Advanced Processing of Hyperspectral Images

A. Plaza*, J. A. Benediktsson†, J. Boardman‡, J. Brazile§, L. Bruzzone¶, G. Camps-Valls‖, J. Chanussot**, M. Fauvel**, P. Gamba††, A. Gualtieri‡‡, J. C. Tilton‡‡ and G. Trianni††

*Department of Computer Science, University of Extremadura, Cáceres, SPAIN
†Department of Electrical and Computer Engineering, University of Iceland, Reykjavik, ICELAND
‡Analytical Imaging and Geophysics, LLC Boulder, Colorado, USA
§Department of Geography, University of Zürich, SWITZERLAND
¶Department of Information and Communication Technology, University of Trento, ITALY
‖Department of Electronics Engineering, University of Valencia, SPAIN
**Laboratoire des Images et des Signaux, LIS/INPG, Saint Martin d'Hères, FRANCE
††Department of Electronics, University of Pavia, ITALY
‡‡NASA's Goddard Space Flight Center, Greenbelt, Maryland, USA

***Abstract*— Hyperspectral imaging offers the possibility of characterizing materials and objects in the air, land and water on the basis of the unique reflectance patterns that result from the interaction of solar energy with the molecular structure of the material. In this paper, we provide a seminal view on recent advances in techniques for hyperspectral data processing. Our main focus is on the development of approaches able to *naturally* integrate the spatial and spectral information available from the data. Special attention is paid to techniques that circumvent the *curse of dimensionality* introduced by high-dimensional data spaces. Experimental results, focused in this work on a specific case-study of urban data analysis, demonstrate the success of the considered techniques. This paper represents a first step towards the development of a quantitative and comparative assessment of advances in hyperspectral data processing techniques.**

I. Introduction

Imaging spectroscopy [1], also known as hyperspectral remote sensing, is concerned with the measurement, analysis, and interpretation of spectra acquired from a given scene (or specific object) at a short, medium or long distance by an (airborne or satellite) sensor. The concept of imaging spectroscopy originated in the 1980's, when Dr. Alexander F. H. Goetz and his colleagues at Jet Propulsion Laboratory began a revolution in remote sensing by designing new instruments such as the Airborne Visible Infra-Red Imaging Spectrometer (AVIRIS) [2], with hundreds of spectral channels and very fine spectral resolution. In this paper, we intend to provide a seminal view on recent, consolidated advances in technologies for efficient processing of hyperspectral imagery. First, we discuss hyperspectral image classification using support vector machines (SVMs), which are one of the most consolidated algorithms for analyzing high-dimensional input spaces where traditional statistical techniques generally under-perform [3], due to the well-known *curse of dimensionality* or Hughes phenomenon [4]. Then, we explore joint spatial/spectral developments, able to exploit *a priori* knowledge about the shape of the objects in the scene and to take advantage of the spatial distribution, thus complementing the wealth of spectral information available from the data. Finally, our contribution outlines the development of innovative techniques for spectral mixture analysis and parallel processing support to overcome the extremely high computational requirements involved in most hyperspectral imaging applications.

II. Classification and segmentation techniques

A. New trends in classification using kernel methods

In this subsection, we first investigate the problem of local variation of spectral energy by the use of scale-invariant kernels. Then, we discuss transductive SVMs as a means to solve ill-posed problems induced by the limited amount of training samples in remote sensing applications. Finally, we introduce a new family of kernels able to include contextual/textural information in the classification process.

