

行政院國家科學委專題研究計畫成果報告

高光譜影像頻帶選取與特徵擷取

計畫編號：NSC89-2211-E-009-084

執行期間：89 年 8 月至 90 年 7 月

主持人：史天元 教授

國立交通大學土木工程學系

一、中文摘要

本研究目的是探討高光譜影像波段選取的方法，文中採用 Price (1990)與 Tu, et al. (1998)波段選取的演算法。實驗數據分別為 2001 年用 GER3700 地面光輻射儀量測草地覆蓋，以及 1992 年 AVIRIS 空載高光譜影像，AVIRIS 資料的地類以農作為主。成果分析顯示，Price 與 Tu, et al.兩種演算法所選出的波段具有相容的結果，草地覆蓋的地面光譜應用 Price 方法重要波段分布在 543~1241nm，Tu, et al.方法重要波段分布在 755~1124nm。AVIRIS 資料 Price 方法所選出的波段，具較重要物理意義者包括 470nm 位置是分辨土壤與植被差異的重要位置；550nm 是綠色植物分辨的重要位置；1280nm 是分辨葉面覆蓋的重要位置。

關鍵詞：波段選取、高光譜影像、特徵萃取

ABSTRACT

The accuracy of classification can be improved by increasing the separability between classes, reducing the dimensionality of data, and choosing an effective classification scheme. Low pass filter, Principal Component Analysis and Minimum Noise Fraction Transformation are studied together with different classification schemes such as the maximum likelihood, Spectral Angle Mapping (SAM) and others. It is found that MNF is more effective than PCA and maximum likelihood is better

than SAM.

KEYWORDS: Band selection, Hyperspectral images, Feature extraction.

二、研究方法與成果

2.1 前言

當代遙測感測器朝向高空間解析度及高光譜解析度兩方向發展。高光譜解析度影像提供更豐富的光譜資訊，舉例而言，JPL 發展的空載高光譜資料 AVIRIS (Airborne visible/infrared imaging spectrometer)，具有 224 個光譜波段，光譜範圍分布在 0.4-2.5 μ m，光譜解析度約 10nm；空間解析度是隨航高而變化，一般約 20 米。AVIRIS 有完整的率定文件 (Green, 1998)。

高光譜解析度提供較細節的光譜特徵，有更好的分類辨識的潛力，但相對的代價是影像資料的訊號雜訊比降低，光譜的波段寬度 (Band width) 越窄其訊號品質越低，二者在感測器的設計上是需取得最佳的選擇。

遙測分類方法，分類精度常與訓練樣區的樣本數目有關，較多的樣本數目有較可靠的分類成果，當光譜資料的光譜解析度提高，光譜的維度增加了，若訓練的樣本數目不變則樣本數目與光譜維度數的比率降低，如此情形，分類方法可靠度亦降低。訓練樣區的取得費用昂貴，比如以地真調查

手段獲取訓練樣本，訓練樣區的總數目會受限於經濟考量以及交通與地形等限制條件，僅有一定的有限樣本數，面臨高光譜資料的分類，在有限的樣本數目，隨著光譜維度的增加，參數估計的可靠性反而面臨降低。高光譜影像分類的精度與以下因素有關 (Hsieh and Landgrebe, 1998)：類別的分離度、訓練樣本的大小、光譜維度、分類方法。

依據高光譜資料，相鄰光譜間高度相關特性，存有大量的多餘性的資訊，由大量多餘資訊的光譜資料選取出特徵資訊，以降低高光譜資料的光譜維度，成為高光譜研究的一個課題。降低維度的方法中分為特徵選取 (Feature selection) 與特徵萃取 (Feature extraction)，特徵選取是從原始的光譜反應值選出重要者成為特徵觀測量 (Price, 1990; Li et al., 1999)。特徵萃取是對原始的光譜反應值進行組合或函數轉換得到另一組特徵資料 (Green et al., 1988; Chang et al., 1999; Roger et al., 1996)。

波段選取的策略可採分離度為評估指標，從不同波段的選取對訓練樣本進行分離度分析，去找最佳的波段，這種策略面臨組合性龐大、計算量高的困難，擇優搜尋的演算法成為研究課題，諸如使用連續向前搜尋法 (Sequential Forward Selection, SFS)、連續向前搜尋法 (Sequential Backward Selection, SBS) 以及擇優評估指標函數等研究，文獻包括 Landgrebe (1997); Young、Fu (1986); Serpico、Bruzzone (2001)。

