Jal Vigyan Sameeksha—a publication of
Indian National Committee on Hydrology
Vol. IV, No. 1, June 1989

Use of Principal Components and Canonical Analysis in Remote Sensing for Classification of Data

Dr. Kunwar K. Rampal
Professor
Department of Civil Engineering
Indian Institute of Technology
KANPUR (U.P.) 208 016, INDIA

Abstract : *Several classification techniques have been used in remote sensing for pattern recognition. These techniques are basically divided into two categories, viz., supervised and unsupervised classification. The classification algorithm identifies the important features of a class, i.e., determines what each class looks like and given the data for a pixel of the imagery, it compares the features of the pixel with the features of each class and then assign the pixel to one of the classes. When multispectral band of data are available, the variances and correlation of spectral responses pattern of different classes are determined from the training data set to identify salient features of the class, to be used in classification.*

Two classification techniques have been analysed and compared in this study. These techniques are the principal components analyses and the canonical analysis. Both of them are preprocessing techniques. The study has revealed that the internal correlation is of the order of 0.77 to 0.97 and only two orthogonal components can explain 97% to 98% of the variance of the original sample. They hence, constitute a data reduction or condensation technique for preprocessing of multi-spectral and TM data.

Introduction

The main objective of classification is as follows : given a set of objects, to assign each object to one of a number of predetermined groups. The most common application is in remote sensing where the categories are classes such as water, built-up areas, vegetation etc. and the objects are the pixel in the image. Classification may also be thought of a labelling problem in which we label each object with a class label.

A commonly employed preprocessing technique is the linear transformation of variables. Consider for example, multispectral data in 'η' channels. Generally there is a correlation between the reflectance values in different channels and these correlation may vary with the groups or classes. The principal components are mutually orthogonal linear transformation of original variables which maximize successively the variance accounted for that component. Geometrically, this is equivalent to fitting to the image data by rotation, translation and scaling, a set of mutually orthogonal and hence uncorrelated axes so that the first axis or principal component clearly explains the greatest amount of variance with successive axes or component accounting for a explaining less and less variance.

Two Dimensional Case

Let us consider, for the sake of simplicity a two dimensional case of say, Band 5 & Band 7 plot of reflectance values for a class such as water with 'X' axis for Band 6 and 'Y' axis for Band 6. (See Fig. 1). Let us rotate now the axis through an angle 'θ', keeping the translation and scale factor unchanged. The relationship between the old and new coordinates is given by :

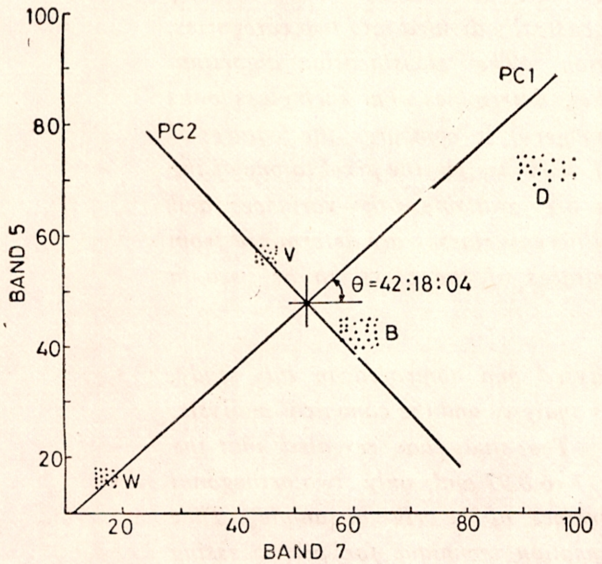

Fig. 1 : Plot of Reflectance Values Vs Wave Length for Bands 5 and 7

$$X' = X\cos\theta + Y\sin\theta \quad (1)$$

$$Y' = -X\sin\theta + Y\cos\theta \quad (2)$$

In matrix this is expressed as :—

$$\begin{pmatrix} X' \\ Y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \quad (3)$$

