Enhancing Hyperspectral Image Quality using Nonlinear PCA

Giorgio Antonino Licciardi, Jocelyn Chanussot, Gabriel Vasile, A. Piscini

HAL Id: hal-01065843
https://hal.science/hal-01065843

Submitted on 18 Sep 2014

ENHANCING HYPERSPECTRAL IMAGE QUALITY USING NONLINEAR PCA

G. A. Licciardi [1], J. Chanussot [1, 2], G. Vasile [1], A. Piscini [3]

(1) GIPSA-Lab, Grenoble Institute of Technology, France
(2) Faculty of Electrical and Computer Engineering, University of Iceland.
(3) Istituto Nazionale di Geofisica e Vulcanologia, Italy
e-mail: Giorgio-Antonino.Licciardi@gipsa-lab.grenoble-inp.fr

ABSTRACT

In this paper, we propose a new method aiming at reducing the noise in hyperspectral images. It is based on the nonlinear generalization of Principal Component Analysis (NLPCA). The NLPCA is performed by an autoassociative neural network that have the hyperspectral image as input and is trained to reconstruct the same image at the output. Thanks to its bottleneck structure, the AANN forces the hyperspectral image to be projected in a lower dimensionality feature space where noise as well as both linear and nonlinear correlations between spectral bands are removed. This process permits to obtain enhancements in terms of hyperspectral image quality. Experiments are conducted on different real hyperspectral images, with different contexts and resolutions. The results are qualitatively and quantitatively discussed and demonstrate the interest of the proposed method as compared to traditional approaches.

Index Terms— nonlinear PCA, noise reduction, image quality, hyperspectral images, NLPCA

1. INTRODUCTION

One of the most important parameters in the design of hyperspectral imagers is the signal-to-noise ratio (SNR), that determine the quality of the images. More in particular, in the process of hyperspectral image acquisition the noise is generated by several factors and some of them can have a negative impact on the processing of hyperspectral images. A high SNR can be obtained using components that are generally expensive, particularly in the case of spaceborne instruments. An alternative cost-effective solution is offered by the use of noise reduction, or de-noising algorithms that operate directly on the images. In the literature many hyperspectral remote sensing image de-noising algorithms have been proposed, most of them combine correlation of spatial and spectral domain. Among de-noising methods, those based on wavelet are widely used [1]. In general the basic procedure of wavelet de-noising is to transform the noisy data cube into wavelet coefficients in the wavelet domain, threshold the wavelet coefficients, and then perform the inverse wavelet transform to obtain the de-noised data cube. In [2] a hybrid approach based on wavelets has been used. The algorithm works in the spectral derivative domain, where the noise level is elevated, and takes advantage of the dissimilarity of the signal regularity in the spatial and spectral domains. Tensor-filters have been used in [3], while a filter based on anisotropic Diffusion is proposed in [4]. The use of Principal Component analysis has been proposed in [5], where the low-energy principal components was de-noised with wavelet shrinkage de-noising processes. However, all those techniques present good results for the removal of white uncorrelated noise, but they usually fail when applied to images with correlated noise [6]. A possible solution to this problem is presented in [7] where a projection in a nonlinear feature space, obtained through the use of Kernel Principal Component Analysis (KPCA), permits to remove most of the noise affecting an image. Recently, a further improvement to the KPCA denoising approach has been obtained by applying classical denoising techniques in the nonlinear feature space obtained using the KPCA [8]. In this paper we propose a novel quality enhancement method for hyperspectral images performed through the use of Nonlinear Principal Component Analysis (NLPCA). The remainder of the paper is organized as follows: section 2 presents a brief description of the NLPCA while in section 3 the results from the conducted tests are reported and discussed. Finally, section 4 concludes the paper and gives some ideas for future works.

2. NONLINEAR FEATURE EXTRACTION

Many methods have been proposed in the literature for the decorrelation of hyperspectral images in order to represent the inherent information content in lower dimensionality domain. One of the most popular feature extraction method for data representation is Principal Component Analysis (PCA) that generates a set of uncorrelated transformed features. Since the components obtained with PCA are ranked in terms of increasing variance, it is possible to suppose that the first components contain most of the information, while the last ones are characterized only by noise. The dimensionality reduction

