TOWARDS THE DEVELOPMENT OF NEXT GENERATION REMOTE SENSING TECHNOLOGY – ERDAS IMAGINE INCORPORATES A HIGHER ORDER FEATURE EXTRACTION TECHNOQUE BASED ON ICA

Chintan A. Shah
University of California Los Angeles, Los Angeles, CA
cashah@ucla.edu

Ian Anderson, Senior Product Manager
Zaiyong Gou, Project Manager
Steven Hao, Principal Scientist
Andrew Leason, Senior Software Engineer
Leica Geosystems Geospatial Imaging, Norcross, GA
ian.anderson@lggi.com

ABSTRACT

Analysis of multi/hyperspectral imagery necessitates a selection of an optimal subset of bands in order to avoid the Hughes phenomena and parameter estimation problems due to interband correlation. This can be achieved by employing feature extraction techniques for significant reduction of data dimensionality. From the perspective of statistical pattern recognition, feature extraction refers to a process, whereby a data space is transformed into a feature space, in which the original data is represented by a reduced number of effective features, retaining most of the intrinsic information content. Feature extraction from remote sensing imagery must also aim to enhance the discriminability of surface materials based on their spectral characteristics. Conventional remote sensing feature extraction techniques such as Principal Component Analysis (PCA) and Minimum Noise Fraction (MNF) model the multi/hyperspectral image data with a multivariate Gaussian distribution. Inadequacy of such algorithms stems from the Gaussian distribution assumption, which is only an assumption rather than a demonstrable property of most remote sensing data. In order to overcome this fundamental limitation, we at Leica Geosystems have developed a feature extraction technique based on Independent Component Analysis (ICA) that exploits the higher order statistical characteristics of multi/hyperspectral imagery. In this paper, we will provide a theoretical formulation of ICA and prove its superiority over PCA and MNF in a mathematical framework. Our illustrations would also demonstrate the potential benefits of employing ICA in improving the performance of several multi/hyperspectral analysis techniques such as spectral unmixing, classification, and anomaly detection.

INTRODUCTION

Any given remote sensing image can be decomposed into several features (Jia and Richards, 1999). The term 'feature' here refers to remote sensing scene objects (e.g. vegetation types, urban materials) with similar spectral characteristics. Therefore, the main objective of a feature extraction technique is to accurately retrieve these features. The extracted features can be subsequently utilized for improving the performance of various remote sensing applications (e.g. classification, target detection, unmixing, etc.).

Data Model

The data model employed in conventional feature extraction techniques is presented in Figures 1 and 2. Each spectral band in Figure 1 is assumed to be a mixture of several features such as those shown in Figure 2. The formulation of each band in the case of three features can be expressed as (Shah, 2003)

$$spectral\ band\left(\lambda\right) = A_1\left(\lambda\right)\cdot feature1 + A_2\left(\lambda\right)\cdot feature2 + A_3\left(\lambda\right)\cdot feature3 \qquad (1)$$

where $A_i\left(\lambda\right)$ is the sensor response to feature *i* at wavelength $\tilde{\lambda}$.

Figure 1. Simulated multispectral imagery.

