



HAL
open science

Linear Versus Nonlinear PCA for the Classification of Hyperspectral Data Based on the Extended Morphological Profiles

Giorgio Licciardi, Prashanth Reddy Marpu, Jocelyn Chanussot, Jon Atli Benediktsson

► **To cite this version:**

Giorgio Licciardi, Prashanth Reddy Marpu, Jocelyn Chanussot, Jon Atli Benediktsson. Linear Versus Nonlinear PCA for the Classification of Hyperspectral Data Based on the Extended Morphological Profiles. IEEE Geoscience and Remote Sensing Letters, 2012, 9 (3), pp.447-451. 10.1109/LGRS.2011.2172185 . hal-00797814

HAL Id: hal-00797814

<https://hal.science/hal-00797814v1>

Submitted on 11 Mar 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

LINEAR VERSUS NONLINEAR PCA FOR THE CLASSIFICATION OF HYPERSPECTRAL DATA BASED ON THE EXTENDED MORPHOLOGICAL PROFILES

Giorgio Licciardi ^{*}, Prashanth Reddy Marpu [†], Jocelyn Chanussot (IEEE Senior Member) ^{*}, Jon Atli Benediktsson (IEEE Fellow) [†]
^{*} *GIPSA-Lab. Grenoble Institute of Technology, Grenoble, France*
[†] *Faculty of Electrical and Computer Engineering, University of Iceland, Reykjavik, Iceland*
E-mail: Giorgio-Antonino.Licciardi@gipsa-lab.grenoble-inp.fr

Abstract—Morphological profiles have been proposed in recent literature, as aiding tools to achieve better results for classification of remotely sensed data. Morphological profiles are in general built using features containing most of the information content of the data, such as the components derived from principal component analysis (PCA). Recently, non-linear PCA (NLPCA), performed by auto-associative neural network, has emerged as a good unsupervised technique to fit the information content of hyperspectral data into few components. The aim of this paper is to investigate the classification accuracies obtained using extended morphological profiles built from the features of non-linear PCA. A comparison of the two approaches has been validated on two different datasets having different spatial and spectral resolution/coverage, over the same ground truth, and also using two different classification algorithms. The results show that the NLPCA permits to obtain better classification accuracies than using linear PCA.

Keywords-Extended Morphological Profiles; Neural Networks; Nonlinear Principal Component Analysis; Classification.

I. INTRODUCTION

Morphological profiles (MP), which combine spectral and spatial information, have been shown to be effective tools for classification of remote sensing data [1] [2] [3] [4] [5] [6]. An MP of a gray-level image (or a feature) can be defined as a sequence generated with the morphological opening by reconstruction and closing by reconstruction operations, using structuring elements of increasing size. An extended morphological profile (EMP) is constructed by stacking the MPs built using different features.

Building EMP from the spectral bands of hyperspectral images (HS) can be not convenient due to their huge number of bands, so a reduction of the number of bands preserving the information contents became important. It was suggested in [4] to build the EMP from the top few components obtained from the principal component analysis (PCA) transformation which retain most of the variance of the image. This approach was successfully applied in classification of hyperspectral images, resulting in better accuracies compared to simply using the spectral information only. Similar

approaches, using combinations of morphological operators have been presented in literature [7] [3]. In particular it has been observed that better classification accuracies can be obtained using the nonlinear features from kernel PCA (KPCA) instead of the features from PCA [8]. In both cases, the derived components are ranked in terms of the amount of variance. This means that the information content is not equally distributed among the components, where the first one is always more relevant than the others. The dimensionality reduction using PCA or KPCA is achieved by discarding the less relevant components. On the other hand, Nonlinear Principal Component Analysis (NLPCA), performed using Autoassociative Neural Networks (AANNs) [9], produces a limited set of components in which the information content tends to be uniformly distributed. The purpose of this paper is to investigate the improvements introduced by using EMP built from NLPCA and comparing it with the results obtained with PCA and KPCA. The paper is organized as follows. In sections II and III the EMP and the NLPCA will be presented, respectively, while a comparison of the classification results obtained using EMP generated from NLPCA and PCA will be presented in section IV. Finally, conclusions are drawn in Section V.

