

Comparison of Support Vector Machine-Based Processing Chains for Hyperspectral Image Classification

Marta Rojas^a, Inmaculada Dópido^a, Antonio Plaza^a and Paolo Gamba^b

^aHyperspectral Computing Laboratory
Department of Technology of Computers and Communications
University of Extremadura, Avda. de la Universidad s/n
10071 Cáceres, Spain

^bTelecommunications and Remote Sensing Laboratory
Department of Electronics
University of Pavia, Via Ferrata, 1
27100 Pavia, Italy

ABSTRACT

Many different approaches have been proposed in recent years for remotely sensed hyperspectral image classification. Despite the variety of techniques designed to tackle the aforementioned problem, the definition of standardized processing chains for hyperspectral image classification is a difficult objective, which may ultimately depend on the application being addressed. Generally speaking, a hyperspectral image classification chain may be defined from two perspectives: 1) the provider's viewpoint, and 2) the user's viewpoint, where the first part of the chain comprises activities such as data calibration and geo-correction aspects, while the second part of the chain comprises information extraction processes from the collected data. The modules in the second part of the chain (which constitutes our main focus in this paper) should be ideally flexible enough to be accommodated not only to different application scenarios, but also to different hyperspectral imaging instruments with varying characteristics, and spatial and spectral resolutions. In this paper, we evaluate the performance of different processing chains resulting from combinations of modules for dimensionality reduction, feature extraction/selection, image classification, and spatial post-processing. The support vector machine (SVM) classifier is adopted as a baseline due to its ability to classify hyperspectral data sets using limited training samples. A specific classification scenario is investigated, using a reference hyperspectral data set collected by NASA's Airborne Visible Infra-Red Imaging Spectrometer (AVIRIS) over the Indian Pines region in Indiana, USA.

Keywords: Hyperspectral image classification, support vector machine (SVM), dimensionality reduction, feature extraction/selection, spatial post-processing.

1. INTRODUCTION

The Hyperspectral Imaging Network (Hyper-I-Net)* is a four-year Marie Curie research training network[†] designed to build an interdisciplinary European research community focused on hyperspectral imaging activities. One of the main activities of Hyper-I-Net is to settle the basis for the definition and testing of a flexible hyperspectral data collection and processing chain, in which individual elements can be integrated in such a way that the resulting chain can be dynamically adapted and reconfigured to satisfy the requirements of different application scenarios with little effort.¹ Since efficient hyperspectral data processing can be a really complex procedure, Hyper-I-Net approaches this problem in the context of a multidisciplinary collaboration so that the proposed activity can benefit from the complementary expertise of partners with focus in heterogeneous disciplines such as sensor design and calibration, pattern recognition, signal and image processing, and Earth observation related

Send correspondence to Antonio J. Plaza:

E-mail: aplaza@unex.es; Telephone: +34 927 257000 (Ext. 51662); URL: <http://www.umbc.edu/rssipl/people/aplaza>

*<http://www.hyperinet.eu>

[†]<http://cordis.europa.eu/mc-opportunities>

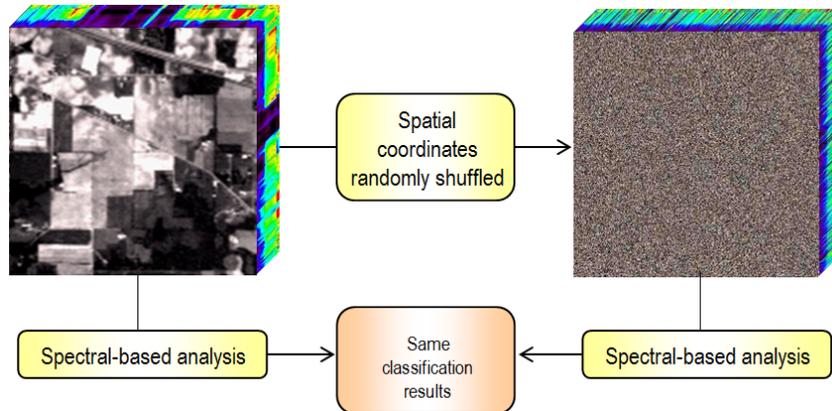


Figure 1. Example illustrating the importance of spatial information in hyperspectral image classification.

products.² The outcome of this joint activity is expected to be a set of hardware/software processing techniques able to deal with the complexity of hyperspectral data in an effective manner.

Generally, a hyperspectral data processing chain may be defined from two perspectives: 1) the provider's viewpoint, and 2) the user's viewpoint, where the first approach generally results in the first part of the chain, and the second approach results in a chain that is placed immediately afterwards. As a result, the first part of the chain comprises data correction activities such as sensor specification, geometric corrections, radiometric calibrations, etc.² Once the data has been pre-processed and geo-corrected, an information extraction process from collected data is needed for the second part of the chain. Several processing steps are available in the literature for this purpose,^{3,4} including data transformation (e.g., for dimensionality reduction), spectral matching (requiring centralized spectral libraries of multiple materials), feature extraction/selection, and data classification. Another important and essential requirement for the user-oriented part of the processing chain is to define precisely characterized and accurately validated high-level products. It is important to note that the elements in the user-oriented chain should be flexible enough to be accommodated not only to different application scenarios, but also to different hyperspectral imaging instruments with varying spatial and spectral resolutions.

