A NEW SEMI-SUPERVISED ALGORITHM FOR HYPERSPECTRAL IMAGE CLASSIFICATION BASED ON SPECTRAL UNMIXING CONCEPTS

Alberto Villa[1,2], *Jun Li*[3,4], *Antonio Plaza*[4] *and José M. Bioucas-Dias*[3]

[1]GIPSA-lab, Signal & Image Dept., Grenoble Institute of Technology - INPG, France.
[2]Faculty of Electrical and Computer Engineering, University of Iceland, Iceland.
[3]Instituto de Telecomunicações, Instituto Superior Técnico, TULisbon, 1900-118, Lisboa, Portugal.
[4]Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura, E-10071 Caceres, Spain.

ABSTRACT

Spectral unmixing is a fast growing area in hyperspectral image analysis. Many algorithms have been recently developed to retrieve pure spectral components (*endmembers*) and determine their abundance fractions in mixed pixels, which dominate hyperspectral images. However, possible connections between spectral unmixing concepts and classification algorithms have been rarely investigated. In this work, we propose a new method to perform semi-supervised hyperspectral image classification exploiting the information retrieved with spectral unmixing. The proposed method integrates a well-established discriminative classifier (multinomial logistic regression) with linear spectral unmixing. Furthermore, the proposed method uses a new active sampling approach which takes into account spatial context when generating new samples. The proposed method is experimentally validated using both simulated and real hyperspectral data sets.

***Index Terms*—** Semi-supervised learning, classification, spectral unmixing, active learning, unlabeled training samples.

1. INTRODUCTION

Hyperspectral image analysis is a very active research area. Two of the most widely used techniques for analyzing these kind of data are *classification* [1], in which each pixel is assumed to be composed of a single spectral component, and *unmixing* [2], in which several components are assumed to interact at sub-pixel levels. Traditionally, both techniques have been applied in separate fashion [3]. However, possible connections between spectral unmixing concepts and classification algorithms offer a new promising direction.

In the context of hyperspectral image classification, supervised techniques have achieved wide acceptance [4]. Due to the limited availability of training samples in remote sensing applications, in which ground-truth information is difficult to obtain in terms of time and finance, semi-supervised learning offers an attractive solution which can take advantage of *unlabeled* samples that can be obtained from a (limited) set of labeled samples available *a priori* without significant effort/cost [5]. We refer to [6] for a literature review. An important aspect in the generation of unlabeled samples is the assumption of consistency, which involves: 1) *spatial homogeneity*, according to which neighboring samples likely belong to the same land-cover class, and 2) *spectral homogeneity*, according to which samples belonging to the same land-cover class exhibit similar spectral properties. In general, most semi-supervised algorithms work under such homogeneity assumptions, i.e., they increment the unlabeled training set by assigning class labels to unlabeled samples based on either spatial or spectral homogeneity properties [5, 7]. However, it is well-known that hyperspectral images are dominated by mixed pixels [2], and hence assigning a single class label to a pixel as a whole may be a potential source of errors if several spectral constituents participate in the associated spectral signature. Therefore, if we simply assign a given label to the selected unlabeled sample, the new label could be harmful to the learning process if the pixel is highly mixed or if the assignment is wrong. In order to exploit the unlabeled information more effectively, soft labels are used in [8] to refine the semi-supervised classification process for hyperspectral imagery.

Inspired by the aforementioned observations, in this work we use spectral unmixing to assist the semi-supervised learning process, where active query selection is performed to infer unlabeled samples based on spatial context. This strategy for active sampling using spatial context represents an innovative contribution with regards to previous work in the literature [9–12]. Our main contributions in this work can be thus summarized as follows:

- We develop a new semi-supervised algorithm that integrates classification and unmixing concepts, represented in this work by multinomial logistic regression (MLR) discriminative classification [11] and linear spectral unmixing, respectively.
- We propose a new active sampling approach for the selection of unlabeled samples which takes into account spatial-contextual information when generating new samples. This approach exploits the heterogeneity of the hyperspectral data. The proposed active sampling approach considers the spectral information from a global viewpoint and spatial information from a local viewpoint.