is then obtained discarding the components with the lowest variance. Moreover, as PCA is a linear method, the resulting components are linearly uncorrelated. However, due to the non completely linear nature of hyperspectral data, relevant information content may be retained in the low variance components and consequently be lost. In order to deal with nonlinearities numerous techniques have been developed. Locally linear embedding (LLE) [9] and Isomap [10] visualize high dimensional data by projecting (embedding) them into a two or three-dimensional space, while self organising maps (SOM) [11] describe data by nonlinear curves and nonlinear planes up to two dimensions. One of the main limitation of these methods is related to the low number of features that can be obtained that may be not sufficient to describe the inherent information of the data. In this regard, KPCA and NLPCA present a higher degree of freedom in terms of feature dimensionality [12]. KPCA first performs a nonlinear transformation of the data into a higher dimensional space, not necessarily of finite dimension, where the projected data is no more affected by nonlinearities. Then a linear PCA can be performed as in the input space [13][12]. However, one of the limitations of the KPCA is related to the non invertible mapping function. In particular, with KPCA only an estimation of the de-noised image can be obtained [14].

2.1. Nonlinear Principal Component Analysis

Similarly to KPCA, NLPCA performs a nonlinear projection of the original data into a feature space. However, a substantial difference between KPCA and NLPCA is that the latter one provides the inverse transformation thus enabling the actual reconstruction of the denoised image. Firstly introduced by Kramer [18], NLPCA is commonly based on multi-layer perceptrons (MLP) and performed by Autoassociative Neural Network (AANN) (or autoencoder) [15]. An AANN is a conventional feedforward NN having sigmoidal activation functions in each node and trained with by the Standard Conjugated Gradient (SCG) [16] in order to minimize the mean sum-of-squares error (MSE). Generally, the nonlinear AANN sees the use of three hidden layers, including an internal bottleneck layer of smaller dimension than either input or output (Fig. 1).

The AANN is trained to perform identity mapping, where the input X has to be equal to the output X'. This means that if the training phase finds an acceptable solution, i.e., a solution that gives an error below a predefined threshold, a good compressed representation of the input must exist in the bottleneck layer. Since there are fewer nodes in the bottleneck layer than in the input/output, the bottleneck nodes must represent or encode the information obtained from the inputs for the subsequent layers to reconstruct the input. The use of autoencoder to project the data into a feature space has been introduced in [17], where an AANN having only the bottleneck layer between input and output was used. In
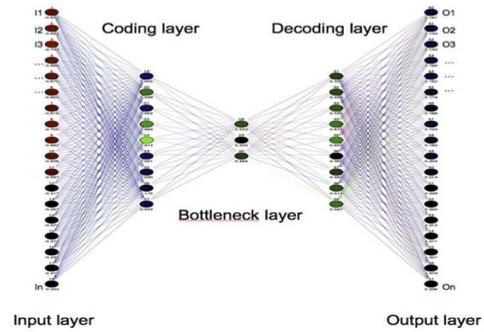

Fig. 1. *A typical AANN structure.*

particular, it has been demonstrated that if the nodes of the bottleneck layer were linear, this would correspond exactly to (linear) PCA. However, changing the activation functions of the bottleneck layer to nonlinear, only linear combinations of the inputs compressed by the activation function could be represented. In order to obtain optimal nonlinear feature extraction through the use of AANN three hidden layers are essential[18]. One of the main difficulties in designing the AANN relies in the selection of the correct number of nodes that minimizes the loss of information produced in the three hidden layers, and in particular in the bottleneck layer. Since the training of the AANN is based on the reconstruction error, the optimal NN topology can be retrieved by varying recursively the number of nodes of the three hidden layers and evaluated the respective MSE errors. However, an AANN is effective only if the number of mapping/demapping nodes is greater then the number of nodes in the bottleneck layer, on the contrary there will not be enough data to effectively extract N nonlinear components. It has to be noted that, as the output has to simply replicate the input, no independent target data are provided, and there is no need to have an a priori knowledge for the implementation of the learning phase. This implies that the AANN training can be performed in a fully automatic way and that all pixels in the image can be considered for this task. In the literature, Nonlinear Principal Component Analysis, and has been proposed as an effective instrument for dimensionality reduction and decorrelation of different types of data. In [19][20] NLPCA has been used to reduce the dimensionality of different hyperspectral images. Compared to linear decorrelation techniques, NLPCA has many advantages. First of all, with the PCA or similar approaches the information content is firstly reprojected in a orthogonal space and then the obtained component are ordered in terms of variance. The compression through PCA can be obtained by discarding the less relevant components in terms of variance. Since this kind of approaches detect only linear correlations among spectral bands, a relevant part of the original information can be retained by the last components and conse-

quently lost during the compression phase. With NLPCA this is not necessary, being the information content compressed by the bottleneck nodes of the AANN. Moreover, in NLPCA the information content is equally distributed among the components [21], avoiding unbalanced components. The advantage of NLPCA over linear decorrelation approaches and its ability to retain all the relevant information in few components allowing an almost ideal reconstruction of the original image has been proved in [22].

