SEASONAL VEGETATION CHANGE DETECTION USING INDEPENDENT COMPONENT ANALYSIS

S. Chitroub [a, *], S. AlSultan [b]

[a] Signal and Images Processing Laboratory, Electronics and Computer Science Faculty, U. S. T. H. B., P. O. Box 32, El – Alia, Bab – Ezzouar, 16111, Algiers, Algeria - s_chitroub@hotmail.com
[b] Al Madinah Remote Sensing Center, AlMadinah RD, Riaydh Alkhabra City, P.O.Box 242 Saudi Arabia - ssulltan@hotmail.com

KEY WORDS: Vegetation analysis, PCA, Independent component analysis, Neural network, Change detection

ABSTRACT:

The seasonal analysis of vegetation can be considered as looking for fundamental redundant information and detecting, at the same time, the natural changes of the vegetative cover undergone by the observed scene. From the statistical point of view, the redundant information can be quantified by the correlation coefficients between the multi-temporal images while the natural changes can be considered as the mutual information between the transition zones of the observed scene. For detecting and emerging the zones of transition and preserving at the same time the zones of vegetation temporal evolution stability, it is interesting to create new images in which the correlation between the images is vanished and the mutual information is minimized. To reach such purpose, we have developed a new approach for seasonal vegetation analysis based on a new statistical multi-variate method called independent component analysis (ICA).

1. INTRODUCTION

Multispectral image processing is a promising tool for the analysis of vegetation in remote sensing imagery, particularly in areas with low vegetation cover (Chen, 1998a ; Chen, 1998b ; Gracia and Ustin, 2001 ; Shabanov et al., 2001; Kogan et al., 2003; Frank and Mentz, 2003). Seasonal vegetation analysis in the absence of land cover change is much more challenging than land cover analysis. A variety of multispectral vegetation indices have been developed in order to detect these changes (Chen, 1998b ; Kogan et al., 2003; Frank and Mentz, 2003). However, these indices are insufficient for seasonal vegetation analysis, especially when the number of spectral bands is important, in which make full use of the available spectral images becomes impossible (Frank and Mentz, 2003). In addition, the developed vegetation indices are limited in the detection of low vegetation cover because of varying background signals (Frank and Mentz, 2003).

The seasonal analysis of vegetation can be considered as looking for fundamental redundant information, which exists between the multi-temporal remote sensing data (acquired for the same scene) and detecting and emerging, at the same time, the natural changes of the vegetative cover undergone by the observed scene. The redundant information characterizes the stability in the vegetation evolution in the areas that are not undergone to the natural changes across the time. The natural changes, however, characterize the transitions across the time between the states of the natural change zones of the scene. From the statistical point of view, the redundant information can be quantified by the correlation coefficients between the multi-temporal images while the natural changes can be considered as the mutual information between the transition zones. For emerging the transition zones and preserving at the same time the zones of vegetation temporal evolution stability, it is interesting to create new images in which the correlation between these images is vanished and the mutual information is minimized. To reach such purpose, we have developed a new approach of seasonal vegetation analysis based on a new statistical multivariate method called independent component analysis (ICA).

ICA is a useful extension of standard Principal Component Analysis (PCA) (Chitroub, et al., 2001 ; Chitroub, et al., 2004 ; Karhunen and Joutsensalo, 1994). As the name implies, ICA is to find the transformation such that the resulting components are as statistically independent from each other as possible. It takes into account of higher order statistical properties and its components are mutually independent with respect to these higher order statistics, thus making ICA more truly independent that PCA (Cardoso, 1999 ; Lee et al., 1999). The ICA model is very suitable for neural network realization (Hyvärinen, 1999 ; Lee et al., 2000). Most application of ICA so far has been on Blind Signal Separation (BSS) of unknown source signals from their linear mixture for which ICA obviously is useful. The use of ICA for images has been much limited. We believe that ICA can be useful in general in image and signal processing (Chitroub, et al., 2004). In this paper, we will demonstrate some potential advantages of ICA in remote sensing study. We are concerned with the seasonal analysis of vegetation. The remainder of this paper is organized as follows. The proposed model is exposed in detail in section 2. Experiments performed on the multi-temporal Landsat-TM images (they cover AlQassim region in Saudi Arabia), are given and commented in section 3. We conclude the paper in the last section.