The effectiveness of the proposed semi-supervised classification algorithm is evaluated via experiments with both simulated and real hyperspectral data sets. Our results indicate that the proposed semi-supervised method combining discriminative classification and spectral unmixing concepts, together with the proposed active sampling approach exploiting spatial-contextual information, can exploit classification and spectral unmixing concepts in simultaneous fashion. The remainder of the paper is organized as follows. Section 2 introduces the proposed semi-supervised algorithm. The mechanism used for unlabeled query selection and active sampling is also presented in this section. Section 3 reports classification results on both

This work was supported by MRTN-CT-2006-035927 and AYA2008-05965-C04-02 projects.

978-1-4577-2201-1/11/$26.00 ©2011 IEEE

simulated and real hyperspectral datasets. Section 4 concludes with some remarks and future research lines.

2. PROPOSED APPROACH

First of all, we briefly define the notations used in this paper. Let $\mathcal{K} \equiv \{1, \ldots, K\}$ denote a set of K class labels; let $\mathcal{S} \equiv \{1, \ldots, n\}$ be a set of integers indexing the n pixels of an image; let $\mathbf{x} \equiv (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ be an image of d-dimensional feature vectors, $\mathbf{y}_i = [y_i^{(1)}, ..., y_i^{(K)}]^T$, and let $\mathbf{y} \equiv (y_1, \ldots, y_n)$ be an image of labels denoting a "1-of-K" encoding of the K classes, where $\mathbf{y}_i$ and y_i are equivalent [*e.g.*, $(\mathbf{y}_i = [0, 0, 1, 0]) \Leftrightarrow (y_i = 3)$]; let L be the number of labeled training samples; and let U be the number of unlabeled training samples. With these definitions in mind, in this section we first introduce the multinomial logistic regression model. Then, we present the proposed semi-supervised algorithm explaining the connections with linear spectral unmixing. Finally, we describe a new active sampling approach which also takes into account the concept of mixed pixel in order provide a more effective sampling.

2.1. Multinomial logistic regression

The original multinomial logistic regression (MLR) [13] models the posterior class probabilities as follows:

$$p(y_i^{(k)} = 1|\mathbf{x}_i, \boldsymbol{\omega}) \equiv \frac{\exp\left(\boldsymbol{\omega}^{(k)}\mathbf{h}(\mathbf{x}_i)\right)}{\sum_{k=1}^{K}\exp\left(\boldsymbol{\omega}^{(k)}\mathbf{h}(\mathbf{x}_i)\right)}, \quad (1)$$

where $\mathbf{h}(\mathbf{x}_i) \equiv [h_1(\mathbf{x}_i), ..., h_l(\mathbf{x}_i)]^T$ is a vector of l fixed functions of the input, often termed features; $\boldsymbol{\omega}$ denotes the regressors and $\boldsymbol{\omega} \equiv [\boldsymbol{\omega}^{(1)^T}, ..., \boldsymbol{\omega}^{(K-1)^T}]^T$. Since the density (1) does not depend on possible translations applied to the regressors $\boldsymbol{\omega}^{(k)}$, in this work we take $\boldsymbol{\omega}^{(K)} = \mathbf{0}$. It should be noted that the function $\mathbf{h}$ may be linear (*i.e.*, $\mathbf{h}(\mathbf{x}_i) = [1, x_{i,1}, ..., x_{i,d}]^T$, where $x_{i,j}$ is the j-th component of $\mathbf{x}_i$) or nonlinear. A kernel function is some symmetric function which offers a mechanism to deal with the nonlinear case, *i.e.*, $\mathbf{h}(\mathbf{x}_i) = [1, K_{\mathbf{x}_i,\mathbf{x}_1}, ..., K_{\mathbf{x}_i,\mathbf{x}_l}]^T$, where $K_{\mathbf{x}_i,\mathbf{x}_j} = K(\mathbf{x}_i, \mathbf{x}_j)$ and $K(\cdot,\cdot)$. Kernels have been largely used in this context since they tend to improve data separability in the transformed space. In this work, we use the Gaussian Radial Basis Function (RBF) kernel: $K(\mathbf{x}, \mathbf{z}) \equiv -\exp(\|\mathbf{x} - \mathbf{z}\|^2/(2\sigma^2))$ kernel, which is widely used in hyperspectral image classification [4].