II. EXTENDED MORPHOLOGICAL PROFILE

In mathematical morphology, one of the most used approaches to analyze spatial inter-pixel dependency is the morphological profile, which has been successfully used to extract spatial information from high spatial resolution images [1]. The idea at the base of the MP is to apply geodesic closing/opening transformations of increasing size, to build a certain set of opening profiles (OP) and closing profiles (CP). The opening/closing profiles P at pixel x of the image f are defined as a p -dimensional vectors:

$$P_i(x) = \gamma_R^{(i)}(x), \forall i \in [0, p] \quad (1)$$

Where $\gamma_R^{(i)}$ can be the opening or closing by reconstruction with a structuring element (SE) of size i .

By grouping the OP, CP and the image $f(x)$, the $(2p+1)$ -dimensional vector is the MP which is defined as:

$$MP(x) = [CP_p(x), \dots, f(x), \dots, OP_p(x)]. \quad (2)$$

It is clear from the representation of MP in 2 that applying MPs directly to the hyperspectral data with huge number of bands leads to a huge increase in the number of features. The stacking of the $q(2p+1)$ MPs obtained with different features (where q is the number of retained components), is called *Extended Morphological Profile* (EMP).

III. NONLINEAR PRINCIPAL COMPONENT ANALYSIS

One of the main difficulties in processing HS images is related to the very high number of spectral bands. Applying any processing technique to each band of the HS image, can lead to a non acceptable increase of the computational time of the entire process. Therefore, it is generally desirable that a reduction in the number of features is achieved without losing the relevant spectral information of the original dataset. In the literature, there exist many methods for representing the information content in lower dimensionality domain, called feature extraction techniques [10]. Two of the most popular feature extraction methods for data representation are Principal Component Analysis (PCA), where a set of uncorrelated transformed features is generated and the Independent Component Analysis (ICA), where a computational method for separating a multivariate signal into additive subcomponents supposing the mutual statistical independence of the non-Gaussian source signals [11]. For these techniques, the dimensionality reduction is obtained by discarding the components with the lowest information content. Also, as most of them are linear methods, the resulting components are linearly uncorrelated but the physical representation of the image may be lost. NLPCA, originally introduced by Kramer [12], is based on a multi-layer perceptron (MLP) commonly referred as (AANN) or as autoencoder [13] [14]. The AANNs are conventional Neural Networks (NNs) featuring feedforward connections and sigmoidal nodal transfer functions, trained by backpropagation algorithm. The particular network architecture used employs three hidden layers, including an internal bottleneck layer of smaller dimension than either input or output. The network is trained to perform identity mapping, where the input has to be equal to the output. Since there are fewer units in the bottleneck layer compared to the output, the bottleneck nodes must encode the information obtained from the inputs for the subsequent layers to reconstruct the input. In such a way, the nonlinear principal components (NLPCs) can be extracted from the bottleneck nodes, after the training of the AANN. The main task in designing the AANN is the selection of the number of nodes minimizing the information losses of the training. This problem was solved by a grid search algorithm varying recursively the number

of nodes and evaluating the respective error. The topology producing the lowest error was then selected. Compared to linear reduction techniques, NLPCA has many advantages. First of all, while linear methods can detect and discard linear correlations among spectral bands, NLPCA detects both linear and nonlinear correlations. Moreover, in NLPCA the information content is equally distributed among the components [15].

In this paper we propose the use of NLPCs to form base images for the EMP. The NLPCs are obtained from an AANN having sigmoidal activation function, trained with Scaled Conjugated Gradient algorithm (SCG). Once trained the AANN, the output of the bottleneck layer will be used as NLPCs and the resulting EMP, will be used as input for the classification task.