1) Use of different kernels: We begin by briefly recalling the general formulation of SVM classifiers. For a two-class problem, we assume that l training samples, $\mathbf{x}_i \in \mathbb{R}^n$, are available with their corresponding labels $y_i = \pm 1$, $S = \{(\mathbf{x}_i, y_i) \mid i \in [1, l]\}$. An optimal (in the linear sense) solution to this classification problem is given by finding the hyperplane that maximizes the margin, *i.e*, the distance to the closest training data points in both classes. Noting $\mathbf{w} \in \mathbb{R}^n$ as the vector normal to the hyperplane and $b \in \mathbb{R}$ as the bias, the hyperplane H_p is defined as $\langle \mathbf{w}, \mathbf{x} \rangle + b = 0$, $\forall \mathbf{x} \in H_p$, where $\langle \mathbf{w}, \mathbf{x} \rangle$ is the inner product between $\mathbf{w}$ and $\mathbf{x}$. If $\mathbf{x} \notin H_p$ then $f(\mathbf{x}) = \langle \mathbf{w}, \mathbf{x} \rangle + b$ is the distance of $\mathbf{x}$ to H_p. By definition, the hyperplane is a linear decision function in the input space. This choice could significantly limit the discriminant ability of the classifier. In order to address non-linear problems while, at the same time, keep the simplicity of linear models, it is possible to project the input data in higher dimensional feature spaces according to a non-linear mapping. While one increases the dimension of the feature space, one could run into computational problems. There is an effective solution, however, for computing inner products in feature spaces which consists of using the so-called *kernel trick*, which allows us to work in a mapped kernel space without knowing explicitly the mapping function Φ but only

0-7803-9510-7/06/$20.00 © 2006 IEEE

the kernel function k, formed by the dot product of mapping functions: $\langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle = k(\mathbf{x}_i, \mathbf{x}_j)$. It is important to note that k must fulfil *Mercer's Theorem* in order to be a valid kernel [5]. Several standard kernels are presented below.

- *Polynomial*. This kernel computes the inner product in the space of all monomials up to degree p as follows: $k_{poly}(\mathbf{x}, \mathbf{z}) = (\langle \mathbf{x}, \mathbf{z} \rangle + \theta)^p$. The parameter θ tunes the weight of the higher order polynomial. High values of θ reduce the relative weight of the higher order.
- *Gaussian Radial Basis Functions*. This kernel is given by $k_{gauss}(\mathbf{x}, \mathbf{z}) = \exp\left(-\gamma \|\mathbf{x} - \mathbf{z}\|^2\right)$. For this kernel, $k_{gauss}(\mathbf{x}, \mathbf{x}) = 1$. The parameter γ tunes the smoothness of the kernel. The Gaussian tends to a constant function for small values of γ, thus inhibiting the learning process.
- *Spectral Angle Mapper*. This kernel is given by $k_{SAM}(\mathbf{x}, \mathbf{z}) = \exp\left(-\gamma \alpha(\mathbf{x}, \mathbf{z})^2\right)$. It is based on the RBF framework and uses the spectral angle mapper (SAM), a scale invariant distance [6] given by $\alpha(\mathbf{x}, \mathbf{z}) = \arccos\left(\frac{\langle \mathbf{x}, \mathbf{z} \rangle}{\|\mathbf{x}\| . \|\mathbf{z}\|}\right)$. It incorporates *a priori* information about the spectrum profile of the hyperspectral data.

2) Exploitation of labeled and unlabeled samples for semi-supervised learning: Despite the proven effectiveness of standard kernel-based methods [7], the high dimensionality of hyperspectral imagery, along with the relatively small number of training samples available and the spatial variability of spectral signatures (which are non-stationary in the 2-D domain of the scene) usually result in *ill-posed* classification problems. In order to address this highly relevant issue, a transductive SVM (TSVM) approach has been recently developed [8]. It is based on the exploitation of *unlabeled* patterns in the framework of a transductive iterative learning process in which the most "informative" *unlabeled* samples are considered. Due to the fact that support vectors bear the richest information (i.e., they are the only patterns that affect the position of the separating hyperplane), unlabeled patterns closest to the margin bounds have the highest probability to be correctly classified among those samples in the margin band. With the above observation in mind, an iterative selection procedure has been designed considering a balanced number of transductive and labeled samples. This approach can greatly enhance classification accuracy, as demonstrated in [8]. In our conference presentation, we will provide experimental results for an improved version of this classifier in which important aspects have been enhanced (e.g., adaptive selection of the regularization parameter, novel model selection strategy).