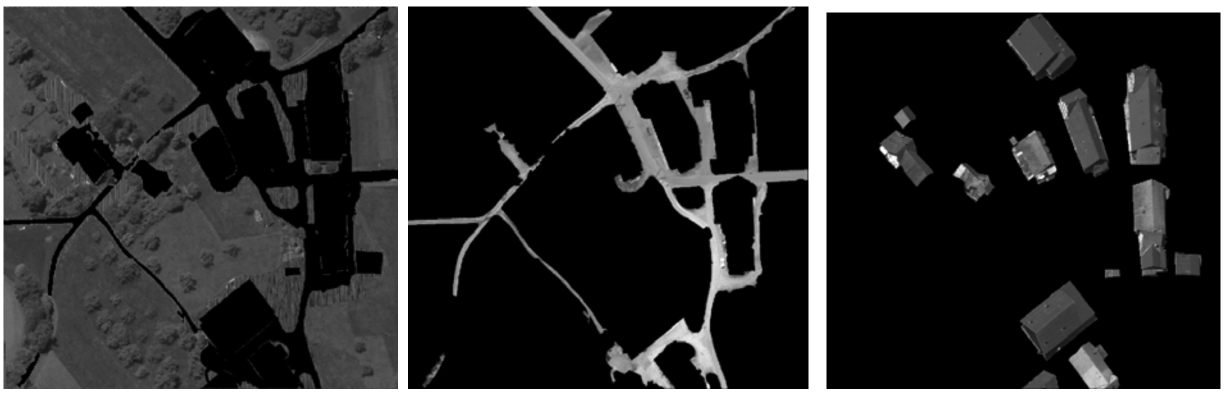


Figure 2. Features corresponding to vegetation, roads, and roof tops.

ICA BASED FEATURE EXTRACTION

PCA (Jia and Richards, 1999) and MNF (Green et al., 1988) are two widely employed feature extraction techniques in remote sensing. These techniques aim to decorrelate the spectral bands to recover the original features. In other words, these techniques perform linear transformation of the spectral bands such that the resulting components are uncorrelated. The results of applying PCA on the spectral bands in Figure 1 are presented in Figure 3. The results in Figure 3 highlight PCA's inability in extracting the original features (shown in Figure 2).

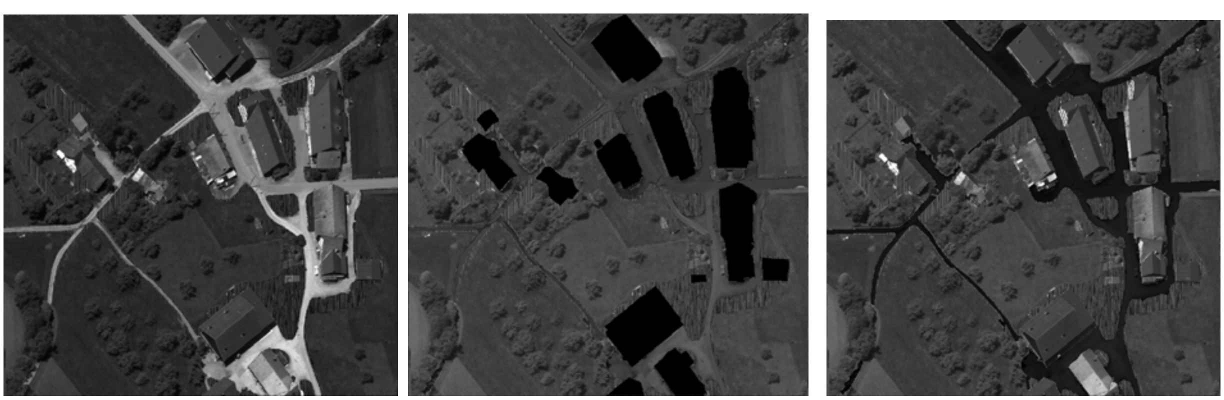


Figure 3. Features extracted using PCA.

ICA based feature extraction technique performs a linear transformation to obtain the independent components (ICs). A direct implication of this is that each component will contain information corresponding to a specific feature. ICs of the spectral bands in Figure 1 are presented in Figure 4. The results indicate that ICA recovers the original features accurately. For a detailed description on the mathematical formulation of ICA refer to (Shah, 2003 and Common, 1994).

