Hyperspectral Image Classification Based on Union of Subspaces

Mahdi Khodadadzadeh, Jun Li†, Antonio Plaza* and José M. Bioucas-Dias,‡*

*Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura, Cáceres, E-10071, Spain.
†School of Geography and Planning, Sun Yat-Sen University, Guangzhou, P. R. China.
‡Instituto de Telecomunicações, Instituto Superior Técnico, , Universidade de Lisboa, Portugal.

***Abstract*—Characterizing mixed pixels is an important topic in the analysis of hyperspectral data. Recently, a subspace-based technique in a multinomial logistic regression (MLR) framework called MLRsub has been developed to address this issue. MLRsub assumes that the training samples of each class live in a single low-dimensional subspace. However, having in mind that materials in a given class tend to appear in groups and the (possible) presence on nonlinear mixing phenomena, a more powerfull model is a union of subspaces. This paper presents a new approach based on union of subspaces for hyperspectral images. The proposed method integrates subspace clustering with MLR method for supervised classification. Our experimental results with an urban hyperspectral image collected by the NSF-funded Center for Airborne Laser Mapping (NCALM) over the University of Houston campus indicate that the proposed method exhibits state-of-the-art classification performance.**

***Index Terms*—Hyperspectral images, subspace-based approaches, subspace clustering, multinomial logistic regression (MLR).**

I. Introduction

The discrimination of different objects on the earth surface can be achieved by processing of the hundreds of continuous narrow spectral bands collected by hyperspectral sensors [1]. Supervised classification in hyperspectral image processing is defined as the task of assigning a unique label to each pixel vector of the image under consideration using the information extracted from labeled training samples a priori. Several techniques have been used to perform supervised classification of hyperspectral data. In particular, using the class probability estimates resulting from a probabilistic classifier in a Markov random field (MRF) framework, allows us to have more accurate classification results by integrating spectral and spatial information [2–4]. Recently, the multinomial logistic regression (MLR) has shown good performance in hyperspectral image classification which models the posterior class distributions in a Bayesian framework [5–8]. Specifically, the integration of a subspace projection method with the MLR algorithm (called MLRsub) has shown significant classification results [3, 5]. The assumption that hyperspectral vectors live in a lowdimensional subspace is strongly linked with the linear mixing model (LMM) [3, 9]. In essence, if each class is associated with a group of materials, then the spectral vectors of this class are convex combinations of the spectral signatures from that class and thus they live in the subspace spanned by those spectral signatures.

There are number of factors which degrade the modeling power of the subspace model. Among these factors, we highlight the the possible presence of nonlinear mixing phenomena, and the typical low spatial resolution of hyperspectral images, which increases the likelihood of having mixed pixels from a number of different groups (clusters) of materials. Under these degrading factors, a better model is that the spectral vectors in given class lie in unions of subspaces [10]. Notice that the single subspace model contains subsets that are not representative of the training set, which is not the case with the union of subspaces.

Exploiting the union of subspaces in an MLR framework for supervised hyperspectral image classification is the main contribution of this paper. For this purpose, we suggest to use a subspace clustering method before the classification in order to divide training samples of each class into multiple subsets regarding to existing subspaces. Subspace clustering refers to the task of finding a multi-subspace representation that best fits high dimensional data samples, i.e. finding the number of subspaces and their dimensions and simultaneously clustering the data into multiple subspaces [11]. In this paper, we introduce a new hyperspectral image classification methodology based on Robust Subspace Clustering (RSC) [11] which, so far as we are aware, has not been applied for hyperspectral image analysis.

II. Methodological Framework

The proposed approach mainly comprises two main steps: 1) subspace clustering of training samples set; 2) subspace projection and probabilistic classification using MLR algorithm. In the following, we present the details of each step for the proposed approach.