3. EXPERIMENTS

In this work we will use the NLPCA approach described above to enhance the quality of two hyperspectral images. The two images have different characteristics in terms of spectral range, spatial/spectral resolution, acquisition mode and consequently in the type of noise. For each image an AANN is trained using 60% of the available pixels. The complete image is then processed by the trained AANN. For sake of comparison, the quality of de-noised images obtained using NLPCA and linear PCA have been evaluated in terms of mean squared error (MSE), Spectral Angle Mapper (SAM) [23] and signal-to-noise ratio (SNR), as defined in [24]. While a high SNR value indicate a low level of noise, the SAM measures the angular distance between spectra. The optimal value is $0°$ but values up to $3°$ are usually considered as good results. A numerical comparison with the KPCA method is not considered since is only possible to have an effective estimation of the de-noised pre-image. However, a qualitative assessment has been carried out comparing the principal components obtained using linear PCA, KPCA and NLPCA, respectively.

3.1. ROSIS-Pavia

A first experiment has been performed using a data set acquired by the ROSIS sensor (0.43 to 0.86 μm with 103 bands) over the University of Pavia, Italy. The image presents several bands, characterized by uncorrelated noise. The image has 340X610 pixels, corresponding to 125.00 samples that have been used to train the AANN. Several configurations of the AANN has been evaluated in order to detect the best topology. This step has been carried out by recursively varying the number of nodes in the three hidden layers and evaluating the Mean Square Error (MSE) between the desired and the obtained output. The network with the lowest value of MSE, having 103 nodes in the input/output layers, 30 nodes in the coding/decoding layers and 4 nodes in the bottleneck layer, has been selected. The same number of components has been selected for the linear PCA. The quality of the reconstructed images, obtained with the two different methods are then evaluated in terms of SAM and SNR. As reported in Table 1, it can be seen that the NLPCA approach returns the best values for both SAM and SNR. From a qualitative point

Fig. 2. *ROSIS dataset: Detail of band 1of the original image(left) and the reconstructed images using 4 components from the PCA (center) and NLPCA (right) approaches, respectively.*

of view, the quality enhancement introduced by the two methods can be appreciated analyzing band 1, that is strongly affected by noise (Fig. 2). In particular, while the PCA-derived image still presents a reduced amount of the original noise, the image obtained using the NLPCA method is more defined and seems to be not affected by any kind of noise. This is also clear in Fig. 3 where the four principal components obtained using PCA, KPCA and NLPCA, respectively, are depicted. While the components obtained with the nonlinear approaches (KPCA and NLPCA) are not affected by any kind of noise, a relevant part of noise is still present in the fourth component of the linear PCA.

	ROSIS		Hyperion	
	PCA	NLPCA	PCA	NLPCA
MSE	$8.1 \cdot 10^5$	$7.68 \cdot 10^5$	$1.59 \cdot 10^{-4}$	$3.24 \cdot 10^{-5}$
SNR	51.81	52.04	25.14	29.79
SAM	0.44	0.16	10,12	10,86

Table 1. Rosis dataset: SAM and SNR (dB) values obtained comparing the original image with the reconstructed ones obtained with PCA and NLPCA, respectively, using 4 components.