IV. EXPERIMENTS

In this section we present results of the proposed approach applied to two different HS images having different spatial and spectral resolution/coverage, over the same ground truth. In both experiments we classified the EMP built from the NLPCs extracted from a HyMap image and from a CHRIS image. HyMap is an airborne 4 spectrometers sensor (VIS, NIR, SWIR1 and SWIR2), providing 128 bands across the reflective solar wavelength region of 0.45-2.5 μm with contiguous spectral coverage (except in the atmospheric water vapor bands) and bandwidths between 15-20 nm (Fig. 1-a). The CHRIS image was acquired in Mode 1 configuration, having 62 spectral bands, with a spatial resolution of 34 m at nadir and a spectral coverage of 0.45-1.03 μm (Fig. 1-b). Both images were acquired over the same area during the ESA - SPectra bARrax Campaigns 2003 (SPARC) campaign (<http://www.uv.es/leo/sparc/>) carried out in Barrax, La Mancha, Spain, from 12 to 14 of July 2003. The Barrax area is mainly used for agricultural cultivations and has been investigated for many years. It is characterized by a flat morphology and large, uniform land-use units, mainly composed by different agricultural types. During the campaign an extensive ground truth was produced (Fig. 1-c) and was used to build the ground truth in this study. The reference classes used for the classification are: Corn, Papaver, Potatoes, Alfalfa, Wheat, Barley, Garlic, Vineyards, Bare soils, Onion and Barley stubbles, resulting in about 60.500 and 2.500 pixels for Hymap and Chris, respectively, equally distributed between training and test sets. To evaluate the effectiveness of the method, the classification was performed by two different algorithms i.e. neural networks (NN) and support vector machines (SVM). A comparison with the classification accuracies obtained using standard PCA and kernel PCA with the EMP, shows the enhancement introduced by the nonlinear principal component analysis. In PCA and KPCA, the dimensionality reduction is performed discarding the features less informative, but while in PCA most of the information content is retained in the first few

features, KPCA requires more components. This means that kernel PCA needs a large number of components, increasing the dimensionality of the data, resulting in a huge number of features when building morphological profiles. Moreover, in KPCA, the choices of the kernel parameter and the sample size to perform kernel PCA are very important and determine these parameters is not an easy task. In particular, for both images, KPCA was performed with 1500 samples, and the kernel parameter was selected as twice the average distance between all the pixels. A tuning of these parameters was not performed because, being strongly dependent on the randomly selected sample set, it will require a further processing step, that cannot be compared with other approaches. The comparison was carried out in terms of (OA) overall accuracy (ratio between the total number of correctly classified samples and total number of test samples), K Kappa coefficient of agreement (percentage of agreement corrected by the amount of agreement that could be expected due to chance alone), and the class accuracy (percentage of correctly classified samples for a given class).

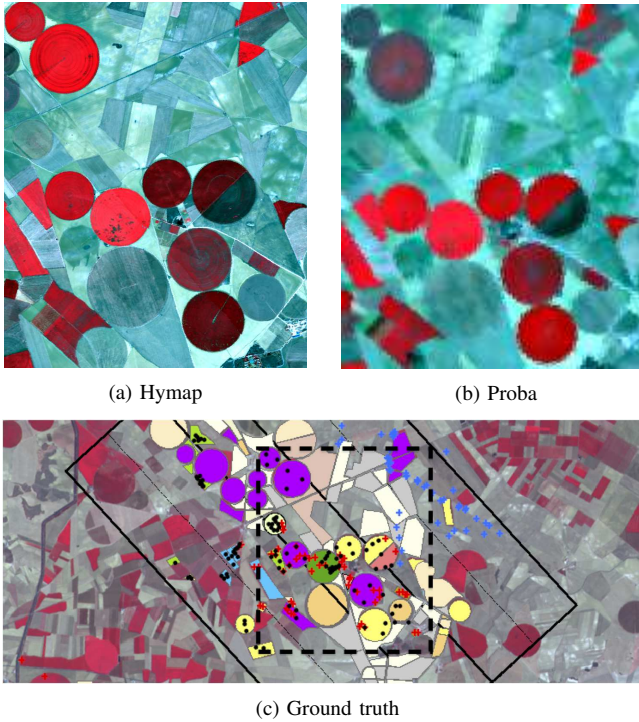


Figure 1. False color RGB of Hymap dataset (a) and CHRIS (b). The map (c) shows the ground truth acquired during the ESA-SPARC campaign.