Let σ^2_X and σ^2_Y be variances of X and Y respectively and $\sigma^2_X{}'$ and $\sigma^2_Y{}'$ be the variances of X' and Y'. By the law of propogation of variances coveriance we have

$$\Sigma' = A\,\Sigma\,A^t \quad (4)$$

where Σ' = variance covariance matrix X', Y'

Σ = variance covariance matrix of X. Y

A = The coefficient matrix =

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

$$\text{Thus}\quad \Sigma' = \begin{pmatrix} \sigma^2_{X'} & \sigma_{X'Y'} \\ \sigma_{Y'X'} & \sigma^2_{Y'} \end{pmatrix} \quad (5)$$

$$\text{and}\quad \sigma^2_{X'} = \sigma^2_X\cos^2\theta + \sigma^2_Y\sin^2\theta + \sin\theta\cos\theta\,(\sigma_{YX} + \sigma_{XY}) \quad (6)$$

$$\sigma^2_{Y'} = \sigma^2_X\sin^2\theta + \sigma^2_Y\cos^2\theta + \sin\theta\cos\theta\,(\sigma_{XY} + \sigma_{YX}) \quad (7)$$

$$\sigma_{X'Y'} = \sigma^2_Y\sin\theta\cos\theta - \sigma^2_X\cos\theta\sin\theta + \sigma_{XY}\cos^2\theta + \sigma_{YX}\sin^2\theta \quad (8)$$

$$\sigma_{Y'X'} = \sigma_{X'Y'} \quad (9)$$

Now, in order to minimize covariance and maximize variance we must have :

$$\sigma_{X'Y'} = \sigma_{X'Y'} = 0 \quad (10)$$

which gives :

$$\tan 2\theta = \frac{2\,\sigma_{XY}}{\sigma_X{}^2 - \sigma_Y{}^2} \quad (10)$$

Thus if the axes are rotated through an angle θ given by the expression (11) above the new variance σ^2_X, and σ^2_Y, will be maximized and covariance $\sigma_{X'X'}$, $\sigma_{,Y'X}$ will vanish, breaking the correlation between X'. Y' and thus making then independent of each other. These new components X', Y' are the First and Second Principal Components (PC's).

In general, if we have more than two dimensional pixel values (reflectance) of image i e. n bands, the principal components can be obtained as follows :

Let Σ be the variance covariance matrix of i class and n bands. Since Σ is symmetric it can be diagonalized. Let γ_i. i = 1, 2. .. ... n be the eigen values of Σ (i.e. the roots of $|(\Sigma - \lambda_i I)| = 0$ in decreasing order i.e. $\lambda_1 \lambda_2 \ldots\ldots \lambda_n$. We assume for the sake of simplicity that there are no degenerate eigen values. Let g_i, i = 1, 2, ... n be the orthonomal base of the eigen vectors of Σ, that satisfy

$$\Sigma, g_i = \gamma_i g_i \quad (12)$$

and let G be the orthogonal matrix that diagonalizes Σ, so that

$$G \Sigma G^{-1} = \lambda \quad (13)$$

where

$$\lambda_{ij} = \lambda_i \delta_{ij}, \delta_{ij} = 0 \text{ if } i \neq j \quad (14)$$

$$\delta_{ij} = 1 \text{ if } i = j$$

Then the transformation

$$Y_i = \sum_{j=1}^{n} G_{ij} X_j \ldots\ldots i=1, \ldots\ldots n \quad (15)$$

gives the required transformation and hence the principle components.

This method, therefore, defines a rotation of n-dimensional coordinate system such that there is data reduction i.e. the data are arranged along axes of decreasing variance. The linear combination of the original data of radiances from the variance of the transformed data is a maximum along the first principal axis.

The percentage of information content of band 'j' is given by :

$$\text{percentage} = \frac{\lambda_{ij}}{\sum_{ij} \lambda_{ij}} \times 100 \quad (16)$$

and the correlation coefficient R_{nm} is given by :

$$R_{nm} = \frac{a_{nm} \cdot (\lambda_{nm})^{1/2}}{(s^2_{nm})^{1/2}} \quad (17)$$

where n = band, m = component

a_{nm} = eigen vector for nth row, jth column

λ_{nm} = eigen value for nth band and jth component

s^2_{nm} = variance for the nth band.