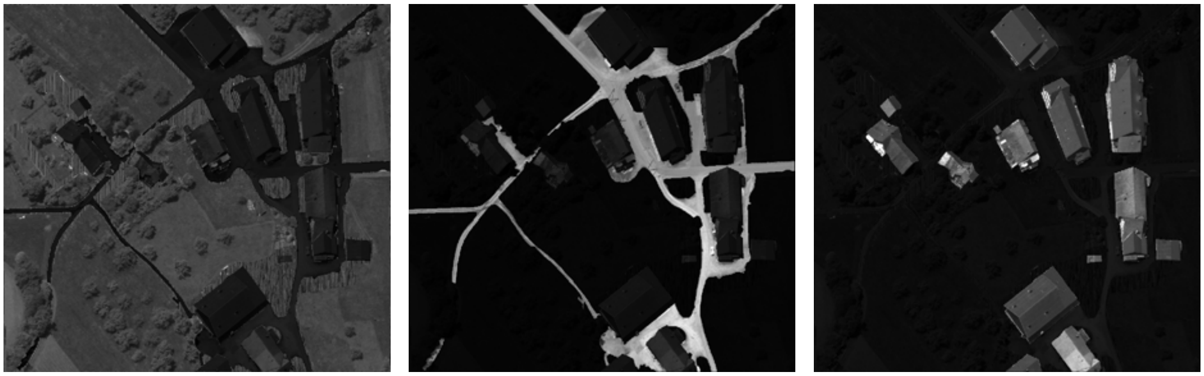


Figure 4. Features extracted using ICA.

IMAGINE ICA USER INTERFACE

The IMAGINE ICA user interface is shown in Figure 5. This interface is similar to the IMAGINE PCA interface and allows the user to specify the desired number of components. As each IC will correspond to a feature, the user should be able to approximate the number of features present in the scene.

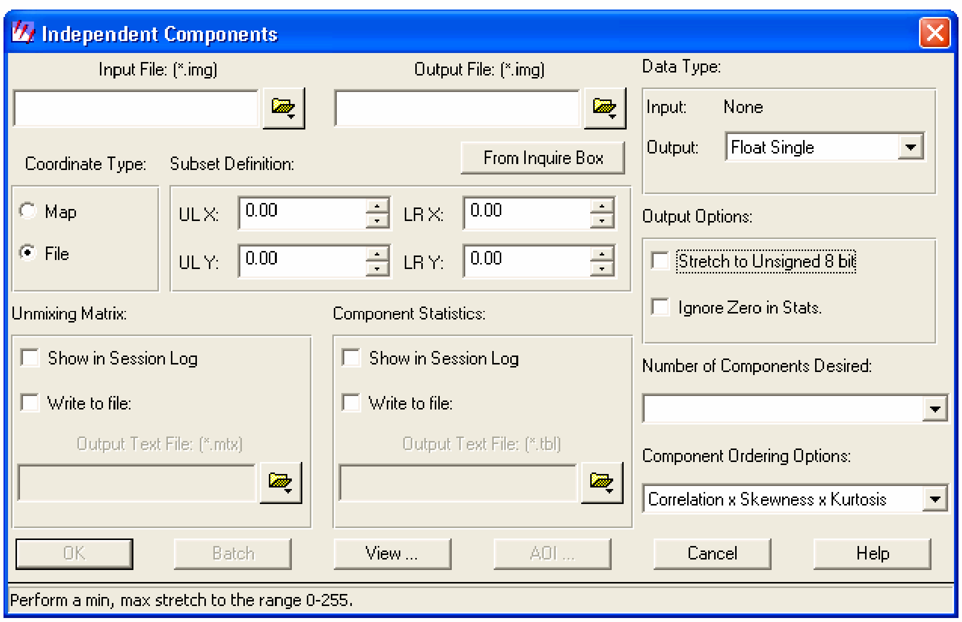


Figure 5. IMAGINE ICA user interface.

It is always desirable to have number of components greater than the number of features in order to ensure that all the features are recovered as ICs. In the case where the number of features happens to be smaller than the number of components, the additional components would contain very little feature information and almost resemble a noisy image. Hence, the user is provided with an option of ordering these ICs so that the noisy images will be the last few components and can be easily eliminated from further analysis.

Component Ordering

The available options for component ordering are:

None. No ordering is applied to the components; as a result, the IC resembling noisy images may appear in an arbitrary order.

Three basic statistical measures provided for component ordering are as follows:

Correlation Coefficient. Correlation coefficient R_{XY} between two images X and Y is defined as

$$R_{XY} = \frac{\sum_{i=1}^{P}\left(X_i - \overline{X}_i\right)\left(Y_i - \overline{Y}_i\right)}{\sqrt{\sum_{i=1}^{P}\left(X_i - \overline{X}_i\right)^2}\sqrt{\sum_{i=1}^{P}\left(Y_i - \overline{Y}_i\right)^2}}, 0 \le \left|R_{XY}\right| \le 1 \qquad (2)$$

In equation (2), P is the total number of image pixels. Correlation coefficient is a measure of similarity between two images; higher the correlation more is the similarity between their pixel values. The ICs are ordered based on their correlation with the spectral bands. ICs with low correlation correspond to noisy images.