Price (1990) 發展一個漸進程序分析光譜值，文中顯示 Price 演算法可用少數的選取波段配合基底函數可重建光譜訊號。Tu et al., (1998) 用 Canonical Analysis 原理，由特徵向量與特徵值決定波段選取的指標。本研究目的是探

討高光譜影像，波段選取的方法，並比較不同特徵選取方法的適用性。

2.2 波段選取方法

2.2.1. Price 演算法波段選取

Price (1990) 提出一個漸近計算程序進行高光譜波段選取，原理如下：

$x^\alpha(\lambda) = (x_1^\alpha, x_2^\alpha, \dots, x_n^\alpha)$ 代表光譜的量測值，是波長的函數所構成光譜曲線， α 為地類， n 為高光譜的波長變數，AVIRIS 資料為例， n 為 224。光譜曲線 $x^\alpha(\lambda)$ 可採 Price 方法分解成為一些基底函數 $\varphi_i(\lambda)$ 的線性組合：

$$x^\alpha(\lambda) \cong \sum_{i=1}^M s_i^\alpha \varphi_i(\lambda) \quad (1)$$

式中， s_i^α 是波段選出後，在該波段位置進行約 $3 \mu m$ 波段寬的積分值，做為基底函數的加權係數。 M 是基底函數的數目，經由 M 個基底函數可重建、近似於光譜曲線。

(1) 選取頻寬波段的中心位置：

應用主軸轉換 (Principal Component Transformation, PCT) 分析的觀念，影像主軸轉換特徵向量的順序是依據變異量排序，第一個特徵向量包含有影像資訊大部分的變異性，特徵向量的數值，可解釋成原始光譜資料加權組合的係數，原始光譜資料經加權組合轉換到主軸影像，係數中最大者意謂該原始波段權重最大。

選波段是選擇一個區間 (比如 $0.3 \mu m$ 或 $0.5 \mu m$ 寬)，這是考慮越窄的頻寬其訊雜比越低，而其頻寬的中心位置波長值，是取第一個特徵向量中最大的係數值，則可對應到原始資料的波段波長值 (λ_i)。

以 λ_i 為中央將光譜資料在區間內 ($\lambda_{min} \sim \lambda_{max}$) 加權平均，得係數 s_i^α 。

(2) 決定基底函數 $\varphi_i(\lambda)$ ：

$$\delta x_i^\alpha(\lambda) = s_i^\alpha \varphi_i(\lambda) \quad (2)$$

式中， $\delta x_i^\alpha(\lambda)$ 是光譜值，當取 10 個基底函數 φ_i ，來近似原始光譜值，則(2)式寫成：

$$x_i^\alpha(\lambda) \cong \sum_{j=1}^{10} s_j^\alpha \varphi_j(\lambda) \quad (3)$$

計算第 11 個基底函數 φ_i ，先算由基底函數近似組含量與光譜值的差值

$\delta x_i^\alpha(\lambda)$ ：

$$\delta x_i^\alpha(\lambda) = x_i^\alpha(\lambda) - \sum_{j=1}^{10} s_j^\alpha \varphi_j(\lambda) \quad (4)$$

重新由 $\delta x_i^\alpha(\lambda)$ 計算特徵向量中最大的係數值，選取頻寬波段的中心位置 λ_i ，再去計算加權平均，得係數 s_i^α 。以此漸近計算將原始 224 波段光譜值分解成 M 個基底函數 φ 的近似重建組合。並選定 M 個特徵波段。

φ 之計算乃對(2)式取最小化得到：

$$\varphi_i(\lambda) = \frac{\sum_{\alpha=1}^N [\delta x_i^\alpha(\lambda) S_i^\alpha]}{\sum_{\alpha=1}^N (S_i^\alpha)^2} \quad (5)$$