3) Integration of contextual/textural information in kernel methods: The good classification performance demonstrated by kernel methods in general (and SVMs in particular [7]) can also be increased by including contextual (or even textural) information in the classifier. For this purpose, a pixel entity $\mathbf{x}_i$ may be redefined simultaneously both in the spectral domain using its spectral content, $\mathbf{x}_i^{\omega} \in \mathbb{R}^{N_\omega}$, and in the spatial domain by applying a feature extraction approach localized in its surrounding area, $\mathbf{x}_i^{s} \in \mathbb{R}^{N_s}$, which yields N_s spatial (contextual) features, e.g. the mean or standard deviation *per* spectral band. These separated entities lead to two different kernel matrices, which can be easily computed using any suitable kernel function that fulfills Mercer's conditions. At this point, one can sum spectral and textural dedicated kernel matrices (k_ω and k_s, respectively), and introduce the cross-information between textural and spectral features ($k_{\omega s}$ and $k_{s\omega}$) in the formulation. This simple methodology yields a full family of composite methods which substantially improve classification [9], as will be detailed at the conference presentation.

B. Integration of spatial and spectral information

As shown above, methodologies for integration of spatial and spectral information can be used via kernel methods to deal with the non-stationary nature of spectral signatures among image classes. This subsection describes further developments in this area. First, we describe how Markov random fields (MRFs) can be used to model spatial information as a complement to spectral-based analysis. Then, we address the extension of mathematical morphology (a classic spatial-based technique) to the spectral domain via the concept of morphological profiles. Finally, we address the concept of segmentation hierarchy and its utilization for joint spatial/spectral analysis of hyperspectral imagery.

1) Markov random fields and artificial neural networks: Spatial characterization of a pixel entity may be performed using MRFs, which are able to model the neighborhood of a pixel as a spatially distributed random process, and then accomplish a regularization process via the minimization of an energy function. Although MRF-based algorithms have already been proposed for hyperspectral data classification [10], in this work we introduce a peculiarity in the methodology which relies on the inclusion of *a priori* knowledge in the standard MRF framework. To do so, we first use a neuro-fuzzy classifier to perform spectral classification alone. The result of this stage is then fed to an MRF-based spatial analysis stage, which is performed jointly with a maximum likelihood spectral-based reclassification stage. In other words, the function to be minimized is computed by integrating the *pattern recognition* capability of the neuro-fuzzy classifier (which has a superior performance on single-pixel classification) with the MRF framework and its spatial/spectral capabilities.

2) Morphological profiles: In another attempt to integrate the rich spatial and spectral information available from hyperspectral data sets, Benediktsson *et al.* [11] have recently proposed the use of extended morphological profiles (EMPs). These profiles rely on classic mathematical morphology operations such as *erosion* and *dilation* [12], which are used to process a binary or grayscale image with a set of known shape, called structuring element (SE). These basic operators can be used to construct *opening* and *closing* operations, which remove from the original image the structures of size less than the SE, but they also modify structures which are still present in the image after the opening/closing. Since this issue may result in the loss of crucial information, opening and closing by reconstruction are generally the tool of choice. With these operators, if the structure of the image cannot

contain the SE then it is totally removed; otherwise it is totally preserved. A morphological profile is a composition of the *opening profile* (OP) and the *closing profile* (CP), so that the OP at the pixel x of a 2-D image I is defined as an n-dimensional vector $OP_i(x) = \gamma_R^{(i)}(x),\ \forall i \in [0, n]$, where $\gamma_R^{(i)}$ is the opening by reconstruction with an SE of a size i and n is the total number of openings. Similarly, the CP at x is defined as an n-dimensional vector $CP_i(x) = \phi_R^{(i)}(x),\ \forall i \in [0, n]$, where $\phi_R^{(i)}$ is the closing by reconstruction with an SE of a size i. We have $CP_0(x) = OP_0(x) = I(x)$. By collating the OP and the CP, the morphological profile of image I is defined as $2n + 1$-dimensional vector $MP(x) = CP_n(x), \ldots, I(x), \ldots, OP_n(x)$. In order to extend this concept to multi-channel imagery, we first need to select the spectral components on which the morphological approach is applied. Previous work has used the full spectral information to construct morphological profiles [13]. In this work, we propose to use the first principal components (PCs) that contain more than a certain percentage of explained variance, and then build the profile on each individual PC. The extended profile will then be a single stacked vector used as an input feature for classification.