2. ICA – BASED METHOD FOR SEASONAL VEGETATION ANALYSIS

In this paper, we demonstrate the usefulness of ICA for seasonal vegetation analysis. For that, a PCA-ICA neural network model is proposed (Figure 1). With the PCA part of the model, the Principal Component (PC) images are decorrelated and consequently the redundant information is annulled between the PC images. With ICA part of the model, we show that in the Independent Component (IC) images the mutual information is reduced compared to the Principal Component (PC) images. This implies that the zones of transition are detected and emerged and, at the same time, the zones of vegetation temporal evolution are preserved in the produced IC images.

2.1 Principal Component Extraction

The PCA-based part (Figure 2) is devoted to extract the PC images. It is based on the simultaneous diagonalization concept of the two matrices Σ_x (input images covariance matrix) and Σ_n (covariance matrix of the noise), via one orthogonal matrix A (Chitroub, et al., 2004). This means that the PC images (Y) are uncorrelated and have an additive noise that has a unit variance. This step of processing allows us making our application coherent with the theoretical development of ICA (Lee et al., 2000).

Based on the well-developed aspects of the matrix theories and computations, the existence of A is proved in (Chitroub, et al., 2004) and a statistical algorithm for obtaining it is proposed. Here, we propose a neuronal implementation of this algorithm (Chitroub, et al., 2001) with some modifications (Figure 2). It is composed of two PCA neural networks that have a same topology. The lateral weights c_j^1, respectively c_j^2 forming the vector C_1, respectively C_2, connect all the first m-1 neurons with the mth one. These connections play a very important role in the model since they work toward the orthogonalization of the synaptic vector of the mth neuron with the vectors of the previous m-1 neurons. The solid lines denote the weights w_i^1, c_j^1, respectively w_i^2, c_j^2, which are trained at the mth stage, while the dashed lines correspond to the weights of the already trained neurons. Note that the lateral weights asymptotically converge to zero, so they do not appear between the already trained neurons. The first network of Figure 2 is devoted to whitening the noise, while the second one is for maximizing the variance given that the noise is being already whitened. Let X_1 be the input vector of the first network. After convergence, the vector X is transformed to the new vector X' via the matrix $U = W_1.\Lambda^{-1/2}$, where W_1 is the weighted matrix of the first network, Λ is the diagonal matrix of eigenvalues of Σ_n and $\Lambda^{-1/2}$ is the inverse of its square root. Next, X' be the input vector of the second network. It is connected to M outputs, with $M \leq N$, corresponding to the intermediate output vector noted X_2. Once this network is converged, the PC images to be extracted (vector Y) are obtained such as: $Y = A^T.X = U.\ W_2.X$, where W_2 is the weighted matrix of the second network. The activation of each neuron in the two parts of the network is a linear function of their inputs. The kth iteration of the learning algorithm, for both networks, is:

$$\begin{aligned} w(k+1) &= w(k) + \beta(k)\left(q_m(k)\mathbf{P} - q_m^2(k)w(k)\right) \\ c(k+1) &= c(k) + \beta(k)\left(q_m(k)\mathbf{Q} - q_m^2(k)c(k)\right) \end{aligned} \qquad (1)$$

where P and Q are, respectively, the input and output vectors of the network. $\beta(k)$ is a positive sequence of learning parameter. The global convergence of the PCA-based part of the model is strongly dependent on the parameter β. The optimal choice of this parameter is well studied in (Chitroub, et al., 2001).

2.2 Independent Component Extraction

The M inputs of the ICA network model (Figure 3) are the PC images. The M output neurons correspond to the IC images (vector Z), then $Z = B.Y$, where B is the separating (or de-mixing) matrix that we want to determine.

ICA can be carried out by using many different methods (Chitroub, et al., 2004 ; Cardoso, 1999 ; Karhunen and Joutsensalo, 1994 ; Lee et al., 1999 ; Hyvärinen, 1999). In this paper, we have used the Informax algorithm to learn the matrix B. Using the concept of differential entropy and the invertible transformation of $Z = B.Y$, the mutual information between the outputs is minimized. This means that finding an invertible transformation B that minimizes the mutual information is approximately equivalent to finding directions in which the mutual information among the output components is minimized. The weight update rule will then be a gradient descent in the direction of maximum joint entropy. The mathematical details of the learning process is out of the scope of this paper and the reader could be consulting, for more details, the following references (Chitroub, et al., 2004 ; Karhunen and Joutsensalo, 1994 ; Lee et al., 1999 ; Lee et al., 2000).

