where m indexes the spectral partitions ($m \in \{1, \cdots, N\}$) and N is the number of partitions, which is equal to the number of classes, according to our interpretation in Fig.1. The final class label for pixel i is determined by $y_i = \arg\max_{k \in \{1,\cdots,N\}} \widehat{p}^{(k)}(i)$, where $\widehat{p}^{(k)}(i)$ is the probability corresponding to class k for a given pixel i, and $\widehat{\mathbf{p}}(i) = \{\widehat{p}^{(1)}(i), \cdots, \widehat{p}^{(N)}(i)\}$.

3. EXPERIMENTAL RESULTS

3.1 Hyperspectral images used in experiments



(a) The ROSIS data (b) Ground-truth (c) The AVIRIS data (d) Ground-truth

Figure 2. Experimental hyperspectral data sets along with their ground-truth.

The first hyperspectral image used in our experiments [see Fig. 2 (a)] was collected by ROSIS over the University of Pavia, Italy. The data set consists of 115 spectral bands between 0.4 and 1.0 μm, with a size of 610×340 pixels. The noisy bands had been removed, yielding 103 spectral bands that are actually used in this work. The ground-truth image in Fig. 2(b) contains 9 ground-truth classes, including 43923 labeled samples (are often separated into 3921 fixed training samples distributed together with the data, and 40002 test samples). The second hyperspectral image used in our experiment is the well-known AVIRIS Salinas* data set, collected by the 224-band AVIRIS sensor over Salinas Valley, California. As displayed in Fig. 2(c), it comprises 512 lines by 217 samples across 204 spectral bands after discarding 20 water absorption bands. As shown in Fig. 2(d),

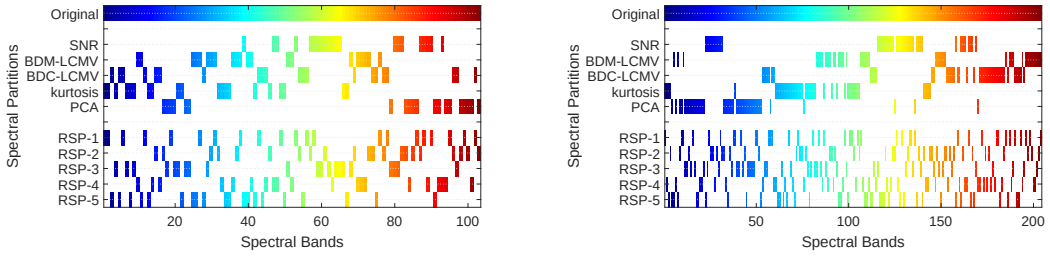
*http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes#Salinas_scene

Table 1. Overall, average and individual class accuracies [%] and κ statistic obtained by the presented classification framework implemented using the MLR classifier with the spectral partitioning methods: RSP and MCB-SP, for the ROSIS Pavia University scene. The results obtained using the original spectral information are also included. Note that RSP is included for comparison. In all cases, only 1% randomly selected training samples have been used.

Class	ROSIS Pavia University		
	ORI	RSP	MCB-SP
Alfalfa	71.53 ± 3.28	72.87 ± 3.38	74.07 ± 2.42
Meadows	62.71 ± 4.32	65.79 ± 4.52	76.72 ± 4.68
Gravel	65.58 ± 5.84	71.66 ± 4.92	75.3 ± 3.95
Trees	88.70 ± 4.73	89.43 ± 4.05	92.77 ± 2.13
Metal sheets	98.65 ± 0.54	98.77 ± 0.40	99.21 ± 0.40
Bare soil	70.53 ± 6.12	73.76 ± 4.91	76.51 ± 5.04
Bitumen	83.67 ± 4.32	89.16 ± 1.78	88.26 ± 2.43
Bricks	66.01 ± 6.50	73.88 ± 4.41	73.96 ± 4.35
Shadows	99.38 ± 0.70	99.56 ± 0.28	99.29 ± 0.51
Overall accuracy	69.45 ± 1.88	72.56 ± 1.88	78.45 ± 1.93
Average accuracy	78.53 ± 1.09	81.65 ± 0.76	84.01 ± 0.77
κ statistic	61.45 ± 2.05	65.18 ± 2.07	72.07 ± 2.24

a total of 54129 pixels are available in the labeled ground-truth, including 16 mutually exclusive classes. In the following experiments with these data sets, we exploit different proportion of randomly selected training samples in total to illustrate the performance of the presented method using limited training sets. We also conduct 20 Monte Carlo runs to ensure statistical significance.

