航測及遙測學刊 第九卷 第三期 第81-96頁 民國93年9月
Journal of Photogrammetry and Remote Sensing
Volume 9, No. 3, September 2004, pp. 81-96

Visualization of Hyperspectral Images for Spectral Analysis

Pai-Hui Hsu[1] Yi-Hsing Tseng[2]

ABSTRACT

Some visualization techniques are used to analyzing and exploring the data set of hyperspectral images. The major objectives of data analysis are to summarize and interpret a data set, describing the contents and exposing important features. For dimensionality reduction of hyperspectral images, visualization can play an important role in illustrating the characteristics of high-dimensional data set. Data projection is one of the common visual ways to get the interesting subsets of the original data, and certain properties of the structures can be preserved as faithfully as possible. An effective visualization tool called statistics images displays the second-order statistics of hyperspectral images as a pseudo colored maps. In addition, the multi-scale approaches such as scale space and wavelet analysis are used to visualize the hyperspectral curve in a time-scale plane. These techniques of visualizations will help us to explore the whole data set and extract useful features for further applications such as data representation classification and compression in the future.

Key Words: Hyperspectral Images, Visualization, Spectral Analysis

1. Introduction

The first step to deal with hyperspectral data is visual interaction with hyperspectral images and their statistical properties (Goetz et al., 1985). The major objectives of visualization are to summarize and interpret a data set, describing its contents, and exposing important features for specified applications. For example, Figure 1(a) provides a direct perceptiveness through the variations of the spectral curves. The a laboratory-like spectral curves describe key bsorption features of materials which can help us to understand the physical or chemical properties of atmospheric variations. The data distributions shown in Figure 1(b) also intuitively reveal the discriminating information among different materials. These two simple examples reveal that visualization may play an important role in analyzing and exploring large amount of data and allows us to apply our perceptual abilities to study the data content.

[1] Associate Research Fellow, National Science and Technology Center for Disaster Reduction
[2] Professor, Department of Geomatics, National Cheng Kung University

Received Date: Sep. 23, 2003
Revised Date: Jan. 14, 2004
Accepted Date: Jan. 16, 2004

However, hyperspectral data can be very subtle in the analysis of visualization (Schowengerdt, 1997) not only due to the limitation of our intuition in multi-dimensional world but also the "curse of dimensionality" of high-dimensional data (Bellman, 1961).

In this study, some graphical methods which have been proposed to visualize the high-dimensional data in the last decade are introduced. In these methods, some apparent data values extracted from original data set are generally mapped into one figure for specified applications. One of the simplest ways for visualizing high-dimensional data is data projection in which the interested subsets are selected from the original data, and certain properties of the structure of the data sets can be preserved as faithfully as possible. Another effective visualization tool called statistics images for second-order statistics of hyperspectral images was proposed by Lee and Landgrebe (1993). In this method, the class covariance is displayed as a pseudo colored maps. In addition, the multi-scale approaches such as scale space (Piech and Piech, 1987; 1989) and wavelet analysis (Hsu, 2000a; 2000b) are used to visualize the hyperspectral curve in a time-scale plane. In these two methods, the original hyperspectral data is transformed into another data space and can be represented by symbolic description. These techniques of visualizations will help us to explore the whole data set and extract useful features for further applications such as data representation classification and compression. Finally, an AVIRIS data set with many different classes is used to demonstrate the ways of visualization and symbolic description.

2. Visualization for Data Projection

Landgrebe (1997) illustrated three representation spaces to view multispectral data quantitatively (Figure 2). The image space shown in Figure 2(a) reveals the spatial information and the relationship of neighboring pixels of a hyperspectral image. The spectral space shown in Figure 2(b) demonstrates the spectral signatures of different classes respectively. And, Figure 2(c) shows the feature space in which a pixel associated with n-band measurements is viewed as a point, i.e., a vector in an n-dimensional space. These representations are still useful for hyperspectral data. In this study, some extended methods of these visualization techniques are described as below subsections to characterize the hyperspectral images.

2.1 Spatial-Spectral Space

Showing data in the RGB image space (Figure 2(a)) directly offers a visual way to understand the spatial variation of the scene and the relationship between an individual pixel and the land cover class it belongs to. Tasks of manual image interpretation are usually carried out in the image space. However, the RGB image only shows the spatial information of three bands represented by Red, Green and Blue colors. The large volume of hyperspectral data permits displaying all the bands sequentially in rapid succession, which is called a "spectral movie" (Schowengerdt, 1997). Detecting unique signatures and previewing rapidly the data before processing are two major benefits of the dynamic visualization.

The spectral slice which is also called spatial spectrogram (Schowengerdt, 1997) combines the spatial information and spectral profiles extracted from a hyperspectral image. Figure 3(a) shows a spectral slice over two mainly apparent patterns corresponding to two different classes. The vertical direction corresponds to the spatial dimension of the image being slices, the horizontal direction corresponds to the spectral dimension, and the pseudo color shows the spectral density (the values of reflectance or radiance). In Figure 3(a), the red color appeared in the near-infrared region of the spectral slice correspond to the pixels belonging to the green vegetations. The dark color of the spectral slice is caused by the water absorptions bands. A hyperspectral image can be also viewed as a 3D image cube, whose face shows the spatial dimensions and the depth presents the spectral band (or wavelength). A 3D image cube can be viewed perspectively as showing the face with the spectral slices of the top row and right column. An example of 3D image cube of the AVIRIS data set is shown in Figure 3(b).

2.2 Spectral Space

The spectral response of a material forms a unique spectral signature of the reflectance or radiance. It can be represented by a function of wavelength. Spectral response of a material, therefore, can be drawn as a spectral curve in the spectral space (Figure 1(a)). Theoretically, each class composed of different material has its own shape and variances of the spectral curve. Some methods like "Spectral Matching" and "Spectral Angle Mapper (SAM)" use this important property to distinguish an unknown spectral curve comparing with a series of pre-labeled spectral curves (Kruse *et al.*, 1993). Figure 4 shows the spectral curves of five different land cover types. Some basic statistics such as mean and standard derivation for each band can be calculated to depict the characteristics of the spectral distribution. The mean curve represents the trend of the spectral variation. The standard deviation curves show the scattering with respect to the mean. The maximum and the minimum values present the range of spectral variation at a particular wavelength. Different spectral signatures can be portrayed on one spectral space for the purpose of comparison. Through interactively plotting the spectral spectrum in the spectral space, one may rapidly locate the pixels having similar spectral signatures (Goetz *et al.*, 1985). This is particularly helpful to create or verify the training areas in the scene for supervised classifications.