2.2.2 Tu, et al. 波段選取演算法

Tu, et al., (1998) 提出高光譜波段選取的策略，首先分析影像 Canonical analysis (CA)，或者主軸分析 (Principal components analysis, PCA)，獲得影像變異量矩陣（亦可取相關係數矩陣）之特徵值（ $\lambda_1, \lambda_2, \lambda_3, \dots$ ）與相應的特徵向量（ w_1, w_2, w_3, \dots ）。其中特徵向量 w_i ，具有 l 個元素（ $w_i = (w_{i1}, w_{i2}, w_{i3}, \dots, w_{il})$ ）， l 為高光譜影像的波段數目。

$$r_{ik} = \sqrt{\lambda_i} (w_{ik} / \|w_i\|) \quad (6)$$

式中 r_{ik} 稱之 Loading Factor，下標 i 為第 i 個特徵值與特徵向量，下標 k 為相應波段位置（ $k=1 \sim$ 波段數目）。

$$\rho_k = \sum_{i=1}^{c-1} r_{ik}^2 \quad (7)$$

式中 ρ_k 稱之辨識力 (Discriminant power)，若 ρ_k 數值越大，意謂第 k 波段具較顯著之訊息，將 ρ_k 排序（ $\rho_1 \geq \rho_2 \geq \dots \rho_l$ ），據以決定出波段選取的次序，辨識力越大者做為選用的波段。需選取多少個波段則以辨識力比率 (Discriminant power probability, DPP) 決定。

$$\text{DPP}_m = \frac{\sum_{k=1}^m \rho_k}{\sum_{k=1}^l \rho_k} \quad (8)$$

當從 l 個波段選出 m 個波段，DPP 乃反應出 m 個波段之辨識力總和與 l 個波段之辨識力總和的比率。

2.3 研究分析與成果

波段選取的方法分別應用於分析 AVIRIS 資料與地面光譜輻射儀量測值。AVIRIS 高光譜影像資料取自美國普渡大學 (Hsieh et al., 1998)，攝像時間是 1992 年 6 月，地點位於美國北印第安那州，覆蓋為農作植物，原始 AVIRIS 影像具有 224 波段，此份資料為 220 波段，光譜範圍 $0.4 \sim 2.5 \mu\text{m}$ ，影像大小 145×145 像元，光譜解析度約 10nm 。另一組資料為 GER3700 地面光譜輻射儀，光譜範圍 $0.35 \sim 2.5 \mu\text{m}$ ，光譜解析度約 1.5nm 。地面光譜輻射儀取像時間在 2001 年 9 月 14 日，下午 2 點，地點在新竹縣工研院內大草坪，量測地類是分布均勻的短草覆蓋。

2.3.1 草地覆蓋的地面光譜量測值

圖 1 是 GER3700 地面光譜輻射儀量

測草地的輻射通量值，圖 2 是相應的輻射通量標準偏差，樣本數目為 20 個量測值，由圖 1 草地的光譜反應高峰分布在 760nm 至 900nm，約在近紅外光段是高峰反應，最大的標準偏差值與相應波段的輻射通量比為 13.9%，圖 3 是同一時間量測標準白板的輻射通量值，由草地輻射通量除以標準白板輻射通量得反射率如圖 4，圖 5 是 20 個量測值反射率的標準偏差。其中，觀察 1388.5nm 附近波段、1805nm~1957nm 標準偏差偏大，反應這些波段受水氣與二氧化碳吸收干擾，標準偏差大於 0.02 的波段主要分布在 723nm 至 1283nm。

圖 6 是應用 Price 演算法，展示前三個光譜基底函數圖形，相應選出的波段如表 1 所示。圖 7 為應用 Tu 演算法計算 Discriminant power 與波長關係圖，重要的波段分布在 755nm 至 1124nm，這與圖 5 反射率的標準偏差圖形相關性很高。

表 1 Price 演算法選取波段

波段序號	波長位置(nm)
153	543
245	679
266	710
312	778
466	1078
484	1241
572	1844 (水氣吸收帶，不採用)

2.3.2 AVIRIS 資料單一地物光譜分析

分析 AVIRIS 資料訓練應本中的生長中期大豆 (Soybean) 資料，資料大小為 15x15 個像元，由圖 8 大豆的光譜

反應高峰分布在 470nm 至 920nm。圖 10 為應用 Tu, et al. 演算法計算 Discriminant power 值，重要的波段分布在 490nm 至 930nm。圖 11~14 是應用 Price 演算法，展示前四個光譜基底函數圖形，與相應的第一個特徵向量 (特徵值最大者) 如圖 15~17。相應選出的波段如表 2 所示。

表 2 Price 演算法選取波段大豆光譜波段

波段序號	波長位置(nm)
8	470
28	670
32	710
53	920
124	1630

2.3.3 AVIRIS 資料光譜分析

整張影像 145x145 個像元，分析光譜值，圖 20 為應用 Tu, et al. 演算法計算 Discriminant power 值。圖 21~23 是應用 Price 演算法，展示前三個光譜基底函數圖形，相應選出的波段如表 3 所示。