A. Hymap dataset

The feature extraction from the HyMap image using AANN was performed by a grid-search algorithm, varying the number of nodes in the bottleneck and in the other two hidden layers looking for the lowest Mean Square Error

(MSE). The optimal solution was found with 6 nodes in the bottleneck layer, corresponding to 6 NLPCs and 55 nodes in the outer hidden layers. A circular SE with a step size increment of 2 was used. Four openings and closings were computed for each component, resulting in a EMP of dimension $9 \times 6 = 54$. As for the PCA and KPCA, the EMPs were constructed using the first components corresponding to more than 99% of the cumulative variance, resulting in 45 and 135 EMP, respectively. Analyzing the confusion matrices in tables I-II and the classification maps in Fig. 2 it is evident that using NLPC to build EMP improves the classification accuracy with both training algorithms. Good accuracies were achieved in all classes except for Alfalfa, that has good accuracy only using NN and NLPCA. This problem raises from the small spectral differences between Alfalfa and Potatoes cultivations that have not been completely synthesized. KPCA reaches good accuracies for all other classes except for Bare soil with SVM. This because of the strong spectral similarity with Barley stubble.

| Feature | Raw | PCA | NLPCA | KPCA |
|-----------------|--------------|--------------|----------------|--------------|
| N. of features | 126 | 5 | 6 | 15 |
| N. of EMP | | 45 | 54 | 135 |
| OA (%) | 75.5792 | 74.1682 | 79.6533 | 73.1162 |
| k | 0.7252 | 0.7090 | 0.7654 | 0.6975 |
| Corn | 99.95 | 99.55 | 99.89 | 99.92 |
| Papaver | 100 | 99.52 | 100 | 100 |
| Potatoes | 96.12 | 99.21 | 99.98 | 100 |
| Alfalfa | 30.95 | 37.21 | 37.39 | 36.25 |
| Wheat | 99.28 | 95.02 | 99.29 | 99.96 |
| Barley | 100 | 99.66 | 99.74 | 99.57 |
| Garlic | 100 | 100 | 96.66 | 100 |
| Vineyards | 97.27 | 98.98 | 97.26 | 95.22 |
| Bare soil | 39.67 | 27.03 | 62.91 | 28.68 |
| Barley stubbles | 99.23 | 99.33 | 74.53 | 97.99 |
| Onions | 99.36 | 98.92 | 100 | 100 |

Table I
CLASSIFICATION RESULTS FOR THE HYMAP DATASET USING SVM CLASSIFICATION ALGORITHM.

B. CHRIS dataset

Following the same procedures used in the previous experiment, an AANN, having 4 nodes in the bottleneck layer and 25 in the outer hidden layers, was used to extract 4 nonlinear principal components from the original 62 bands. Also in this case a circular SE with a step size increment of 2 was used and four openings and closings were computed for each component. The resulting dimensionality of EMP was $9 \times 4 = 36$. The 99% of the cumulative variance of the PCA was retained by the first 4 components, resulting in a dimensionality of the EMP of 36 while KPCA needs 15 components, corresponding to 135 EMP. The results reported in tables III-IV and in Fig. 3, show once again that the best performances were obtained using NLPCs to build the EMP for both NN and SVM classifications. Compared to the HyMap experiments, it is evident that the highest accuracies