On the other hand, an important consideration in any hyperspectral image processing chain is the inclusion of spatial information in the analysis.⁵⁻⁷ Specifically, one of the distinguishing properties of hyperspectral data is the multivariate information coupled with a two-dimensional (pictorial) representation amenable to image interpretation.⁸ Subsequently, spectral-based classification chains may not be able to accurately model spatial dependencies in the scene. Recent efforts in the literature have demonstrated that hyperspectral image classification can greatly benefit from an integrated framework in which both the spectral information and the spatial arrangement of pixel vectors are taken into account. An example of this situation is given in Fig. 1, in which a hyperspectral data cube is modified by randomly permuting the spatial coordinates of the pixel vectors, thus removing the spatial correlation. In both scenes, the application of a spectral-based classification method would yield the same analysis results while a spatial-spectral classifier could incorporate more effectively the spatial information present in the original scene into the analysis process.

In this work, we adopt a user-oriented perspective and further explore the suitability of defining a flexible hyperspectral processing chain in two different application domains, namely, urban environment monitoring and vegetation mapping. These are complex problems which may serve as adequate case studies to demonstrate the validity of our approach in a real application context, using a limited number of processing steps for a preliminary assessment focused on dimensionality reduction, feature selection/extraction (possibly including spatial information), data classification, and spatial post-processing. Specifically, our processing chains are all based on the support vector machine (SVM) classifier,^{9,10} which has demonstrated an excellent ability to classify hyperspectral data sets using limited training samples.¹¹ A hyperspectral data set collected by NASA's Airborne Visible Infra-Red Imaging Spectrometer (AVIRIS) over the Indian Pines region in Indiana, USA is used to illustrate our findings.

The remainder of the paper is organized as follows. Section 2 describes the individual modules that will be used to form our proposed hyperspectral data processing chains. Section 3 describes several different processing chains that will be evaluated as possible solutions for the considered problem. Section 4 validates the previously introduced processing chains using the aforementioned hyperspectral data set. Finally, section 5 concludes with some remarks and hints at plausible future research lines.

2. PROCESSING MODULES

This section is intended to provide a general overview of the processing modules that will be used for the second part (i.e., user-oriented) of the hyperspectral processing chain. Usually, these techniques are subdivided into different groups, such as spectral analyses, spatial analyses, and spatial/spectral analyses.¹ In our preliminary approach towards the development of a flexible hyperspectral processing chain, the following techniques have been considered:

- *Dimensionality reduction.* In hyperspectral imaging, there is clear need for methods that can reduce the dimensionality of the data to the right subspace without losing the original information that allows for the separation of classes.^{3,4} Standard spectral-based transformations such as the principal component analysis (PCA) and minimum noise fraction (MNF) will be used to transform input data to a dataset in a new uncorrelated coordinate system.
- *Feature selection.* Another interesting approach to reduce input data dimensionality has been the selection of the most highly relevant spectral bands for data exploitation. Along with traditional approaches, a new criterion based on maximum band separation index in terms of entropy is considered to select the most relevant input bands prior to the analysis.
- *Feature extraction.* One of the distinguishing properties of hyperspectral data, as collected by available imaging spectrometers, is the multivariate information coupled with a two dimensional pictorial representation amenable to image interpretation. However, feature extraction from hyperspectral data has been traditionally carried out without incorporating information on the spatially adjacent data. In this work, multiscale texture features¹² and morphological features⁸ are tested to accurately characterize spatial information jointly with spectral information. A novel approach based on using spectral unmixing concepts (endmember extraction plus fractional abundance estimation) for feature extraction prior to classification is also considered.
- *Classification.* Often, standard supervised classifiers, for high dimensional data like hyperspectral images, require large volumes of training data, which have to be obtained by costly ground truth measurements. In this work, we resort to a classifier such as the SVM which has demonstrated its ability to perform accurately in scenarios dominated by limited training samples, thus providing a good compromise between the extremely large dimensionality of hyperspectral data and the often limited availability of training data, due to the difficulties to obtain such training information in real scenarios as a matter of time and finance. To investigate the potential of spatial information within the classification procedure, the previously generated classification images were reprocessed using a spatially-aware approach based on a fuzzy neural network structure.

The impact of the processing modules above is objectively quantified in this work by implementing different chains made up of different combinations of such modules. In the following section, we propose several different chains that will be substantiated by experiments using real hyperspectral data.

3. HYPERSPECTRAL PROCESSING CHAINS

A general, user-oriented hyperspectral processing chain can be described by a small group of interconnected processing modules or blocks. In fact, after pre-processing the input data set (e.g., via radiometric calibrations, geometric or atmospheric corrections, etc.) relevant information can be extracted from the scene. Assuming that we start from a corrected data set, three principal steps can be used to define a simple, user-oriented

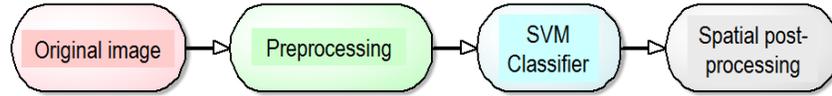


Figure 2. Generic processing chain.

processing chain: data transformation and feature extraction, feature selection, and classification (see Fig. 2). It should be noted that the representation given in Fig. 2 is a very general one, in which each module can be implemented in very different ways. For instance, feature extraction may be constituted by a spectral analysis procedure (e.g., PCA or MNF) or by a spatial-oriented processing (such as textures, semivariograms), or even by multi-scale analysis (differential morphological profiles). On the other hand, feature selection provides best combinations of bands and/or features for a given problem, but the combination of these features is complex and thus, fast and efficient algorithms are necessary. Finally, classification may also be accomplished by different techniques, including approaches in the spectral domain (e.g., spectral angle-based classification), in the spatial domain (context-aware classifiers), or even in both domains (object-oriented classification, morphology-based classifiers). In this work, we have selected representative techniques to generally describe each processing module. Specifically, feature extraction is described by PCA and MNF. Moreover, other techniques such as texture and morphological analysis, or linear spectral unmixing, are included in some chains to obtain best features from different perspectives (spatial information, mixture information). On the other hand, the feature selection step is supplied by separation index analysis, which permits an efficient selective procedure to choose best subsets, starting from high-dimensional data. Finally, in all cases we include a post-processing stage which performs spatial re-processing using a spatially-aware approach based on a fuzzy neural network structure to increase the spatial consistency of the final classification results.¹³ In the following we outline the six hyperspectral processing chains considered in our study:

- **Processing chain #1.** This chain comprises feature selection from the original hyperspectral image using separation index analysis, followed by spectral-based SVM classification on the retained features, and spatial post-processing.
- **Processing chain #2.** This chain comprises feature extraction from the original hyperspectral image using the PCA transformation, followed by spectral-based SVM classification on the extracted features, and spatial post-processing.
- **Processing chain #3.** This chain comprises feature extraction from the original hyperspectral image using the MNF transformation, followed by spectral-based SVM classification on the extracted features, and spatial post-processing.
- **Processing chain #4.** This chain comprises feature extraction from the original hyperspectral image using linear spectral unmixing¹⁴ (endmember extraction using the N-FINDR method¹⁵ plus abundance estimation using fully constrained least squares linear spectral unmixing¹⁶), followed by spectral-based SVM classification and spatial post-processing.
- **Processing chain #5.** This chain comprises feature extraction from the original hyperspectral image using the MNF transformation, followed by texture-based feature extraction¹² in the MNF space, followed by spectral-based SVM classification and spatial post-processing.
- **Processing chain #6.** This chain comprises feature extraction from the original hyperspectral image using the MNF transformation, followed by morphological-based feature extraction⁸ in the MNF space, followed by spectral-based SVM classification and spatial post-processing.

In all the aforementioned processing chains, the number of features selected/extracted was varied in experiments (different tests were performed and we report those which resulted in the best overall results in terms of classification accuracy). Future work should comprise an exploration of appropriate mechanisms for automatically determining the optimal number of features to be retained for classification experiments.

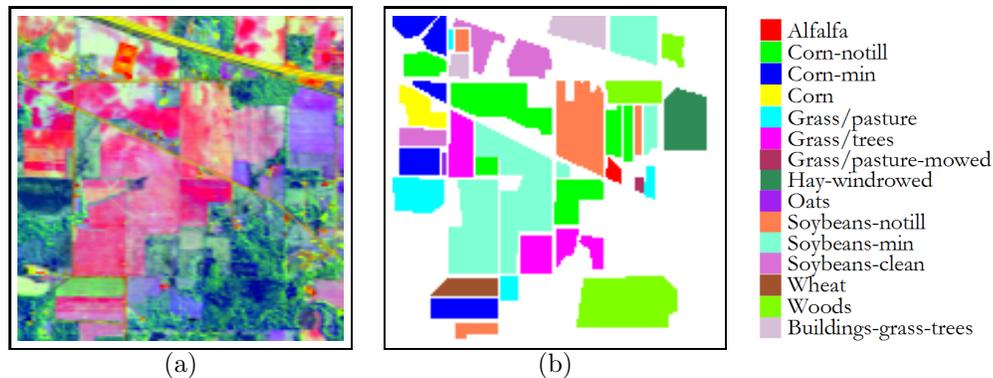


Figure 3. (a) False color composition of the AVIRIS Indian Pines scene. (b) Ground truth-map containing 15 mutually exclusive land-cover classes (right).

4. EXPERIMENTAL RESULTS

4.1 Hyperspectral data set

The data set used in our experiments was collected by the AVIRIS sensor over the Indian Pines region in Northwestern Indiana in 1992. This scene, with a size of 145 lines by 145 samples, was acquired over a mixed agricultural/forest area, early in the growing season. The scene comprises 202 spectral channels in the wavelength range from 0.4 to 2.5 μm , nominal spectral resolution of 10 nm, moderate spatial resolution of 20 meters by pixel, and 16-bit radiometric resolution. After an initial screening, several spectral bands were removed from the data set due to noise and water absorption phenomena, leaving a total of 164 radiance channels to be used in the experiments. For illustrative purposes, Fig. 3(a) shows a false color composition of the AVIRIS Indian Pines scene, while Fig. 3(b) shows the ground-truth map available for the scene, displayed in the form of a class assignment for each labeled pixel, with 15 mutually exclusive ground-truth classes. These data, including ground-truth information, are available online[‡], a fact which has made this scene a widely used benchmark for testing the accuracy of hyperspectral data classification algorithms.

4.2 Experimental design

Before describing the results obtained in experimental validation, we first briefly describe the adopted supervised classification system. Firstly, depending on the considered chain, relevant features for classification are selected/extracted from the original image. The resulting features are used to train an SVM classifier in which four types of kernels: linear, polynomial, Gaussian RBF, and sigmoid were used (since the best results were always obtained using the Gaussian RBF kernel,⁹ we only report results obtained using this configuration). Specifically, the SVM was trained different training subsets (all comprising 10% of the available ground-truth) and then evaluated with the remaining test set. Each experiment was repeated ten times, and the mean accuracy values were reported to guarantee the statistical significance of the results. Kernel parameters were optimized in all experiments by a grid search procedure. In essence, the SVM classification is based on the notion of fitting an optimal separating hyperplane between classes by focusing on the training samples that lie at the edge of the class distributions, the support vectors. All of the other training samples are effectively discarded as they do not contribute to the estimation of hyperplane location. In this way not only is an optimal hyperplane fitted, in the sense that it is expected to have a large degree of generalizability, but also a high accuracy may be obtained with the use of a small training set.