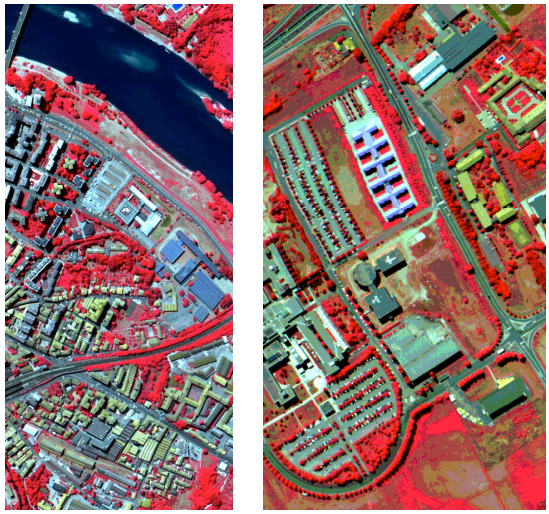
3) Use of segmentation hierarchies: A segmentation hierarchy is a set of several segmentations of the same image at different levels of detail, in which the segmentations at coarser levels of detail can be produced from simple merges of regions at finer levels of detail [14]. Segmentation hierarchies may be formed through a region growing approach to image segmentation. In hyperspectral imaging, spatially adjacent regions iteratively merge through a specified merge selection process based on spectral criteria. The hierarchical segmentation (HSEG) algorithm allows spatially non-adjacent regions to merge, controlled by a threshold parameter based on previous merges of spatially adjacent regions. HSEG also includes a method for selecting the most *significant* iterations from which the segmentation result is saved to form an output segmentation hierarchy. A recursive approximation (RHSEG) will be used in experiments to alleviate the heavy computational demands introduced by merging of spatially non-adjacent regions.

III. Experiments

A. Hyperspectral data sets

Although results with a variety of data sets collected by different sensors (including AVIRIS) have been obtained and will be presented at the conference, this section focuses exclusively on a case study based on urban hyperspectral data over the town of Pavia, Italy. The data were collected by the ROSIS sensor, operated by DLR (German Aerospace Agency), with spectral coverage ranging from 0.43 to 0.86 μm. The data are very fine spatial resolution (1.3-meters per pixel) and contains a total of 115 spectral bands. Two different subsets of the full data set are considered in the experiments:

1) *Subset #1.* This subset, with 492 by 1096 pixels in size, was collected over Pavia city centre, Italy. It contains 102 spectral channels after removal of noisy bands [see Fig. 1(a) for a color composite]. Nine ground-truth classes were considered in experiments: *Asphalt*, Meadow, *Gravel*, *Trees*, *Metal sheet*, *Bare soil*, *Bitumen*, *Bricks* and *Shadow*.
2) *Subset #2.* This subset, with size of 610 by 340 pixels, is centered at the University of Pavia. It contains 103 spectral channels after removal of noisy bands [see Fig. 1(b)]. Nine ground-truth classes were considered in experiments: *Water*, *Trees*, *Grass*, *Parking lot*, *Bare Soil*, *Asphalt*, *Bitumen*, *Tiles* and *Shadow*.
3) *Subset #3.* Finally, a superset of the scene in Fig. 1(a) which included a dense residential area (located west of the imaged area) was also used.



(a) (b)

Fig. 1. ROSIS urban hyperspectral data: Subset #1 (a) and Subset #2 (b).

