Unsupervised Hyperspectral Image Analysis Using Independent Component Analysis (ICA)

Shao-Shan Chiang[1] Chein-I Chang[1] Irving W. Ginsberg[2]
[1]Remote Sensing Signal and Image Processing Laboratory
Department of Computer Science and Electrical Engineering
University of Maryland Baltimore County, Baltimore, MD 21250
[2]Remote Sensing Laboratory, U.S. Department of Energy, Las Vegas, Nevada 89191

ABSTRACT

In this paper, an ICA-based approach is proposed for hyperspectral image analysis. It can be viewed as a random version of the commonly used linear spectral mixture analysis, in which the abundance fractions in a linear mixture model are considered to be unknown independent signal sources. It does not require the full rank of the separating matrix or orthogonality as most ICA methods do. More importantly, the learning algorithm is designed based on the independency of the material abundance vector rather than the independency of the separating matrix generally used to constrain the standard ICA. As a result, the designed learning algorithm is able to converge to non-orthogonal independent components. This is particularly useful in hyperspectral image analysis since many materials extracted from a hyperspectral image may have similar spectral signatures and may not be orthogonal. The AVIRIS experiments have demonstrated that the proposed ICA provides an effective unsupervised technique for hyperspectral image classification.

I. INTRODUCTION

In the past years [1], linear spectral mixture analysis (LSMA) has been widely used for endmember unmixing. It models a pixel in an image scene as a linear mixture of materials with relative abundance concentrations. Two restrictions are generally applied to LSMA. One is that complete knowledge about materials must be given *a priori*. In many practical applications, obtaining such *a priori* information is usually difficult if not impossible. To relax this requirement, an unsupervised method to generate material information from the image data is needed. The second restriction is that the abundance fractions of endmembers are unknown, non-random constants that can be estimated by statistical methods such as least squares estimation. Due to noise and atmospheric effects, however, the mixture of apparent (i.e., observed) material abundance fractions may vary pixel-by-pixel and consequently can be viewed as a random process resulting from an apparently random composition of multiple spectra for distinct materials plus random noise. Therefore, it is more realistic to assume that the abundance fractions of materials in a pixel are random quantities rather than unknown constants. In order to appropriately represent such a random linear mixture, the abundance fraction of each material must be viewed as a random signal source. Under this circumstance, it requires the prior knowledge of their probability distributions. This further complicates the problem. Independent Component Analysis (ICA) [2] seems to provide a feasible approach to solving this random abundance mixture problem. ICA is an unsupervised source separation process. It differs from Principal Components Analysis (PCA) in many aspects. Unlike PCA which only requires the second order statistics, ICA looks for components which are *statistically independent*, which is a much stronger condition than *uncorrelated*. As a result, it requires statistics of orders higher than second order. In addition, ICA components are not necessarily geometrically orthogonal. The most important difference is that ICA needs a linear model to describe data while PCA does not. Therefore, ICA is not a generalization of PCA. However, the requirement of linear model is exactly what we are interested in for ICA. Assume that the abundance fraction of each material is an unknown random signal source. Then the source mixing-model considered in ICA can be directly applied to LSMA, in which case ICA can be an effective means to solve random abundance fractions for the linear mixture model used in LSMA.

II. INDEPENDENT COMPONENT ANALYSIS (ICA)

Suppose that L is the number of spectral bands. Let $\mathbf{r}$ be an $L \times 1$ column pixel vector in a multispectral or hyperspectral image where the bold face is used for vectors. Let $\mathbf{M}$ be an $L \times p$ endmember signature matrix denoted by $[\mathbf{m}_1\, \mathbf{m}_2 \cdots \mathbf{m}_p]$ where $\mathbf{m}_j$ is an $L \times 1$ column vector represented by the signature of the j^{th} endmember of materials resident in the pixel $\mathbf{r}$ and p is the number of endmembers in the pixel. Let $\boldsymbol{\alpha} = (\alpha_1\, \alpha_2 \cdots \alpha_p)^T$ be a $p \times 1$ abundance column vector associated with $\mathbf{r}$ where α_j denotes the fraction of the j^{th} signature present in the pixel vector $\mathbf{r}$. Assume that the spectral signatures of the p

endmembers in the pixel vector $\mathbf{r}$ are linearly mixed and that $\boldsymbol{\alpha}$ is an unknown constant vector. In this case, the spectral signature of a pixel vector $\mathbf{r}$ can be represented by the linear regression model.

$$\mathbf{r} = \mathbf{M}\boldsymbol{\alpha} + \mathbf{n} \,, \tag{1}$$

where **n** is noise that can be interpreted as measurement error. One drawback of LSMA is that the signature matrix **M** must be known *a priori*.