Different spectral curves can be plotted together in one spectral space for comparison. In Figure 5, five different spectral signatures are superimposed together. It is easy to compare the magnitude of the reflectance values and the variations of the spectral curve of different materials in the superimposed spectral space. The spectra can also be stacked with an offset vertically to allow interpretation (Figure 6).

2.3 Scatter Plots

The Scatter plot is one of the oldest and most commonly used methods to project high-dimensional data to a 2D space. A multispectral image is viewed with $n(n-1)/2$ scatter plots of pair-wise parallel projection, where n is the dimensionality of the multispectral image. Each scatter plot provides a general impression of the spectral correlation between two selected bands from a data set. Figure 7 shows a 2D scatter plot of five different land-cover classes on band 20 and 50 of an AVIRIS data set. One may instinctively obtain the correlation between these two selected bands. The pixels of one particular material class will distribute closely in a scatter plot. Thus the characteristics of classes can be interpreted based on statistical pattern recognition (Swain and Davis, 1978). For example, a two-dimensional Gaussian density can be used to characterize the distribution for each class, and classification can be performed based on the decision boundaries in a 2D scatter plot. In Figure 7, the distribution of the "highway" class is far away from the other distributions of vegetation classes. Therefore, it is easy to determine the decision boundary between "highway" and other vegetations. Furthermore, a 2D scatter plot can be easily extended to a 3D scatter plot of three selected bands. Figure 8 shows a 3D scatter plot on band 10, 20, and 50. The 3D scatter plot can be rotated interactively to show an appropriate view in a viewing program.

In order to inspect the correlations between bands of three or more, the scatterplot matrix can be used to represent all of the 2D scatter plots simultaneously. Figure 9 shows a scatterplot matrix containing all the pairwise scatter plots of band 10, 20, 50 and 128 on a single page. The diagonal plots of the scatterplot matrix are simply 45-degree lines since data is plotting band i verse band i. This reveals a point of view in terms of showing the univariate distribution of the variable. Figure 9 signals a linear relationship between band 10 and 20 indicating that there exists redundancy between these two bands. This provides a basic idea of dimensionality reduction for hyperspectral images. In addition, separability analysis can be applied in each scatter plot to select the salient bands which are helpful for the classification of hyperspectral image. For instance, the largest average separability of the five classes is apparent in the scatter plot of band 50 and 128.

When the number of bands is larger than 3, spectra can also be thought of as points in an n-dimensional scatterplot, where n is the number of bands (Boardman, 1993). The coordinates of the points in n-space consist of "n" values that are simply the spectral radiance or reflectance values of a given pixel. The distribution of these points in n-space can be used to estimate the number of spectral endmembers and their pure spectral signatures (Boardman *et al.*, 1995). Figure 10 shows an 8-dimensional Visualization of the AVIRIS data which is created by the n-Dimensional Visualizer™ tool provided by Environment for Visualizing Images (ENVI) software. The distributions of these points in n-space can be used to estimate the number of spectral endmembers and their pure spectral signatures.

2.4 Parallel Coordinates

The idea of using parallel coordinates as a visualization method for multi-dimensional data was firstly proposed by Inselberg (1981). In traditional Cartesian coordinates such as the scatter plot shown in Figure 7 and Figure 8, all axes are mutually perpendicular. In Parallel coordinates, as implied by the name, all axes are parallel to one another and equally spaced. Thus a specific point in n-dimensional space can be represented by n Y-axis values which are connected as a polyline along the X-axis. Figure 11 shows an example of the parallel coordinates using five different materials from AVIRIS data set. In Figure 11(a), the original spectral values are put in each axis. In Feature 11(b), the spectral values are normalized to the range of 0 to 1 for each axis. Some variations will be exaggerated in the normalized axes. Data structures, such as points, lines or hyperplanes in high dimensional space $\mathbf{R}^n$ can be projected onto this two-dimensional graph for visualizing analytic and synthetic geometry in $\mathbf{R}^n$ (Inselberg, 1990). Viewed as a whole, the polylines in parallel coordinates might well exhibit some apparent patterns which could possibly be associated with inherent correlation of the data sets. For example, the lines of a particular materials cluster together in parallel coordinates, indicating a degree of correlation. The near-parallel lines between the adjacent axes indicate a high correlation, and the ×-shaped structure between adjacent axes indicates an inversely correlated.

Compared with scatter plots, the number of dimensions that can be visualized by parallel coordinates is rather large, however it becomes more difficult to perceive structures or clusters as the axes get closer to each other. The main difficulty of the parallel coordinates is that large data set can cause difficulty in cluster interpretation (Ward, 1994; Fau *et al.*, 1999). Also, relationships between adjacent axes are easier to perceive than between non-adjacent axes.

Because the large band number of hyperspectral spectral data, it is not possible to plot the whole data in a parallel coordinate. In fact, the parallel coordinate is a simplified spectral space with selected bands. However, one may use this technique to inspect the results of feature selection or extraction visually. For example, in Figure 11(b), the Chaparral and the Broadleaf Evergreen Forest appeared a symmetrical shape can be easily separated with the selected features.

3. Visualization for Data Transformation

It would be helpful to analyze and quantify the characteristics of hyperspectral data when the data set is transformed into some mathematical or conceptual representation spaces in which one may inspect the data set from different viewpoints. The type of data transformation may be linear or non-linear that a set of new data set would be created to reveal spectral information as much as possible. In this section, we firstly introduce the statistics images which represent the second order statistics of hyperspectral images using pseudo colors. As a glance, one may subjectively perceive how each band is correlated and easily compare the variances between difference classes. In addition, the results of data transformation may also be represented by way of symbolic description. Spectral fingerprint is a familiar example of symbolic description for the detection of absorption features (Piech and Piech, 1987). Finally, a new method which produces the similar results to fingerprints using wavelet transform was introduced (Hsu, 2000a; 2000b).

3.1 Statistics Images

The covariance or correlation matrix of a set of hyperspectral data can be displayed as a pseudo colored image, so that one can subjectively perceive how much the bands are correlated and easily compare the statistics of different classes (Lee and Landgrebe, 1993). Figure 12 shows the pseudo images of covariance and correlation matrices corresponding to two different classes respectively: the grassland and highway surfaces. In Figure 12(a) and 12(b), the mean, the positive and negative standard deviations, and the maximum and minimum of the spectra are displayed together in the spectral space. The color of the covariance matrices shown in Figure 12(c) and 12(d) indicates the degree of covariance between different bands. It can be seen that the variances of the spectra of grassland have the largest values in the portion from red light to near-infrared (dark red color from band 33 to band 74). Oppositely, the dark blue colors appear in the portion from band 98 to 222 indicate small variances between the visible-near-infrared (VNIR) and short-wave infrared (SWIR) bands. On the other hand, the variances of the spectra of road surface are nearly homogenous but the patterns formed by different spectral regions are still apparent. The correlation matrices shown in Figure 12(e) and 12(f) indicate that adjacent bands are highly correlated in a hyperspectral image, i.e., the information content is highly redundant. It is, therefore, reasonable to perform a feature extraction before the analysis of hyperspectral images. In addition, the significant difference of the correlation matrix between different classes provides the classifier useful information to distinguish a class from others.