3.2 Experimental results with the hyperspectral data sets



(a) Spectral partitions(ROSIS) (b) Spectral partitions(AVIRIS)

Figure 3. Spectral partitions obtained by our presented spectral partitioning (SP) method and the random spectral partitioning (SP) method from the ROSIS Pavia University data(a) and the AVIRIS Salinas data(b). In all the plots, the x-axis denotes the set of original spectral bands, while the y-axis represents the group of selected bands (each horizontal line displays one group of selected bands).

In this subsection, we tested our spectral partitioning method with the two previous introduced data sets. First of all, we used different limited proportions of randomly selected training samples from the labeled reference data for training (and the rest for testing). We employ random spectral partitioning (RSP), which randomly reassigning the spectral bands into multiple subgroups. For simplicity, the experimental results are obtained by using the same conditions for all groups. Empirically, we set the number of spectral partitions as five that resulting in about 20 spectral bands of each partition for the ROSIS Pavia University data and 40 spectral bands for the AVIRIS Salinas data. The state-of-the-art classifier, multinomial logistic regression (MLR),[34] is used to validate our presented method.

Figs. 3(a) and 3(b) plot the obtained multiple spectral partitions. The spectral partitioning results obtained by random spectral partitioning (RSP) are shown following the results of our presented method. From Fig. 3, we can observe that each band priority criterion-based partition takes a band subset of much lower dimensionality than the original number of bands. And in comparison with the random spectral partitioning (RSP), the band priority criterion-based spectral partitions display strong preference of the spectral bands that are relevant to a specific criterion. This indicates a higher diversity of the spectral partitions via our presented method than

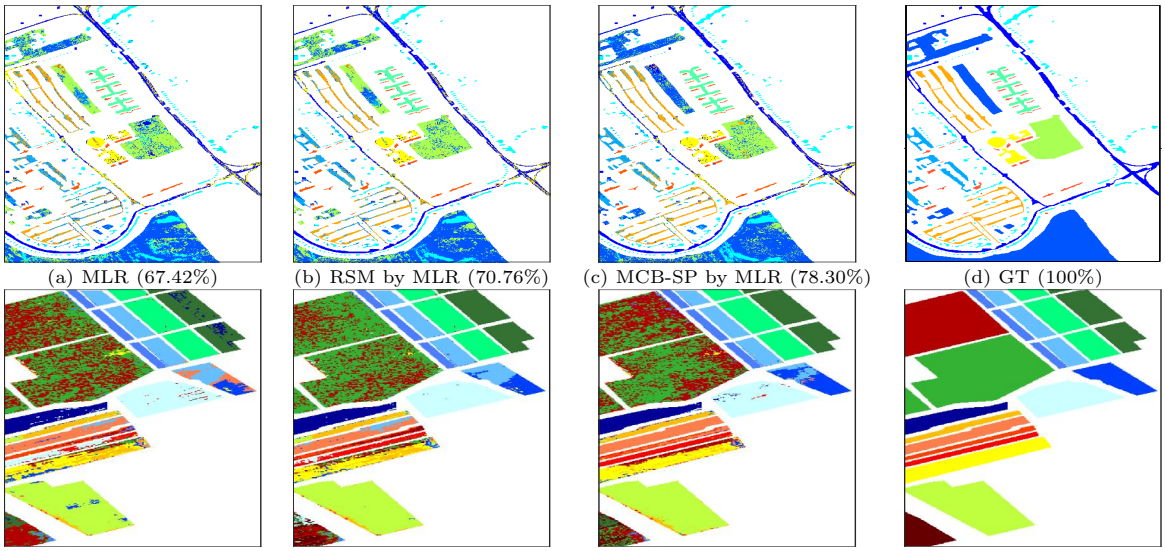
Table 2. Overall, average and individual class accuracies [%] and κ statistic obtained by the presented classification framework implemented using the MLR classifier with the spectral partitioning methods: RSP and MCB-SP, for the AVIRIS Salinas scene. The results obtained using the original spectral information are also included. Note that RSP is included for comparison. In all cases, only 1% randomly selected training samples have been used.

