

정준벡터분석에 의한 混合畵素 解析技法에 關한 研究 A Technique for Mixed Pixel Extraction by Canonical Vector Analysis

박 민 호*
Park, Min-Ho

要 旨

인공위성데이터로부터 보다 자세한 정보를 취득하기 위해 혼합화소를 해석하는 방법에 대해 연구하였다. 본 연구에서는 2가지 토지피복만 혼합되어 있는 경우로 한정하였다. 정준상관분류기법에서 생성되는 정준벡터를 분석하여 혼합화소를 분류해 낼 수 있었으며, 그 기준으로 정준벡터의 요소인 정준가중치 2개의 상대적 비율을 역치로 사용하였다. 9월 1일을 전후한 TM 데이터의 다리과 물 항목에 대한 분류의 경우에, 혼합화소를 가장 적절히 분류하는 역치는 4.0으로 결정되었다. 즉 정준가중치사이의 비율이 4.0이상이면 단일피복화소이며, 4.0이하이면 혼합화소로 간주하게 된다. 정준가중치의 분포에 의해 대략적인 토지피복 구성비율도 추정할 수 있다. 실험영역에 대한 혼합화소 추출의 정확도는 90%로서 높은 수준이었다. 따라서 정준벡터분석에 의한 혼합화소 분류방법은 효용성이 있다고 판단된다.

ABSTRACT

To achieve more accurate information from satellite image data, a research on a technique for mixed pixel extraction has been produced. The mixed pixels with only two land covers have been experimented. By analyzing canonical vector in canonical correlation classification, the mixed pixels have been classified. The ratio of the two canonical weighted values-the elements of canonical vector have been used as a threshold to discriminate mixed pixels. In case of the classification for the mixed pixels of bridge and water class in TM data before or after the 1st of September, the threshold for the optimal classification of the mixed pixels is 4.0. That is, if the ratio of the two canonical weighted values is less than 4.0, the pixel is a mixed pixel. Also, using the distribution of canonical weighted values, the constitution percentages of land covers within one mixed pixel can be approximately deducted. The accuracy of mixed pixel extraction for experimental area is 90% and quite acceptable. Conclusively, a technique for mixed pixel extraction by canonical vector analysis is effective.

1. 序 論

분류기법이 고도화 되면서 인공위성 데이터로 부터 보다 정확하고 많은 정보를 취득하고자 하는 노력이 계속되고 있다. 인공위성 데이터는 래스터형식의 데이터로 제공되므로 각각의 화소는 일정면적을 갖게된다. 이때 각 화소가 나타내는 토지피복특성은 한가지 성분만을 가질 수도 있지만 여러 토지피복특성이 혼합되어 있을 수도 있다. 예를 들면 첫째, 두 가지 이상의 토지피

복 경계를 포함하는 화소를 들 수 있고, 둘째, 인공위성 센서의 해상도가 낮아서 한 화소가 차지하는 영역이 넓어 한 화소내에 여러 토지피복항목이 존재하는 경우를 들 수 있다. 이러한 다양한 토지피복상태를 갖는 한 화소를 하나의 토지피복항목으로 결정함은 무리이며 정보의 손실을 가져오게 된다. 혼합화소를 정확하게 분석해 내는 것은 원격탐사분야의 다수의 전문가들에게 상당한 관심사이며, 최근까지 여러 접근방법으로 연구되어져 왔다. 즉 영상분류를 통하여 토지피복정보를 얻는 과정에서 정보의 손실을 가능한 한 줄여 보다 많은 정보를 취득하기 위해서, 여러 토지피복들이 한 화소에

*목포대학교 지적학과 전임강사

함께 존재하는 화소들을 혼합화소로 분류해 내는 다양한 방법(화소분할방법이라고도 함)들이 시도되어 왔다. 화소분할방법은 크게 2가지로 분류되며, 모델베이스의 Textural Analyses와 통계적처리과정 방법이 있다.⁹⁾ 모델베이스 Textural Classifiers는 식물분류에 적용하기 위해 Jackson(1979)에 의해 개발된 방법으로, Texture와 판별항목을 분석하기 위해 혼합화소내에서 토지피복의 비율을 명확히 모델화하는 분류과정이다.¹⁶⁾ 화소분할 알고리즘의 두 번째 형태는 개별적인 혼합화소내의 기지의 토지피복에 대한 비율을 평가하는 통계적과정으로 이루어진다. 비율을 정하는 데에는 대표적으로 4가지 통계적 방법을 들 수 있다.

(1) 단순 비중평균법(Marsh *et al.*, 1980)¹⁷⁾

(2) 최대우도분류(Horwitz *et al.*, 1971; Work and Gilmer, 1976; Foody and Cox, 1991)^{14,15,19)}

(3) 선형판별분류(Marsh *et al.*, 1980; Chhikara, 1985)^{13,17)}

(4) 선형화귀법(Richardson *et al.*, 1975; Marsh *et al.*, 1980; Foody and Cox, 1991)^{14,17,18)}

이 외에도 Wang(1990)이 Fuzzy Classification에 의해 혼합화소를 분류해 내었으며,⁷⁾ Foschi(1993)는 Moving-Window 알고리즘을 제안하여 혼합화소의 토지피복비율을 평가하였고,⁹⁾ Karaska 등은 Thematic Mapper 영상에 대해 혼합화소 처리방법인 Applied Analysis Spectral Analytical Process(AASAP)를 적용하여 수종을 구분해 내었다.⁸⁾ 또한 국내에서도 박민호 등(1995)이 혼합화소의 해석을 위해 퍼지시그모이드 함수와 역전파 신경망을 이용한 방법을 제안하였다.²⁾

본 논문에서도 위성영상으로부터 토지피복정보를 취득함에 있어 정보의 손실을 줄이기 위해, 개념과 분석의 기본원리는 같지만 이제까지는 사용해 보지 않은 다른 변수를 사용하여 혼합화소를 분류하는 방법에 대해 연구하였다. 즉 정준상관분류 기법에서 생성되는 정준벡터를 분석하여, 혼합화소를 해석하는 방법에 대해 연구를 수행하였다. 이 방법의 핵심은 정준상관분류기법에서 유사도 판별 기준값으로 쓰이는 정준가중치의 상대적 비율을 분석하는 것이다. 이로부터 혼합화소를 판별할 수 있는 기준을 마련하였고, 화소분할을 위한 토지피복의 비율을 평가하는데 있어 기존의 4가지 통계적 방법 외에 정준벡터에 의한 방법도 가능성을 제시하였다.