3.2 Spectral Fingerprints

A plot of fingerprints is a way of symbolic description developed by Piech and Piech (1987; 1989) of the absorption bands in hyperspectral data. The fingerprints are created based on a scale space filtering of the hyperspectral data. A scale space image is a set of progressively smoothed version of the original spectral curve which is usually calculated by the Gaussian function. Fingerprints are then simply obtained from the zero-crossings of the second derivative of the scale space image. As the smoothing scale increase, the number of zero-crossings is reduced until only a dominant spectral shape remains. The net result of this scale space analysis is thus a sequence of closed arches. For the purpose of quantification of the fingerprint, each closed arch can be described by a triplet $(\sigma, \lambda_L, \lambda_R)$, where σ is the scale at which the arch closes, λ_L is the location of the left arc, and λ_R is the location of the right arc. Each triplet describes a spectral absorption feature and contains a measure of importance of the spectral feature within the triplet.

The triplet descriptions demonstrate that the symbolic representation can make both gross and fine distinctions between spectral features. Figure 13 shows the results of three fingerprints corresponding to three different kinds of vegetations. Clearly, the fingerprint appears as a pattern of closed arches. Each arch is composed of a pair of zero-crossing contours which corresponds to a spectral absorption feature. The representation is compact, selects and quantifies features related to absorption bands, ranks features according to their strength, and is capable of reliably extracting small features (Piech and Piech, 1987).

3.3 Wavelet-Based Fingerprints

A new method which produces similar fingerprints using wavelet transform was proposed by Hsu and Tseng (2000a, 2000b). The wavelet transform has the ability to focus on the localized signal structures using a zooming procedure. When the scale s of wavelet function varies from its maximum to zero, the decay of the wavelet coefficients $Wx(u,s)$ characterizes the regularity of x in the neighborhood of u. This characteristic of multi-scale analysis is very similar to the fingerprint method which is based on the scale space filtering. This provides us the essential idea for the hyperspectral data visualization and the detection of the absorption features from the reflectance spectrum automatically.

In the wavelet-based method, the scale-space image is computed directly using the continuous wavelet transform (CWT) instead of the derivative computation following the convolution by Gaussian filters. By detecting the positions of modulus maxima of wavelet coefficients calculated using the first or second derivative of Gaussian function, the fingerprints corresponding to the absorption features can be easily delineated. Finally, a smaller number of absorption features can be both automatically processed and physically interpreted. Besides, the method is also helpful for spectral analysis to reduce the dimensionality of hyperspectral data. Figure 14(a) and 14(b) respectively show the wavelet coefficients represented by a pseudo colored map using the 1^{st} and 2^{nd} derivative of Gaussian function. Figure 14(c) and 14(d) show the maximum lines of wavelet coefficients. By detecting the positions of the local maxima from the coarse to fine scale, we can obtain the positions of the absorption feature of a hyperspectral curve. The major advantage of this method is that the localized signal structures of absorption features can be accurately characterized by the zooming procedure of CWT. Furthermore, a fast algorithm can be used for the computation of the wavelet transform in order to avoid the computation expense of direct convolution (Mallat, 1999).

3. Conclusions

The high spectral resolution characteristic of hyperspectral sensors preserves important aspects of the spectrum and makes differentiation of different materials on the ground possible. However, due to the high dimensionality and high correlation between spectral bands, traditional analysis approaches do not applicable to process the hyperspectral images. It is necessary to analyze the whole data set before the processing of hyperspectral data. The main objective of data analysis is to summarize and interpret a data set, describing the contents and exposing important features. Visualization plays an important role in data exploring and analysis. In this paper, some projection methods are used to explore and analyze the hyperspectral images. The features can be selected intuitively from the scatterplots. The endmembers can also be obtained from the *n*-dimensional visualization. In addition, the multi-scale approaches, such as the finger-prints and wavelet transform are used to visualize the hyperspectral curve in a time-scale plane. The useful features such as the absorption features then can be explored and extracted for further applications.

Acknowledgment

The authors would like to thank the National Science Council of Republic of China, for support of this research project: NSC 91-2211-E-006-091.

References

Bellman, R., 1961. *Adaptive Control Processes: A Guided Tour*, Princeton University Press.

Boardman, J. W., 1993. Automating Spectral Unmixing of AVIRIS Data Using Convex Geometry Concepts, *Proceedings of Summeries of the Fourth Annual JPL Ariborne Geoscience Workshop*, Pasadena, CA, 1:11-14.

Boardman, J. W., F. A. Kruse, and R. O. Green, 1995. Mapping Target Signatures via Partial Unmixing of AVIRIS Data, *Proceedings of Summaries of Fifth JPL*

Airborne Earth Science Workshop, Pasadena, CA, 1:23-26.

Fau, Y.-H., M. O. Ward, and E. A. Rundensteiner, 1999. Hierarchical Parallel Coordinates for Exploration of Large Data Sets, *Proceedings of IEEE Conerence on Visualization '99*, San Francisco, CA, pp. 43-50.

Goetz, A. F. H., G. Vane, J. E. Solomon, and B. N. Rock, 1985. Imaging Spectrometry for Earth Remote Sensing, *Science*, 228:1147-1153.

Hsu, P.-H. and Y.-H. Tseng, 2000a. Wavelet-based Analysis of Hyperspectral Data for Detecting Spectral Features, *Proceedings of XIXth Congress of ISPRS,* vol. XXXIII, Amsterdam, Netherlands, pp. 61-68.

Hsu, P.-H. and Y.-H. Tseng, 2000b. Multiscale Analysis of Hyperspectral Data Using Wavelets for Spectral Feature Extraction, *Proceedings of 21th ACRS*, Taipei.

Inselberg, A., 1981. N-Dimensional Graphics, Part I – Lines and Hyperplanes, *Technical Report G320-2711*, IBM Los Angeles Scientific Center, Los Angeles, CA.

Inselberg, A., 1990. A Tool for Visualizing Multi-dimensional Geometry, *Proceedings of IEEE Visualization 1990*, pp. 361-378.