3.2. Hyperion

The same approach has been applied to a Hyperion image (220 hyperspectral bands from 0.4 to 2.5 μm) acquired in 2008 over the Campi Flegrei area, North-West of Naples, Italy. From the original dataset only 155 spectrally unique bands have been retained, discarding the bands without relevant information [25]. In this case the de-noising of the image is extremely challenging because several kinds of noise are present at the same time. In particular, the image has been acquired from a satellite, meaning that the atmospheric contribution has a relevant role in the noise contribution of the image. Moreover, the image is also strongly affected by striping, a spatially correlated noise. The considered image consists of 100X100 pixels, and also this time we trained the AANNs using 60% of the total number of pixels of the image (6000). As in the previous experiment, an iterative training of the AANN has been performed in order to find the best network configuration. The best topology has been found to
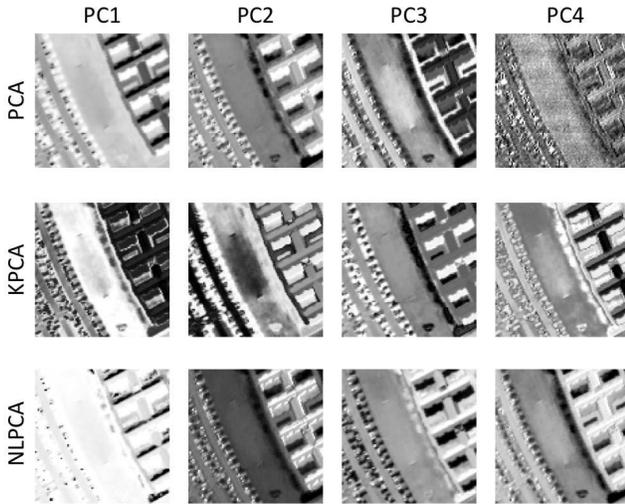
Fig. 3. *ROSIS dataset: Detail on four principal components obtained using PCA, KPCA and NLPCA, respectively.*
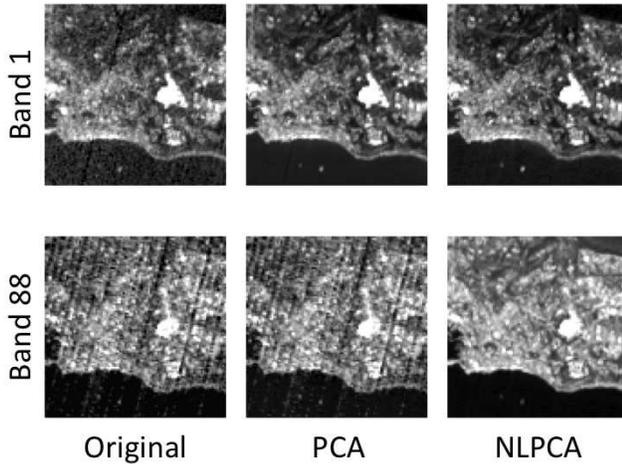

Fig. 4. *Hyperion dataset: Details of band 1 and 88 of the original image(left) and the reconstructed images using 9 components from the PCA (center) and NLPCA (right) approaches, respectively.*
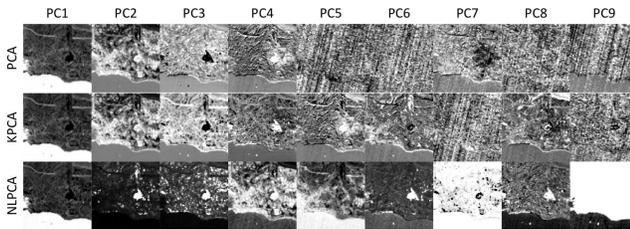

Fig. 5. *Hyperion dataset: Detail on the first nine principal components obtained using PCA, KPCA and NLPCA, respectively..*

be 155-50-9-50-155. Table 1 reports the quality enhancement obtained with NLPCA and PCA, respectively in terms of MSE, SAM and SNR. As it can be noted, NLPCA presents the best values in all the three indexes. The improvement in terms of quality can be also deduced from a qualitative analysis of the reconstructed images, in particular from bands 1 where the presence of additive noise is evident and band 88 that is strongly affected by striping, as reported in Fig. 4. In the image reconstructed using the PCA approach the additive noise has been effectively filtered, but the stripe is still present. On the other hand, the NLPCA-reconstructed image both kind of noise are no more present. From this point of view, being a nonlinear approach, the NLPCA can achieve better results compared to linear methods in terms of image enhancement. Analyzing the components depicted in Fig 5 is possible to qualitatively appreciate the enhancement introduced by the proposed method. In particular, noise can be clearly noticed in component 5 and 6 of the PCA transformation, and on components 7 and 9 of the KPCA transformation. On the other hand, in the 9 components obtained with the NLPCA, the noise is not present or not relevant.

4. ACKNOWLEDGEMENTS

5. CONCLUSIONS

This paper presented a novel approach for the enhancement of the quality of hyperspectral data. The proposed approach is based on NLPCA, defined as the nonlinear generalization of the standard PCA performed by autoassociative neural networks. Real images presenting different kinds of noise have been used to evaluate the ability of the NLPCA technique to effectively suppress noise. MSE, SAM and SNR have been used to quantify the improvement in image quality introduced by the proposed method and the obtained results have been compared with those obtained using the linear PCA. The results demonstrate that while linear PCA is able to reduce the influence of gaussian noise, NLPCA permits to remove also correlated noise. A further qualitative assessment has been exploited analyzing the components obtained using PCA, KPCA and NLPCA. Even if KPCA and NLPCA should return similar components, in one case the KPCA was not able to detect all the data structure in the data. Moreover, the reprojection into the original space is not always possible with KPCA. In any case, both qualitative and quantitative results demonstrated a good performance of the NLPCA approach for noise suppression. Further analysis will be conducted in order to evaluate the effectiveness of the proposed approach in de-noising other type of images featuring different kind of noise (i.e. SAR speckle).