2. 混合畫素抽出 알고리즘

2.1 概念

정준상관분류과정에서 혼합화소로 판단하는 기준은 다음과 같다.

토지피복항목의 표본데이터(P)					개별화소(Q)	
Class 밴드	1	2	p	Pixel 밴드	1
1	x ₁₁	x ₁₂	x _{1p}	1	y ₁
2	x ₂₁	x ₂₂	x _{2p}	2	y ₂
.
7	x ₇₁	x ₇₂	x _{7p}	7	y ₇

위의 작성된 표는 정준상관분류를 화소별로 수행하기 위한 기본자료구조가 된다.^{3,4)} 토지피복항목의 표본데이터인 p개의 변수(예측변수)는 항공사진과 원래의 수치영상에서 識別이 분명하고 지형도에서 분류항목이 명확한 대상들에 대한 training data이며, 자료행렬(data matrix)의 각 요소값(x_{nq})은 각 밴드별로 training area별 화소값들에 대한 평균값 또는 대표값을 Z-score를 사용하여 변환한 값이다. 위성영상에 대한 정준상관분류에서 기준변수로는 분류하고자 하는 수치영상의 화소 1개씩만을 사용한다. 즉, 크기(row, column)=(m, n)인 직사각형 영역 위성영상의 경우, 좌측상단에서부터 우측하단 방향으로 순서대로 화소별 분류를 실시한다. 따라서 정준상관분류의 전체 수행과정이 총 화소 개수만큼 반복 실시된다. 이상과 같은 방법은 기준변수와 예측변수간의 정준상관관계에 의해 산출되는 정준벡터를 이용하여 직접 분류를 수행하는 영상분류 기법이다. 기준변수군의 변수는 항상 1개이므로, 새로 생성되는 정준변량 U_i의 정준벡터 a_i는 정준가중치가 항상 1로 고정된다.^{5,10)} 예측변수군의 변수는 training area의 개수에 따라 결정되며, 정준변량 V_i의 정준벡터 b_i는 벡터요소로서 training area 개수만큼의 정준가중치를 갖게 된다.^{5,10)} 이 중 가장 큰 정준가중치값으로 각 화소가 귀속되게 된다. 즉 가장 큰 정준가중치를 갖는 피복이 각화소의 토지피복으로 결정된다. 정준가중치는 각 토지피복으로 귀속될 수 있는 성향의 크기이므로, 정비례하지는 않지만 각 화소가 갖는 토지피복 성향의 비중에 볼 수도 있다. 따라서 정준가중치의 상대적인 비율에 의해 혼합화소를 판단하게 되며, 이 값

의 분포가 화소를 분할할 수 있는 기준이 된다.

2.2 方法

2.2.1 유의성 검정

혼합화소를 해석하기 전에 각 화소에 대해 예측변수 군내의 어느 하나의 토지피복과 정준상관관계가 유의한지 알아보기 위해 고유치에 대한 유의성을 검정해야 한다. 유의성 검정을 통과하지 못하면 이 화소를 예측 변수군의 토지피복과는 다른 피복특성을 가진 화소로 간주하여, 추출하고자 하는 혼합화소에 대한 해석과정을 수행하지 않게 된다. 즉 미분류화소로 구분한다. 유의성 검정을 통과한 화소는 정준벡터를 사용하여 토지 피복을 결정하게 된다.

본 연구에서 수행되는 유의성 검정은 다음과 같은 근사식을 이용하여 처리된다.^{6,10)}

$$-\left(n-1-\frac{1}{2}(p+q+1)\right) \ln \prod_{i=1}^q (1-\rho_i^2) > \chi_{pq}^2(\alpha) \quad (1)$$

유의수준 α 에서 자유도 pq를 갖는 카이제곱분포함수의 값이 위의 식을 만족하면 귀무가설을 기각하므로 정준상관관계가 유의한 것이며, 따라서 한 개의 토지피복 또는 혼합화소로 결정되게 된다. 본 연구에서는 $n=7$, $p=2$, $q=1$ 이며, 유의수준 $\alpha=1\%$ 에 자유도 $pq=2$ 일 때 카이제곱함수의 값 $\chi^2=9.21$ 을 사용하여 계산된 고유치 $\rho_i^2=0.9$ 를 기준으로 검정을 수행한다. 원래 식 (1)은 n 값이 큰 경우 근사적으로 자유도 pq를 갖는 카이제곱 분포를 따른다고 정의되지만, 본 연구에서 $n=7$ 의 경우에 적용하였을 때도 대략적이기는 하나 수용할 만한 결과를 가져옴을 알 수 있었다. 즉 동작대교와 수역만으로 이루어진 실험영역 외에 산림, 논, 나대지가 포함된 5가지 토지피복으로 이루어진 다른 영상데이터에 대해서도 물과 다리영역으로 구성된 2가지 예측변수로 각 화소에 대한 유의성검정을 수행한 결과 전체영역의 97% 이상이 올바르게 분류되었다. 즉 물과 다리는 거의 100% 가깝게 정확히 분류되었고 산림, 논, 나대지의 97% 이상이 유의수준 1%에서의 허용고유치인 0.9보다 훨씬 작은 값을 가지며 정확히 미분류화소로 분류되었다.