Kruse, F. A., A. B. Lefkoff, and J. B. Dietz, 1993. Expert System-based Mineral Mapping in Northern Death Valley, California/Nevada Using the Visible/Infrared Imaging Spectrometer (AVIRIS), *Remote Sensing of Environment*, 44: 309-336.

Landgrebe, D. A., 1997. On Information Extraction Principles for Hyperspectral Data (A White Paper), School of Electrical & Computer Engineering, Purdue University, West Lafayette.

Lee, C. and D. A. Landgrebe, 1993. Analyzing High-Dimensional Multisepctral Data, *IEEE Transactions on Geoscience and Remote Sensing*, 31:792-800.

Mallat, S., 1999. *A Wavelet Tour of Signal Processing*, Second Ed., Academic Press, San Diego.

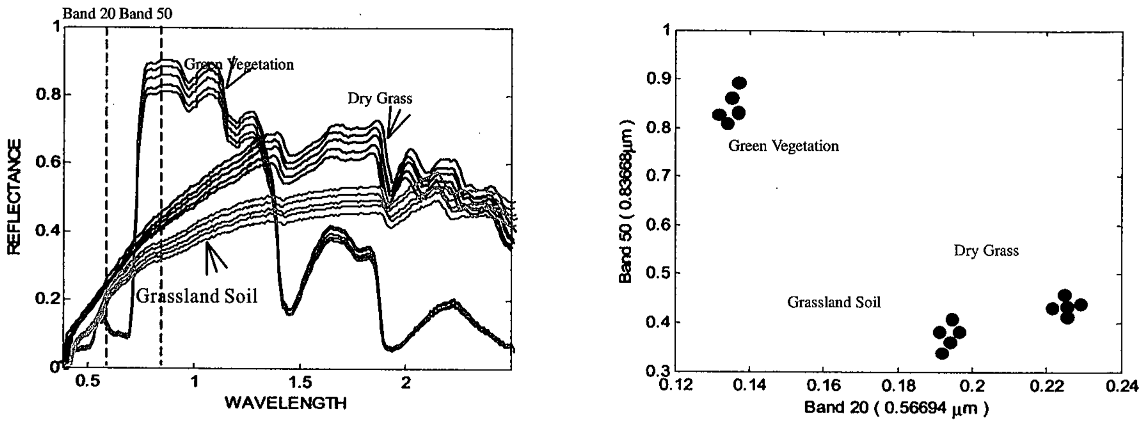
Piech, M. A. and K. R. Piech, 1987. Symbolic Representation of Hyperspectral Data, *Applied Optic*, 26: 4018-4026.

Piech, M. A. and K. R. Piech, 1989. Hyperspectral Interactions: Invariance and Scaling, *Applied Optic*, 28: 481-489.

Schowengerdt, R. A., 1997. Remote Sensing, Models and Methods for Image Processing, Second Ed., Academic Press, San Diego.

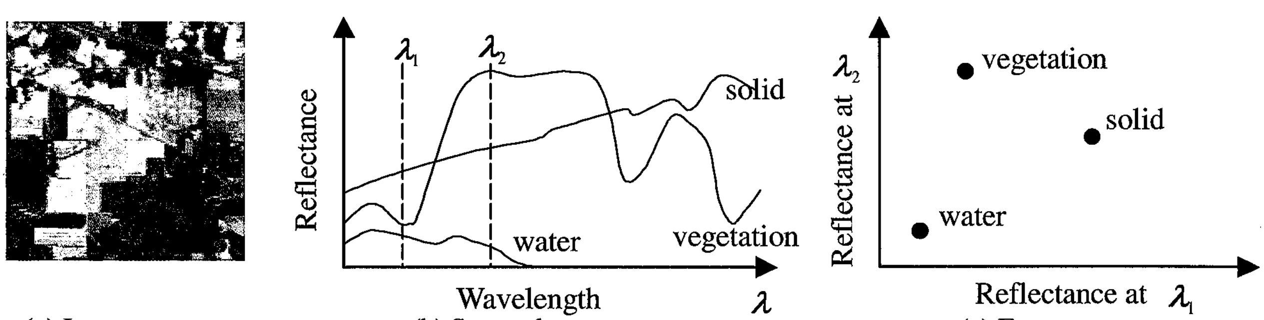
Swain, P. H. and S. M. Davis, 1978. *Remote Sensing: The Quantitative Approach.* McGraw-Hill, New York.

Ward, M. O., 1994. XmdvTool: Integrating Multiple Methods for Visualizing Multivariate Data, *Proceedings of IEEE Conference on Visualization '94*, Washington, DC, pp. 326-333.



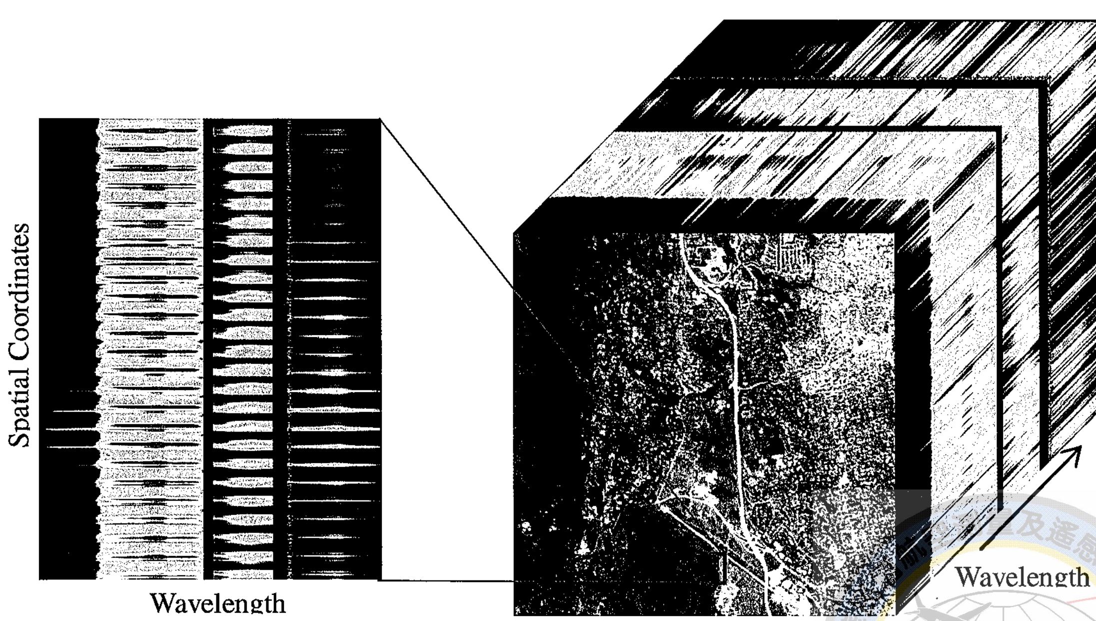
(a) Reflectance spectra (b) 2D scatter plot