4.3 Results and discussion

In the following, we evaluate the performance of the different chains in the task of classifying the AVIRIS Indian Pines scene. In order to illustrate the performance of the different chains, we have decided to show actual classified scenes instead of overall accuracies, despite the fact that each classification experiment was repeated ten times to guarantee the statistical significance of our results. This is because we feel that visual interpretation

[‡]<http://dynamo.ecn.purdue.edu/biehl/MultiSpec>

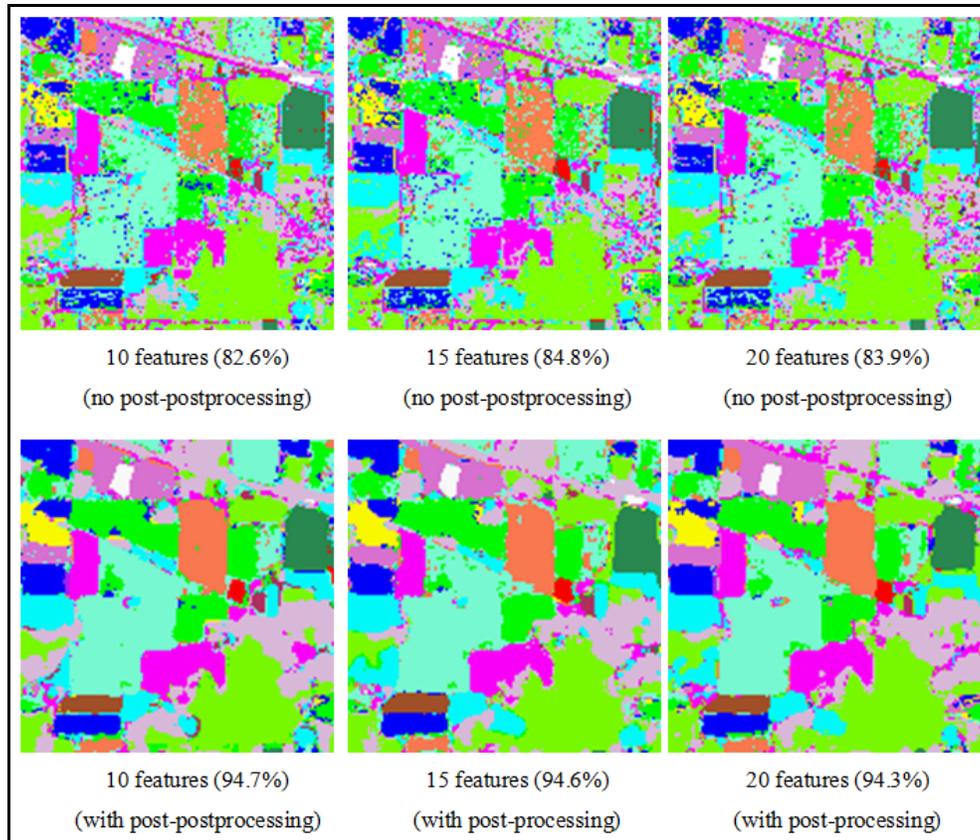


Figure 4. Classification results obtained by processing chain #1.

is also very important in analyzing the quality of classification results. With this observation in mind, in the following we discuss the performance achieved by the different processing chains in the task of classifying the AVIRIS Indian Pines scene and finally we discuss the results globally prior to concluding the paper with some observations and remarks.

4.3.1 Results obtained by processing chain #1

Fig. 4 shows the most relevant results obtained with processing chain #1. Specifically, three different results are displayed for the feature selection stage (10, 15 and 20 selected features, respectively) and spatial post-processing results are also displayed in all cases. As shown by Fig. 4, it can be seen that SVMs generalize quite well: with only 10% of training pixels per class, at least 80% overall classification accuracy is reached in all cases. Interestingly, classification accuracies decreased when the number of selected features were increased from 15 to 20, but it is also worth noting that the classification results are quite similar in both cases. This confirms the fact that SVMs are not significantly affected by the Hughes phenomenon, in particular, when performing feature selection/extraction prior to classification. Finally, it can be noticed that spatial post-processing significantly increased classification results in all cases, increasing the overall classification accuracy in approximately 10% or more in all considered cases. This is due to the spatial homogeneity of the ground-truth classes in Fig. 3(b). It should also be noticed that the final classification maps after spatial post-processing appear a bit degraded from a spatial point of view, but classification scores improve due to the spatial consistency of the resulting classes which increases the similarity with regards to the ground-truth.

4.3.2 Results obtained by processing chain #2

The main difference between processing chain #2 and processing chain #1 is that the feature selection stage in processing chain #1 is replaced by a PCA-based feature extraction stage. Fig. 5 shows the most relevant results obtained with processing chain #2. Specifically, three different results are displayed for the PCA-based