2.2.2 혼합화소 해석

분류할 각 화소는 혼합화소의 해석을 고려하지 않을 때에는, 열벡터 \mathbf{b} 의 원소 값 중 가장 큰 값(max1)이 해

당되는 피복항목으로 결정된다. 토지피복이 복합되어 있는 혼합화소를 해석하는 방법은 다음과 같다. 열벡터 \mathbf{b} 의 원소 값 중 가장 큰 값(max1)을 두 번째 큰 값(max2)으로 나눈 값이 일정값(역치)보다 크면 화소의 피복상태를 가장 큰 값이 해당되는 피복항목으로 결정하고, 역치보다 작으면 최소한 2개 이상의 피복이 분류 화소의 영역에 존재한다고 판단하여 혼합화소(mixed pixel)로 지정한다. 역치는 $\text{ratio}=(\text{max1}/\text{max2}>1)$ 의 값을 최소한 0.1씩 증가 또는 감소시켜 적용하여 분류된 각 결과와, 영상에 대한 시각적 분광패턴에 의해 혼합 화소로 판단되는 지역을 서로 비교하면서 施行錯誤를 거쳐 가장 적절하다고 판단되는 값으로 결정한다. 일반적으로 혼합화소란 한 화소내에서 분류항목 하나가 차지하는 면적이 대략 90% 이내일 때를 말한다.²⁾ 본 연구에서도 그 기준에 따라 정준가중치의 크기를 분석하였다. 여기서 90%란 정준가중치의 값이 0.9를 의미하는 것은 아니며, 정준벡터에서 정준가중치들의 상대적인 크기와 정준가중치의 요소 개수에 따라 달리 판단해야 한다. 그림 1은 혼합화소를 해석하는 프로그램과 정의이다.

3. 實驗 및 評價

3.1 實驗 領域

실험대상영역은 동작대교를 중심으로 21행x11열을 갖는 직사각형 면적을 선택하였다. 다양한 토지피복지역에서는 혼합화소를 확인·평가할 수 있는 지상실체 데이터를 얻어내는 것이 거의 불가능하였기 때문에, 물과 동작대교영역만을 포함하는 2가지 피복항목으로 구성된 영상데이터로 실험을 수행하였다. 이렇게 함으로써 물과 동작대교의 인접부분에서 물과 대교의 혼합화소가 존재할 수도 있음을 명백히 판단할 수가 있게 되었다. 즉 화소의 디지털값의 분포만으로도, 물도 아니고 대교부분도 아닌 경우는 분명히 혼합화소임을 판단할 수 있도록 실험영역을 2가지 토지피복형태로 선택하였다. 여기서 다리영역은 일반적으로는 폭이 좁아 독립적인 토지피복으로 나타날 수 있을 지에 대해 우려될 수 있으나, 동작대교는 다른 대교들에 비해 그 폭이 약 10 m 이상 넓어서 폭이 한 화소 이상으로 나타나므로 물의 반사특성과 잘 구별되어 존재한다. 보충실험을 위해 산림, 농경지, 도시역, 수역, 나대지 등의 5가지 토지

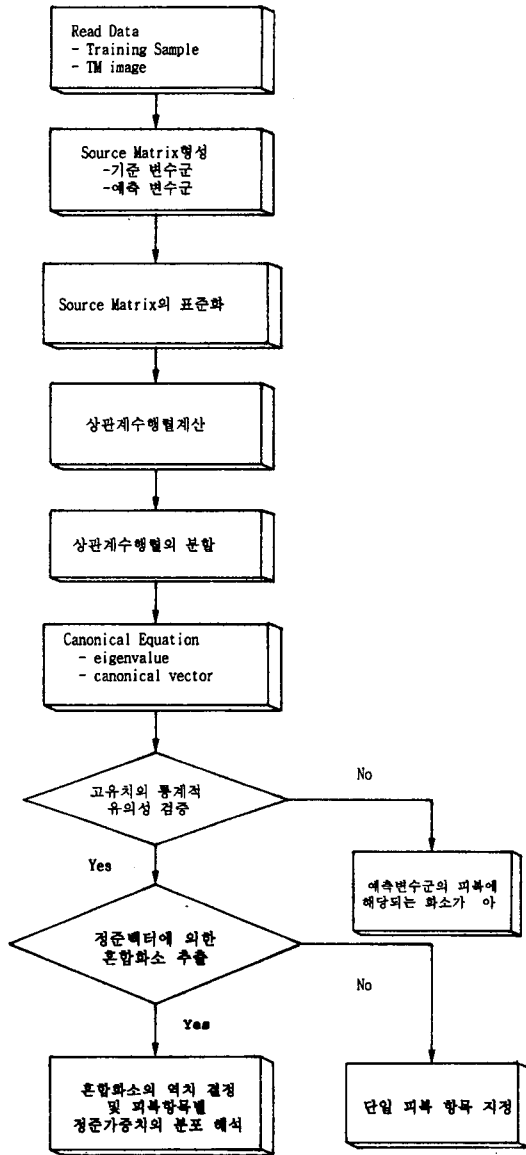


그림 1. 혼합화소해석 흐름도

피복이 함께 존재하는 영상에 대해서도 혼합화소분류를 실행하였다.

2가지 토지피복은 다음과 같은 성질을 가진 대상물이다.

① Bridge: 동작대교로서 연장 1330 m, 폭 40 m의 강교이며 전철교가 복합 시공되어 있어 다른 대교들 보다 폭이 좀 더 크다. 일반적인 경우는 도시역에 포함되는 대상물로 콘크리트와 아스팔트가 섞여 있는 도로가 갖는 반사특성을 가지고 있다.

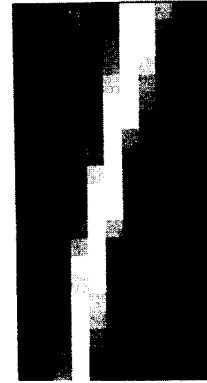


사진 1. 실험영역의 자연색 영상

② Water: 한강의 물을 지칭함.

인공위성데이터는 LANDSAT-5 TM(Thematic Mapper) (1996. 9. 1)을 사용하였다. 실험대상영역의 위성데이터를 Natural Color로 나타내면 사진 1과 같다.