表 3 Price 演算法選取整張影像之光譜波段

波段序號	波長位置(nm)
8	470
16	550
29	680
32	710
35	740
42	810
70	1090
89	1280
123	1500
155	1820

三、結論與建議

本研究所採用的兩個演算法做高光譜波段的選取，基本原理皆以影像之

變異數矩陣為基礎，由特徵向量分析出含有主要資訊的波段，此二演算法的波段選取結果具有相容的成果。其中，Price 演算法可選出少於 10 個的重要波長位置，地面光譜儀量測草地，Price 演算法近紅外光部分重要的波段分布在 710nm、778nm、1078nm、1241nm。Price 演算法分析 AVIRIS 資料選出 470nm 位置是分辨土壤與植被差異的重要位置；550nm 是綠色植物分辨的重要位置；1280nm 是分辨葉面覆蓋的重要位置。

本研究進一步方向擬採用搜尋技術，求出最佳的波段組合，進一步與目前成果比對，並做進一步波段選取的物理意義解釋。

四、參考文獻

1. Benediktsson, J. A., Sveinsson, R., Amason, K., 1995. Classification and feature extraction of AVIRIS data. *IEEE Trans. Geosci. Remote Sensing*. 33(5): 1194-1205.
2. Kim, B., Landgrebe, D. A., 1991. Hierarchical classifier design in high-dimensional, numerous class cases. *IEEE Trans. Geosci. Remote Sensing*. 29(4): 518-528.
3. Curran, P. J., Dungan, J. L., 1989. Estimation of signal-to-noise: A new procedure applied to AVIRIS Data. *IEEE Trans. Geosci. Remote Sensing*. 27(5): 620-628.
4. Chang, C.-I., Du, Q., 1999. Interference and noise-adjusted principal components analysis. *IEEE Trans. Geosci. Remote Sensing*. 37(5): 2387-2396.
5. Green, R. O. et al., 1998. Image spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS). *Remote Sens. Environ.* 65: 227-248.
6. Green, A. A., Berman, M., Switzer, P., Craig, M. D., 1988. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Trans. Geosci. Remote Sensing*. 26(1): 65-74.
7. Haertel, V., Landgrebe, D. A., 1999. On the classification of classes with nearly equal spectral response in remote sensing hyperspectral image data. *IEEE Trans. Geosci. Remote Sensing*. 37(5): 2374-2386.
8. Hsieh, P.-F., Landgrebe, D., 1998. Classification of high dimensional data. Ph.D. Dissertation, School of Electrical and Computer Engineering, Purdue University.
9. Jimenez, L.O., Landgrebe, D. A., 1999. Hyperspectral data analysis and supervised feature reduction via projection pursuit. *IEEE Trans. Geosci. Remote Sensing*. 37(6): 2653-2667.
10. Landgrebe, D., Biehl, L., 1997. An introduction to Multispec. <http://dynamo.ecn.purdue.edu/~biehl/MultiSpec/documentation.html>.
11. Li, Z.-L., Becker, F., Stoll, M. P., Wan, Z., 1999. Evaluation of six methods for extracting relative emissivity spectra from thermal infrared images. *Remote Sens. Environ.* 69: 197-214.
12. Price, J. C., 1990. On the information content of soil reflectance spectra, *Remote Sens. Environ.* 33, 113-121.
13. Price, J. C., 1997. Spectral band selection for visible near infrared remote sensing: spectral-spatial resolution tradeoffs. *IEEE Trans. Geosci. Remote Sensing*. 35(5): 1277-1285.
14. Roger, R. E., 1996. Principal components transform with simple, automatic noise