FIGURE 1 (a) Reflectance spectra of green vegetation, dry grass, and grassland soil. (b) Three distributions corresponding to the three materials plotted in a scatter-plot



(a) Image space (b) Spectral space (c) Feature space

FIGURE 2. Three data space for representing multispectral data (Landgrebe, 1997)



(a) Spectral slice (b) 3D image cube

FIGURE 3. A hyperspectral image shown in the spatial-spectral space

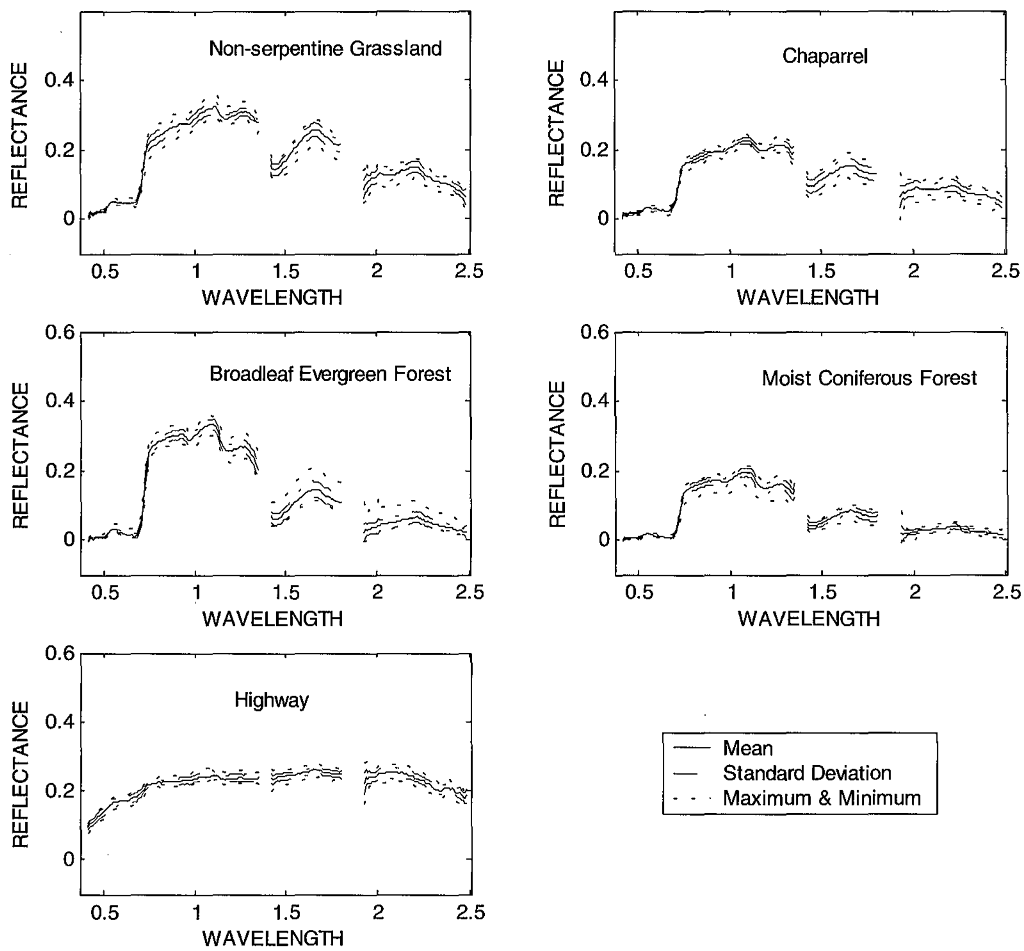


FIGURE 4. Five spectral signatures of an AVIRIS data set.

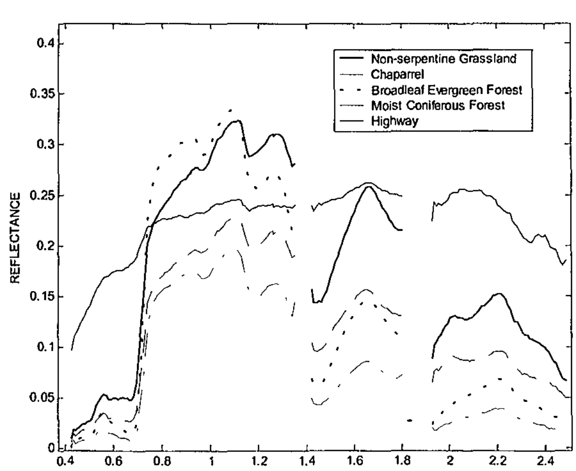


FIGURE 5. Superimposed spectral data in the spectral space

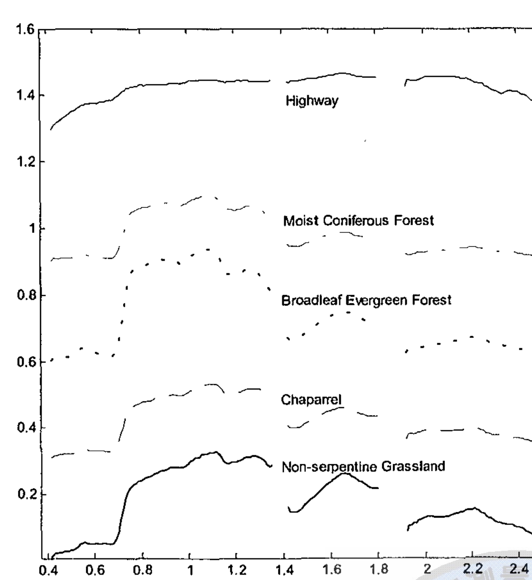


FIGURE 6. Stacking spectral data in the spectral space

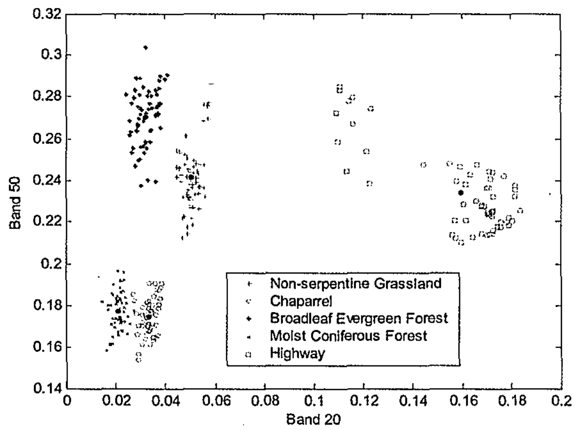


FIGURE 7. Five spectral distributions in a 2D scatter plot.