Skewness. Skewness is a measure of lack of symmetry in an image histogram. The skewness of an image X is defined as

$$skewness_X = \frac{1}{P-1}\frac{\sum_{i=1}^{P}\left(X_i - \overline{X}_i\right)^3}{\sigma_X^3}, 0 \le \left|skewness_X\right| < \infty \qquad (3)$$

Kurtosis. Kurtosis is a measure of peakedness or flatness of an image histogram.

$$kurtosis_X = \frac{1}{P-1}\frac{\sum_{i=1}^{P}\left(X_i - \overline{X}_i\right)^4}{\sigma_X^4} - 3,\ 0 \le \left|kurtosis_X\right| < \infty \qquad (4)$$

An image with a normally distributed histogram has zero skewness and kurtosis.

Combinations of above three measures can also be employed for component ordering.

$\boldsymbol{skewness_X \times kurtosis_X}$.

$\boldsymbol{R_{XY} \times skewness_X \times kurtosis_X}$

Negentropy. Negentropy of an image is defined as (Common, 1994)

$$negentropy_X = \frac{1}{12}\left(skewness_X\right)^2 + \frac{1}{48}\left(kurtosis_X\right)^2 \qquad (5)$$

and is proportional to its skewness and kurtosis.

Lower values of skewness/kurtosis/negentropy correspond to ICs resembling noisy image.

BAND GENERATION FOR MULTISPECTRAL IMAGERY

It must be noted that the number of desired components cannot exceed the number of spectral bands in the imagery. This should not be of concern when processing hyperspectral imagery, where the number of spectral bands is significantly higher compared to the number of features in the scene. However, feature extraction from multispectral imagery necessitates the generation of additional spectral bands as explained below.

A linear combination of the original spectral bands will not lead to additional information required for ICA feature extraction from multispectral imagery. Therefore, the user is required to generate additional spectral bands through non-linear operations such as $\log\left(X_i\right)$, $\sqrt{X_i}$, X_i^2, $\frac{X_i}{X_j}$, $X_i \cdot X_j$ (where X_i is a multispectral band and $i \ne j$).

Figure 6 provides with the IMAGINE ICA procedural overview.

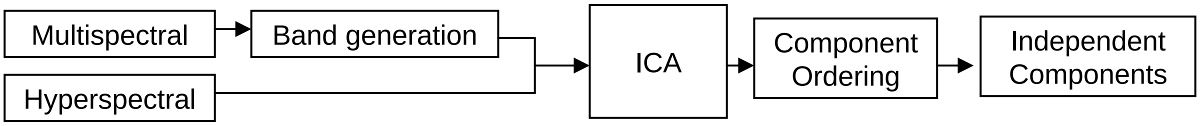


Figure 6. IMAGINE ICA procedural overview.

ICA FEATURE EXTRACTION: REMOTE SENSING APPLICATIONS

Visual Analysis

ICs can used to improve the visual interpretability through component color coding. Color composite image formed using the ICs are provided in Figures 7.

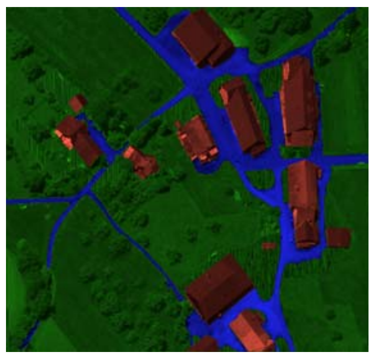


Figure 7. False color composite formed using ICs for improved visual interpretability.

Similar approach can be used for enhanced visual interpretability of hyperspectral imagery employing the ICs.

Resolution Merge

Improved integration of imagery at different spatial resolution can be attained by substituting a high spatial resolution image for an IC followed by an inverse transformation.