| Feature | Raw | <i>PCA</i> | <i>NLPCA</i> | <i>KPCA</i> |
|-----------------|--------------|------------|----------------|--------------|
| N. of features | 126 | 5 | 6 | 15 |
| N. of EMP | | 45 | 54 | 135 |
| OA (%) | 79.6533 | 72.5309 | 81.9068 | 74.7217 |
| <i>k</i> | 0.7654 | 0.6912 | 0.7930 | 0.7147 |
| Corn | 99.89 | 99.55 | 99.73 | 99.48 |
| Papaver | 100 | 99.52 | 99.95 | 98.94 |
| Potatoes | 99.98 | 99.21 | 99.98 | 86.99 |
| Alfalfa | 37.39 | 37.51 | 75.15 | 27.06 |
| Wheat | 99.26 | 95.02 | 94.25 | 99.70 |
| Barley | 99.74 | 99.66 | 91.47 | 43.10 |
| Garlic | 96.66 | 100 | 99.64 | 99.64 |
| Vineyards | 99.81 | 98.98 | 99.18 | 93.29 |
| Bare soil | 39.67 | 27.03 | 79.14 | 82.27 |
| Barley stubbles | 99.33 | 68.76 | 75.57 | 99.97 |
| Onions | 100 | 98.92 | 98.66 | 97.96 |

Table II
CLASSIFICATION RESULTS FOR THE HYMAP DATASET USING A NN CLASSIFICATION ALGORITHM.

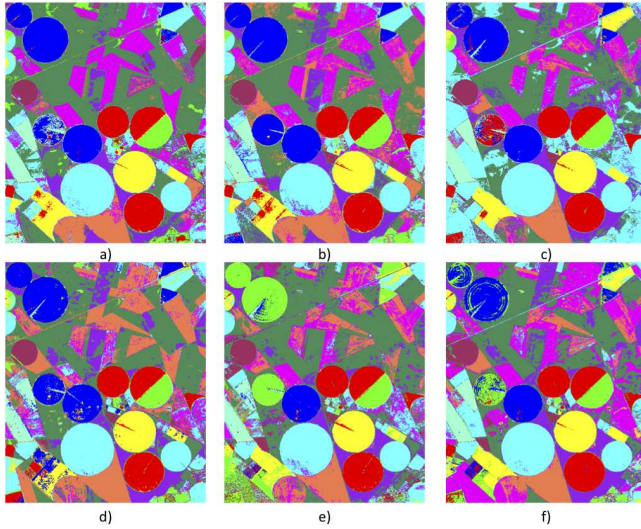


Figure 2. Classification results obtained from the HyMap image using SVM classification algorithm on EMPs built from PCA (a), NLPCA (b) and KPCA (c), and using NN classification algorithm on EMP built from PCA (d), NLPCA (e) and KPCA (f). The color map is as follows: **Corn**, **Papaver**, **Potatoes**, **Alfalfa**, **Wheat**, **Barley**, **Garlic**, **Vineyards**, **Bare soil**, **Barley stubble**, **Onions**.

are obtained with the CHRIS data. Because the low spatial resolution of the CHRIS data is more suited to the chosen class types. The ground truth pixels in the CHRIS image are related to the same land cover type and hence have more uniform values than those from HyMap. This effect, on the other hand, produced poor results in some cases. In particular NLPCA and KPCA approaches show poor results for the classification of Barley stubble class. This problem is related to the classification algorithm and can be explained analyzing the spectral signature of pixels of Barley stubble class, that is very similar to the bare soil signature. This leads alternatively SVM and NN to consider Barley stubble as Bare soil.

| Feature | Raw | <i>PCA</i> | <i>NLPCA</i> | <i>KPCA</i> |
|-----------------|--------------|------------|----------------|-------------|
| N. of features | 62 | 4 | 4 | 15 |
| N. of EMP | | 36 | 36 | 135 |
| OA (%) | 78.6342 | 73.8019 | 85.2636 | 70.0080 |
| <i>k</i> | 0.7513 | 0.6945 | 0.8277 | 0.6525 |
| Corn | 100 | 100 | 100 | 31.77 |
| Papaver | 100 | 100 | 100 | 100 |
| Potatoes | 100 | 100 | 100 | 99.17 |
| Alfalfa | 75.46 | 72.2 | 77.87 | 65.57 |
| Wheat | 100 | 100 | 100 | 100 |
| Barley | 100 | 100 | 40.00 | 100 |
| Garlic | 79.89 | 79.84 | 96.12 | 50.37 |
| Vineyards | 74.89 | 69.36 | 49.36 | 100 |
| Bare soil | 100 | 78.69 | 100 | 100 |
| Barley stubbles | 100 | 68.76 | 61.16 | 32.34 |
| Onions | 100 | 50.37 | 100 | 95.14 |