6. REFERENCES

[1] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

[2] H. Othman and S. E. Qian, "Noise reduction of hyperspectral imagery using hybrid spatial-spectral derivative-domain wavelet shrinkage," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp. 397–408, 2006.

[3] D. Letexier and S. Bourennane, "Noise removal from hyperspectral images by multidimensional filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 7, pp. 2061–2069, 2008.

[4] Wang Y., Niu R., and Yu X., "Anisotropic diffusion for hyperspectral imagery enhancement," *IEEE Sensors J.*, vol. 10, no. 3, pp. 469–477, 2010.

[5] G. Chen and S. E. Qian, "Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 3, pp. 973–980, 2011.

[6] Goossens B., Pizurica A., and Philips W., "Removal of correlated noise by modeling the signal of interest in the wavelet domain," *Image Processing, IEEE Transactions on*, vol. 18, no. 6, pp. 1153–1165, June 2009.

[7] S. Mika, B. Schlkopf, A. Smola, K. R. Muller, M. Scholz, and G.Ratsch, "Kernel pca and de-noising in feature spaces," *Advances in Neural Information Processing Systems*, vol. 11, pp. 536–542, 1999.

[8] A. R. Teixeira, A. M. Tom, and and E. W. Lang K. Stadlthanner, "Kpca denoising and the pre-image problem revisited," *Dig. Signal Process.*, vol. 18, no. 4, pp. 568–580, 2008.

[9] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 6, pp. 2323–2326, 2000.

[10] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 6, pp. 2319–2323, 2000.

[11] T. Kohonen, "Self-organizing maps," *Springer*, 2001.

[12] B Scholkopf, A. Smola, and K. R. Mller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, vol. 10, no. 5, pp. 1299–1319, 1998.

[13] R. Muller, S. Mika, G. Ratsch, K. Tsuda, and B. Scholkopf, "An introduction to kernel-based learning algorithms," *IEEE Trans. Neural Networks*, vol. 12, no. 2, pp. 181–202, 2001.

[14] T. Y. Kwok and I. W. H. Tsang, "The pre-image problem in kernel methods," *IEEE Trans. Neural Networks*, vol. 15, no. 6, pp. 1517–1525, 2004.

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

[16] M. F. Moller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural Networks*, vol. 6, pp. 525–533, 1993.

[17] Sanger T. D., "Optimal unsupervised learning in a single-layer linear feedforward neural network," *Neural Networks*, vol. 2, 1989.

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

[19] G. Licciardi and F. Del Frate, "Pixel unmixing in hyperspectral data by means of neural networks," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 49, pp. 4163–4172, 2011.

[20] G. Licciardi, M. M. Khan, J. Chanussot, A. Montanvert, L. Condat, and C. Jutten, "Fusion of hyperspectral and panchromatic images using multiresolution analysis and nonlinear pca band reduction," *EURASIP Journal on Advances in Signal Processing*, vol. 4885, no. 8, pp. 347–358, 2012.

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

[22] G. Licciardi, M. M. Khan, and J. Chanussot, "Fusion of hyperspectral and panchromatic images: a hybrid use of indusion and nonlinear pca," *ICIP 2012, Orlando, Florida, USA, September 30 October 3 2012*, 2012.

[23] F.A. Kruse, A. B. Lefkoff, J. B. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon, and A. F. H. Goetz, "The spectral image processing system (sips) - interactive visualization and analysis of imaging spectrometer data," *Remote Sensing of Environment, Special issue on AVIRIS*, vol. 44, pp. 145–163, 1993.

[24] J. E. Fowler and J. T. Rucker, "3d wavelet-based compression of hyperspectral imagery," *Hyperspectral Data Exploitation: Theory and Applications - John Wiley and Sons Inc*, 2007.

[25] B. Datt, T. R. Mc Vicar, T. G. Van Niel, D. L. B. Jupp, and J. S. Pearlman, "Preprocessing eo-1 hyperion hyperspectral data to support the application of agricultural indexes," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 6, pp. 1246–1259, 2003.