Spectral Unmixing

ICA can be employed for linear spectral unmixing when the user has no prior information regarding the spectral response of features present in the scene. $A_i(\lambda)$ and the i^{th} $feature$ (in equation 1) estimated by ICA correspond to the spectral response and abundance respectively of the i^{th} $feature$ in spectral band at wavelength λ ◉ Hence, ICA extracted features can be further analyzed for identifying the proportion of each feature in a pixel.

Shadow Removal

High spatial resolution multispectral images necessitate shadow removal to facilitate improved feature analysis. By employing band combinations (e.g. band ratio) for band generation, the spectral difference between the shadow and the shadow occluded features can be enhanced. ICs obtained from these bands would recover the shadow in one of the components.

Land Use/ Land Cover Classification

ICs can be further analyzed based on their spectral, textural and contextual information in order to obtain improved thematic map.

Multi Temporal Data Analysis

Since feature based change detection techniques necessitate extraction of features with high accuracy, ICs are well suited in the analysis of multi temporal data.

Anomaly/Target Detection

In cases where there is no prior information regarding material of the target features present in the scene, spectra from libraries can not be used for detecting them. ICA, however, when employed for such applications, anomalous features (i.e. features with spectral response significantly different from other features present in the scene) would be contained in the independent components. These components can be further analyzed for improved anomaly/target detection.

ICA FEATURE EXTRACTION: RESULTS AND ANALYSIS

Spectral Unmixing

Application of ICA for spectral unmixing can be easily explained with the help of illustrations in Figure 8. It depicts three original gray level images. For each mixed image in Figure 9, a pixel $p(x,y)$ can be expressed as the weighted sum of pixel $p_i(x,y)$ in original image *i*,

$$p(x,y) = a \cdot p_1(x,y) + b \cdot p_2(x,y) + c \cdot p_3(x,y) \qquad (6)$$

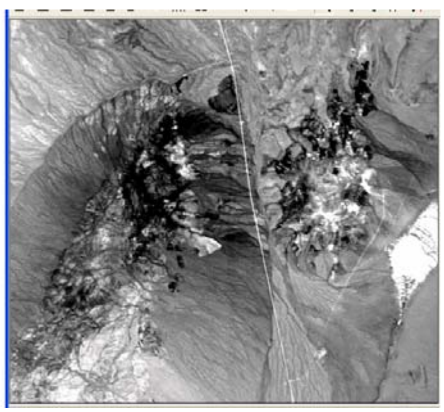
The independent components as estimated by ICA from the mixed images are shown in Figures 10. A visual analysis of these images reveals that given only the mixed images, ICA can successfully recover the original (unknown) images.

Figure 8. Original images.

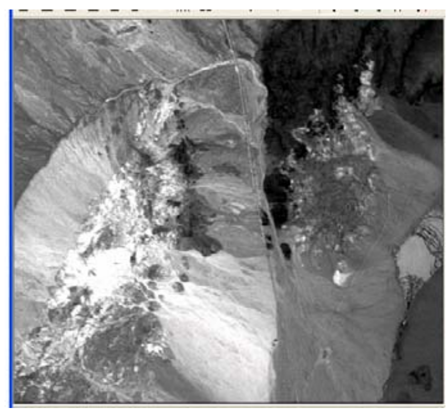
Figure 9. Linear mixtures of the original images.

Figure 10. ICA solution – estimates of the original images.

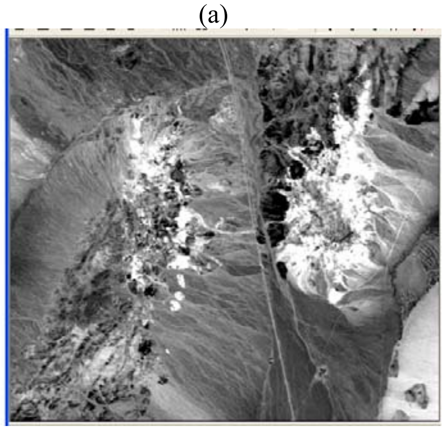
Similarly the images in Figure 11 show the ICs obtained by applying ICA on the AVIRIS cuprite Nevada imagery.