A. Subspace Clustering

The first step of the proposed procedure consists in performing subspace clustering to find a collection of lower-dimensional subspaces fitting the available training set. Recently, several subspace clustering algorithms have been developed. However, most of them are working under restrictive conditions [11]. Here, we adopt the Robust Subspace

978-1-4799-6652-3/15/$31.00~ ©2015 IEEE

Clustering (RSC) [11] algorithm, which, as our experiments showed, is well suited for clustering of hyperspectral data. Based on the ideas from geometric functional analysis, the RSC method can accurately recover the underlying subspaces under minimal requirements on their orientation, and on the number of samples per subspace [11]. The RSC method is an extension of the Sparse Subspace Clustering (SSC) method [12] to cluster noisy data, that is always an important issue in hyperspectral image clustering. The SSC method applies spectral clustering to an adjacency matrix, obtained by sparsely representing each data point in terms of all the other data points through l_1-minimization; whereas, the RSC algorithm replaces the l_1-minimization step in SSC by an l_1-penalized least squares step (the so-called LASSO) and successfully performs subspace clustering under Gaussian noise.

B. Subspace-based MLR classifier

Let $S \equiv \{1, 2, ..., n\}$ is the set of integers indexing the n pixels of a hyperspectral image, $\mathbf{x} \equiv \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ denote the input hyperspectral image, where $\mathbf{x}_i = [\mathbf{x}_{i1}, \mathbf{x}_{i2}, ..., \mathbf{x}_{id}]^T$ denotes a spectral vector associated with an image pixel $i \in S$ and d is the number of spectral bands. Let $\mathbf{y} \equiv (\mathbf{y}_1, \ldots, \mathbf{y}_n)$ denote an image of labels, $\mathbf{y}_i \equiv [y_{i1}, y_{i2}, \ldots, y_{ic} \ldots, y_{iK}]^T$, where K is the number of classes, $y_{ic} \in \{0, 1\}$, for $c = 1, \ldots, K$ and $\sum_c y_{ic} = 1$. Furthermore, let $\mathcal{D}^{(c)}$ represent the set of training samples with label c and $\mathcal{D} \equiv \{\mathcal{D}^{(1)}, \ldots, \mathcal{D}^{(K)}\}$ represent the complete training set.

In general, the MLR models the posterior class probabilities as follows [13]:

$$p(y_{ic} = 1|\mathbf{x}_i,\ \boldsymbol{\omega}) \quad = \quad \frac{\exp({\boldsymbol{\omega}^{(c)}}^T \boldsymbol{\phi}(\mathbf{x}_i))}{\sum_{c=1}^{K} \exp({\boldsymbol{\omega}^{(c)}}^T \boldsymbol{\phi}(\mathbf{x}_i))}, \qquad (1)$$

where $\boldsymbol{\omega}^{(c)}$ is logistic regressors vector for class c, and $\boldsymbol{\omega} \equiv [{\boldsymbol{\omega}^{(1)}}^T, \ldots, {\boldsymbol{\omega}^{(K)}}^T]^T$; $\boldsymbol{\phi}(\mathbf{x}) \equiv [\phi_1(\mathbf{x}), \ldots, \phi_m(\mathbf{x})]$ is a vector of m fixed functions of the input, often termed as features.

In [3, 5], the subspace for the class c, for $c = 1, 2, \ldots, K$, is estimated via eigenanalysis of the spectral vectors available in the set $\mathcal{D}^{(c)}$. The respective subspace is represented by the orthogonal matrix $\mathbf{U}^{(c)}$ holding on its columns an orthogonal basis computed from the sample correlation matrix of those vectors.

As already stated, in these work we assume the spectral vectors $\mathbf{x}_i \in \mathcal{D}^{(c)}$ live in an union of subspace to be learnt using the RSC method. The output of RSC consists in a partition of $\mathcal{D}^{(c)}$ into $L^{(c)}$ subsets. That is, for each c, we obtain the collection of sets $\mathcal{D}_i^{(c)}$ such that $\mathcal{D}_i^{(c)} \cap \mathcal{D}_j^{(c)} = \emptyset$ for $i \neq j$ and $\mathcal{D}^{(c)} = \mathcal{D}_1^{(c)} \cup \mathcal{D}_2^{(c)} \cup \ldots \cup \mathcal{D}_{L^{(c)}}^{(c)}$.