Class	AVIRIS Salinas scene		
	ORI	RSP	MCB-SP
Alfalfa	84.97 ± 16.73	96.40 ± 4.51	95.33 ± 6.66
Corn-notill	92.46 ± 10.17	96.25 ± 4.65	94.97 ± 5.57
Corn-mintill	67.00 ± 11.60	72.03 ± 9.64	85.50 ± 10.64
Corn	90.23 ± 17.14	94.36 ± 13.33	96.29 ± 7.73
Grass/pasture	94.06 ± 5.29	87.98 ± 13.21	89.09 ± 10.34
Grass/trees	98.55 ± 1.68	98.58 ± 1.06	97.39 ± 2.08
Grass/pasture-mowed	93.89 ± 11.43	98.15 ± 2.20	99.69 ± 0.18
Hay-windrowed	53.77 ± 16.40	52.94 ± 15.64	52.49 ± 12.52
Oats	82.81 ± 17.69	85.55 ± 10.91	94.91 ± 4.79
Soybeans-notill	41.51 ± 17.52	38.35 ± 16.70	62.75 ± 18.85
Soybeans-mintill	68.95 ± 16.35	72.35 ± 13.32	79.49 ± 12.27
Soybeans-cleantill	71.41 ± 17.91	68.79 ± 13.99	83.99 ± 13.53
Wheat	67.69 ± 22.87	55.03 ± 14.74	69.12 ± 21.19
Woods	80.39 ± 19.62	84.00 ± 8.85	91.72 ± 2.94
Bldg-Grass-Tree-Drives	58.72 ± 17.38	58.31 ± 15.80	64.38 ± 12.67
Stone-stell towers	1.24 ± 5.56	17.14 ± 12.69	39.81 ± 15.82
Overall accuracy	70.12 ± 4.65	71.33 ± 3.38	76.92 ± 2.06
Average accuracy	71.73 ± 4.30	73.51 ± 2.74	81.06 ± 2.73
κ statistic	66.86 ± 5.07	68.16 ± 3.69	74.40 ± 2.26

performing the partitioning via simply randomly arranging. In all cases of both compared methods, the selection of partitions allows us to obtain multiple views of the original data. Also, the classification learning process on each spectral partition can be independently performed, thus can be straightforwardly parallelized, as indicated by the framework in Section 2.1 and the classification ensemble strategy in Section 2.2.

Tables 1, 2 displays the OA, AA, κ, and individual accuracies obtained after respectively using only 1% and 0.1% randomly selected samples for training and the rest for testing [see Fig. 2(b),(d)]. Several observations can be drawn from Table 1, 2. First of all, consistent observations for both used data sets have been obtained. Specifically, an increase of about 7% to 9% in classification OA is obtained by our presented multiple band priority criterion-based spectral partitioning (MCB-SP) methods in comparison with those using all the spectral bands. And also compared with randomly spectral partitioning(RSP), about 6% increase in classification OA is obtained for both data sets. This observation suggests the effectiveness of our newly proposed methods for multiple classifier-based feature learning based on multiple views that are provided by spectral partitions. Specifically, relevant improvements are obtained for the individual classes in contrast with the RSP method, which indicates that each spectral partition generates stronger classification output when motivated by a band priority criterion in our presented framework. Furthermore, the significant increase in OA than RSP suggests that our presented method is able to provide enough diversity that better drives the classification ensemble process.