In this section, an ICA is described. It is also based on model (1), but does not require the prior knowledge of **M**. In addition, it assumes that the *p* abundance fractions, $\alpha_1, \alpha_2, \cdots, \alpha_p$, are unknown random signal sources instead of unknown constants as assumed in model (1). However, in this case we also need to make three additional assumptions on the abundance vector $\boldsymbol{\alpha} = (\alpha_1 \, \alpha_2 \cdots \alpha_p)^T$:

(i) The endmember signature matrix **M** is full rank, that is, the *p* material endmember signature vectors, $\mathbf{m}_1, \mathbf{m}_2, \cdots, \mathbf{m}_p$ must be linearly independent.

(ii) The *p* abundance fractions $\alpha_1, \alpha_2, \cdots, \alpha_p$ are mutually statistically independent.

(iii) At most one of the *p* abundance fractions $\alpha_1, \alpha_2, \cdots, \alpha_p$ is Gaussian. In order to use ICA for our application in hyperspectral image analysis, the mixing matrix is the **M** in model (1) and the unknown signal sources to be separated are *p* random abundance fraction sources denoted by $\alpha_1, \alpha_2, \cdots, \alpha_p$. ICA finds a $p \times L$ separating matrix **W** to unmix the $\alpha_1, \alpha_2, \cdots, \alpha_p$ from **r** via the equation

$$\hat{\boldsymbol{\alpha}}(\mathbf{r}) = \mathbf{W}\mathbf{r} \,, \tag{2}$$

where $\hat{\boldsymbol{\alpha}}(\mathbf{r}) = (\hat{\alpha}_1(\mathbf{r}), \hat{\alpha}_2(\mathbf{r}), \cdots, \hat{\alpha}_p(\mathbf{r}))^T$ is the estimated abundance vector based on **r** and is used to unmix the *p* independent random abundance fractions $\alpha_1, \alpha_2, \cdots, \alpha_p$.

Under the above assumptions, the estimate of the i^{th} abundance fraction α_i may appear as any component $\hat{\boldsymbol{\alpha}}(\mathbf{r})$ because changing order of components in $\hat{\boldsymbol{\alpha}}(\mathbf{r})$ does not affect their statistical independence. In order to simplify infomax criterion used in ICA, Comon [3] introduced an alternative criterion, referred to as contrast functions that maximize the higher order statistics of the data given by

$$\max_{\mathbf{W}, E\left[\hat{\boldsymbol{\alpha}}(\mathbf{r})\hat{\boldsymbol{\alpha}}(\mathbf{r})^T\right] = \mathbf{I}} \left\{ \psi(\mathbf{W}) = \sum_{j=1}^{p} E\left[\hat{\alpha}_j(\mathbf{r})^m\right] \right\} \quad \text{for} \quad m \geq 3 \,, \tag{3}$$

where **I** is the $p \times p$ identity matrix. This constrained maximization problem is equivalent to maximizing the following cost function

$$J(\mathbf{W}) = \psi(\mathbf{W}) - \frac{\lambda}{2} \sum_{i,j=1}^{p} E\left[\hat{\alpha}_i(\mathbf{r})\hat{\alpha}_j(\mathbf{r}) - \delta_{ij}\right]^2 \,. \tag{4}$$

From (4), a learning algorithm to generate the separating matrix **W** can be from

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu E\left[g(\hat{\boldsymbol{\alpha}}(\mathbf{r}))\mathbf{r}^T\right] - \eta\left(E\left[\hat{\boldsymbol{\alpha}}(\mathbf{r})\hat{\boldsymbol{\alpha}}(\mathbf{r})^T\right] - \mathbf{I}\right)\hat{\boldsymbol{\alpha}}(\mathbf{r})\mathbf{r}^T \,, \tag{5}$$

where μ and η are learning parameters.