- adjustment. *Int. J. Remote Sensing*. 17(14): 2719-2727.
15. Roberts, D. A., Gardner, M., Church, R., Ustin, S., Scheer, G., Green, R. O., 1998. Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models. *Remote Sens. Environ.* 65: 267-279.
 16. Serpico, S. B., Bruzzone, L., 2001. A new search algorithm for feature selection in hyperspectral remote sensing images. *IEEE Trans. Geosci. Remote Sensing*. 39(7): 1360-1367.
 17. Switzer, P., Green, A. A., 1984. Min/max autocorrelation factors for multivariate spatial imagery. Dept. of Statistics, Stanford University, Tech. Rep. 6.
 18. Tu, T.-T., Chen, C.-H., Wu, J.-L., Chang, C.-I., 1998. A fast two-stage classification method for high-dimensional remote sensing data. *IEEE Trans. Geosci. Remote Sensing*. 36(1): 182-191.
 19. Tung, F., Ledrew, E., 1987. Application of principal components analysis to change detection. *Photogrammetric Engineering and Remote Sensing*. 53(12): 1649-1658.
 20. Tadjudin, S., Landgrebe, D., 1998. Classification of High Dimensional Data with Limited Training Samples. Ph.D. Dissertation, School of Electrical and Computer Engineering, Purdue University.
 21. Young, T. Y., Fu, K. S., 1986. Handbook of pattern recognition and image processing. College of engineering, Uni. Of Miami, Coral Gables, Florida.