Canonical Analysis

The objective of canonical analysis (MC. Murty, 1976) is to derive a linear transformation that will emphasize the difference amongst the pattern sample belonging to different categories. The theoretical formulation is given in Bow (1985) and may be compared with that of James (1985).

Data used in Study

The data collected for the study are as follows :–The area is covered by one Landsat-4 scene. Tha index map is supplied by NRSA Hyderabad. A scene is located by its path number and row number. The area selected is located by the Path Number 142 Row Number : 042. The Landsat data, corresponding to the aforesaid scene are obtained in the form af CCT (1600 BPI) containing reflectance values for bands 1, 2, 3 and 4. But the corresponding positive films (1 : 1 m) and paper prints (1 : 1/2 m) could not be obtained from NRSA during the course of this study. The toposheets 63 K and 630 covering the area were also available for the study.

Classification using Principal Components

Using the principles explained in earlier, and with two principal components (i.e., p=2) the category wise means and standard deviations of the principal components as well as the 95% symmetrical confidence limits have been calculated and they are shown in Table 1. The confidence limits for different categories for the first principal component indicate that water and dry soil have non-overlapping confidence limits and built up areas and

Table 1 ; Categorywise Means, Standard Deviation and 95% Confidence Limits

Principal Components 1

Sl. No.	Category	Means (μ)	Standard Deviation (σ)	95% Confidence (μ—1.96)	Limits (μ+1.96)
1.	Water	— 5.6336	0.4147	— 6.4465	— 4.8209
2.	Built-up Areas	—16.9809	0.6099	—18.1586	—15.6803
3.	Vegetation	—15.5893	0 5313	—16.6306	—14.5479
4.	Dry Soil	—27.3697	1.5082	—30.3258	—24.4137
	Principal Component 2				
1.	Water	— 3.0213	0.2830	— 3.5761	— 2.4667
2.	Built-up Areas	—12.2285	0.6894	—13.5797	—10 8773
3.	Vegetation	— 5.5449	0.3990	— 6.3270	— 4 7629
4.	Dry Soil	—17.5357	1.7006	—20.8689	—14.2025

vegetation have overlapping confidence limits among themselves, but disjointed from water and dry soil. The confidence limits for the different categories in second principal component indicate a disjoint set for built-up areas and vegetation. Accordingly, a layered or decision tree approach for classification as shown in Fig. 2, is valid. This can also be used for preparing colour composites with enhanced discrimination.

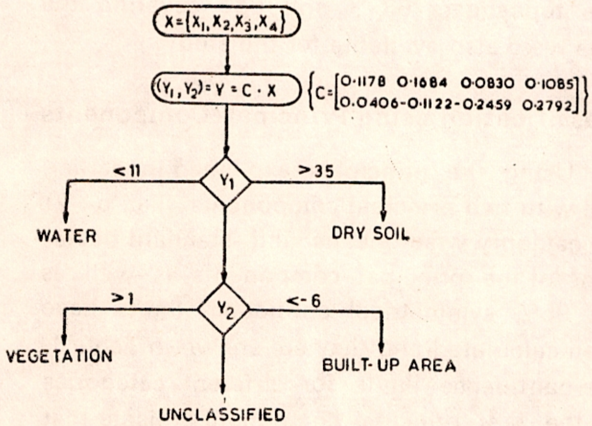

Fig. 2 : Decision tree for Layered Classification of Canonical Components

The eigen values show that the first two eigen values contributing to 85.69 and 12.38 percent of the variance respectively are significant, the third one contributing to 1.92% is small and the last one with 0.01% is trivial. Accordingly, it is identified that p = 2 and 2 canonical components are required.

Classification using Canonical Analysis

The transformation matrix derived from canonical analysis can be used to transform original data and for subsequent classification as in the case of principal component analysis.

Assuming two canonical components, i.e., p = 2, the category wise means and standard deviation of the component as well as 95% symmetrical confidence limits have been calculated and are shown in Table 2.