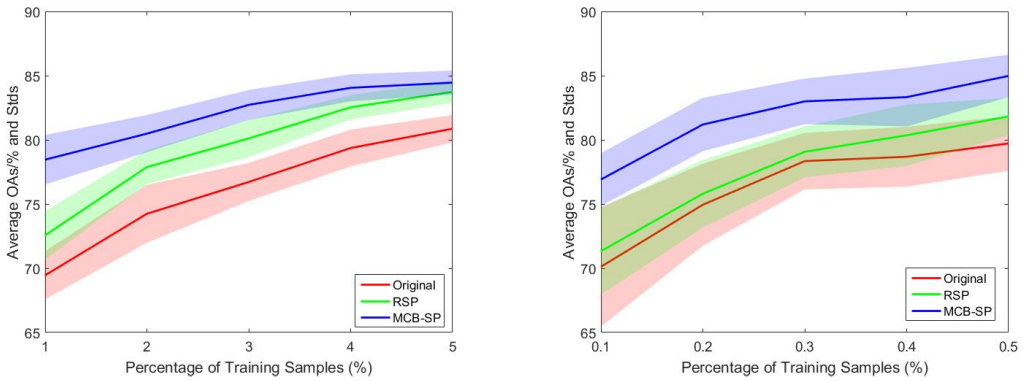
In order to better illustrate this, Fig. 4 displays the classification maps obtained for different methods that are from one of the 20 Monte Carlo runs in Tables 1, 2. It is remarkable that by both using random spectral partitioning (RSP) and our presented spectral partitioning methods, different increases of the individual classes are achieved in comparison with using all the spectral bands of the original image. This indicates that via spectral partitioning, the problem between the imbalanced number of limited training samples and high spectral dimensionality are well attacked. Furthermore, for both used hyperspectral data sets, our presented method provides better classification results when compared with random spectral partitioning (RSP), implying its potential of both making use of the band priority criteria for each partition and meanwhile generating high diversity among the criterion-associated partitions. To conclude this section, we evaluate the sensitivity of the compared methods to different percentages of training samples and obtain the average accuracy statistics and the corresponding standard deviation under 20 Monte Carlo runs. The results are displayed in Fig 5. First of



(a) MLR (67.42%) (b) RSM by MLR (70.76%) (c) MCB-SP by MLR (78.30%) (d) GT (100%)

(e) MLR (69.58%) (f) RSM by MLR (72.72%) (g) MCB-SP by MLR (77.80%) (h) GT (100%)

Figure 4. Classification maps obtained by the proposed classification framework, using the original spectral information (a, e), with spectral partitions by using RSP (b,d), and with the spectral partitions obtained by our proposed spectral partitioning (MCB-SP) approach (c,f). Maps (e,f) are corresponding collected ground truth. The first row shows the classification maps for the ROSIS Pavia University data while the second row displays the maps for the AVIRIS Salinas data. These maps are obtained from one of the 20 Monte Carlo runs, whose OAs are close to the average statistics in Table 1, 2.



(a) ROSIS Pavia University (b) AVIRIS Salinas

Figure 5. Overall classification accuracies (as a function of the number of training samples) obtained by the proposed classification framework (with the original spectral information and with RSP) for the ROSIS Pavia University scene (a) and for the AVIRIS Salinas scene (b). The solid lines represent the average of 20 Monte Carlo runs, whereas the colored area around the lines represent the standard deviation around the mean. Both plots are obtained by using the MLR classifier.

all, it is remarkable that our presented spectral partitioning methods provide highly competitive results in this experiment. The processing time consumptions reported on Table 3, indicate that our methods need about 4 to 17 times more computation than the classification of the original image and about the same consumption of time than the random spectral partitioning. These results can be improved by straightforward parallelization, as discussed in previous experiments.

4. CONCLUSIONS AND FUTURE RESEARCH LINES

In this paper, we presented a new multiple band priority criteria based spectral partitioning method for hyperspectral image classification. The proposed method is shown to be effective in the task of exploiting the

Table 3. Processing times of different methods for ROSIS Pavia University scene and AVIRIS Salinas scene.

Time/sec	ROSIS Pavia University (training samples by percentage /%)				
	1	2	3	4	5
Ori	1.14 ± 0.09	2.98 ± 0.16	5.83 ± 0.22	9.39 ± 0.3	13.62 ± 0.43
RSP	17.49 ± 0.8	38.72 ± 1.35	67.76 ± 1.8	96.44 ± 2.33	129.01 ± 2.26
MCB-SP	17.39 ± 0.77	38.37 ± 1.27	67.2 ± 1.97	95.81 ± 2.82	127.29 ± 3.22
Time/sec	AVIRIS Salinas (training samples by percentage /%)				
	0.1	0.2	0.3	0.4	0.5
Ori	0.97 ± 0.09	1.4 ± 0.07	2.09 ± 0.09	2.66 ± 0.07	3.13 ± 0.08
RSP	3.84 ± 0.1	5.68 ± 0.17	8.21 ± 0.34	10.64 ± 0.33	12.52 ± 0.12
MCB-SP	4.34 ± 0.09	6.18 ± 0.12	8.63 ± 0.17	11.2 ± 0.25	13.19 ± 0.12