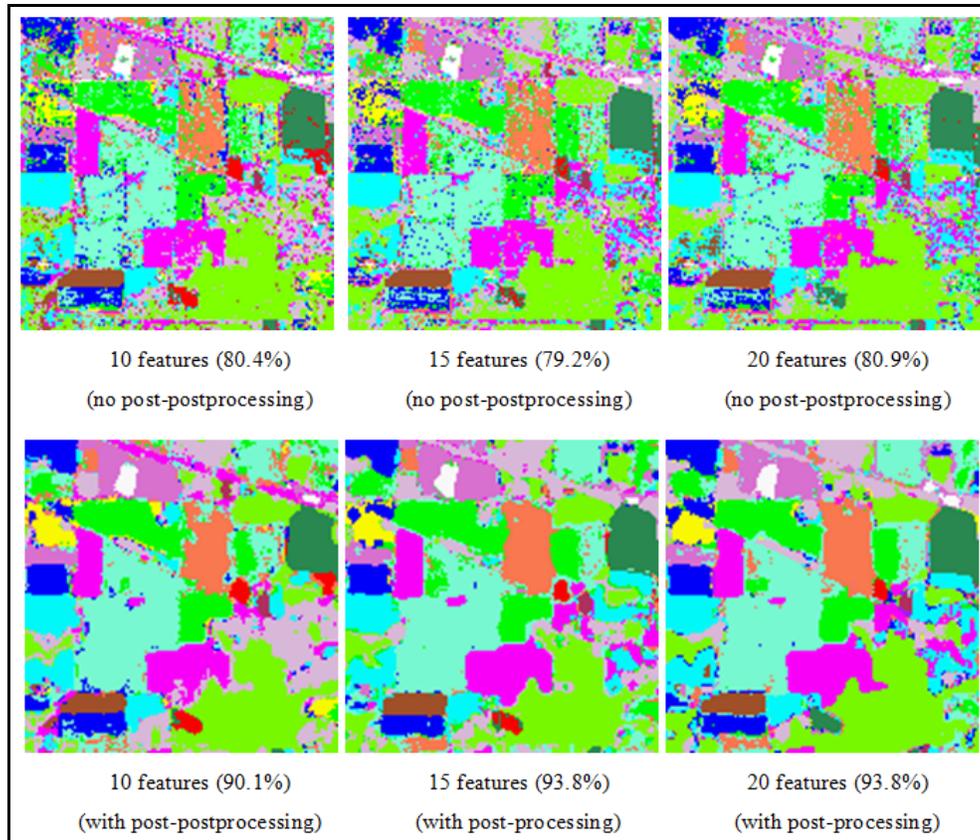


Figure 5. Classification results obtained by processing chain #2.

feature extraction stage (10, 15 and 20 extracted features, respectively) and spatial post-processing results are also displayed in all cases. As shown by Fig. 5, the results obtained by the PCA feature extraction module cannot improve those achieved by using band selection in processing chain #1. Again, spatial post-processing significantly increased classification results in all cases, increasing the overall classification accuracy in 10% or more in all considered cases. This confirms that spatial post-processing is very important in order to enhance the comparison with the ground-truth classes in Fig. 3(b), although it is also noticed that the final classification maps after spatial post-processing exhibit some spatial distortion.

4.3.3 Results obtained by processing chain #3

The main difference between processing chain #3 and processing chain #2 is that the PCA-based feature extraction stage is replaced by an MNF-based feature extraction stage. While the PCA orders the transformed features in terms of variance, the MNF orders the transformed features in terms of signal-to-noise ratio (SNR) and performs a better characterization of the noise present in the input hyperspectral image when obtaining the transformed features. Fig. 6 shows the most relevant results obtained with processing chain #3. Specifically, three different results are displayed for the MNF-based feature extraction stage (10, 15 and 20 extracted features, respectively) and spatial post-processing results are also displayed in all cases. As shown by Fig. 6, the results obtained by the MNF feature extraction significantly improve those obtained using the PCA transform in processing chain #2 and band selection in processing chain #1, achieving 91.2% classification accuracy with 10 extracted features without spatial post-processing. This represents a significant increase in classification accuracy with regards to the first two considered chains. When spatial post-processing is applied, classification results increase to 96% or more for 15 and 20 features, which is a very significant result taking in mind the complexity of this scene, dominated by mixed pixels and very similar spectral classes. Interestingly, it can be visually appreciated that the use of MNF transform as feature extraction helps reducing the spatial distortion in the final classified maps after spatial post-processing. Interestingly, while in the results without spatial post-processing it