2.2. Proposed semi-supervised algorithm

In order to control the machine complexity and, thus, its generalization capacity, following the SMLR algorithm introduced in [14], we model $\boldsymbol{\omega}$ as a random vector with Laplacian density $p(\boldsymbol{\omega}) \propto \exp(-\lambda\|\boldsymbol{\omega}\|_1)$. Therefore, estimating the regressors $\boldsymbol{\omega}$ amounts to computing the maximum a posteriori (MAP) estimate given by:

$$\widehat{\boldsymbol{\omega}}_{\text{MAP}} = \arg\max_{\boldsymbol{\omega}} \quad \ell_L(\boldsymbol{\omega}) + \ell_U(\boldsymbol{\omega}) + \log p(\boldsymbol{\omega}), \quad (2)$$

where $\ell_L(\boldsymbol{\omega})$ is the log-likelihood function over the labeled information given by:

$$\ell_L(\boldsymbol{\omega}) = \sum_{i=1}^{L}\left(\sum_{k=1}^{K} y_i^{(k)}\boldsymbol{\omega}^{(k)}\mathbf{x}_i - \log\sum_{k=1}^{K}\exp(\boldsymbol{\omega}^{(k)}\mathbf{x}_i)\right), \quad (3)$$

and $\ell_U(\boldsymbol{\omega})$ is the log-likelihood function over the unlabeled information which has the same structure as (3) given by:

$$\ell_U(\boldsymbol{\omega}) = \sum_{i=L+1}^{U}\left(\sum_{k=1}^{K} \widehat{y}_i^{(k)}\boldsymbol{\omega}^{(k)}\mathbf{x}_i - \log\sum_{k=1}^{K}\exp(\boldsymbol{\omega}^{(k)}\mathbf{x}_i)\right), \quad (4)$$

where $\widehat{y}_i^{(k)}$ is the missing variable of the unlabeled information. Following [8], where a soft sparse MLR model is proposed for semi-supervised classification with the capacity of dealing with unlabeled information under a soft constraint, in this work we use the fractional abundances obtained from linear spectral unmixing as soft labels, where the regressors are efficiently learnt by the LORSAL algorithm [12, 15]. The use of fractional abundances exploits the mixed nature of hyperspectral data. In this work, we assume that all labeled samples are spectrally pure and use them as endmembers for a subsequent spectral unmixing process in which the fractional abundances of each endmember are estimated using the fully constrained least squares (FCLS) algorithm [16]. Our preliminary results using this strategy are promising, but we are aware that our assumption regarding the spectral purity of all labeled samples may not hold in some cases. To address this issue, in the future we can consider a more standardized spectral unmixing chain based on: 1) estimating the number of *endmembers* in the hyperspectral data [17]; 2) identifying the spectral signatures of such endmembers [18–20], and 3) estimating the fractional abundance of each endmember in each pixel of the scene [16].

2.3. Active learning

In this work, we propose a new active sampling approach which takes into account spatial-contextual information. This represents an innovation with regards to previous approaches in the literature, such as: 1) the mutual information (MI)-based criterion [21, 22], which maximizes the mutual information between the MLR regressors and the class labels when performing the sampling; 2) the breaking ties (BT) algorithm [9], proposed to achieve diversity in the sampling thus alleviating the bias in the MI-based sampling; and 3) a recently developed modified BT (MBT) algorithm [12], which promotes more diversity in the selection process as compared with the BT criterion.