information contained in the spectral partitions that are, respectively, associated with specific criteria. This is mainly done to address the Hughes phenomenon by means of a multiple classifier system, while avoiding the elimination of relevant spectral bands in the original hyperspectral image that may be useful for the discrimination of the classes. In our experiments, a multiple criteria pool of measuring band priority is constructed based on five well-established band selection techniques: SNR, BDM-LCMV, BDC-LCMV, kurtosis and PCA. Each spectral partition is one-by-one associated with one of the five criteria, thus the obtained spectral partitions keep most of the relevant spectral bands from the original image and provide different criterion-based views (understood as low-dimensional partitions) for multiple feature learning, thus addressing the potential problems associated to the limited availability of training samples. Our experiments illustrated reasonable advantages in classification accuracy achieved by our presented spectral partitioning framework with three well-known hyperspectral images: the ROSIS Pavia University data and the AVIRIS Salinas data. Furthermore, the experimental classification results indicate that the key for a successful spectral partitioning lies in the capability to generate a group of spectral partitions with diverse views of the original hyperspectral image. In our future work, we will focus on exploring the potential of measuring and make use of the diversity of a group of spectral partitions by multiple criteria. The potential of employing different classifiers will also be considered to enlarge the diversity of the multiple views generated through spectral partitions.

REFERENCES

[1] Plaza, A. and Chang, C.-I., [*High performance computing in remote sensing*], Taylor & Francis: Boca Raton, FL (2007).

[2] Bioucas-Dias, J., Plaza, A., Camps-Valls, G., Scheunders, P., Nasrabadi, N., and Chanussot, J., "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geoscience and Remote Sensing Magazine* **1**, 6–36 (2013).

[3] Plaza, A., Martinez, P., Plaza, J., and Perez, R., "Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations," *IEEE Transactions on Geoscience and Remote Sensing* **43**(3), 466–479 (2005).

[4] Serpico, S., D'Inca, M., Melgani, F., and Moser, G., "Comparison of feature reduction techniques for classification of hyperspectral remote sensing data," in [*International Symposium on Remote Sensing*], 347–358, International Society for Optics and Photonics (2003).

[5] Melgani, F. and Bruzzone, L., "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Transactions on Geoscience and Remote Sensing* **42**(8), 1778–1790 (2004).

[6] Fauvel, M., Dechesne, C., Zullo, A., and Ferraty, F., "Fast forward feature selection for the nonlinear classification of hyperspectral images," *arXiv preprint arXiv:1501.00857* (2015).

[7] Byrne, G., Crapper, P., and Mayo, K., "Monitoring land-cover change by principal component analysis of multitemporal landsat data," *Remote Sensing of Environment* **10**(3), 175–184 (1980).

[8] Hyvärinen, A., Karhunen, J., and Oja, E., [*Independent component analysis*], vol. 46, John Wiley & Sons (2004).

[9] Ma, L., Crawford, M., and Tian, J., "Local manifold learning-based-nearest-neighbor for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing* **48**(11), 4099–4109 (2010).

[10] Lu, Y., Lai, Z., Fan, Z., Cui, J., and Zhu, Q., "Manifold discriminant regression learning for image classification," *Neurocomputing* **166**, 475–486 (oct 2015).

[11] Bioucas-Dias, J. and Nascimento, J., "Hyperspectral subspace identification," *IEEE Transactions on Geoscience and Remote Sensing* **46**(8), 2435–2445 (2008).

[12] Chen, W., Huang, J., Zou, J., and Fang, B., "Wavelet-face based subspace lda method to solve small sample size problem in face recognition," *Int. J. Wavelets Multiresolut Inf. Process.* **07**, 199–214 (mar 2009).