Table III
CLASSIFICATION RESULTS FOR THE CHRIS DATASET USING SVM CLASSIFICATION ALGORITHM.

| Feature | Raw | <i>PCA</i> | <i>NLPCA</i> | <i>KPCA</i> |
|-----------------|------------|------------|----------------|---------------|
| N. of features | 62 | 4 | 4 | 15 |
| N. of EMP | | 36 | 36 | 135 |
| OA (%) | 89.1342 | 70.4872 | 93.3706 | 74.2259 |
| <i>k</i> | 0.8694 | 0.6647 | 0.9209 | 0.7094 |
| Corn | 100 | 100 | 100 | 99.89 |
| Papaver | 100 | 100 | 100 | 100.00 |
| Potatoes | 95.80 | 100 | 82.09 | 99.96 |
| Alfalfa | 74.74 | 32.62 | 100 | 37.39 |
| Wheat | 100 | 100 | 99.34 | 98.87 |
| Barley | 100 | 38.57 | 61.43 | 99.74 |
| Garlic | 100 | 100 | 92.25 | 96.66 |
| Vineyards | 86.19 | 46.38 | 94.86 | 99.55 |
| Bare soil | 83.72 | 100 | 100 | 39.67 |
| Barley stubbles | 100 | 76.86 | 26.45 | 74.53 |
| Onions | 100 | 50.37 | 99.26 | 100 |

Table IV
CLASSIFICATION RESULTS FOR THE CHRIS DATASET USING A NEURAL NETWORK CLASSIFICATION ALGORITHM.

V. CONCLUSIONS

This paper presented a novel classification approach with two main issues: a feature extraction method based on NLPCA as a tool which is able to maintain the information content of hyperspectral remote sensing imagery into few components, and the construction of EMP with the NLPCs, to include spatial information in the classification task. Comparisons in terms of classification accuracies with standard PCA and KPCA approaches, using a SVM and a NN classifiers, demonstrates that NLPCA extracts more informative features and does not suffer from the noise contained in the HS data. The poor results obtained by KPCA can be explained by the fact that the sample size may not be enough, and also by the fact that kernel PCs are more influenced by noise than the other components. Moreover kernel PCA results in a large number of features, thus increasing the dimensionality of the data, which increases many times when building morphological profiles, allowing

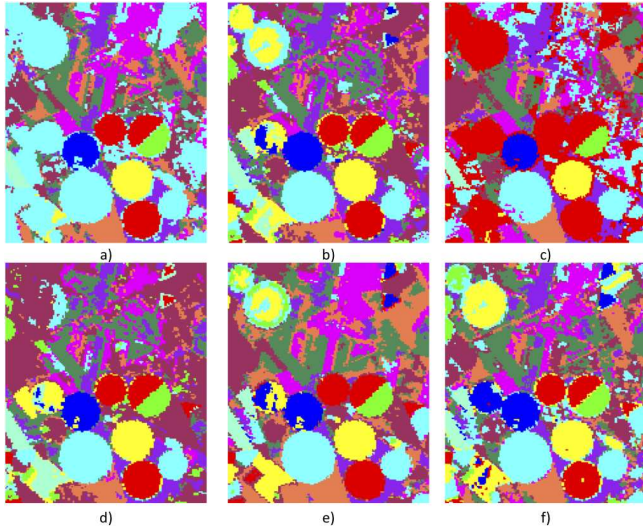


Figure 3. Classification results obtained from the CHRIS image using SVM classification algorithm on EMPs built from PCA (a), NLPCA (b) and KPCA (c), and using NN classification algorithm on EMP built from PCA (d), NLPCA (e) and KPCA(f). The color map is as follows: Corn, Papaver, Potatoes, Alfalfa, Wheat, Barley, Garlic, Vineyards, Bare soil, Barley stubble, Onions.