3.2 適用 및 分析

3.2.1 Training areas 統計的 數值

분류할 토지피복의 각 training area의 평균값이 예측 변수군의 변수로 구성된다. 수역의 training areas는 직사각형 영역으로 선택되었다. 다리영역은 그 형태상 길게 존재하고 폭이 매우 작을 뿐만 아니라, 실험되는 영역 내에서 차지하는 화소수도 수십 개에 불과하므로 확실한 다리를 나타내는 부분을 직사각형으로 절출해내는 것은 불가능하였다. 그러므로 다리의 중심부에서 분명한 다리의 특성을 갖는 화소를 하나만 선택하여 training area로 하였다. 본 연구에서 선택된 training areas는 다리 영역의 경우 실험영역의 21행 4열에 위치한 화소이며, 수역은 실험영역비갈에서 확실한 물을 나타내는 부분 중 135개의 화소를 직사각형으로 절출하였다. Training areas의 통계적 수치는 표 1과 같다.

3.2.2 실험영역 위성데이터의 화소값 분포

정준상관분류에서 기준변수군의 변수는 위성영상의 개별화소이다. 실험영역이 총 231개의 화소로 이루어져 있으므로 총 231번의 정준상관분석이 실행된다. 즉 231개의 서로 다른 정준상관계수(혹은 고유치)와 정준벡터가 계산된다. 실험영역의 231개 화소의 수치는 표 2와 같다. 화소값의 분포를 보면 실험영역의 중심열인 6열의 데이터를 기준으로 바로 우측부에 동작대교일 것이라 짐작되는 화소들의 분포를 볼 수 있다. 따라서

표 1. Training areas의 統計數值

화소수=1							
Bridge							
Band	1	2	3	4	5	6	7
Minimum	109	48	53	42	51	145	30
Mean	109	48	53	42	51	145	30
Standard	-	-	-	-	-	-	-
Maximum	109	48	53	42	51	145	30
화소수=135							
Water							
Band	1	2	3	4	5	6	7
Minimum	100	42	43	27	15	137	5
Mean	103.32	44.08	44.56	27.80	16.92	138.61	7.47
Standard	0.92	0.68	0.58	0.53	0.98	0.95	0.97
Maximum	105	46	46	29	19	142	10

다리영역과 수역의 경계부분에 위치한 화소들이 혼합 화소일 가능성이 예측된다.

3.2.3 지상실제데이터 작성

표 2에서 진한글씨는 각 화소의 영역이 거의 90% 이상 Bridge일 것으로 예상되는 화소들이며, 외곽선으로 처리된 화소들은 약 50~90% 사이에서 수역과 다리영역이 혼합되어 있는 화소들이다. 이는 7개 밴드들의 화소값의 분포 패턴을 기준으로 하여 구분한 것으로서 정확도 평가시에 지상실제 데이터(ground truth data)로 활용할 수 있을 것이다. 즉 화소의 분광적 패턴을 시각적으로 인식하여 각 분류항목을 판별한 것으로서,¹¹⁾ 판별에 대한 정확도는 100% 정확할 수는 없음을 밝힌다. 예를 들면 실험영역의 3, 4행과 7, 8행에서 다리영역을 표현하는 화소가 각 행당 연속 2개 이상씩 존재하는 것으로 판별되었다. 그러나 화소 1개의 해상도가 28.5 m 이므로 화소 2개로 표현되는 동작대교의 폭은 57 m이고, 실제 동작대교의 폭이 약 40 m이므로 엄밀하게는 화소 2개가 전부 다리영역을 나타내서는 안 된다. 이것은 연구에서 사용된 데이터가 기하보정에서의 resampling 과정도 수행되지 않았으며, 다리가 영상데이터 내에서 수직형태로 위치되지 않고 경사지게 위치한다는 것을 감안한다고 해도 명확히 설명되어질 수 없다. 아마도 Bulk 보정의 효과로 인해, 동작대교의 폭이 실제보다 약간 큰 영상데이터로 취득된다고 이해하는 것이 가능한 설명이 될 수 있지 않을까 생각된다. 그러므로 동작대교는 약 60 m의 폭을 갖는 것으로 나타나며, 실제 폭보다는 약 1.5배 크게 표현된다고 판단할 수 있다. 일반적으로 위성영상에 대한 지상실제데이터를 취득

하는데 있어서는 현지조사 등 많은 어려움이 따른다. 특히 지형도와 항공사진을 통해 얻어지는 지상실제데이터의 부정확성을 고려해 볼 때, 화소의 분광패턴을 시각적으로 인식하여 판별하는 방법이 오히려 더 정확한 지상실제데이터를 얻을 수 있을 것으로 판단된다. 다만 본 연구에서는 토지피복상태의 특성상 동작대교의 폭을 대략 60 m 정도로 간주하는 것을 바탕으로 하여 실험을 수행한다. 지상실제데이터를 작성하는 화소의 시각적 분광패턴의 기준은 다음과 같다.

실험 영역의 토지피복이 다리와 수역으로 구성되어 있으므로 물의 밴드별 분광특성을 판별 기준으로 삼는다. 물을 다른 토지피복들과 잘 구분해 줄 수 있는 밴드는 TM의 밴드별 특성에서 알 수 있는 바와 같이 4, 5, 7밴드이다.¹²⁾ 이 중에서 5밴드의 수치패턴을 1차적인 판별기준으로 삼고, 2차적으로는 5밴드와 7밴드의 수치패턴을 동시에 고려하여 판별기준으로 삼는다. 즉 전체 7가지 밴드의 수치패턴에서 볼 때, 5밴드의 수치가 50 이상이면 거의 분명히 다리영역이며, 50이 안되는 경우라도 5밴드의 수치가 45 이상이며 7밴드가 27 이상이면 다리영역으로 보는 것이 타당하다고 판단된다. 수역은 5밴드의 수치가 25 이하인 경우로 간단히 결정된다. 단 5밴드의 수치가 25보다 크더라도 7밴드의 수치가 12이하이면 수역으로 보는 것이 타당한 것으로 판단된다. 이상과 같은 기준에 들지 않는 화소들은 명확히 다리도 수역도 아니므로 혼합화소일 것으로 예측한다. 이상과 같은 기준으로 지상실제데이터를 작성하면 사진 2와 같다. 토지피복별로 보면 다리영역 30개, 혼합화소 30개, 수역 171개로 구성된다.