[13] Zhu, Z., Jia, S., He, S., Sun, Y., Ji, Z., and Shen, L., "Three-dimensional Gabor feature extraction for hyperspectral imagery classification using a memetic framework," *Information Sciences* **298**, 274–287 (mar 2015).
[14] Ceccato, P., Gobron, N., Flasse, S., Pinty, B., and Tarantola, S., "Designing a spectral index to estimate vegetation water content from remote sensing data: Part 1: Theoretical approach," *Remote Sensing of Environment* **82**(2), 188–197 (2002).
[15] Odermatt, D., Gitelson, A., Ernesto, V. B., and Schaepman, M., "Review of constituent retrieval in optically deep and complex waters from satellite imagery," *Remote sensing of environment* **118**, 116–126 (2012).
[16] Qian, Y., Yao, F., and Jia, S., "Band selection for hyperspectral imagery using affinity propagation," *IET Computer Vision* **3**(4), 213–222 (2009).
[17] Jia, S., Ji, Z., Qian, Y., and Shen, L., "Unsupervised band selection for hyperspectral imagery classification without manual band removal," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **5**(2), 531–543 (2012).
[18] Lombardo, V., Musacchio, M., and Buongiorno, M., "Error analysis of subpixel lava temperature measurements using infrared remotely sensed data," *Geophysical Journal International* **191**(1), 112–125 (2012).
[19] Liu, Y., Li, J., Plaza, A., Bioucas-Dias, J., Cuartero, A., and Garcia, P., "Spectral partitioning for hyperspectral remote sensing image classification," in [*2014 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*], 3434–3437 (July 2014).
[20] Liu, Y., Li, J., and Plaza, A., "Spectrometer-driven spectral partitioning for hyperspectral image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **9**, 668–680 (Feb 2016).
[21] Schapire, R. and Singer, Y., "Improved boosting algorithms using confidence-rated predictions," *Machine learning* **37**(3), 297–336 (1999).
[22] Kawaguchi, S. and Nishii, R., "Hyperspectral image classification by bootstrap adaboost with random decision stumps," *IEEE Transactions on Geoscience and Remote Sensing* **45**(11), 3845–3851 (2007).
[23] Ho, T., Hull, J., and Srihari, S., "Decision combination in multiple classifier systems," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **16**(1), 66–75 (1994).
[24] Woźniak, M., Graña, M., and Corchado, E., "A survey of multiple classifier systems as hybrid systems," *Information Fusion* **16**, 3–17 (2014).
[25] Benediktsson, J., Chanussot, J., and Fauvel, M., "Multiple classifier systems in remote sensing: from basics to recent developments," in [*Multiple Classifier Systems*], 501–512, Springer Berlin Heidelberg (2007).
[26] Di, W. and Crawford, M., "Active learning via multi-view and local proximity co-regularization for hyperspectral image classification," *IEEE Journal of Selected Topics in Signal Processing* **5**(3), 618–628 (2011).
[27] Wang, H., Nie, F., and Huang, H., "Multi-view clustering and feature learning via structured sparsity," in [*Proceedings of the 30th International Conference on Machine Learning (ICML-13)*], 352–360 (2013).
[28] Chang, C.-I., [*Hyperspectral data processing: algorithm design and analysis*], John Wiley & Sons (2013).
[29] Bajcsy, P. and Groves, P., "Methodology for hyperspectral band selection," *Photogrammetric Engineering & Remote Sensing* **70**(7), 793–802 (2004).
[30] Keshava, N., "Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries," *IEEE Transactions on Geoscience and Remote Sensing* **42**(7), 1552–1565 (2004).
[31] Sarhrouni, E., Hammouch, A., and Aboutajdine, D., "Band selection and classification of hyperspectral images using mutual information: An algorithm based on minimizing the error probability using the inequality of fano," in [*2012 International Conference on Multimedia Computing and Systems (ICMCS)*], 155–159, IEEE (2012).
[32] Sun, K., Geng, X., Ji, L., and Lu, Y., "A new band selection method for hyperspectral image based on data quality," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **7**(6), 2697–2703 (2014).
[33] Chang, C.-I. and Wang, S., "Constrained band selection for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing* **44**(6), 1575–1585 (2006).
[34] Li, J., Bioucas-Dias, J., and Plaza, A., "Semi-supervised hyperspectral classification using active label selection," in [*SPIE Europe Remote Sensing*], 74770F–74770F, International Society for Optics and Photonics (2009).