B. Quantitative analysis

1) Kernel methods: The scene labeled as Subset #1 was selected for experiments with kernel methods. The training set consisted of a total of 5535 samples, while the test set consisted of 103505 samples. Small training sets were randomly extracted from the training set, composed of 10, 20, 40, 60, 80 and 100 pixels by class, respectively. The SVMs were then trained with each of these training subsets and then evaluated using the whole test set. Each experiment was repeated five times and the mean accuracy values were reported. Three kernels were used: polynomial, Gaussian RBF and SAM. Kernel parameters were adjusted empirically to maximize the estimated overall accuracy, which was computed using a fivefold *cross validation*. Table I summarizes the results obtained. Overall, we can conclude that SVMs are not drastically affected by the high dimensionality of training samples, and good generalization performance is obtained: with only 10 training pixels per class, more than 90% accuracy was reached by all considered kernels.

2) Markov random fields and artificial neural networks: The use of a MRF model for contextual classification of

TABLE I

CLASSIFICATION ACCURACIES (%) FOR DIFFERENT SVM KERNELS.

Training Set Size		10	20	40	60	80	100	All
RBF	OA	93.85	94.51	94.51	94.71	95.36	95.29	96.45
	AA	88.76	91.00	92.66	92.04	93.24	93.39	95.08
	Kappa	0.90	0.91	0.92	0.91	0.92	0.92	0.94
Poly	OA	92.34	92.77	94.20	94.07	94.29	94.81	96.03
	AA	87.87	88.91	91.74	92.41	92.31	93.35	94.91
	Kappa	0.87	0.88	0.90	0.90	0.90	0.91	0.93
SAM	OA	93.32	93.87	93.79	94.23	94.40	94.54	95.56
	AA	86.36	88.64	91.26	91.67	91.89	92.61	94.26
	Kappa	0.89	0.90	0.90	0.90	0.90	0.91	0.93

TABLE II

CLASSIFICATION ACCURACIES (%) FOR DAFE/MRF AND ARTMAP.

	DAFE/MRF	ARTMAP
OA	97.27	97.29
Water	99.04	99.71
Trees	91.27	93.19
Grass	97.09	94.80
Parking lot	76.83	71.99
Bare soil	93.33	93.36
Asphalt	99.65	81.87
Bitumen	88.45	96.42
Tiles	98.33	99.98
Shadow	99.86	99.93

spectral and spatial features is illustrated using Subset #3. It should be noted that a feature reduction step based on discriminant analysis feature extraction (DAFE) [3] was applied prior to MRF-based spatial characterization to increase spectral separability. For sake of comparison, results obtained by a neuro-fuzzy spatial/spectral classifier are also reported. Table II shows that the two considered approaches result in almost the same classification accuracies, mainly because they both provide integration of spatial details into the classification based on comparison of the spectra. It should be noted that the DAFE/MRF technique was applied using a fixed 3×3 pixel window. In experiments, we observed that using a wider window for the whole data set resulted in lower accuracies because the spatial resolution in border areas was not fully exploited. In turn, we expect that an adaptive procedure based on using wider windows in homogeneous areas and smaller windows at the borders between regions may result in more accurate classification results.

3) Morphological profiles: In order to illustrate the performance of morphological approaches, we use Subset #2. The training set consisted of a total of 3921 samples, while the test set consisted of 400002 samples. PCA was first applied to the original data set, and 3 PCs were selected (representing 99% of the cumulative variance). Morphological profiles with 10 openings/closings by reconstruction (using a disk with step size increment of 1 as SE) were then constructed for all training samples. As a result, each profile was made up of 21 bands and the final EMP, constructed using three principal components, consisted of 63 bands. These extended profiles were then used to train an SVM classifier with RBF kernel, in which parameters were tuned using a five-fold cross validation.

TABLE III

CLASSIFICATION ACCURACIES (%) FOR SVMS WITH THE ORIGINAL SPECTRAL INFORMATION AND EXTENDED MORPHOLOGICAL PROFILES.