표 2. 실험영역 위성영상데이터의 화소값

행	열	밴드	1	2	3	4	5	6	7	8	9	10	11	1	2	3	4	5	6	7	8	9	10	11
1	1	107	107	106	106	105	106	111	112	107	105	106	11	104	103	103	106	108	110	105	102	103	104	104
	2	45	45	45	46	45	45	48	49	46	45	45	44	44	44	45	48	48	45	44	44	44	44	
	3	47	46	46	47	46	48	54	56	50	46	46	45	44	43	45	53	54	47	44	45	44	44	
	4	29	29	29	29	28	32	40	41	33	28	29	27	27	28	31	40	40	31	27	27	28	28	
	5	20	21	22	21	20	24	42	55	38	22	20	16	17	17	22	42	53	34	16	17	17	17	
	6	138	139	140	141	143	146	147	147	147	147	146	141	142	143	144	145	145	145	143	142	140	140	
	7	9	10	10	10	10	13	26	31	19	10	10	8	8	9	12	25	28	15	7	8	8	6	
2	1	106	105	106	107	108	106	111	111	106	103	104	12	104	102	102	106	110	109	105	102	102	103	103
	2	45	44	45	46	44	45	49	49	45	45	45	44	45	44	45	48	48	44	42	43	44	44	
	3	47	46	47	47	46	49	55	54	48	45	45	46	45	44	47	55	53	46	43	43	44	44	
	4	28	29	29	30	28	33	42	39	31	29	28	27	28	27	33	41	37	29	27	28	29	28	
	5	19	20	21	22	19	23	46	53	30	20	20	16	18	16	23	46	51	30	17	26	17	17	
	6	138	138	140	141	143	144	144	145	144	143	143	141	143	144	145	145	145	144	143	141	140	140	
	7	8	8	10	10	9	13	29	29	15	11	10	8	8	8	18	29	25	13	8	8	7	7	
3	1	105	103	107	103	105	106	111	110	105	102	101	13	104	103	101	108	111	107	101	101	104	104	103
	2	45	44	44	45	45	46	50	48	44	43	45	43	44	44	46	49	47	44	43	44	44	43	
	3	46	46	46	45	44	49	56	53	47	45	44	45	45	44	48	56	51	45	43	44	45	45	
	4	29	29	29	29	28	34	43	38	29	28	28	27	28	27	35	42	35	28	27	28	28	27	
	5	19	20	20	21	17	24	52	53	24	18	18	16	16	14	27	50	47	23	17	16	17	16	
	6	139	140	141	142	144	144	144	144	143	142	141	142	143	144	145	146	145	144	143	142	140	140	
	7	9	8	10	10	7	14	32	28	12	9	8	8	8	8	19	33	22	10	7	8	7	7	
4	1	104	104	105	102	103	108	112	109	105	103	102	14	104	103	102	109	110	105	103	104	104	104	103
	2	44	44	44	45	44	46	49	47	43	43	44	44	45	43	47	49	45	43	43	44	45	43	
	3	45	45	45	45	45	50	56	50	44	45	44	44	44	44	51	57	49	44	43	44	45	45	
	4	29	29	29	28	28	36	43	36	27	28	28	27	27	27	38	44	33	27	28	27	27	27	
	5	20	20	19	20	17	28	53	48	20	18	18	17	17	15	31	55	44	18	18	17	16	16	
	6	140	141	143	144	145	145	145	144	143	141	141	142	143	145	146	146	145	144	143	141	140	139	
	7	9	9	9	10	8	17	33	27	11	10	9	8	8	8	18	35	23	7	8	6	8	7	
5	1	104	104	105	105	103	109	112	107	104	104	105	15	103	103	102	109	110	105	104	103	103	101	103
	2	44	44	44	45	45	47	49	46	44	44	44	43	44	45	48	49	45	44	43	43	44	44	
	3	45	45	46	46	45	51	55	50	45	45	45	45	45	45	51	55	48	44	44	44	44	44	
	4	28	29	29	28	29	38	42	34	28	29	29	28	28	30	39	42	31	26	28	27	27	27	
	5	19	19	20	19	19	35	54	43	21	21	20	16	17	19	37	53	37	15	15	17	17	16	
	6	140	141	143	144	145	145	145	144	142	141	140	142	143	145	146	146	145	144	143	141	140	139	
	7	8	9	9	10	11	21	32	23	11	11	10	9	9	11	22	33	20	9	8	5	8	7	
6	1	103	103	105	105	105	110	111	106	104	104	104	16	103	104	106	110	109	104	104	107	104	101	103
	2	44	44	45	44	45	48	48	46	44	44	44	43	43	44	47	48	44	44	44	44	44	44	
	3	45	45	46	46	47	53	54	49	45	45	45	45	45	46	52	54	48	45	44	44	45	45	
	4	28	29	29	27	31	41	41	32	28	29	28	28	29	32	40	41	31	28	28	28	28	28	
	5	18	18	19	17	22	43	53	36	20	19	19	17	18	22	40	51	34	16	15	16	17	16	
	6	140	141	143	144	145	145	145	143	142	141	140	143	144	145	146	146	145	143	142	141	140	139	
	7	7	8	9	9	12	24	31	20	11	10	9	8	9	12	22	30	18	9	8	7	8	7	
7	1	103	103	105	105	107	110	110	105	102	102	102	17	104	103	105	110	108	103	103	106	104	102	104
	2	44	44	44	44	45	48	47	44	44	43	44	44	43	45	48	47	44	43	44	44	44	44	
	3	45	45	44	43	47	54	53	46	44	44	45	45	44	47	53	52	47	44	44	44	45	45	
	4	28	28	28	27	33	42	39	30	28	28	27	28	29	32	41	41	32	28	28	28	28	27	
	5	18	17	18	15	23	48	51	28	19	17	17	17	19	23	42	52	33	19	17	16	17	16	
	6	140	141	143	144	145	145	144	143	141	140	139	143	145	146	146	145	144	143	142	141	140	139	
	7	8	8	8	8	12	28	29	18	10	9	9	7	8	11	23	30	19	10	9	8	8	7	
8	1	104	105	106	103	107	110	107	104	101	104	105	18	104	102	105	109	106	104	103	103	102	103	104
	2	44	44	45	43	45	49	46	43	43	44	44	44	43	45	48	47	45	45	44	44	44	44	
	3	45	45	46	44	48	55	51	45	44	44	44	44	45	48	54	52	46	45	46	45	45	44	
	4	28	28	28	26	35	43	36	28	28	27	27	27	29	33	42	38	30	28	28	28	28	28	
	5	19	17	18	15	26	54	51	23	17	15	17	16	18	25	47	51	28	19	18	17	17	17	
	6	140	142	143	144	145	145	144	142	141	140	139	144	146	146	146	145	144	142	141	140	140	139	
	7	9	8	7	7	14	33	26	13	6	7	7	6	7	11	27	31	18	9	9	8	8	8	
9	1	105	105	104	102	107	112	106	103	106	104	103	19	104	103	107	110	105	103	106	102	102	103	104
	2	44	44	44	42	46	50	45	43	43	44	44	44	44	45	48	46	44	45	45	44	43	42	
	3	45	45	44	43	49	57	50	44	45	44	44	44	45	49	56	51	46	45	44	44	44	44	
	4	28	27	28	27	37	45	35	27	28	27	27	27	28	34	45	35	28	28	27	28	28	28	
	5	18	15	15	15	30	58	47	16	17	14	18	16	17	28	53	51	21	18	19	16	17	16	
	6	140	141	143	144	145	145	144	142	141	140	139	144	146	146	146	145	143	141	141	140	140	140	
	7	8	7	7	7	22	35	21	9	6	8	7	6	8	11	33	31	12	7	9	7	7	10	
10	1	105	105	105	104	110	112	106	102	104	104	103	20	103	104	106	110	105	103	106	104	101	102	103
	2	44	44	45	44	47	49	45	43	43	44	44	44	43	46	48	45	44	44	45	45	43	42	
	3	45	44	45	46	52	56	48	44	44	45	44	44	46	50	56	49	45	45	44	44	44	45	
	4	27	28	28	29	38	42	33	27	28	27	27	27	29	36	44	33	28	29	27	28	28	28	
	5	17	17	16	20	37	56	40	18	17	16	19	16	18	31	53	47	16	19	17	15	16	16	
	6	141	142	143	144	145	145	144																