Using the concept of differential entropy and the invertible transformation of $Z = B.Y$, the mutual information between the outputs is:

$$I(\mathbf{z}) = \sum_{i=1}^{M} H(z_i) - H(\mathbf{y}) + log(|det\ \mathbf{B}|) \qquad (2)$$

where $H(Z_i)$ are the marginal entropies of the outputs and $H(Z)$ is the joint entropy of Z. By constraining z_i to be uncorrelated and of unit variance, this implies that: $det\ E(\mathrm{z.z}^T) = 1$. As the negentropy is a measure of non-Gaussianity, that is:

$$J(\mathrm{z}) = H(\mathrm{z}_{Gaussian}) - H(\mathrm{z}) \qquad (3)$$

So the mutual information and negentropy differ only by a constant that does not depend on B and the sign, that is:

$$I(\mathrm{z}) = C - \sum_{i=1} z \qquad (4)$$

which means that finding an invertible transformation B that minimizes the mutual information is approximately equivalent to finding directions in which the sum of non-Gaussianities of z_i is maximized. Maximizing the joint entropy $H(Z)$ can approximately minimize the mutual information among the output components:

$$z_i = g_i(v_i) \qquad (5)$$

where $g_i(v_i)$ is an invertible monotonic non-linearity and $V = B.Y$. If the mutual information among the outputs is zero, the mutual information before the non-linearity must be zero as well since the nonlinear transfer function does not introduce any dependencies. Thus, the relation between z_i, v_i, and $g_i(v_i)$ is such as:

$$p(z_i) = p(v_i) / \left|\partial g_i(v_i) / \partial v_i\right| \qquad (6)$$

By this relationship, *g(V)* must be chosen so that its derivative approximately forms a probability distribution function for the sources to be recovered. The only remaining parameters to adapt are the synaptic weights that can be found by maximizing *H(Z)* with respect to *B*. The weight update rule will then be a gradient descent in the direction of maximum joint entropy.

More computationally efficient approaches have been proposed in (Lee et al., 2000), the reader can obtain more of the mathematical details in (Chitroub et al., 2006). If we define the term score function $\varphi(\mathbf{v})$ as:

$$\varphi(\mathrm{v}) = (\partial p(\mathrm{v})/\partial \mathrm{v})/p(\mathrm{v}) \tag{7}$$

then an efficient weight update is:

$$\Delta \mathrm{B} \propto \left(\mathrm{I} - \varphi(\mathrm{v}).\mathrm{v}^T\right)\mathrm{B} \tag{8}$$

The form of $\varphi(\mathbf{v})$ plays a crucial role because it is function of the transfer and therefore a function of the source estimate. For the sub-Gaussian sources, the form of $\varphi(\mathbf{v})$ is such as:

$$\varphi(\mathrm{v}) = \mathrm{v} - tanh(\mathrm{v}) \tag{9}$$

where $tanh(.)$ is the hyperbolic tangent. For the super-Gaussian sources, $\varphi(\mathbf{v})$ takes the form:

$$\varphi(\mathrm{v}) = \mathrm{v} + tanh(\mathrm{v}) \tag{10}$$

The switching between the sub-Gaussian and super-Gaussian learning rule gives the following learning rule for the ICA – part model (Lee et al., 1999):

$$\Delta \mathrm{B} \propto \left(\mathrm{I} - \mathrm{K}.tanh(\mathrm{v}).\mathrm{v}^T - \mathrm{v}.\mathrm{v}^T\right)\mathrm{B} \tag{11}$$

K is a *N*-dimensional diagonal matrix with elements $sign(k_4(v_i))$. $k_4(v_i)$ is the kurtosis of the source estimate v_i. The switching parameter $k_4(v_i)$ can be derived from the general stability analysis of separating solutions (Cardoso, 1996 ; Hyvärinen, 1999).

3. EXPERIMENTAL RESULTS

We present in this section our preliminary results. More detailed study and more completed results are under development. They will be subject of the future works for publication. A real multi-temporal data provided by the Landsat-TM are used to evaluate the proposed method. The data were acquired over the AlQassim region in Saudi Arabia (140x235pixels) during April and June 1994. The fifth bands of the two sets of data are shown in Figure 4. The first three extracted PC images are given in Figure 4. The first components have the best image quality (contrast). Figure 5 shows the extracted IC images. These images are different to the PC images. In the PC images the correlation is vanished and consequently the redundant information is minimized. In these images the zones of vegetation temporal evolution stability are well mapped since they are characterized by the variance of the PC image. This variance is maximized in the first PC images. While the contrast between the input spectral images, which characterize the differences between the spectral bands, is mapped the last PC images that are much noised and consequently it is not possible to overcome the information about the natural changes undergone by the observed scene.