III. EXPERIMENTS

The hyperspectral data used are Airborne visible infrared imaging spectrometer (AVIRIS) data extracted from a scene of the Lunar Crater Volcanic Field in Northern Nye County, Nevada (Figure 1). Water bands and low signal noise ration bands have been removed from the data, reducing the images data from 224 to 158 bands. There are five target signatures of interest, cinders, rhyolite, playa (dry lakebed), vegetation and shade.

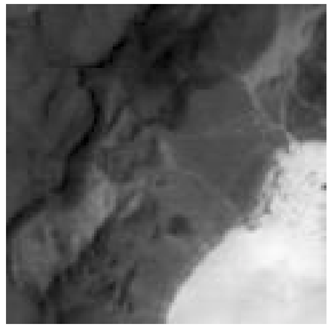


Figure 1

Since there is no prior knowledge about the number of target signatures, we first assume that there is a large number of materials, p = 158. Skewness and kurtosis were used in (4) as criteria. Our experiments showed that skewness performed better than kurtosis. So, only skewness results were given in this paper. Since very little information was found in all images after the 9^{th} component, Fig. 2 shows only the first 9 component images, labeled (a-i). The targets cinders, rhyolite and shade were extracted in Figures 2(a), 2(d) and 2(i), respectively, while the vegetation was picked up in Figures 2(c) and 2(e), and playa (dry lakebed) was shown up in Figures 2(b) and 2(f-h). This experiment showed that if p is taken too large, it would classify as target materials those spectral variations produced by mixing. This is illustrated in Fig. 2, vegetation was classified in two separate components in Figures 2(c) and 2(e). Similarly, due to a large coverage of the dry lakebed, the different abundance fractions of the playa were detected and classified in four separate images in Figures 2(b) and 2(f-h).

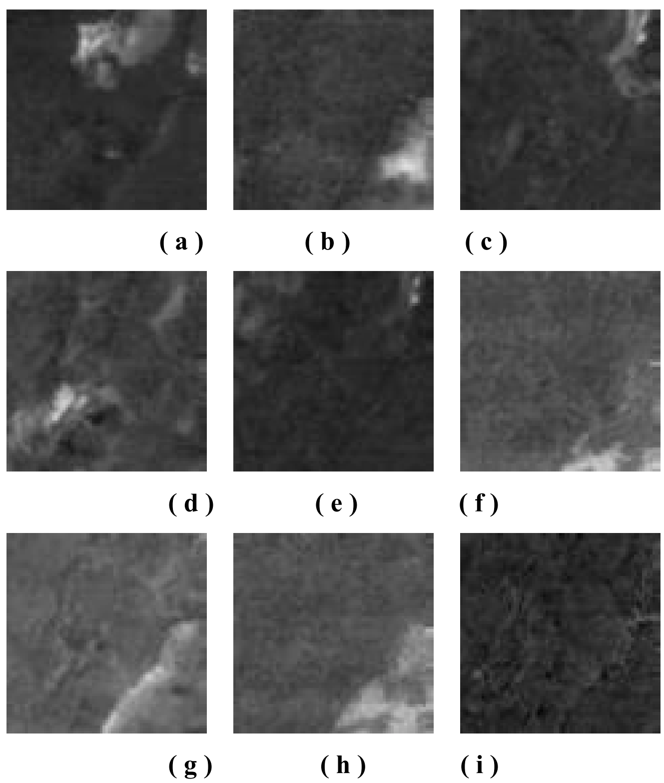
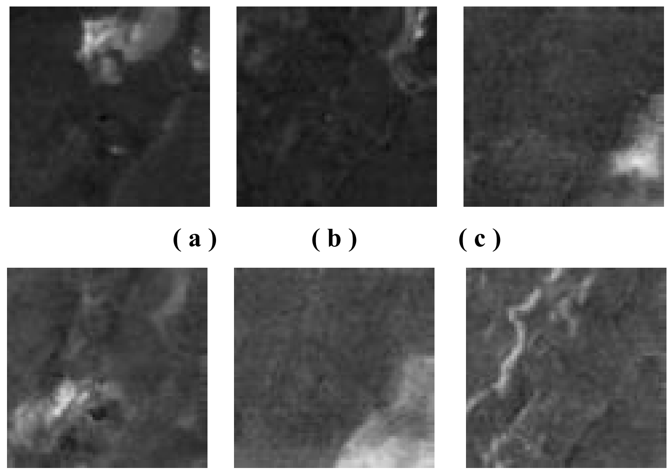


Figure 2

However, if we chose p = 12, only the first six component images labeled by (a-f) contained the information shown in Figure 3. The cinders, vegetation, rhyolite and shade were extracted in Figures 3(a), 3(b), 3(d) and 3(f), while the playa was detected in Figures 3(c) and 3(e). Also in this case, the vegetation was classified in only one image, and the number of images in which the playa was detected was cut down from 4 to 2.