The obtained collection of subspaces is exploited by including the norms of the projection of the spectral vectors onto the subspaces estimated by RSC. More concretely, we propose the following feature vector:

$$\begin{aligned} \boldsymbol{\phi}(\mathbf{x}_i) \quad = \quad & [\|\mathbf{x}_i\|^2,\ \|\mathbf{x}_i^T \mathbf{U}_1^{(1)}\|^2,\ \ldots,\ \|\mathbf{x}_i^T \mathbf{U}_{L^{(1)}}^{(1)}\|^2, \\ & \ldots, \|\mathbf{x}_i^T \mathbf{U}_1^{(K)}\|^2,\ \ldots,\ \|\mathbf{x}_i^T \mathbf{U}_{L^{(K)}}^{(K)}\|^2]^T, \end{aligned} \qquad (2)$$

where, $\mathbf{U}_l^{(c)}$, $c = 1, \ldots, K$ and $l = 1, \ldots, L^{(c)}$ are orthogonal matrices holding basis for the subspaces spanned by the sets $\{\mathbf{x}_i \in \mathcal{D}_l^{(c)}\}$.

Similarly to [8], we apply the LORSAL algorithm [14] to estimate the regressors for the proposed subspace-based MLR classifier. The pseudocode of the proposed algorithm, referred as MLRusub, is shown in Algorithm 1.

Algorithm 1 MLRusub (union of subspaces MLR)

1: **Input:** $\mathcal{D}$, $\mathbf{x}$
2: **Output:** $\widehat{\mathbf{y}}$
for $c := 1$ to K **do**
  3: $\{\mathcal{D}_l^{(c)}\}_{l=1}^{L^{(c)}} = \mathbf{RSC}(\mathcal{D}^{(c)})$ ($*$ subspace clustering $*$)
  for $l := 1$ to $L^{(c)}$ **do**
    4: $\mathbf{U}_l^{(c)} = \mathbf{sub}(\mathcal{D}_l^{(c)})$ ($*$ subspace computation $*$)
  end for
end for
5: $\mathbf{U} \equiv \{\mathbf{U}_1^{(1)}, \ldots, \mathbf{U}_{L^{(1)}}^{(1)}, \ldots, \mathbf{U}_1^{(K)}, \ldots, \mathbf{U}_{L^{(K)}}^{(K)}\}$
6: $\hat{\boldsymbol{\omega}} = \mathbf{LORSAL}(\mathbf{U}, \mathcal{D})$
7: $\hat{\mathbf{y}} \equiv \mathbf{MLR}(\mathbf{x}, \hat{\boldsymbol{\omega}})$

As shown in Algorithm 1, for the input, we have the hyperspectral data $\mathbf{x}$ and the training set $\mathcal{D}$. The classification objective is the image of classes labels $\widehat{\mathbf{y}}$. In steps 3 and 4, we use the RSC method to cluster the class dependent training sets, and estimate the individual subspaces for every cluster. Then in step 5, the union of subspaces is obtained and in step 6, we use LORSAL for classification as LORSAL is able to manage linear/nonlinear features. Finally, in step 7, we obtain the image of class labels following model (1).

III. Experimental Results

In this section, we use a newly released urban hyperspectral image, University of Houston, to evaluate the proposed approach. This data set was acquired on June 23, 2012 by the NSF-funded Center for Airborne Laser Mapping (NCALM) over the University of Houston campus and its neighboring area, which was distributed for the 2013 Data Fusion Contest of the IEEE Geoscience and Remote Sensing Society (GRSS). The hyperspectral image has 144 bands in the 380-1050 nm spectral region and spatial resolution 2.5 m. The image size in pixels is 349×1905. Fig. 1(a) shows a false color composite of the image, while Fig. 1(b) shows 15 classes of interest. In the original data set, 2832 samples were used for training and 12197 samples were used for testing [see Fig. 1]. Detailed information about this image can be found in [15].

To have a complete comparison, the classification results are reported for the three probabilistic classifiers: support vector machine (SVM) with radial basis function (RBF) kernel [16], MLRsub [3] and MLRsub_{mod} [5]. Concerning the classifiers, we optimized the related parameters. Figs. 1 (d)-(g) shows the obtained classification maps. As can be seen, using union of subspaces the MLR$usub$ classifier gives higher classification accuracies than the other subspace based MLR classifiers. For instance, the proposed approach obtained an overall accuracy

of 82.94% which is 11.44% higher than the result obtained by the MLRsub algorithm. More importantly, in the right part of the image where a large cloud shadow is present, the performance improvements reported for the proposed method are quite significant. For example for the class "Highway" in the cloud-covered region, we can see a significant improvement in the obtained classification result.