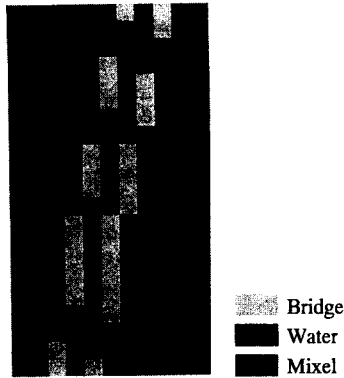


사진 2. 지상실제데이터 영상

3.2.4 정준벡터 생성

혼합화소를 구분하지 않고 정준가중치가 가장 큰 쪽을 토지피복으로 결정하는 일반적인 정준상관분류를 수행한 결과는 사진 3과 같이 나타난다. 실험대상영역은 유의수준 5%(고유치; $\lambda=0.7763$)에서 미분류 화소가 추출되지 않았다. 이것은 실험영역이 물과 다리영역으로만 구성되어 있으므로 예측되는 당연한 결과이다.

정준상관분류를 수행하여 계산된 정준벡터는 표 3과 같다. 표 3에서는 실험영역인 21행 11열 전체 화소에 대하여 정준벡터를 나타내었다. 각 화소의 정준벡터는 2개씩의 정준가중치를 벡터요소로 포함하고 있으며, 여기서 첫 번째 값은 다리를 나타내며 두 번째 값은 수역을 나타낸다. 표 3에서도 진한글씨의 수치는 다리영역을 나타내며, 외곽선으로 처리된 것은 혼합화소이고 나머지는 수역을 나타낸다. 이러한 구분은 표 2에서와는 달리 전적으로 정준벡터 요소에 의존하여, 지상실제



사진 3. 혼합화소를 추출하지 않을 때의 정준상관분류 결과

데이터와 가장 가까운 형태를 이루는 순간의 정준가중치 사이의 비율을 기준으로 이루어진 것이다.

3.2.5 혼합화소 해석

혼합화소를 분류해 내기 위한 역치를 결정하는 과정은 다음과 같이 정리된다.

① 대상영역 TM영상의 분광수치에 대해 시각적 패턴인식을 이용하여 지상실제데이터를 작성한다. 각 토지피복은 물론이며, 혼합된 피복상태도 구별하여 작성한다.

② 본 연구에서는 2가지 토지피복만 혼합되어 있는 화소의 분석에만 한정한다. 따라서 예측변수군에는 분류하고자 하는 혼합화소의 2가지 토지피복 각각의 training area의 밴드별 평균수치만 기록된다. 이러한 방법의 타당성을 확인하기 위해 예측변수군에 5가지 토지피복의 데이터를 사용하여 실험영역에 적용하여 보았다. 산출된 결과는 2가지를 사용한 것에 비해 해당 토지피복수치외의 나머지 3개의 토지피복수치의 영향을 받게 되어 혼동되는 경향이 나타났으며, 결과적으로 분류결과가 정확하지 못하였다.

③ 각 화소별로 정준상관분류 과정에서 산출되는 정준가중치 중에 큰 값을 작은 값으로 나눈 비(ratio)를 구한다. 단 작은 값이 음인 경우는 절대값으로 하여 계산한다. 정준벡터의 요소를 보면 정준가중치의 총합은 항상 1을 유지하는 것을 알 수 있다. 따라서 이 때의 음의 값은 항상 거의 0에 가까운 값이므로 혼합화소가 아닌 100% 독립적인 피복으로 이루어진 화소임을 알 수 있다.