However, in IC images the mutual information between the PC images are minimized and so the natural changes, which can be considered as the mutual information between the transition zones of the PC images, are emerged. In the IC images, the zones of vegetation temporal evolution stability are also preserved as they are in PC images. This can be quantified by computing the image of the vegetation stability zones and the image of the vegetation transition zones (natural changes) of the scene both from the first and the second extracted IC images (Figure 6).

4. CONCLUSION

In this paper, we have presented a new method for seasonal vegetation analysis. The method is based on the emergent technique, which is the independent component analysis (ICA). The basic idea is to consider that the natural change undergone by the observed scene can be detected and emerged if the mutual information that exists between the images is minimized. This task can be realised by the compound neural network model PCA-ICA. We have presented here the main outlines of the theoretical analysis of such model. More complicated mathematical development of the proposed method cannot be given here and it will be subject of the future works for publication. Although the preliminary results, given in this paper, needed to be evaluated by an objective and mathematical criteria, they are, however, of interesting in the sense that the extracted IC images are very informative about the surface state of the observed scene, concerning the stability and transition zones of the vegetation, compared to the extracted PC images.

REFEFRENCES

Cardoso, J. F. Laheldm B., 1996. Equivariant adaptive source separation. *IEEE Transactions on Signal Processing*, 45(2), pp. 434-444.

Cardoso, J. F., 1999. High-order contrasts for independent component analysis. *Neural Computation*, 11(1), pp. 157-192.

Chen, Z., Elvidge, C. D. and Groeneveld, D. P., 1998a. Monitoring seasonal dynamics of arid land vegetation using AVIRIS data. *Remote Sensing of Environment*, 65(2), pp. 255-266.

Chen, Z., Elvidge, C. D. and Groeneveld, D. P., 1998b. *Vegetation change detection using high spectral resolution vegetation indices*. In Remote Sensing Change Detection: Environmental Monitoring Applications and Methods, C. D. Elvidge and E. R. Lunettat, Eds. Ann Arbor Press, Ch. 11, pp. 181-190.

Chitroub, S., Houacine, A. and Sansal, B., 2001. Neuronal principal component analysis for an optimal representation of multispectral images. *Intelligent Data Analysis, International Journal*, 5(5), pp. 385-403.

Chitroub, S., Houacine, A. and Sansal, B., 2004. PCA-ICA neural network model for POLSAR images analysis. In *Proceedings of IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP'04)*, Montreal, Canada. Vol. V, pp. 757-760.

S. Chitroub, 2006. *Two ICA Approaches for SAR Image Enhancement Part II: Independent Component Analysis of POLarimetric Synthetic Aperture Radar (POLSAR) Images. Bayesian Approach*. Signal and Image Processing for Remote Sensing", Edited by Prof. C.H. Chen, University of Massachusetts Dartmouth, Publisher: Taylor & Francis, CRC Pressm pp. 655-675.

Frank, M. and Menz, G., 2003. Detecting Seasonal Cseasonal changes in a semi-arid environment using hyperspectral vegetation indices. *Presented at the 3rd EARSel Workshop on Imageing Spectroscopy*, Herrsching.

Gracia, M. and Ustin, S. L., 2001. Detection of interannual vegetation responses to climatic variablity using AVIRIS data in a coastal savanna in California. *IEEE Transactions on Geoscience and Remote Sensing*, 39(7), pp. 1480-1490.

Hyvärinen, A., 1999. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10(3), pp. 626-634.

Karhunen, J. and Joutsensalo, J., 1994. Representation and separation of signals using nonlinear PCA type learning. *Neural Networks*, 7, pp. 113-127.

Kogan, N., Gitelson, A., Edige, Z., Spivak I., and Lebed, L., 2003. AVHRR-based spectral vegetation index for quantitative assessment of vegetation state and productivity: Calibration and validation. *Photogrammetric Engineering and Remote Sensing*, 69(8), pp. 899-906.

Lee, T. W., Girolami, M., and Sejnowski, T. J. , 1999. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Computation*, 11, pp. 417-441.

Lee, T. W., Girolami, M., Bell, A. J. and Sejnowski, T. J. , 2000. A Unifying information-theoretic framework for independent component analysis, *Neural Networks*, 39, pp. 1-21.

Shabanov, N. V., Zhou, L., Knyazikhin, Y., Ranga, B. and Tucker, C. J., 2001. Analysis of interannual changes in northern vegetation activity observed in AVHRR data from 1981 to 1994. *IEEE Transactions on Geoscience and Remote Sensing*, 39(7), pp. 1480-1490.