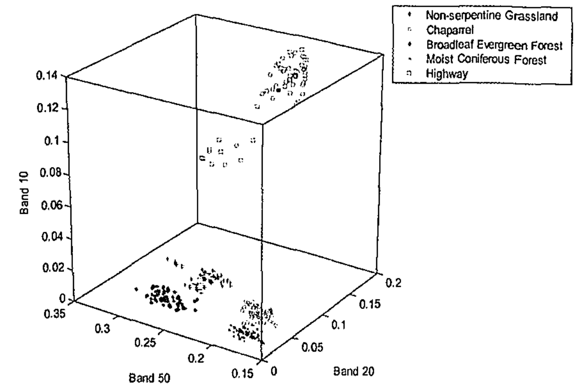


FIGURE 8 Five spectral distributions in a 3D scatter plot.

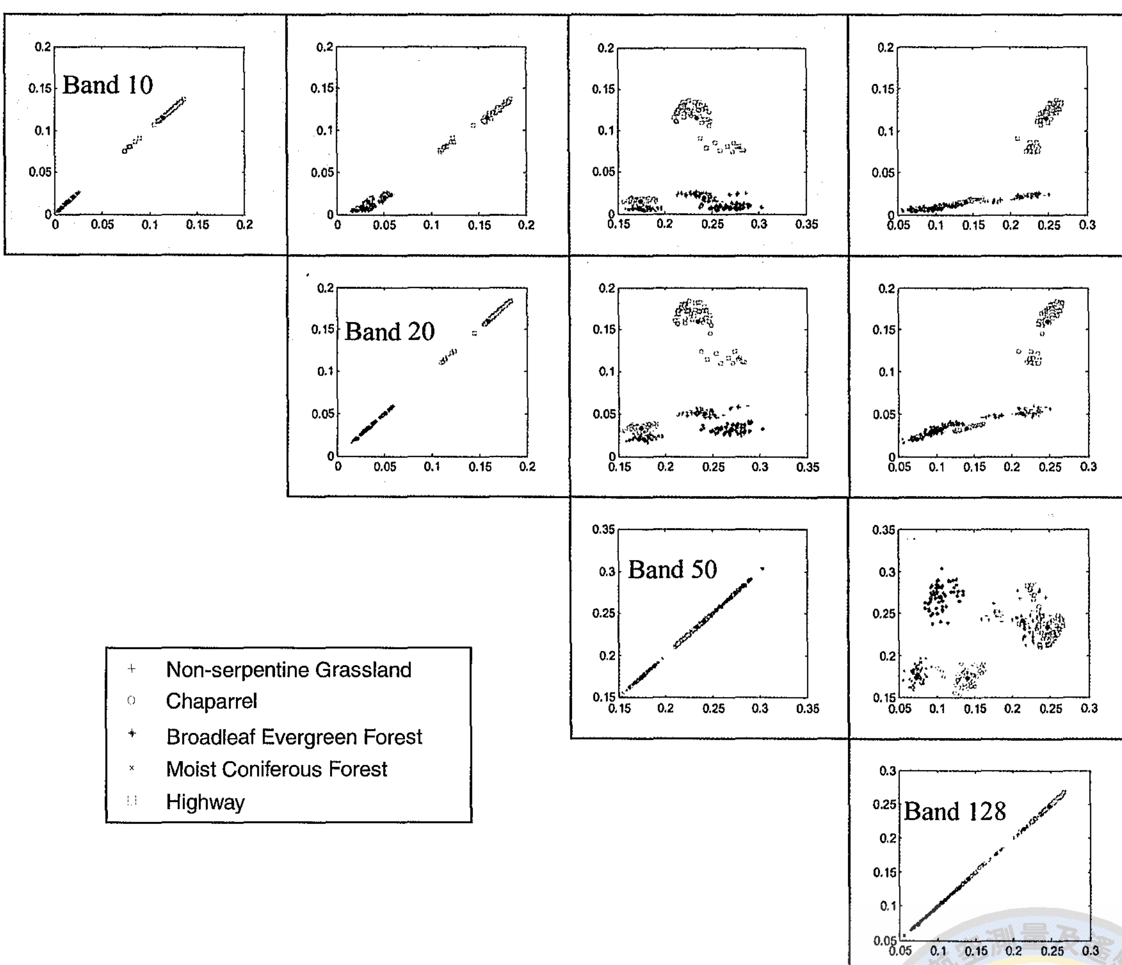


FIGURE 9. Four dimensional scatterplot matrix

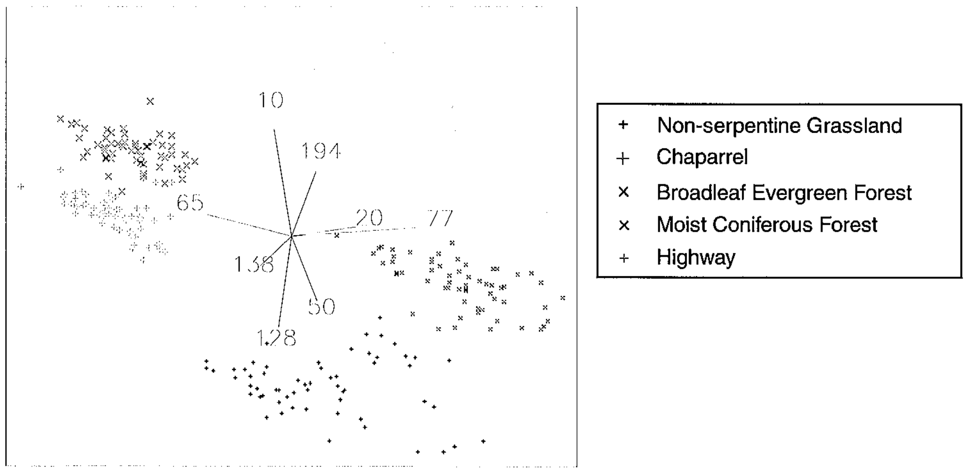
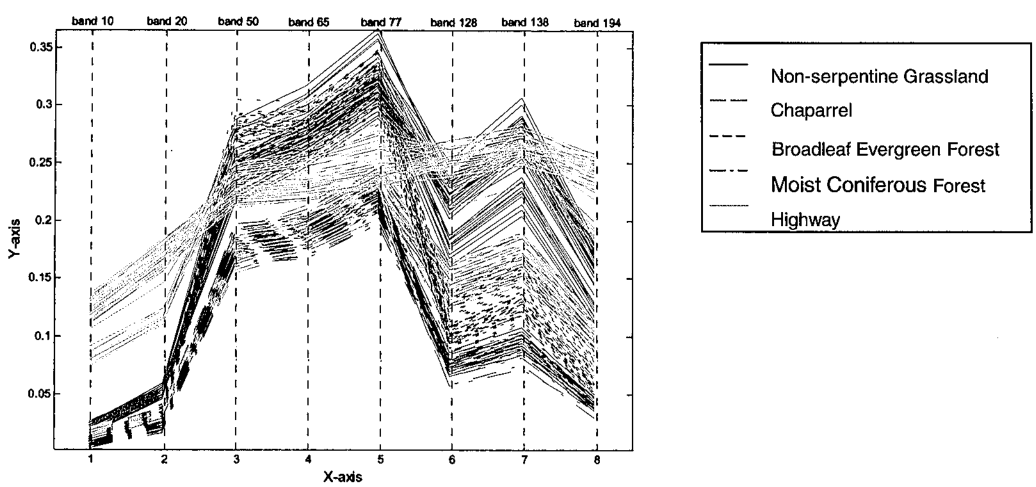
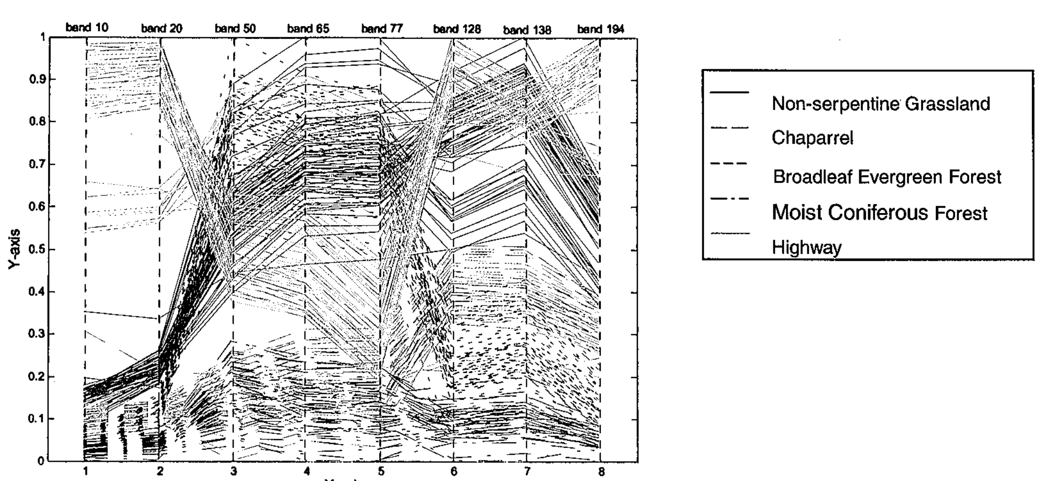