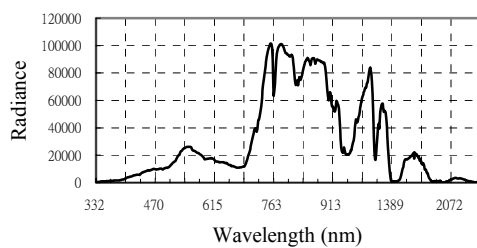


圖 1 GER 草地光譜反應

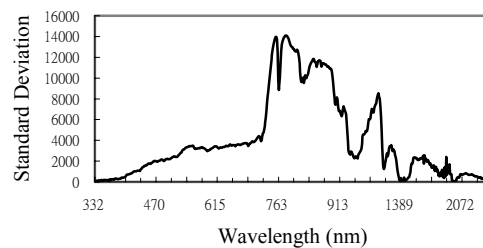


圖 2 GER 草地光譜反應標準偏差

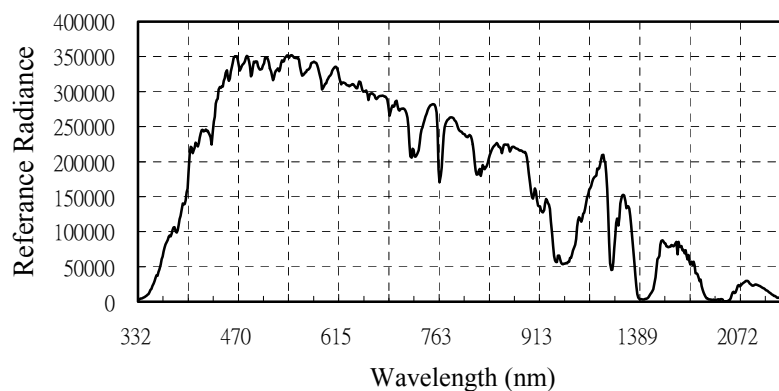


圖 3 GER 參考白板光譜反應

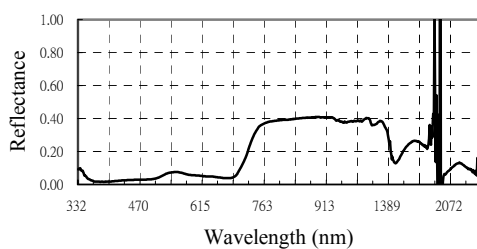


圖 4 GER 草地反射率

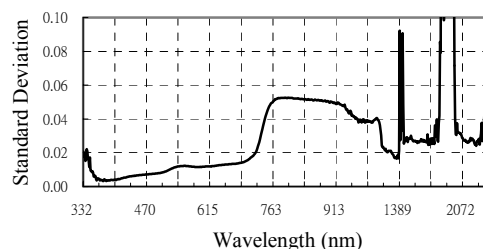


圖 5 GER 草地反射率標準偏差

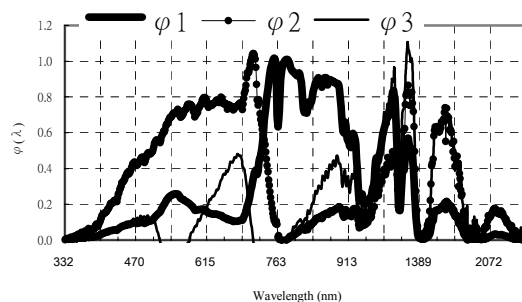


圖 6 基底函數圖

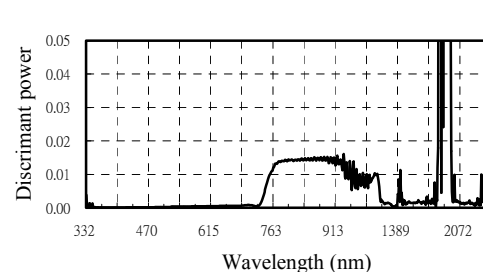


圖 7 Discriminant power 值與波長關係

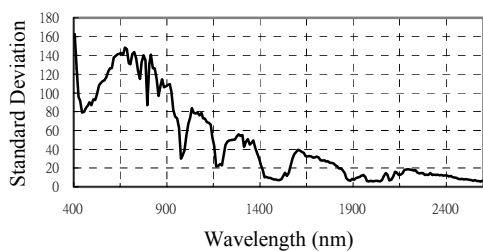


圖 8 大豆光譜反應標準偏差

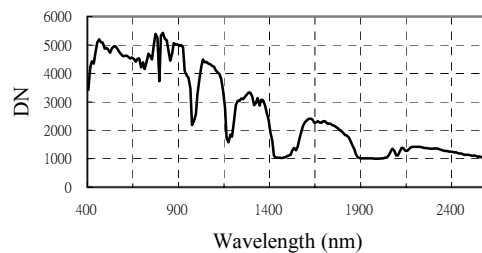


圖 9 大豆光譜反應

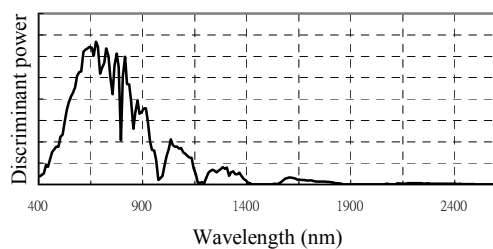


圖 10 大豆 Discriminant power 值

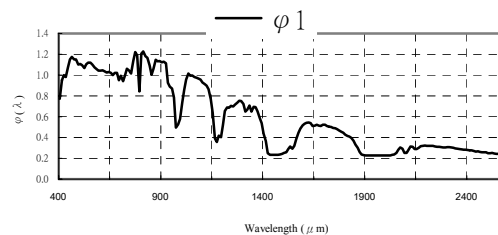


圖 11 大豆第一個基底函數

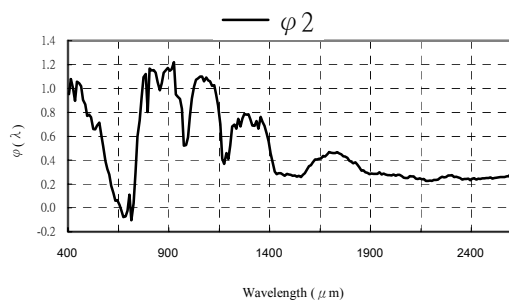