④ ③에서 구한 비가 임의의 역치(즉 1.1, 1.2, 1.3, ..., 4.0, 4.1, 4.2, ... 등)보다 작을 때 그 화소를 혼합화소로 결정한다.

⑤ ①과 ④에서 얻어진 결과를 비교하여 거의 일치하는 경우를 정준가중치 비의 역치로 결정하고, 그 때의 화소들을 혼합화소로 결정한다.

위의 과정에서 정준가중치 사이의 비율에 따른 분류 결과는 표 4와 같이 정리된다. 지상실제데이터와 비교할 때, 가장 비슷한 형태는 역치가 3.7 또는 4.0으로 결정될 때일 것으로 예상된다. 명확한 결정을 위해 정확도 평가를 수행한 결과는 표 5와 같이 정리된다. 표 5는 지상실제데이터와 역치별로 분류된 결과를 비교하기 위해, 지상실제데이터의 각 화소수치에 10을 곱한 후 분류된 결과수치를 감한 값들의 화소개수를 나타낸 것

표 3. 정준상관분류에 의한 각 화소의 정준벡터

행 번호	1	2	3	4	5	6	7	8	9	10	11
1	0.0642	-0.0001	0.0465	-0.0146	0.0233	0.1423	0.6234	1.0847	0.5982	0.1284	0.0461
	1.0630	1.0001	-0.9543	1.0143	0.9771	0.8598	0.3808	- 0.0864	0.4060	0.8736	0.9547
2	-0.0923	-0.0241	-0.0013	0.0087	0.0021	0.0677	0.7209	1.0266	0.3403	0.0428	0.0410
	1.0905	1.0237	1.0013	0.9915	0.9979	0.9335	0.2827	- 0.0271	0.6636	0.9579	0.9597
3	-0.0648	0.0038	0.0138	0.0593	-0.0458	0.0751	0.9326	1.0650	0.1670	0.0146	-0.0322
	1.0636	0.9962	0.9864	0.9416	1.0450	0.9261	0.0685	- 0.0662	0.8354	0.9857	1.0316
4	0.0212	0.0321	0.0188	0.0575	-0.0161	0.2018	1.0066	0.9539	0.1120	0.0059	-0.0056
	0.9791	0.9684	0.9816	0.9435	1.0158	0.8010	- 0.0067	0.0468	0.8897	0.9942	1.0055
5	-0.0099	-0.0002	0.0359	0.0023	0.0321	0.4125	1.0568	0.7814	0.0922	0.0753	0.0235
	1.0097	1.0002	0.9647	0.9978	0.9685	0.5918	- 0.0578	0.2216	0.9093	0.9259	0.9769
6	-0.0443	-0.0347	-0.0198	-0.0248	0.0858	0.6444	1.0618	0.5287	0.0596	0.0049	-0.0049
	1.0435	1.0340	1.0195	1.0243	0.9156	0.3596	- 0.0630	0.4757	0.9414	0.9952	1.0048
7	-0.0394	-0.0602	0.0030	-0.0589	0.0997	0.8440	1.0185	0.3469	0.0320	-0.0248	-0.0691
	1.0387	1.0591	0.9970	1.0579	0.9019	0.1583	- 0.0189	0.6571	0.9686	1.0243	1.0678
8	-0.0049	-0.0541	-0.0572	-0.0435	0.1819	1.0764	1.0818	0.1934	-0.0258	-0.1207	-0.0713
	1.0048	1.0531	1.0562	1.0427	0.8207	- 0.0779	- 0.0834	0.8093	1.0253	1.1183	1.0700
9	-0.0447	-0.1262	-0.0930	-0.0093	0.3144	1.1730	0.9381	-0.0308	-0.0535	-0.1476	-0.0341
	1.0438	1.1237	1.0912	1.0092	0.6894	- 1.1767	0.0629	1.0302	1.0525	1.1446	1.0335
10	-0.0588	-0.0394	-0.1115	0.0494	0.4989	1.0995	0.6837	0.0311	-0.0288	-0.0994	0.0041
	1.0577	1.0386	1.1094	0.9514	0.5055	- 0.1015	0.3201	0.9694	1.0283	1.0975	0.9960
11	-0.0879	-0.0278	-0.0036	0.1083	0.6318	1.0451	0.5006	-0.0506	-0.0426	-0.0591	-0.0687
	1.0863	1.0273	1.0035	0.8934	0.3723	- 0.0459	0.5038	1.0497	1.0418	1.0580	1.0674
12	-0.1028	-0.0275	-0.0348	0.1358	0.7659	0.9780	0.3975	0.0479	-0.0357	-0.0679	-0.0614
	1.1008	1.0270	1.0342	0.8663	0.2373	0.0224	0.6067	0.9529	1.0350	1.0666	1.0603
13	-0.0529	-0.0694	-0.0948	0.2165	0.9074	0.8811	0.1822	0.0220	-0.0683	-0.0788	-0.0776
	1.0519	1.0681	1.0930	0.7865	0.0941	0.1208	0.8204	0.9784	1.0670	1.0773	1.0761
14	-0.0305	-0.0406	-0.0316	0.2636	1.1042	0.8820	0.0439	0.0441	-0.0511	-0.1235	-0.0890
	1.0299	1.0399	1.0311	0.7398	- 0.1062	0.1198	0.9569	0.9567	1.0502	1.1210	1.0873
15	-0.0519	-0.0330	0.0287	0.4680	1.0470	0.6310	-0.0600	-0.0623	-0.0296	-0.0449	-0.0982
	1.0510	1.0324	0.9718	0.5363	- 0.0479	0.3731	1.0589	1.0612	1.0291	1.0441	1.0964
16	-0.0141	0.0243	0.1154	0.5745	0.9946	0.5405	-0.0671	-0.1067	-0.0843	-0.0665	-0.1198
	1.0139	0.9762	0.8864	0.4298	0.0055	0.4639	1.0659	1.1046	1.0827	1.0653	1.1175
17	-0.0453	0.0787	0.1170	0.6079	1.0922	0.5188	0.0745	-0.0371	-0.0796	-0.0690	-0.1156
	1.0444	0.9225	0.8849	0.3963	- 0.0940	0.4856	0.9267	1.0364	1.0781	1.0678	1.1133
18	-0.0494	0.0409	0.1610	0.8103	1.0917	0.3355	-0.0042	-0.0383	-0.0690	-0.0716	-0.0705
	1.0485	0.9598	0.8414	0.1924	- 0.0934	0.6684	1.0042	1.0376	1.0678	1.0702	1.0692
19	-0.0494	-0.0054	0.2315	1.0428	1.1537	0.0962	-0.0623	0.0089	-0.0908	-0.0373	-0.0331
	1.0485	1.0053	0.7716	- 0.0436	- 0.1569	0.9053	1.0611	0.9913	1.0891	1.0366	1.0325
20	-0.0361	0.0308	0.3117	1.0380	1.0235	-0.0646	-0.0156	-0.0749	-0.1328	-0.0506	-0.0503
	1.0354	0.9697	0.6920	- 0.0387	- 0.0239	1.0634	1.0153	1.0735	1.1301	1.0496	1.0494
21	-0.0121	0.0605	0.4786	1.0000	0.6845	-0.1051	-0.0627	-0.0991	-0.0819	-0.0159	-0.0274
	1.0119	0.9405	0.5258	0.0000	0.3193	1.1031	1.0616	1.0972	1.0803	1.0156	1.0269