The confidence limit for the first component show, as in the case of principal components analysis that water and dry soil have non overlapping confidence limits and built-up areas and vegetation have ovarlapping con-

Table 2 : Categorywise Means, Standard Deviation and 95% Confinence Limits

Canonical Components 1

Sl. No.	Category	Means (μ)	Standarp Deviation (σ)	95% Confidence (μ—1.96)	Limits (μ+1.96)
1.	Water	9 2909	0 6871	7.9442	10 6376
2.	Built-up Areas	24.4806	0.7893	23.2891	25.6721
3.	Vegetation	27.1588	0.6079	25.6112	28.7058
4.	Dry Soil	38.5298	1.5914	35.4106	41.6489
	Canonical Component 2				
1.	Water	— 0.9558	—0.4020	— 1.7437	— 0.1679
2.	Built-up Areas	2.6330	0.8693	1.2224	4.0436
3.	Vegetation	— 7.9891	0.7197	— 9.6929	— 6.1872
4.	Dry Soil	0.3251	1.6013	— 2.8134	3 4636

fidence limits among themselves but disjointed from water and dry soil. The confidence limits for different categories in 2nd component also indicate in a similar fashion a disjointed set for built up areas and vegetaion. Accordingly, a decision tree layered classification is shown in Fig. 3.

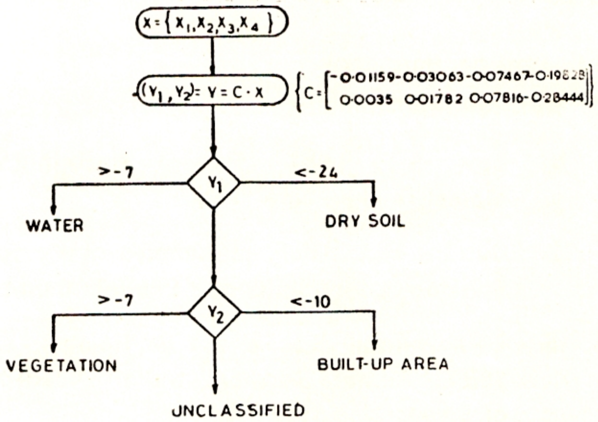

Fig. 3 : Decision tree for Layered Classification of Principal Components

omparison of Principal Components nalysis and Canonical Analysis

Both the principal components analysis and canonical analysis are orthogonal linear transformation procedure suitable for a condensation of multichannel information into fewer components. The components are orthogonal to one another and are hence uncorrelated. Principal components analysis significantly maximises the variance explained by the components, while canonical analysis maximizes the ratio of among the class variance to within the class variance sequentially.

A comparison of results of eigen values in principal components analysis indicate that the canonical components explain a slightly larger percentage of the variance (98.07 vs 97.5) than that of principal components.

Results of Classification

The pixel wise classification of data may be presented either as colour composites in terms of principal components or canonical components, or as coded or grey imageries. The results of canonical analysis are used for classification of a part of the scene used in the study with the following corner points.

Fig. 4 : A portion of line printer hard copy

Sl. No.	Longitude			Latitude		
1.	82	13	12	26	56	24
2.	84	22	48	26	37	48
3.	81	57	00	25	22	48
4.	84	05	24	25	04	48

The coded imagery obtained from this step is shown in Fig. 4. The result are considered quite satisfactory. It is, hence concluded that canonical analysis is a powerful computer oriented decision tree or layered classification technique, eminently suitable for categorisation of remote sensed data.

Acknowledgement

The author acknowledges the work of Shri Sharad Varshney for collection of data in this paper and as in Varshney (1987).

References

1. Bow, S.T., 1984. Pattern Recognition, Marcel Dekker, New York.
2. Curran, P.J., 1985. Principles of Remote Sensing, Longmen Group Limited, London.
3. Deekshatulu, B.L. and Y.S. Rajan., (ed) 1984. Remote Sensing, Indian Academy of Sciences, Bangalore.
4. James, M., 1985. Classification Algorithms John Wiley and Sons. New York.
5. Varshney, S., "Linear Transformations in Remote Sensing Classification", M. Tech. Thesis, July 1987, Department of Civil Engineering, Indian Institute of Tecbnology Kanpur (U.P.), India.