If p was further reduced to 8, Figure 4 shows the first 5 component images wherein the cinders, vegetation and rhyolite were detected and classified in Figures 4(a), 4(b) and 4(d). The playa was still classified into two separate images in Figures 4(c) and 4(e). In this case, only five component images were found to contain information, of which two were used to classify the playa. As a result, no component image could be spared to classify the shade.

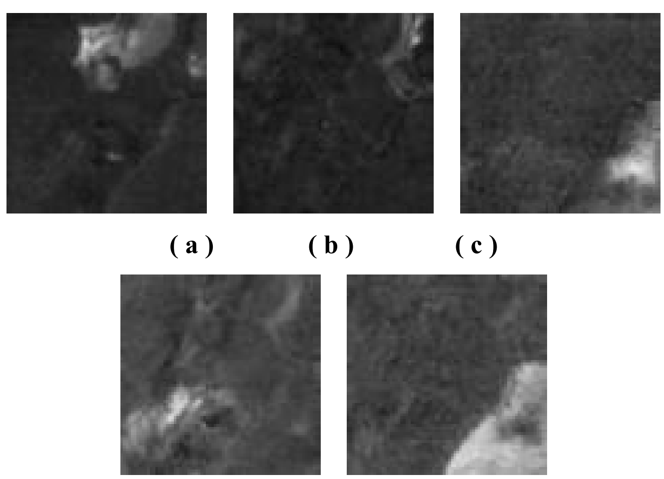
These three experiments demonstrated that the value of p to be used is crucial for unsupervised image analysis. This issue is closely related to the determination of the intrinsic dimensionality of images, and has been investigated in [5].



(a) (b) (c)

(d) (e) (f)

Figure 3



(a) (b) (c)

(d) (e)

Figure 4

IV. CONCLUSION

Recently, ICA has received considerable interest in hyperspectral image analysis [4]. This paper presented an ICA approach to hyperspectral image analysis which is different from the commonly used ICA approach. First, the separating matrix **W** is not necessarily a square matrix of full rank. Second, the mixing matrix **M** is not necessarily orthogonal. Third, the learning algorithm was designed based on the independency of the material abundance fractions not the matrix **W**. These three advantages have been shown by experiments to be very useful in hyperspectral image classification. However, since ICA separates unknown signal sources rather than estimates the strengths of the signals, it is very effective in detection and classification, but not quantification.

ACKNOWLEDGMENTS

The authors would like to thank Bechtel Nevada under contract No. DE-AC08-96NV11718 through the U.S. Department of Energy for their support and Dr. J.C. Harsanyi for providing the AVIRIS data.

REFERENCES

[1] J.B. Adams, M.O. Smith, and A.R. Gillespie, "Image spectroscopy: interpretation based on spectral mixture analysis," *Remote Geochemical Analysis: Elemental and Mineralogical Composition*, edited by C.M. Pieters and P.A. Englert, Cambridge University Press, pp. 145–166, 1993.

[2] T.W. Lee, *Indpendent Component Analysis: Theory and Applications*, Boston: Kluwer Academic Publishers, 1998.

[3] P. Comon, "Independent component analysis, A new concept?", *Signal Processing*, vol. 36, pp. 287–314, 1994.

[4] C.H. Chen and X. Zhang, "Independent component analysis for remote sensing study," *EOS/SPIE Symposium on Remote Sensing, Conference on Image and Signal Processing for Remote Sensing V*, SPIE vol. 3871, Florence, Italy, pp. 150–158, September 20-24, 1999.

[5] C.-I Chang and Q. Du, "A noise subspace projection approach to determination of intrinsic dimensionality for hyperspectral imagery," *EOS/SPIE Symposium on Remote Sensing, Conference on Image and Signal Processing for Remote Sensing V*, SPIE vol. 3871, Florence, Italy, pp. 34–44, September 20–24, 1999.