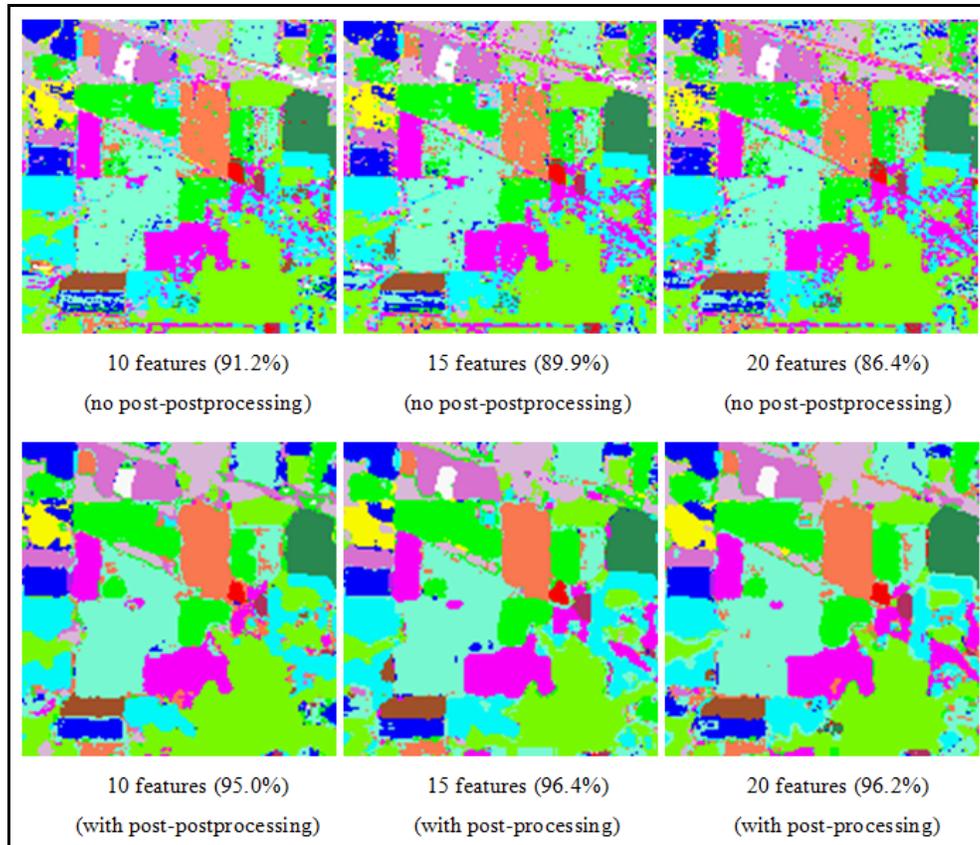


Figure 6. Classification results obtained by processing chain #3.

seems that 10 features are more appropriate than 15 or 20, the best overall results after spatial post-processing are obtained for both 15 and 20 MNF features, which indicates that the number of retained components is a topic deserving future research and that should be evaluated in light of the impact of possible spatial post-processing applied to the classification results.

4.3.4 Results obtained by processing chain #4

The main difference between processing chain #4 and processing chains #2 and #3 is that the feature extraction stage is replaced by linear spectral unmixing. The rationale for using an unmixing-based feature extraction approach presents some distinctive features with regards to PCA and MNF-based feature extraction. First, it provides useful information for classification in hyperspectral analysis scenarios with moderate spatial resolution, since the sub-pixel composition of training samples can be used as part of the learning process of the classifier. Second, it can model the non-stationary behavior of the spectral signatures of land-cover classes in the spatial domain of the scene, since (possibly disjoint) regions belonging to the same class are represented by the same spectral signature, and the variations related with different cover proportions or illumination conditions are modeled via the abundance estimation process inherent in spectral unmixing. Third, the components estimated by the proposed feature extraction strategy exhibit physical meaning. A final distinctive feature is that it does not penalize classes which are not relevant in terms of variance or signal-to-noise ratio (SNR). In our experiments, we have performed linear spectral unmixing for feature extraction strategy using the following steps: 1) extracting the purest spectral signatures (endmembers¹⁷) from the scene using the N-FINDR,¹⁵ a volume-based fully automatic algorithm; and 2) estimating the abundance fractions associated to the spectral endmembers using fully constrained linear spectral unmixing.^{16, 18, 19}

Fig. 7 shows the most relevant results obtained with processing chain #4, which correspond to extraction of 15 endmembers (based on a previous estimation of the number of endmembers in the data^{20, 21}) and estimation

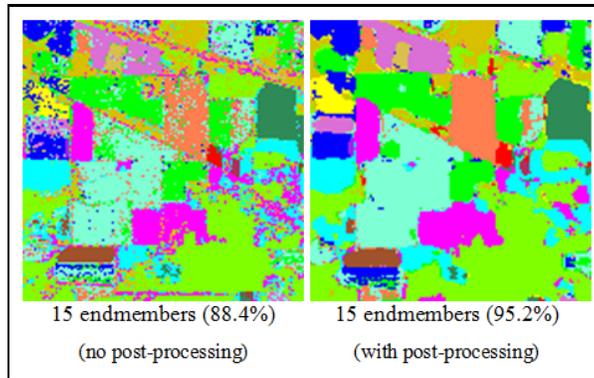


Figure 7. Classification results obtained by processing chain #4.

of their fractional abundances, which are then used as input features to the SVM-based classification process. Spatial post-processing is also optionally applied. As shown by Fig. 7, the results obtained by the linear spectral unmixing-based feature extraction are similar to those reported for the MNF-based feature extraction. Spatial post-processing increases the quality of classification results in experiments. We believe that this strategy can still be further enhanced by testing additional algorithms and fractional abundance estimation methods, as well as more robust methodologies which can automatically estimate the number of endmembers in the scene and include the spatial information during the search for spectral endmembers and not only during the post-processing stage. The most interesting property of unmixing-based feature extraction is that the resulting features have physical meaning as opposed to those derived by PCA or MNF. In future research efforts, we will continue exploring this strategy to perform feature extraction prior to classification of hyperspectral data.

4.3.5 Results obtained by processing chain #5

The main difference between processing chain #5 and processing chains #3 is that texture features are extracted from the MNF space prior to the application of the SVM-based classifier. In this work, we have used up to 18 classic texture features (mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, different correlation measures, sum average, sum variance, sum entropy, difference variance, difference entropy, skewness and kurtosis). Due to the increase in data dimensionality resulting from the application of texture feature extraction to the first 4 MNF components (resulting in $4 \times 18 = 72$ features) we applied a feature selection stage to avoid increasing the dimensionality significantly. Fig. 8 shows the most relevant results obtained with processing chain #5, in which the best results reported correspond to the selection of 10 and 15 features, respectively. Spatial post-processing results are also displayed in all cases. As shown by Fig. 8, the inclusion of texture features does not produce an improvement of the results already reported for processing chain #3 based exclusively on the MNF transform. This indicates that the considered texture features were not relevant for the classifier, which provided lower classification accuracies than in the MNF space. A similar observation applies to the results with spatial post-processing, which cannot improve those obtained for the processing chain #3. In future work, we will conduct a more detailed investigation of additional texture features that could bring more significance to the results originally provided by processing chain #3 due to the potential of well-defined texture features to improve spatial characterization. However, in our current set of experiments the considered texture features did not seem to bring additional information which is relevant to improve the overall classification accuracy.