	Spectral Information	Extended Morphological Profile
OA	80.99	85.22
AA	88.28	90.76
Kappa	76.16	80.86
Asphalt	83.71	95.36
Meadow	70.25	80.33
Gravel	70.32	87.61
Tree	97.81	98.37
Metal sheet	99.41	99.48
Bare Soil	92.25	63.72
Bitumen	81.58	98.87
Brick	92.59	95.41
Shadow	96.62	97.68

Table III shows the classification results for the test set. It is clear that the use of morphological profiles instead of the original spectral information resulted in significantly improved classification results for most classes. This was not the case, however, for the *Bare soil* class. One possible explanation is that three components only can preclude an accurate modeling of the original spectral content. In spite of this fact, results in Table III indicate that EMPs provide a good way to perform joint *spatial/spectral* classification of hyperspectral images.

4) Segmentation hierarchies: A key concern in applying RHSEG to hyperspectral data is selecting an appropriate dissimilarity function. Currently available functions in RHSEG are all measures of the difference of the region mean (computed spectrally) from the variance of spectral signatures in the region. Therefore, a measure of the appropriateness of a dissimilarity function would be the classification accuracy obtained when each ground-truth pixel is assigned to the nearest region (as per a particular dissimilarity function) when the region means are initially modeled as the means of pixels labeled as a particular ground cover class. Using the above criterion, RHSEG was used to explore the effect of adding spatial information. When input parameter $spclust\_wght$ was set to 1.0 (which means that spatial information is *not* taken into account) the overall classification accuracy achieved for Subset #1 with a band summed mean squared error (BSMSE) criterion was 90.5%. However, setting of $spclust\_wght = 0.1$ (which means that spatial information is incorporated) resulted in an overall accuracy of 97.7%. Using the same parameter settings above, RHSEG was now used to process Subset #1 without initialization with the ground-truth data, but with each pixel being a separate region. The overall accuracy was 96.5% (the *Meadows* class was not well separated, with 36% accuracy). In doing the evaluation, the coarsest hierarchical segmentation level that separated all (or most) ground cover classes was selected, and the region segments were assigned to a ground cover class if a plurality of the pixels in the region were covered by the ground-truth for that particular class.

IV. DISCUSSION

Although the main body of this paper has been focused on classification and segmentation techniques, several important issues could not be addressed due to space considerations. Two

of the most relevant aspects, which will be fully developed at the conference, are the mixed nature of hyperspectral pixels and the need for computationally efficient algorithm implementations.

A. Spectral mixture analysis

Spectral mixture analysis (or unmixing) has been an alluring exploitation goal since the earliest days of imaging spectroscopy. No matter the spatial resolution, in natural environments the spectral signature for a nominal pixel is invariably a mixture of the signatures of the various materials found within the spatial extent of the ground instantaneous field view. In hyperspectral imagery, the number of spectral bands usually exceeds the number of spectral mixture components and the unmixing problem is cast in terms of an overdetermined system of linear equations in which, given the correct set of spectral endmembers, the actual unmixing to determine apparent pixel fractions is a simple and straightforward numerical process. Since each observed spectral signal is the result of an actual mixing process, the driving abundances must obey two rather common-sense constraints. First, all abundances must be non-negative. Second, the sum of abundances for a given pixel must be unity. However, it is the derivation and validation of the correct suite of endmembers that has remained a challenging and elusive goal for the past twenty years. Perhaps the best known approach for deriving spectral endmembers is the Pixel Purity Index method [15], embedded in the ENVI software package. Although this method has shown considerable promise, much further work still remains, especially in the area of exploiting the joint spatial and spectral nature of the data. In our conference presentation, we will develop new endmember extraction methodologies for joint spatial/spectral exploitation of the inherent n-dimensional convexity of mixed pixels.