the classification to be prone to the Hughes effect. A better result in terms of accuracies using the KPCA features could be obtained by selecting the best KPCA features This required a further preprocessing steps that was not taken into account being outside of the investigation of this study. A further assessment of the NLPCA method was made up by using two different HS images acquired on the same area on the same days and thus sharing the same ground truth. Both classifiers produced results of the EMP built with the NLPCA that significantly outperforms those obtained with the EMP with the PCA. The principal conclusion is that NLPCA, extracting more useful features from a HS image, permits to obtain better classification accuracies than using linear feature extraction approaches. So, it is reasonable to say that the features obtained with AANN are more reliable compared to standard PCA and kernel PCA for the purpose of classification with morphological profiles.

ACKNOWLEDGMENT

The hyperspectral data used in this paper are provided by ESA under the project 6519.

REFERENCES

[1] Pesaresi M. Benediktsson J. A., "A new approach for the morphological segmentation of high-resolution satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 2, pp. 309–320, 2001.

[2] Plaza A. Martinez P. Plaza J. Perez R., "Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations," *IEEE*

Trans. Geosci. Remote Sens., vol. 43, no. 3, pp. 466–479, 2005.

[3] Fauvel M. Benediktsson J. A. Chanussot J. Sveinsson J. R., "Spectral and spatial classification of hyperspectral data using svms and morphological profile," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3804–3814, 2008.

[4] Benediktsson J. A. Palmason J. A. Sveinsson J. R., "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 480–490, 2005.

[5] A. Plaza, J. A. Benediktsson, J. 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*, 2009.

[6] G. Licciardi, F. Pacifici, D. Tuia, S. Prasad, T. West, F. Giacco, Ch. Thiel, J. Inglada, E. Christophe, J. Chanussot, and P. Gamba, "Decision fusion for the classification of hyperspectral data: Outcome of the 2008 grs-s data fusion contest," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 11, pp. 3857–3865, 2009.

[7] Tuia D. Ratle F. Pozdnoukhov A. Camps-Valls G., "Multisource composite kernels for urban-image classification," *IEEE Geosci. Remote Sens. Letters*, vol. 7, no. 1, pp. 88–92, 2010.

[8] M. Fauvel, J. Chanussot, and J.A. Benediktsson, "Kernel principal component analysis for the classification of hyperspectral remote-sensing data over urban areas," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, 2009.

[9] Kramer M. A., "Nonlinear principal component analysis using autoassociative neural networks," *American Institute of Chemical Engineers*, vol. 37, pp. 233–243, 1991.

[10] B. Serpico, M. DInca, F. Melgani, and G. Moser, "Comparison of feature reduction techniques for classification of hyperspectral remote sensing data universal image quality index," *Proc. SPIE Image Signal Process. Remote Sens.*, vol. 4885, no. 8, pp. 347–358, 2003.

[11] C. Jutten and J. Héroult, "Blind separation of sources, part i: an adaptive algorithm based on neuromimetic architecture," *Signal Processing*, vol. 24, pp. 1–10, 1991.

[12] M. A. Kramer, "Nonlinear principal component analysis using autoassociative neural networks," *AICHE J.*, vol. 37, pp. 233–243, 1991.

[13] C. Bishop, "Neural networks for pattern recognition," *Oxford Univ. Press, London, U.K.*, 1995.

[14] M. Scholz, F. Kaplan, L. Guy, J. Kopka, and J. Selbig, "Non-linear pca: a missing data approach," *Bioinformatics*, vol. 21, pp. 3887–3895, 2005.

[15] G. Licciardi, F. Del Frate, and R. Duca, "Feature reduction of hyperspectral data using autoassociative neural networks algorithms," *IEEE Int. Geoscience and Remote Sensing Symp. IGARSS 2009*, vol. 1, pp. 176–179, 2009.