4.3.6 Results obtained by processing chain #6

The main difference between processing chain #6 and processing chains #5 is that morphological features (instead of texture features) are extracted from the MNF space prior to the application of the SVM-based classifier. In this work, we have used a series of morphological opening and closing by reconstruction operations (using disk-shaped structuring elements of increasing radius) applied to the first 4 MNF components (resulting in $4 \times 18 = 72$ features as it was the case with texture features). Due to the large dimensionality of the resulting feature space, we applied a feature selection stage. Fig. 9 shows the most relevant results obtained with processing chain #6, in which

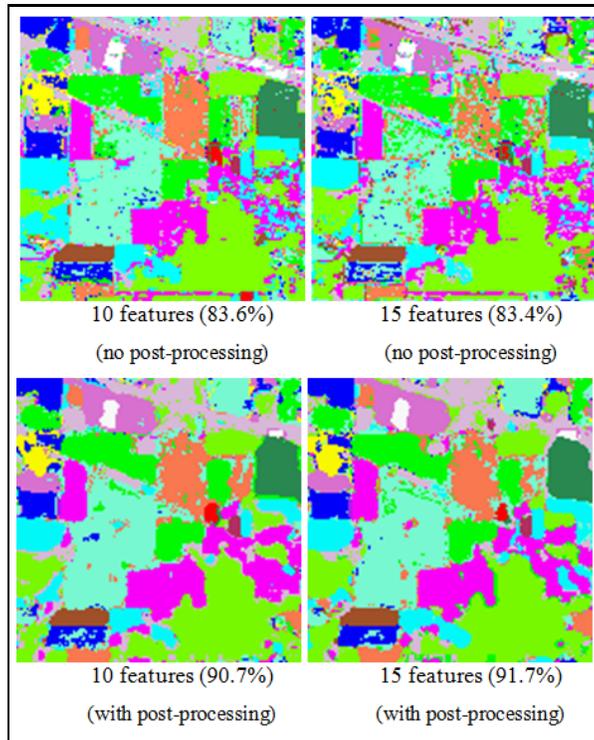


Figure 8. Classification results obtained by processing chain #5.

Table 1. Best classification results (in percentage overall accuracy) achieved for the AVIRIS Indian Pines scene by the different processing chains tested, with and without spatial post-processing.

	Chain #1	Chain #2	Chain #3	Chain #4	Chain #5	Chain #6
No spatial post-processing	84.8	80.9	91.2	88.4	83.6	91.4
With spatial post-processing	94.7	93.8	96.4	95.2	91.7	96.1

the best results reported correspond to the selection of 10 and 15 features, respectively. Spatial post-processing results are also displayed in all cases. As shown by Fig. 9, the inclusion of morphological features (processing chain #6) improves the classification results significantly with regards to the case in which texture features are used (processing chain #5). However, if we compare the classification results with morphological features with the case in which no such features are extracted from the MNF space (processing chain #3) the results are very similar. This indicated that morphological feature extraction on the MNF space cannot significantly improve the results already obtained for the processing chain #3. A similar observation can be made when analyzing the results with spatial post-processing, which indicates that the inclusion of spatial information appears to be more relevant in the post-processing stage than in the feature extraction stage (via spatial-spectral morphological features). In any event, the inclusion of spatial information introduces moderate improvements and we expect further developments in this area by appropriately selecting additional morphological features prior to the SVM-based classification stage.

4.3.7 Summary

To conclude this section, we summarize in Table 1 the best classification results achieved for the AVIRIS Indian Pines scene by the different processing chains tested, with and without spatial post-processing. As shown by Table 1, the best overall classification scores (without spatial post-processing) were achieved by processing chain #3 (which uses the MNF transform for feature extraction) and by processing chain #6 (which applies morphological feature extraction on the MNF space prior to classification), although processing chain #4 based on using spectral unmixing for feature extraction provided comparable results but with a more physically meaningful feature extraction stage. When spatial post-processing was included, again processing chains #3 and #6 achieved

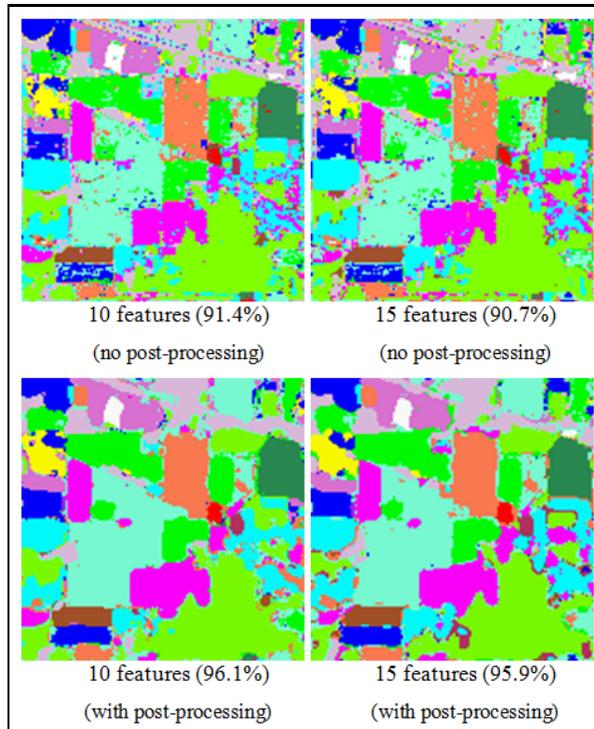


Figure 9. Classification results obtained by processing chain #6.

the best overall accuracies, with classification results above 96% which are really remarkable taking in mind the complexity of this reference hyperspectral scene. Again, processing chain #4 provided results which are comparable to the best obtained cases. The results obtained by processing chain #6 should be analyzed under the observation that they are based on the MNF space which already provided comparable results in processing chain #3 without the application of morphological descriptors.