B. Parallel processing support

While advances in data processing techniques such as integrated spatial/spectral developments hold great promise for hyperspectral image analysis, they also introduce new processing challenges. Several applications exist, however, where having the desired information calculated quickly enough for practical use is highly desirable. With the recent explosion in the amount and dimensionality of hyperspectral imagery, parallel processing is expected to become a tool of choice in many remote sensing missions, especially with the advent of low-cost systems such as commodity clusters. Along with the continuous doubling (roughly, every two years) in commodity processing capability, the number of problems that can be feasibly solved on parallel or distributed resources is ever increasing. The best situation arises when a problem can be decomposed into a non-trivial number of independent subproblems. This kind of problem can be mapped to any parallel/distributed implementation and achieve good scalability. While several hyperspectral image processing techniques can be tackled under the above assumption (e.g., mathematical morphology-based approaches), many other data processing techniques are not so *embarrasingly parallel*. In our conference presentation, we will explore the performance of parallel implementations of discussed algorithms using high performance parallel computers such as NASA's Thunderhead (http://thunderhead.gsfc.nasa.gov) or the University of Zürich's Matterhorn (http://www.matterhorn.unizh.ch).

V. Conclusions

Earth observation-related imaging spectroscopy has been transformed in less than 30 years from being a sparse research tool into a commodity product available to a broad user community. As a result, there is an emerging need for standardized hyperspectral data processing chains, able to assimilate consolidated analysis techniques and to take advantage of last-generation computing environments. In this paper, we have taken a necessary first step towards the understanding and assimilation of advanced, high-level hyperspectral image processing techniques. New trends in algorithm design, including the integrated use of spatial and spectral information and the appropriate exploitation of limited training samples, have been specifically addressed and discussed.

References

[1] A. F. H. Goetz, G. Vane, J. E. Solomon, and B. N. Rock, "Imaging spectrometry for earth remote sensing," *Science*, vol. 228, pp. 1147–1153, 1985.

[2] R. O. Green et al., "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (aviris)," *Remote Sensing of Environment*, vol. 65, pp. 227–248, 1998.

[3] D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*, John Wiley and Sons, 2003.

[4] G. F. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Trans. Inform. Theory*, Jan. 1968.

[5] B. Scholkopf and A. J. Smola, *Learning with Kernels*, MIT Press, 2002.

[6] G. Mercier and M. Lennon, "Support vector machines for hyperspectral image classification with spectral-based kernels," in *Geoscience and Remote Sensing Symposium*. IGARSS Proceedings, July 2003.

[7] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 6, pp. 1351– 1362, June 2005.

[8] L. Bruzzone, M. Chi, and M. Marconcini, "A novel transductive svm for the semisupervised classification of remote-sensing images," *IEEE Trans. Geosci. Remote Sensing*, In press 2006.

[9] G. Camps-Valls, L. Gomez-Chova, J. Muñoz-Marí, J. Vila-Francés, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Geosci. Remote Sensing Letters*, vol. 3, no. 1, pp. 93–97, Jan 2006.

[10] P. Gamba, F. Dell'Acqua, A. Ferrari, J. A. Palmason, J. A. Benediktsson, and K. Arnasson, "Exploiting spectral and spatial information in hyperspectral urban data with high resolution," *IEEE Geosci. Remote Sensing Letters*, vol. 1, no. 4, pp. 322–326, Oct 2004.

[11] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geosci. Remote Sensing*, vol. 42, no. 3, pp. 480–491, Mar. 2005.

[12] P. Soille, *Morphological Image Analysis: Principles and Applications*, Germany:Springer-Verlag, 2003.

[13] A. Plaza, P. Martinez, J. Plaza, and R. Perez, "Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations," *IEEE Trans. Geosci. Remote Sensing*, vol. 43, no. 3, pp. 466–479, Mar. 2005.

[14] J. C. Tilton, "Image segmentation by region growing and spectral clustering with a natural convergence criterion," in *Proceedings of IGARSS 1998*, 1998, pp. 1766–1768.

[15] J. Boardman, F. A. Kruse, and R. O. Green, "Mapping target signatures via partial unmixing of aviris data," in *Summaries of JPL Airborne Earth Science Workshop*, 1995.