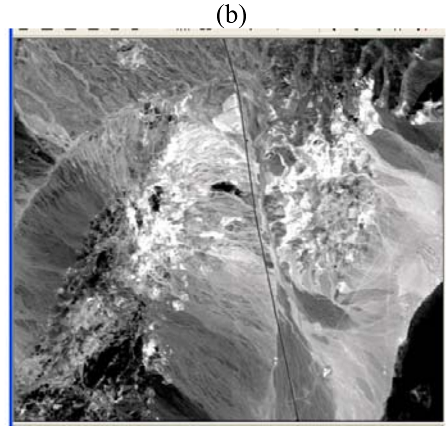
(a)



(b)



(c)



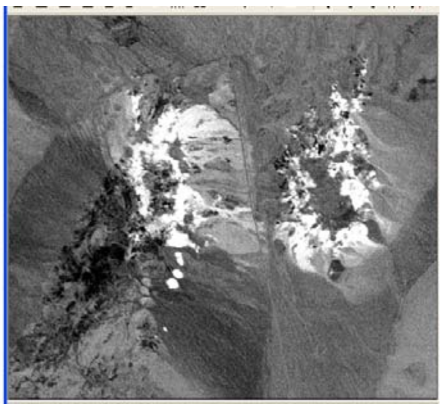
(d)

Figure 11. Components obtained by applying ICA on the AVIRIS cuprite Nevada imagery.

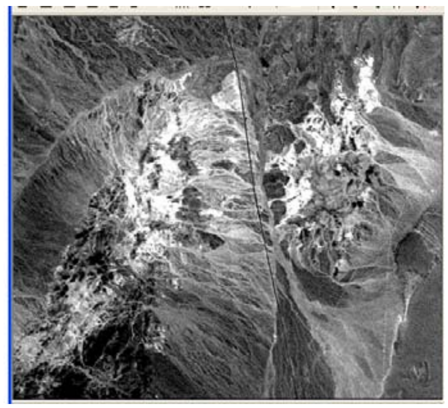
Figure 12 presents the results obtained by performing supervised linear unmixing of original cuprite data (without any preprocessing) using the following endmembers: (a) Stonewall Playa, (b) Kaolinite, (c) Alunite, and (d) Buddingtonite.

(a)

(b)



(c)



(d)

Figure 12. Abundances obtained by performing supervised linear unmixing of original cuprite data using the following endmemebrs: (a) Stonewall Playa, (b) Kaolinite, (c) Alunite, and (d) Buddingtonite.

In order to demonstrate the high correspondence between the ICs and the abundances that were obtained using the original imagery (and not ICs), a correlation coefficient metric was employed. The resulting correlation coefficients were found to be as high as **0.74**, **0.79**, **0.62**, and **0.69** for Stonewall Playa, Kaolinite, Alunite, and Buddingtonite respectively. The results indicate that ICs are highly suitable for spectral unmixing from remote sensing imagery.

Shadow Removal

Figure 14 shows the results of applying ICA on the ADS40 multispectral imagery shown in Figure 13.

Figure 13. True color composite of ADS40.

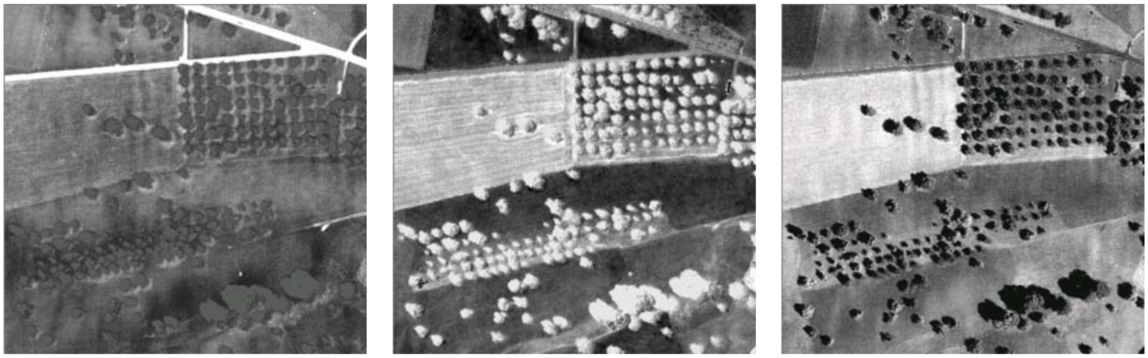


Figure 14. Components obtained using ICA on the ADS40 imagery.