The basic strategy of our proposed active sampling approach can be simply summarized as follows. Let us assume that a certain pixel $\mathbf{x}_i$ has been assigned a label $y_i = k$, with $k \in \{1, \ldots, K\}$ (*i.e.*, the pixel is a labeled training sample). Let us denote by $\mathcal{N}(i)$ the set of neighboring pixels (in spatial sense) of $\mathbf{x}_i$. Let us now consider a neighboring pixel $\mathbf{x}_j \in \mathcal{N}(i)$. If the classification label estimated by MLR for $\mathbf{x}_j$ is different than the class label of $\mathbf{x}_i$, *i.e.*, $\widehat{y}_j = c$ with $c \in \mathcal{K}$ and $c \neq k$, then we increment the unlabeled set by adding pixel $\mathbf{x}_j$. However, it is likely that several neighboring pixels meet the aforementioned sampling criterion. To address this issue, in this work we adopt a two-step scheme to select the new unlabeled sample. First, we consider a 4-connected pixel neighborhood around each labeled sample $\mathbf{x}_i$ to form the set $\mathcal{N}(i)$. Second, we run MBT to select the most informative sample in $\mathcal{N}(i)$ which is added to the set of unlabeled samples. As a result, in our proposed sampling strategy MBT is run in a local neighborhood, so that unlabeled samples are expected to be extracted from class boundaries in spectral space (by means of the application of MBT) but also from spatial boundaries defining transition areas between land-cover classes (since the MBT is applied in local spatial neighborhoods). As will be shown by experiments, such local application of MBT has the potential to increase the effectiveness of the sampling process.

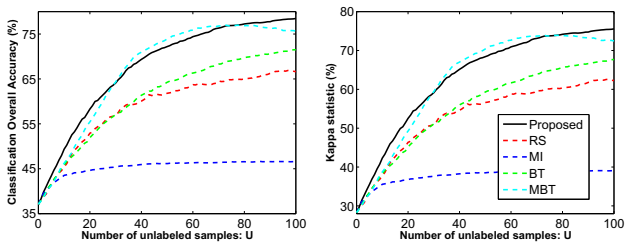


Fig. 1. Overall classification accuracy and κ results (as a function of the number of unlabeled samples) obtained for the simulated data set by the proposed semi-supervised algorithm, using different active sampling methods and $L = 144$ labeled samples.

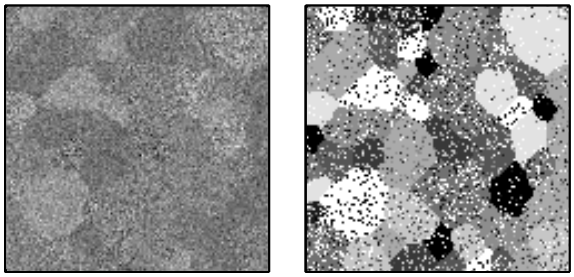
3. EXPERIMENTAL RESULTS

3.1. Experiments with simulated images

A simulated data set has been generated with 10 classes simulated from spectral signatures obtained from the U.S. Geological Survey (USGS) digital spectral library[1], where the spatial information is generated with a multilevel logistic (MLL) model [12, 23] and the feature vectors are generated according to a linear mixture model. For each objective class in the simulated image, we assign 90% abundance of one endmember, and the remaining 10% abundance is established by uniformly mixing the endmembers associated to the remaining 9 classes. The total size of the simulated image is 120×120 pixels. Zero-mean Gaussian noise with covariance $\sigma^2 I$ ($\sigma = 0.5$) is added to the simulated scene. In order to evaluate our newly introduced active sampling strategy, four algorithms developed for labeled query selection are used for comparative purposes, including MI, BT, MBT and a simple random selection (RS) strategy. As performance metrics we consider the overall accuracy (OA) and the kappa statistic (κ), where each value of the aforementioned performance indicators reported in this work is obtained as the average of 10 Monte Carlo runs.