圖 12 大豆第二個基底函數

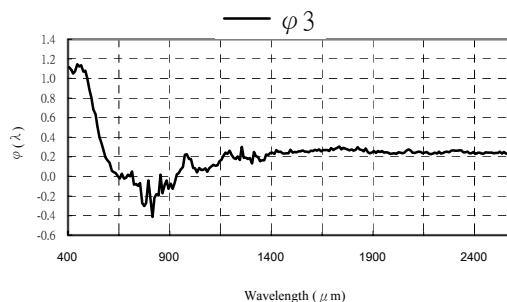


圖 13 大豆第三個基底函數

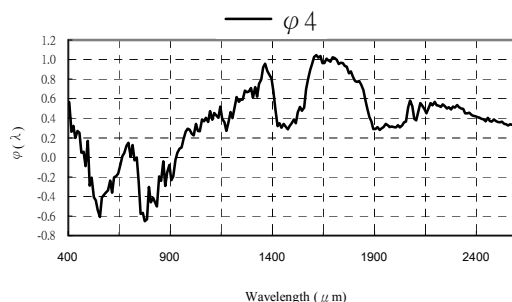


圖 14 大豆第四個基底函數

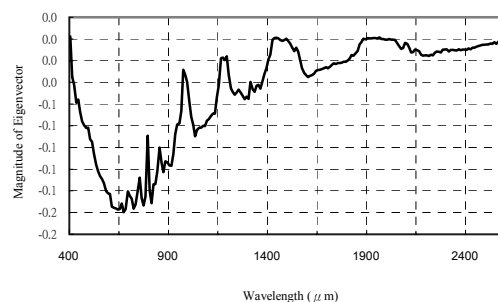


圖 15 $\phi 1$ 相應第一個特徵向量

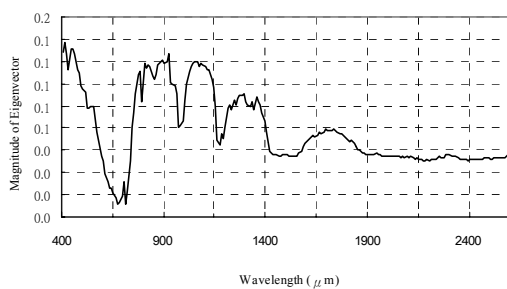


圖 16 ϕ_2 相應第一個特徵向量

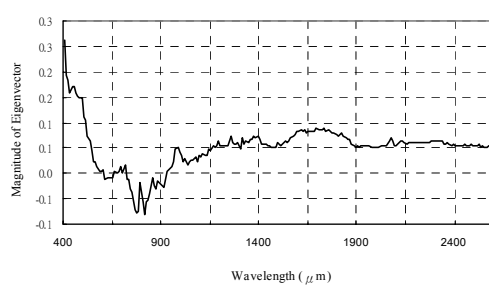


圖 17 ϕ_3 相應第一個特徵向量

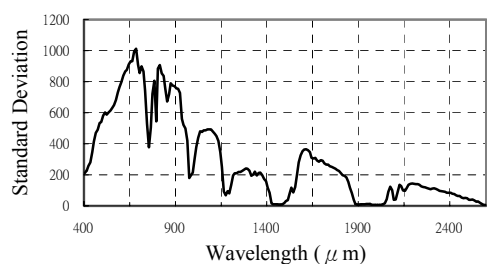


圖 18 AVIRIS 光譜反應標準偏差

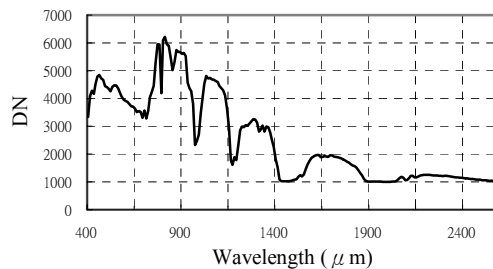


圖 19 AVIRIS 光譜反應值

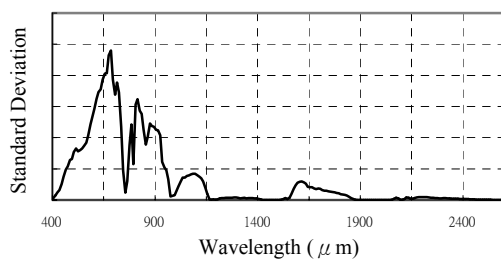


圖 20 AVIRIS Discriminant power 值

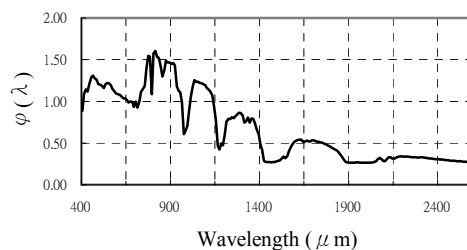


圖 21 AVIRIS 第一個基底函數

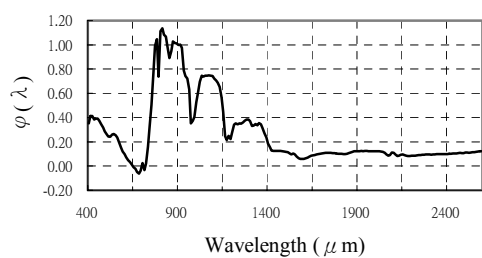


圖 22 AVIRIS 第二個基底函數

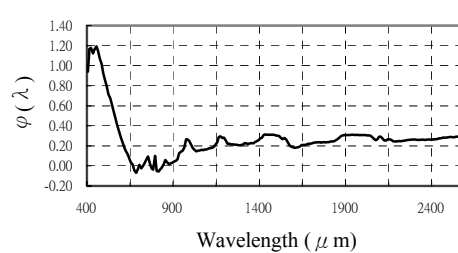


圖 23 AVIRIS 第三個基底函數