IV. Conclusion and Remarks

In this paper a new classification method based on union of subspaces was proposed for characterizing mixed (linear and nonlinear) pixels in hyperspectral images. For this purpose, we exploited a subspace clustering method to partition training samples obtained for each class to several subsets and then the multinomial logistic regression algorithm was used to learn the posterior probability distributions from the spectral information of each subset, using a subspace projection. Our experimental results with a new urban hyperspectral image collected by the NSF-funded Center for Airborne Laser Mapping (NCALM) over the University of Houston campus showed that the proposed method exhibits state-of-the-art classification performance as compared to other widely used methods.

V. ACKNOWLEDGEMENT

The authors would like to thank the Hyperspectral Image Analysis Lab at the University of Houston for providing the data set used in this study.

References

[1] C.-I. Chang, *Hyperspectral Data Exploitation: Theory and Applications.* New York: Wiley-Interscience, 2007.

[2] J. Bioucas-Dias, F. Condessa, and J. Kovacevic, "Alternating direction optimization for image segmentation using hidden markov measure field models," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2014, pp. 90 190P–90 190P.

[3] J. Li, J. Bioucas-Dias, and A. Plaza, "Spectral-spatial 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.

[4] Y. Tarabalka, M. Fauvel, J. Chanussot, and J. A. Benediktsson, "SVM- and MRF-based method for accurate classification of hyperspectral images," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 4, pp. 736–740, Oct. 2010.

[5] M. Khodadadzadeh, J. Li, A. Plaza, and J. Bioucas-Dias, "A subspace based multinomial logistic regression for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, 2014, in press.

[6] I. Dopido, J. Li, P. Marpu, A. Plaza, J. Bioucas Dias, and J. Benediktsson, "Semisupervised self-learning for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing,*, vol. 51, no. 7, pp. 4032–4044, July 2013.

[7] J. Li, J. Bioucas-Dias, and A. Plaza, "Semi-supervised hyperspectral image segmentation using multinomial logistic regression with active learning," *IEEE Trans. Geosci. Remote Sens*, vol. 48, no. 11, pp. 4085–4098, 2010.

[8] ——, "Hyperspectral image segmentation using a new bayesian approach with active learning," *IEEE Trans. Geosci. Remote Sens*, vol. 49, no. 10, pp. 3947–3960, Oct. 2011.

[9] J. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, 2012.

[10] M. Soltanolkotabi and E. J. Cands, "A geometric analysis of subspace clustering with outliers," *The Annals of Statistics*, vol. 40, no. 4, pp. 2195–2238, Aug. 2012.

[11] M. Soltanolkotabi, E. Elhamifar, E. J. Candes *et al.*, "Robust subspace clustering," *The Annals of Statistics*, vol. 42, no. 2, pp. 669–699, 2014.

[12] E. Elhamifar and R. Vidal, "Sparse subspace clustering," in *IEEE Conference on Computer Vision and Pattern Recognition*, June 2009, pp. 2790–2797.

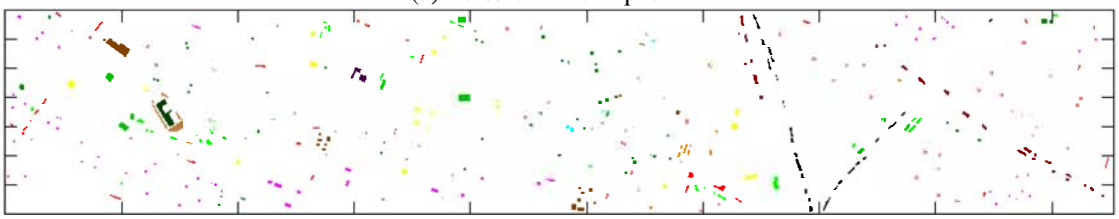
[13] D. Bohning, "Multinomial logistic regression algorithm," *Ann. Inst. Stat. Math.*, vol. 44, no. 1, pp. 197–200, 1992.

