[9-72]

다중분광 영상데이터의 주성분변환 및 주성분 기반 영상분류

서 용 수

동서대학교 정보시스템공학부

seoyong@dongseo.ac.kr

Principal Component Transformation of the Multispectral Image Data and Principal-Components-Based Image Classification

Yong-Su Seo

School of Information System Engineering, Dongseo University

Abstract

Advances in remote sensing technologies are resulting in the rapid increase of the number of spectral channels, and thus, growing data volumes. This creates a need for developing faster techniques for processing such data. One application in which such fast processing is needed is the dimension reduction of the multispectral data. Principal component transformation is perhaps the most popular dimension reduction technique for multispectral data. In this paper, we discussed the processing procedures of principal component transformation. And we presented and discussed the results of the principal component transformation of the multispectral data. Moreover principal components image data are classified by the Maximum Likelihood method and Multilayer Perceptron method.

1. 서 론

다중분광 영상데이터는 여러개의 파장범위로 분리된 여러 개의 분광대역별로 수집되고 있다. Landsat TM 영상데이터의 경우 분광대역 수가 7개이나, 일반적으로 다중분광 영상데이터는 수 개에서 수십 개의 분광대역을 가진다.

Hyperspectral 영상데이터의 경우에는 수십 개에서 수백 개의 분광대역을 가진다[1],[2]. 다중분광 영상데이터의 각 대역들 간에는 서로 상관관계가 높다. 이렇게 상관관계가 높다는 것은, 각 대역들간에 정보의 중복성(redundancy)이 높음을 의미하고, 중복성이 높은 원래의 분광 영상데이터들을 가지고 영상을 분석하는 것은 처리시간 및 처리비용의 증가로 인하여 비효율적이다. 분광 대역 수를 줄이기 위한 차원 축소기법(dimension reduction)으로 널리 사용되는 것이 주성분변환 법이다[1]-[3]. 다중분광 영상데이터의 주성분변환은 영상분류[4], 영상디스플레이[5], 변화검출, 영상융합[6] 등의 여러 분야에 응용되고 있다.

본 논문에서는 주성분변환 처리방법에 대해 논한 후 Landsat TM 데이터를 사용하여 주성분변환 결과를 분석하는데 목적이 있다. 또한 주성분 영상데이터를 최대유사법과 신경회로망을 이용한 다층 퍼셉트론법[7],[8]으로 분류하고 결과를 분석하는데 있다.

2. 주성분변환 방법

다중분광 영상데이터의 분광 대역수가 n개 일 때, 이들 데이터(B_1, B_2, ... , B_n)로부터 계산되는 공분산행렬(Covariance Matrix) C는 식(1)과 같이 표현할 수 있다.

$$C = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix} \qquad 식(1)$$

이때, 행렬의 대각선상의 원소 σ_{ii}는 대역 i의 분산 값이고, 비대각선상의 원소 σ_{ij}는 대역 i와 대역 j간의 공분산값이다.

공분산행렬 C로 표현된 다음 식(2)를 특성방정식이라고 하며, 이 방정식을 풀이하면 n개의 근(λ_1, λ_2, .., λ_n)을 구할 수 있고, 이들 근을 공분산행렬의 고유값(eigenvalue)이라 한다. n개의 고유값들을 계산한 후

$$|C-\lambda I| = 0 \qquad 식(2)$$

이때, I는 단위행렬, $|C-\lambda I|$는 행렬식 크기 순으로 나열한다. k번째의 고유값 λ_k에 대한 고유벡터 K_k는 다음 식(3)을 풀이하여 계산한다. 이 방정식을 n개의 고유값들에 대해서 풀이하면 n개의

$$(C - \lambda_k I)\, K_k = 0 \qquad 식(3)$$

이때, k는 1, 2, ..., n열 벡터(K_1, K_2, .., K_n)들을 구할 수 있고, 이들 벡터를 고유벡터(eigenvector)라 한다. 각 고유벡터를 표현하면 다음 식(4)와 같다. 이는 n개의 각 고유벡터들을

$$K_1 = \begin{bmatrix} K_{11} \\ K_{21} \\ \vdots \\ K_{n1} \end{bmatrix},\quad K_2 = \begin{bmatrix} K_{12} \\ K_{22} \\ \vdots \\ K_{n2} \end{bmatrix},\ \ldots\ldots\ K_n = \begin{bmatrix} K_{1n} \\ K_{2n} \\ \vdots \\ K_{nn} \end{bmatrix} \qquad 식(4)$$

다음 식(5)과 같이 구성한 것을 주성분 변환행렬이라고 하며 W로 나타낸다.

$$W = \begin{bmatrix} K_1^T \\ K_2^T \\ \vdots \\ K_n^T \end{bmatrix} = \begin{bmatrix} K_{11} & K_{21} & \ldots & K_{n1} \\ K_{12} & K_{22} & \ldots & K_{n2} \\ \vdots & \vdots & & \vdots \\ K_{1n} & K_{2n} & \ldots & K_{nn} \end{bmatrix} \qquad 식(5)$$

이 변환행렬 W을 이용해서 아래 식(6)과 같이 변환하는 것을 주성분변환이라 한다.

$$\begin{bmatrix} PC_1 \\ PC_2 \\ \vdots \\ PC_n \end{bmatrix} = \begin{bmatrix} K_{11} & K_{21} & \ldots & K_{n1} \\ K_{12} & K_{22} & \ldots & K_{n2} \\ \vdots & \vdots & & \vdots \\ K_{1n} & K_{2n} & \ldots & K_{nn} \end{bmatrix} \cdot \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{bmatrix} \qquad 식(6)$$