FIGURE 10. ENVI's *N*-Dimensional Visualizer™ with eight bands

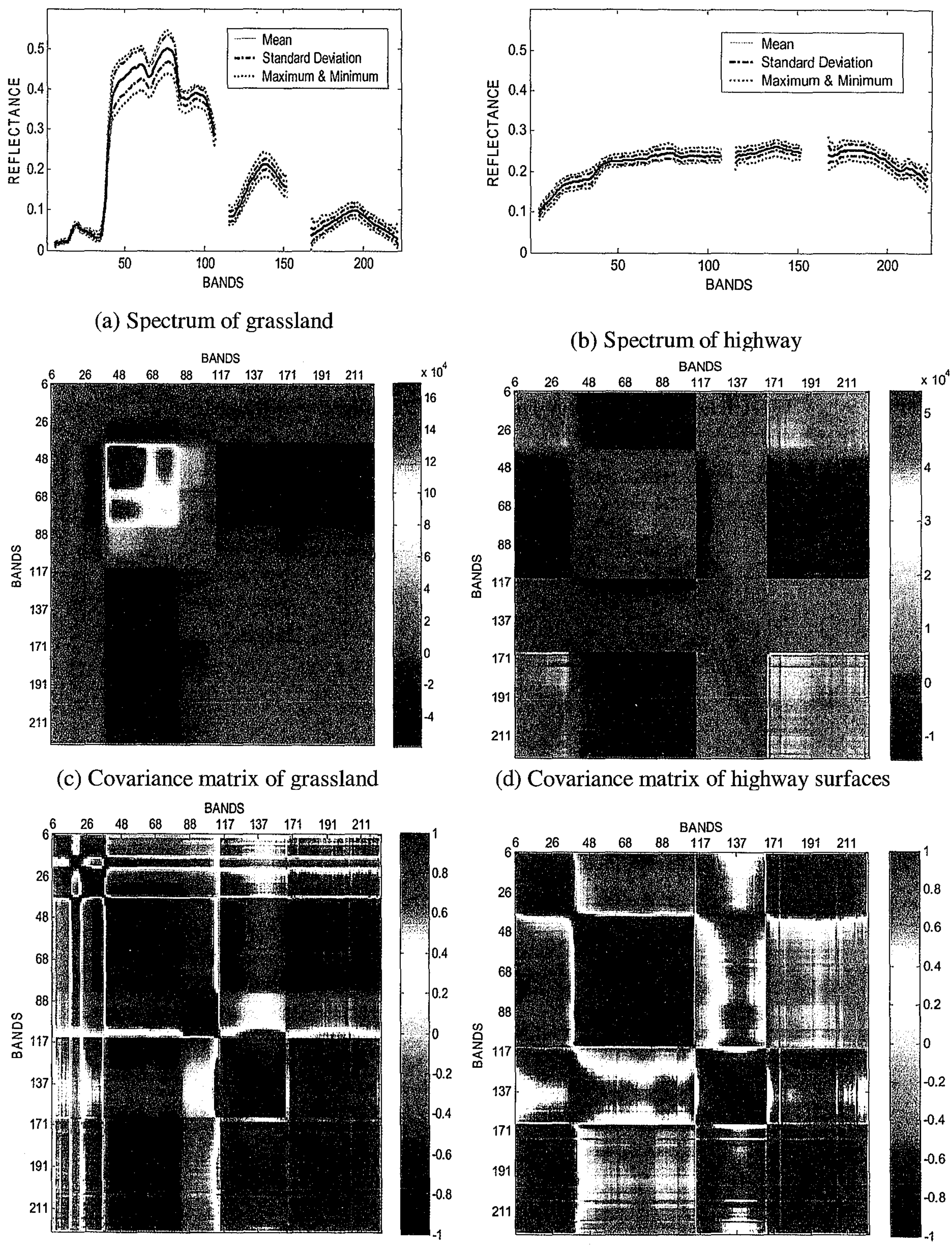


(a) Parallel coordinates without normalized axes



(b) Parallel Coordinates with normalized axes

FIGURE 11. Five spectral distributions in the parallel coordinates.