Fig. 1 illustrates the obtained OA and κ results, as a function of the number of unlabeled training samples. The proposed results are based on a total of 144 labeled samples (which represent 1% of the ground truth image) which are used for training purposes, whereas the remaining samples are used as the validation set. The plots in Fig. 1 indicate how the OA and κ increase as additional (unlabeled) samples are incorporated. Several conclusion can be obtained from Fig. 1. First and foremost, it can be seen that the inclusion of unlabeled training samples in the proposed semi-supervised algorithm allows us to significantly improve the classification results provided by the original (supervised) method using only labeled samples. Second, the proposed sampling strategy provides results which are comparable or superior to MBT. Finally, it is also worth noting that the results obtained by the MI algorithm are worse than those achieved by RS. This is expected, as MI focuses on the most complex areas in spectral space. For additional details about the performance of the MI, BT and MBT active sampling approaches for labeled query selection, we refer to our previous work [12]. For illustrative purposes, Fig. 2 shows simulated classification map obtained by the proposed semi-supervised algorithm with the proposed active sampling approach.

[1]The USGS library of spectral signatures is available online: http://speclab.cr.usgs.gov.



(a) Original image (Band 10) (b) Proposed sampling (79.57%)

Fig. 2. Classification map and OA result (in the parentheses) obtained –for a random Monte Carlo run– by the proposed semi-supervised algorithm for the simulated images with $L = 144$ and $U = 100$.

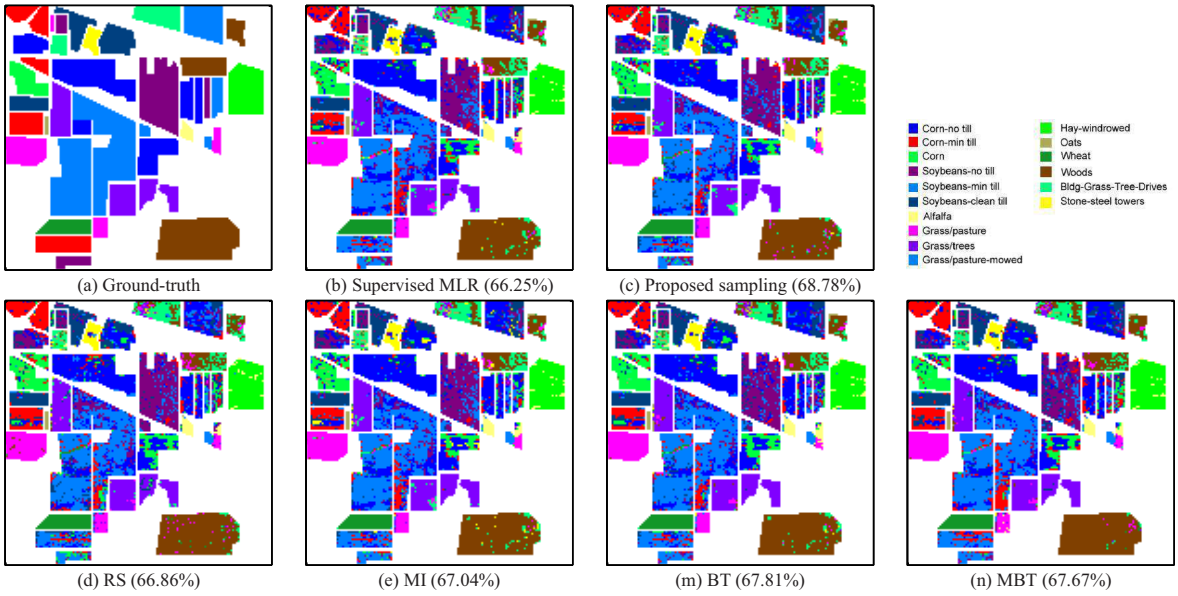
Table 1. Overall, average, individual class accuracies [%] and κ statistic obtained for the AVIRIS Indian Pines data set.