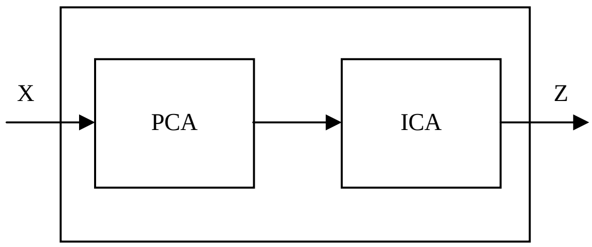

Figure 1. PCA-ICA neural network model

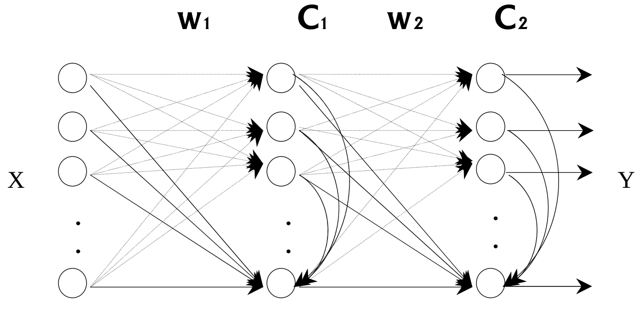

Figure 2. PCA-based part of the model

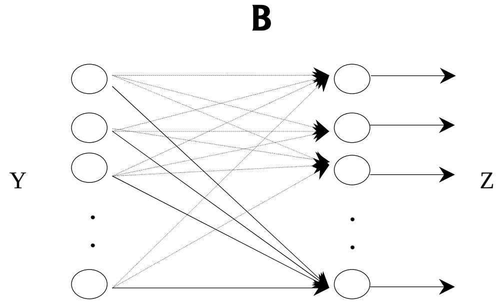

Figure 3. ICA-based part of the model

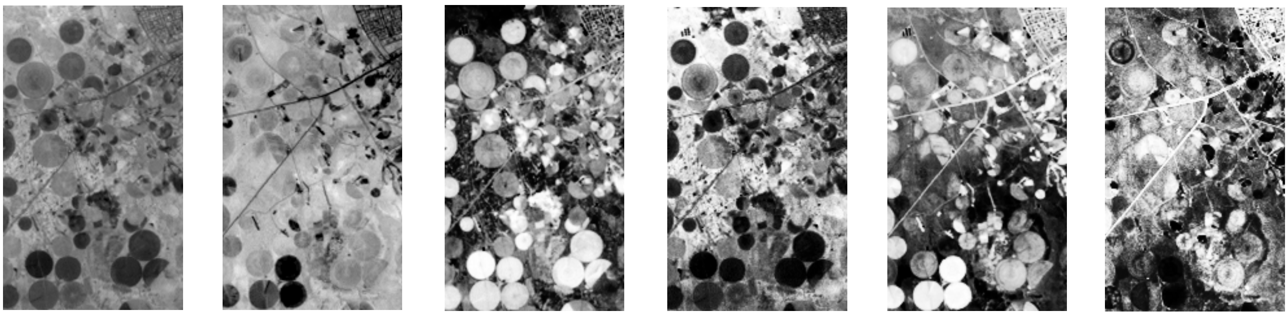

Figure 4. (a) The fifth spectral band of April 94, (b) The fifth spectral band of June 94, (c) The first PC of April 94 data set, (d) The second PC of April 94 data set, (e) The first PC of June 94, (f) The second PC of June 94

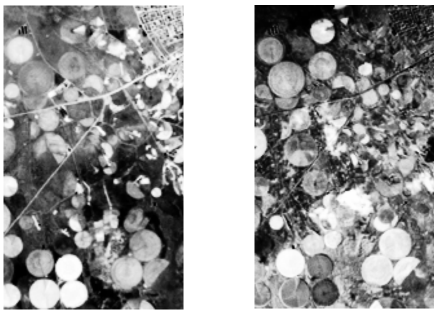

Figure 5. The first and the second IC images

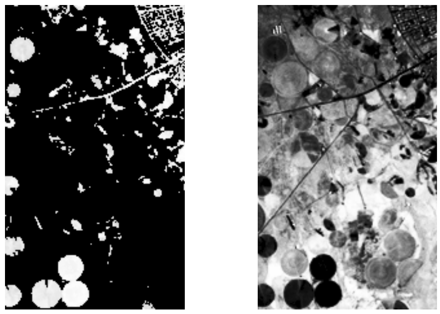

Figure 6. The stability and transition zones of the vegetation