(a) Spectrum of grassland

(b) Spectrum of highway

(c) Covariance matrix of grassland

(d) Covariance matrix of highway surfaces

(e) Correlation matrix of grassland

(f) Correlation matrix of highway surfaces

FIGURE 12. Pseudo colored images of covariance and correlation matrices for grassland and highway surfaces.

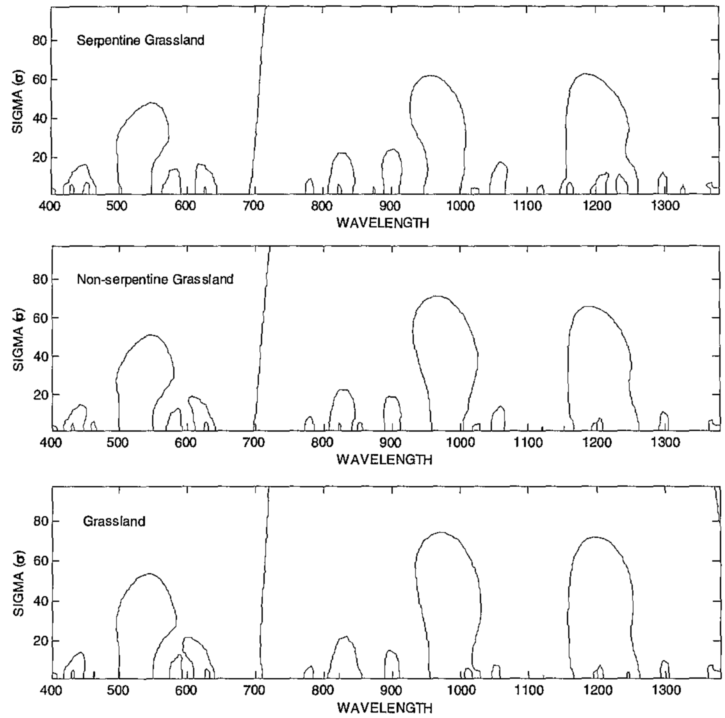
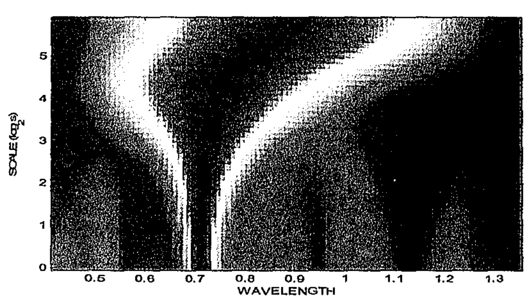
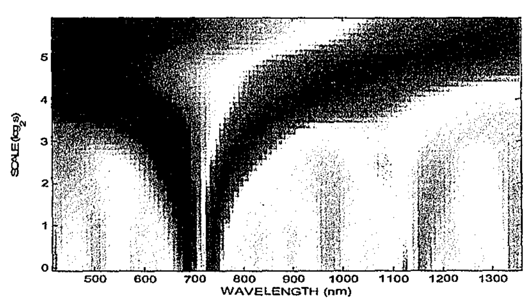


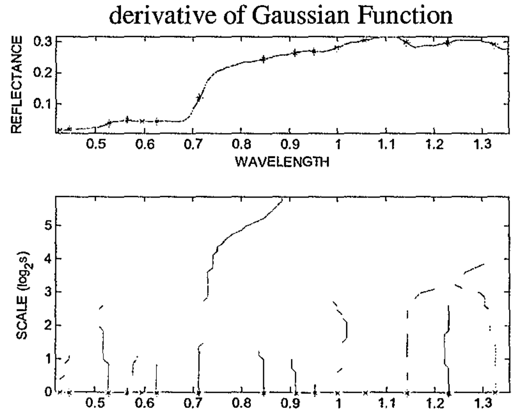
FIGURE 13. The fingerprints of three different kinds of vegetations



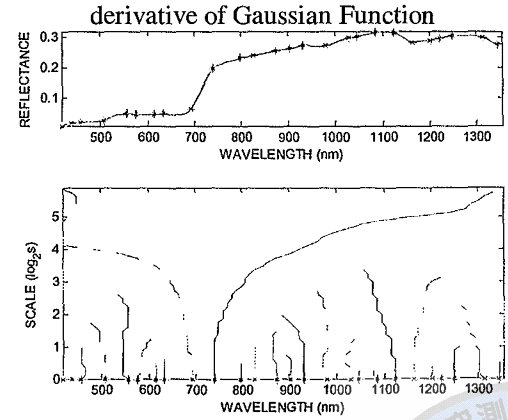
(a) Wavelet coefficients using the 1st derivative of Gaussian Function



(b) Wavelet coefficients using the 2nd derivative of Gaussian Function



(c) The modulus maximum line of wavelet coefficients



(d) The modulus maximum line of wavelet coefficients

FIGURE 14. The results of wavelet transform of Grassland and its corresponding maximum lines using the 1st and 2nd derivative of Gaussian function

高光譜影像視覺化及其分析

徐百輝[1] 曾義星[2]

摘要

資料分析的主要目的在於歸納、詮釋及描述資料之內涵，並凸顯出某些重要的特徵，例如欲降低高維度資料之維度時，必須先分析不同類別資料的分佈特性，此時資料的視覺化將扮演一個重要的角色。本文利用各種不同的視覺化技術來檢視與分析高光譜影像，其中資料投影為最常用的視覺化方法之一，其基本原理係將高維度資料投影至人眼可辨識的二維空間中，其結果將有助於我們保存重要的資料結構，並萃取出感興趣之資料特徵。另外，高光譜影像之二階統計量亦可以利用虛擬色彩表示成所謂的統計量影像。此外，本文亦採用多尺度分析方法，如尺度空間法及小波分析法，於時頻空間中顯示高光譜影像的吸收特徵。這些視覺化方法都將有助於我們探索整個資料集，未來並可針對不同的應用，如資料表示、影像分類及影像壓縮等萃取出相對應的有用特徵。

關鍵詞：高光譜影像、視覺化、光譜分析

[1] 國家災害防救科技中心副研究員

[2] 國立成功大學測量與空間資訊學系教授

收到日期：民國92年09月23日
修改日期：民國93年01月14日
接受日期：民國93年01月16日