It can be seen that each of the ICs contain more than one feature and this can be attributed to the fact that the number of features in the imaged scene is greater than the number of spectral bands. Hence, a non-linear band combination technique must be employed to generate additional spectral bands. In this image, since the vegetation is occluded by shadow, a band generation technique should enhance the difference in the spectral response of these two features. Hence, band ratios as well as log operators are employed to generate additional spectral bands. Figure 15 shows one of the ICs corresponding to shadow obtained by applying ICA on the imagery with additional spectral bands, whereas a color composite formed using three ICs is shown in Figure 16.

Figure 15. IC corresponding to shadows.

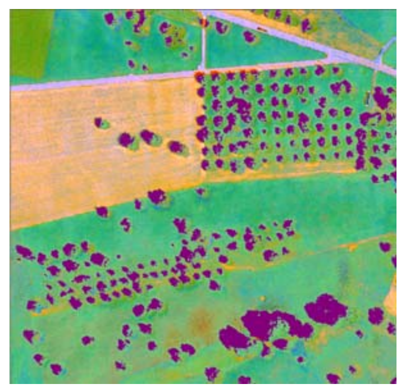


Figure 16. False color composite obtained using ICs.

Land Use/ Land Cover Classification

Figure 17 shows the true and false color composite images formed using the spectral bands of the hyperspectral imagery acquired by Hyperion.

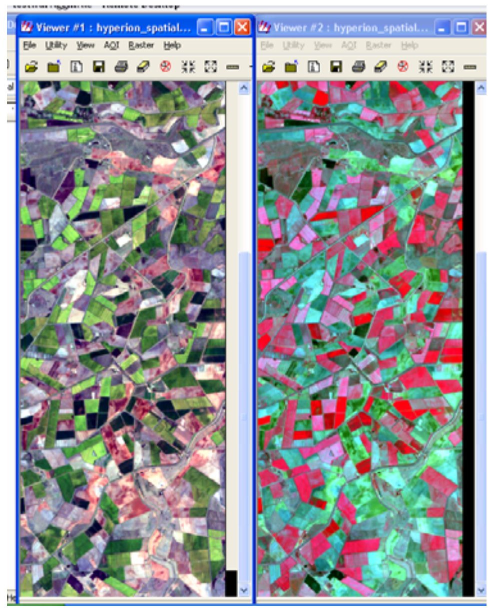


Figure 17. False color composite images formed using the spectral bands of Hyperion.

The imagery is acquired over an agricultural area, and is highly suitable for demonstrating the application of ICA in improving the performance of land cover classification. Figure 18 is the thematic map obtained by classifying the pixels into one of the 6 classes through *K*-means clustering algorithm.

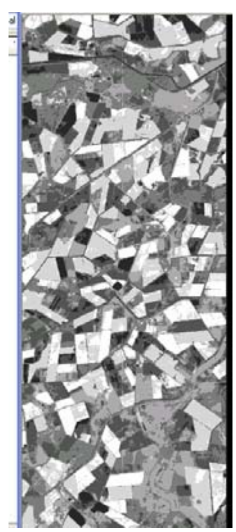


Figure 18. Thematic map obtained by classifying the pixels into one of the 6 classes through *K*-means clustering algorithm.

The map shows significant misclassification in several areas with homogenous land cover types and this can be attributed to the Hughes phenomena (Hughes, 1968) as well as parameter estimation problems due to interband correlation (Shah, 2003). In Figure 19, some of the ICs obtained by performing ICA based feature extraction on this imagery are shown. Features in IC2 (shown in Figure 19) correspond to high vegetation, those with high spectral response in the green and NIR wavelengths (Figure 17).