Class	# samples	$U = 0$	$U = 100$				
			Proposed	RS	MI	BT	MBT
Alfalfa	54	83.85	86.41	85.13	84.87	84.87	85.13
Bldg-grass-tree-drives	1434	59.32	61.00	57.56	58.65	56.59	60.78
Corn	834	52.66	56.11	52.58	52.32	54.02	55.76
Corn-no till	234	79.77	79.82	80.23	75.89	78.22	80.14
Corn-min till	497	79.96	80.85	79.54	79.52	77.16	79.36
Grass/pasture	747	90.16	94.32	89.92	92.23	92.95	92.05
Grass/pasture-mowed	26	88.18	88.18	89.09	86.36	89.09	92.73
Grass/tree	489	93.04	94.26	91.81	93.48	94.16	92.24
Hay-windrowed	20	100	100	100	100	100	100
Oats	968	62.11	65.56	66.72	62.46	65.73	64.27
Soybeans-no till	2468	47.45	51.81	51.85	50.95	50.79	48.22
Soybeans-min till	614	62.52	67.01	69.15	61.90	66.68	69.55
Soybeans-clean till	212	99.29	99.70	99.39	99.70	99.54	99.49
Stone-steel towers	1294	83.60	84.82	84.63	83.85	85.56	89.21
Wheat	380	56.63	52.25	52.99	55.89	51.64	47.04
Woods	95	87.62	92.75	91.25	89.50	92.37	91.87
OA		64.92	69.00	66.86	67.61	67.81	67.67
AA		77.36	77.99	76.96	77.24	77.55	77.20
κ		60.82	64.98	62.67	63.49	63.74	63.58

3.2. Experiments with real images

The well-known AVIRIS Indian Pines scene was used in our real data experiments. These data were collected over Northwestern Indiana in June 1992 [1], and contains 145×145 pixels and 220 spectral bands. A total of 20 bands were removed prior to experiments due to noise and water absorption in those channels. The ground-truth data contains 16 mutually exclusive classes, and a total of 10366 labeled pixels. This image is a classic benchmark to validate the accuracy of hyperspectral image analysis algorithms and constitutes a challenging problem due to the significant presence of mixed pixels in all available classes, and also because of the unbalanced number of available labeled pixels per class.

Table 1 illustrates the individual, overall (OA), average (AA) accuracies and the κ values obtained by the proposed semi-supervised algorithm for $L = 240$ labeled samples (meaning 15 samples per class) and no unlabeled samples (represented in the table as $U = 0$) and for $U = 100$ unlabeled samples. Notice the good performance achieved by the proposed semi-supervised algorithm, which outperforms the supervised algorithm in all cases by using unlabeled information. Furthermore, the proposed active sampling approach obtained the best results in terms of OA, κ and AA scores. It is also noticeable that the proposed active sampling approach leads to unbiased improvements in the individual accuracies for all the considered classes. For illustrative purposes, Fig. 3 shows some of the classification maps obtained in the considered experiments.



(a) Ground-truth (b) Supervised MLR (66.25%) (c) Proposed sampling (68.78%)

(d) RS (66.86%) (e) MI (67.04%) (m) BT (67.81%) (n) MBT (67.67%)

Fig. 3. Classification maps and OA results (in the parentheses) obtained –for a random Monte Carlo run– by the proposed semi-supervised algorithm with the AVIRIS Indian Pines image by using $L = 240$ labeled samples (meaning 15 samples per class) and $U = 100$ unlabeled samples obtained by different active sampling strategies. Supervised MLR denotes the algorithm without any unlabeled information.