[14] J. Bioucas-Dias and M. Figueiredo, "Logistic regression via variable splitting and augmented lagrangian tools," Instituto Superior Tcnico, Lisbon, Portugal, Tech. Rep., 2009.

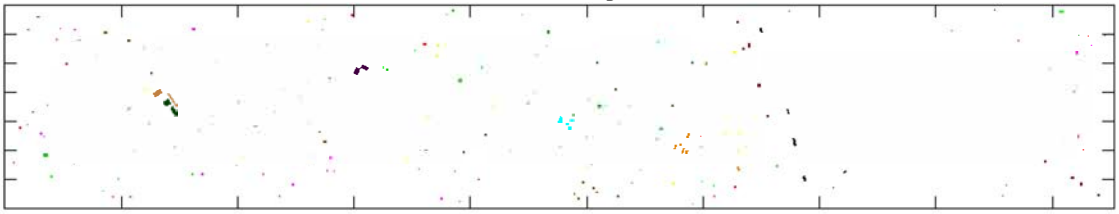
[15] C. Debes, A. Merentitis, R. Heremans, J. Hahn, N. Frangiadakis, T. van Kasteren, W. Liao, R. Bellens, A. Pizurica, S. Gautama, W. Philips, S. Prasad, Q. Du, and F. Pacifici, "Hyperspectral and lidar data fusion: Outcome of the 2013 grss data fusion contest," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 2405–2418, June 2014.

[16] C. Chang and C. Lin. (2009) LIBSVM: A library for support vector machines. [Online]. Available: http://www.csie.ntu.edu.tw/ cjlin/libsvm

(a) False color composition

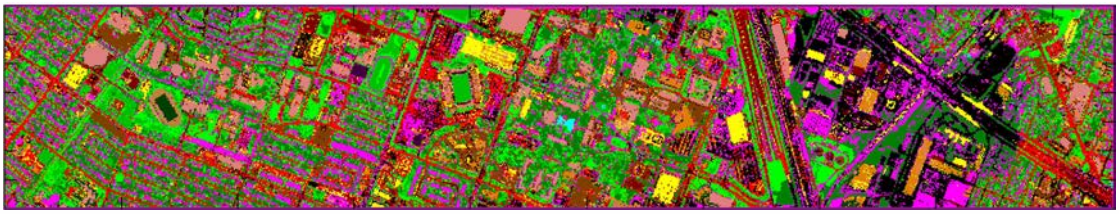


(b) Test samples

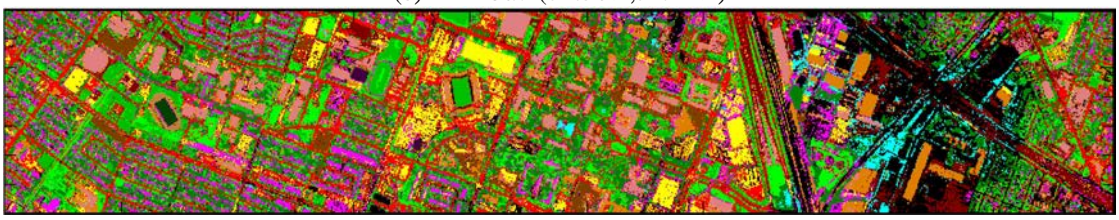


(c) Training samples

(d) SVM (80.49%,83.37%)

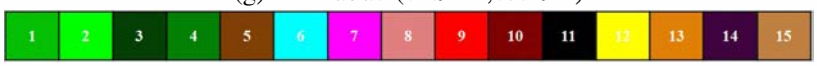


(e) MLRsub (71.50%,74.44%)



(f) MLRsub_{mod} (79.86%,82.51%)

(g) MLR$usub$ (82.94%,85.19%)



(h) Labels color: 1-Healthy grass, 2-Stressed grass, 3-Synthetic grass, 4-Trees, 5-Soil, 6-Water, 7-Residential, 8-Commercial, 9-Road, 10-Highway, 11-Railway, 12-Parking Lot 1, 13-Parking Lot 2, 14-Tennis Court, 15-Running Track

Fig. 1. University of Houston data set, classification maps and overall and average classification accuracies (in the parentheses) obtained by different methods.