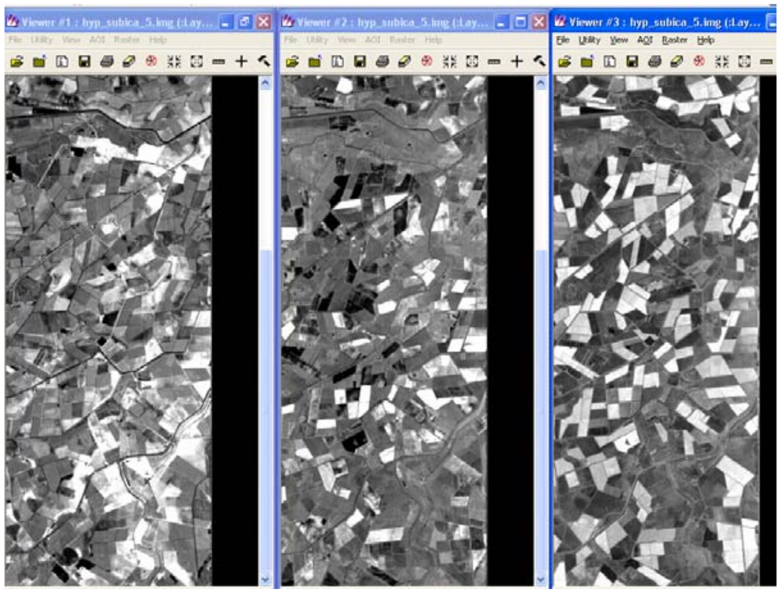


Figure 19. Some of the ICs obtained by performing ICA based feature extraction on Hyperion imagery.

IC3, however, corresponds to features that have high spectral response in the green wavelength only. Finally, the thematic map obtained by *K*-means clustering on the ICs is shown in Figure 20. The thematic map obtained by performing classification on 7 ICs (Figure 20) shows significant improvement over the thematic map obtained from the original data (Figure 18) in the areas of homogenous land cover.

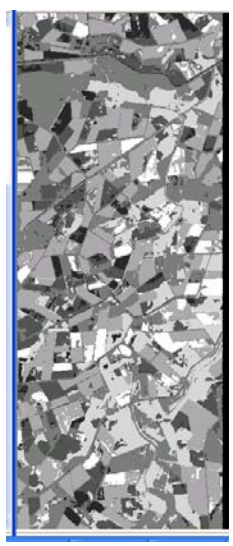


Figure 20. Thematic map obtained by *K*-means clustering on the ICs.

CONCLUSIONS

Analysis of multi/hyperspectral imagery necessitates a selection of an optimal subset of bands in order to avoid the Hughes phenomena and parameter estimation problems due to interband correlation. This can be achieved by employing feature extraction techniques for significant reduction of data dimensionality. Conventional remote sensing feature extraction techniques such as PCA and MNF model the multi/hyperspectral image data with a multivariate Gaussian distribution and are inadequate as most remote sensing data do not follow a Gaussian distribution. In order to overcome this fundamental limitation, we at Leica Geosystems have developed a feature extraction technique based on Independent Component Analysis (ICA) that exploits the higher order statistical characteristics of multi/hyperspectral imagery. Our illustrations demonstrated the potential benefits of employing ICA in improving the performance of several multi/hyperspectral analysis techniques such as spectral unmixing, classification, and anomaly detection. Furthermore, it was shown that the use of meaningful band combinations for non-linear band generation in case of multispectral imagery would improve the performance of ICA in extracting features. For example, employing a Normalized Differential Vegetation Index (NDVI) as an additional band would certainly enhance the performance of ICA in extracting vegetation features.

REFERENCES

Common, P (1994). Independent component analysis, a new concept? *Signal Processing*, *36*: 287-314.

Green, A. A., M. Berman, P. Switzer, and M. D. Craig (1988).A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Trans. Geosci. Remote Sensing*, *26*: 65-74.

Hughes, G. F. (1968). On the mean accuracy of statistical pattern recognition. *IEEE Trans. Inform. Theory*, *14:* 55-63.

Jia, X. and J. A. Richards (1999). *Remote Sensing Digital Image Analysis: An Introduction*. 3rd ed., Berlin, NY: Springer Verlag.

Shah, C. A. (2003). ICA Mixture Model Algorithm For Unsupervised Classification of Multi/Hyperspectral Imagery. Masters Thesis, Syracuse University, Syracuse, NY.