5. CONCLUSIONS AND FUTURE RESEARCH

In this paper, we have evaluate the performance of different processing chains for remotely sensed hyperspectral image classification. A challenging classification problem, based on a reference hyperspectral data set collected by AVIRIS over the Indian Pines region, has been thoroughly investigated. One of the main observations of our study is that the MNF transform provides very competitive results as a feature extraction strategy, despite the fact that it has been mainly used in the context of spectral unmixing (not classification scenarios, in which the PCA transform has been commonly one of the tools of choice). Although our experiments in this work indicate that the MNF has the potential to outperform most other feature extraction approaches considered in this work, we believe that unmixing-based approaches have good potential for feature extraction and may provide competitive results to this strategy, since they have the potential to extrapolate concepts from mixed-pixel classification into a standard full-pixel classification scenario, thus addressing the non-stationary behavior of spectral signatures throughout spatially disjoint classes. Future work should comprise experiments with additional scenes, techniques and configurations in order to prove the generality of our compared approaches. Computationally efficient implementations of the considered processing chains are being currently developed, in line with our recent proposal of an efficient parallel implementation of the SVM classifier in commodity clusters of computers.²²

6. ACKNOWLEDGEMENT

This work has been supported by the European Community's Marie Curie Research Training Networks Programme under reference MRTN-CT-2006-035927 (HYPER-I-NET). Funding from the Spanish Ministry of Science

and Innovation (HYPERCOMP/EODIX project, reference AYA2008-05965-C04-02) is gratefully acknowledged.

REFERENCES

1. A. Plaza, J. A. Benediktsson, J. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, J. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," *Remote Sensing of Environment* **113**, pp. 110–122, 2009.
2. M. E. Schaepman, S. L. Ustin, A. Plaza, T. H. Painter, J. Verrelst, and S. Liang, "Earth system science related imaging spectroscopy – an assessment," *Remote Sensing of Environment* **113**, pp. 123–137, 2009.
3. D. A. Landgrebe, *Signal theory methods in multispectral remote sensing*, John Wiley and Sons, Hoboken, NJ, 2003.
4. J. A. Richards, "Analysis of remotely sensed data: the formative decades and the future," *IEEE Trans. Geoscience and Remote Sensing* **43**, pp. 422–432, 2005.
5. P. Gamba, F. Dell'Acqua, A. Ferrari, J. A. Palmason, and J. A. Benediktsson, "Exploiting spectral and spatial information in hyperspectral urban data with high resolution," *IEEE Geoscience and Remote Sensing Letters* **1**, pp. 322–326, 2004.
6. J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geoscience and Remote Sensing* **42**, pp. 480–491, 2005.
7. Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques," *IEEE Transactions on Geoscience and Remote Sensing* **47**, pp. 2973–2987, 2009.
8. P. Soille, *Morphological image analysis: principles and applications*, Springer-Verlag, Berlin, 2003.
9. G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing* **43**, pp. 1351–1362, 2005.
10. L. Bruzzone, M. Chi, and M. Marconcini, "A novel transductive svm for the semisupervised classification of remote sensing images," *IEEE Trans. Geoscience and Remote Sensing* **44**, pp. 3363–3373, 2006.
11. G. M. Foody and A. Mathur, "Toward intelligent training of supervised image classifications: directing training data acquisition for svm classification," *Remote Sensing of Environment* **93**, pp. 107–117, 2004.
12. R. M. Haralick, "Statistical and structural approaches to texture," *Proceedings of the IEEE* **67**(5), pp. 786–804, 1979.
13. P. Gamba and F. Dell'Acqua, "Increased accuracy multiband urban classification using a neuron-fuzzy classifier," *International Journal of Remote Sensing* **24**, pp. 827–834, 2003.
14. J. B. Adams, M. O. Smith, and P. E. Johnson, "Spectral mixture modeling: a new analysis of rock and soil types at the Viking Lander 1 site," *Journal of Geophysical Research* **91**, pp. 8098–8112, 1986.
15. M. E. Winter, "Algorithm for fast autonomous spectral endmember determination in hyperspectral data," in *Imaging Spectrometry V*, M. R. Descour and S. S. Shen, eds., *Proceedings of SPIE* **3753**, pp. 266–275, 1999.
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* **39**, pp. 529–545, 2001.
17. A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing* **42**(3), pp. 650–663, 2004.
18. C. L. Lawson and R. J. Hanson, *Solving least squares problems*, Prentice-Hall Englewood Cliffs, NJ, 1974.
19. C.-I. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*, Kluwer Academic/Plenum Publishers: New York, 2003.
20. C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing* **42**(3), pp. 608–619, 2004.
21. J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Transactions on Geoscience and Remote Sensing* **46**(8), pp. 2435–2445, 2008.
22. J. Munoz-Mari, A. Plaza, J. A. Gualtieri, and G. Camps-Valls, "Parallel implementations of svm for earth observation," in: *Parallel Programming, Models and Applications in Grid and P2P Systems* **17**, pp. 292–312, 2009.