4. CONCLUSION

In this work, we proposed a new semi-supervised algorithm for hyperspectral image classification which integrates a well-established discriminative classifier (multinomial logisic regresion) with linear spectral unmixing. Furthermore, we have described a new active sampling approach based on spatial-contextual information. Our experimental results, conducted with simulated and real data sets, reveal that the proposed method can benefit from the proposed integrated framework. Future work should comprise experiments with additional data sets and comparisons with other semi-supervised algorithms with active learning.

5. REFERENCES

[1] D. A. Landgrebe, *Signal theory methods in multispectral remote sensing*. Hoboken, NJ: John Wiley and Sons, 2003.

[2] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 44–57, 2002.

[3] A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," *Remote Sensing of Environment*, vol. 113, pp. 110–122, September 2009.

[4] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, pp. 1351–1362, 2005.

[5] G. Camps-Valls, T. Bandos, and D. Zhou, "Semi-supervised graph-based hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, pp. 3044–3054, Oct 2007.

[6] X. Zhu, "Semi-supervised learning literature survey," Computer Sciences, University of Wisconsin-Madison, Tech. Rep. 1530, 2005.

[7] D. Tuia and G. Camps-Valls, "Semi-supervised hyperspectral image classification with cluster kernels," *IEEE Geoscience and Remote Sensing Letters.*, vol. 6, no. 2, pp. 224 –228, 2009.

[8] J. Li, J. Bioucas-Dias, and A. Plaza, "Semi-supervised hyperspectral classification using soft labels," in *IEEE GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS'11), submitted*, 2011.

[9] T. Luo, K. Kramer, D. B. Goldgof, S. Samson, A. Remsen, T. Hopkins, and D. Cohn, "Active learning to recognize multiple types of plankton," *Journal of Machine Learning Research*, pp. 589–613, 2005.

[10] D. Tuia, F. Ratle, F. Pacifici, M. F. Kanevski, and W. J. Emery, "Active learning methods for remote sensing image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 7, pp. 2218–2232, 2009.

[11] J. Li, J. Bioucas-Dias, and A. Plaza, "Semi-supervised hyperspectral image segmentation using multinomial logistic regression with active learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, pp. 4085–4098, 2010.

[12] ——, "Hyperspectral image segmentation using a new bayesian approach with active learning," *IEEE Transactions on Geoscience and Remote Sensing (submitted)*, 2010.

[13] D. Böhning, "Multinomial logistic regression algorithm," *Annals of the Institute of Statistical Mathematics*, vol. 44, pp. 197–200, 1992.

[14] B. Krishnapuram, L. Carin, M. Figueiredo, and A. Hartemink, "Sparse multinomial logistic regression: Fast algorithms and generalization bounds," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 6, pp. 957–968, 2005.

[15] J. Bioucas-Dias and M. Figueiredo, "Logistic regression via variable splitting and augmented lagrangian tools," Instituto Superior Técnico, TULisbon, Tech. Rep., 2009.

[16] D. Heinz and C.-I. Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, pp. 529–545, 2001.

[17] J. Bioucas-Dias and J. Nascimento, "Hyperspectral subspace identification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 8, pp. 2435–2445, 2008.

[18] J. M. Bioucas-Dias and M. A. T. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2010 2nd Workshop on*, 2010, pp. 1 –4.

[19] J. Bioucas-Dias and A. Plaza, "Hyperspectral unmixing: Geometrical, statistical and sparse regression-based approaches," in *SPIE Remote Sensing Europe, Image and Signal Processing for Remote Sensing Conference*, 2010.

[20] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, pp. 650–663, 2004.

[21] D. Mackay, "Information-based objective functions for active data selection," *Neural Computation*, vol. 4, pp. 590–604, 1992.

[22] B. Krishnapuram, D. Williams, Y. Xue, A. Hartemink, L. Carin, and M. Figueiredo, "On semi-supervised classification," in *Proc. 18th Annual Conference on Neural Information Processing Systems*, Vancouver, Canada, 2004.

[23] S. Geman and D. Geman, "Stochastic relaxation, gibbs distribution, andthe bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721–741, 1984.