이다.¹⁾ 즉 파일 A*10-파일 B의 행렬연산이 이루어진다. 피복별 수치는 Bridge=1, Water=2, Mixel=3으로 정의되어 있다. 파일간 연산에서 산출될 수 있는 수치는 7, 8, 9, 17, 18, 19, 27, 28, 29의 9가지이다. 9, 18, 27은 빠르게 분류된 것을 나타내며, 7은 다리영역이 혼합화

소로 결정된 것이고, 29는 혼합화소로 생각됐던 화소가 다리로 분류된 것을 표현한다. 나머지 수치도 같은 방식으로 생각하면 된다. 수치의 의미를 쉽게 알 수 있도록 하기 위해 7을 BM, 29를 MB로 표시하기로 한다. 예를 들면 BM에서 앞의 B가 지상실체데이터의 다리

표 4. 역치에 따른 분류결과

(단위 : 화소개수)

역치																											
	2.5	2.6	2.7	2.8	2.9	3.0	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	4.0	4.1	4.2	4.3	4.4	4.5	4.6	4.7	4.8	4.9	5.0	
Bridge	34	33	33	32	32	31	31	31	31	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Mixel	23	24	25	26	27	28	29	30	31	32	34	35															
Water	174	174	173	173	172	172	171	170	169	169	167	166															

표 5. 지상실제데이터와 역치별 분류결과와의 비교

(단위 : 화소개수)

역치	2.5	3.0	3.5	3.6	3.7	4.0	4.2	4.5	5.0
7(BM)	0	1	2	2	2	2	2	3	3
8(BW)	0	0	0	0	0	0	0	0	0
9(BB)	30	29	28	28	28	28	28	27	27
17(WM)	0	0	1	1	1	1	2	2	4
18(WW)	171	171	170	170	170	170	169	169	167
19(WB)	0	0	0	0	0	0	0	0	0
27(MM)	23	24	24	25	26	27	27	27	27
28(MW)	3	2	2	2	1	0	0	0	0
29(MB)	4	4	4	3	3	3	3	3	3

를, 뒤의 M이 분류된 결과로서의 혼합화소를 의미한다. 이 수치들에 의해 분류에 대한 정확도와 오분류의 상황을 판단할 수 있다.

혼합화소를 가능한 한 정확히 추출할 때의 역치는 대략 4.0으로 결정된다. 지상실제데이터와 가장 가까운 형태인 역치 4.0에서의 분류결과를 대표적으로 사진 4에 나타내었다. 이 결정은 다리영역, 수역, 혼합화소를 평균적으로 가장 정확히 추출할 때를 기준으로 한다. 표 5에서 보면 혼합화소 30개 중에 27개가 정확히 분류

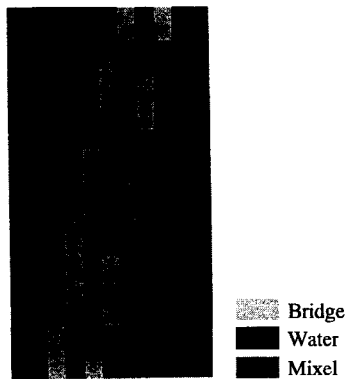


사진 4. 혼합화소 추출결과(역치 4.0)

되고, 다리가 28개, 수역이 170개일 때 가장 분류가 정확히 된 상태이며, 오분류된 화소의 개수도 6개로 가장 작은 상태이므로 가장 적절한 역치라고 판단된다. 전체 정확도가 가장 좋을 때의 혼합화소 구분의 정확도는 90%이다. 역치가 상승할 때 다리영역과 수역은 정확도가 감소하며, 혼합화소는 반대로 증가함을 알 수 있다. 이 두 가지 상황이 각각 급격히 변화하는 점이 역치 4.0이다. 4.0보다 훨씬 크게 되면 결국 혼합화소만의 정확도를 100%가 되도록 할 수 있으나 이 때는 다리와 수역의 정확도가 심하게 떨어지므로 의미가 없다. 역치 4.0은 정준벡터의 요소가 0.8과 0.2로 구성되어 있을 때이다. 이 값은 각 피복의 비중이 80%와 20%를 의미하는 것은 아니며, 본 연구에서는 대략 90%와 10%일 때의 상태를 표현하는 값으로 추정된다.

이상의 과정에 의해 결정된 역치는 일부 대상영역에 대한 실험값이므로 모든 위성영상에 절대적으로 적용된다고 주장하기는 어려우나, 이와 같은 실험을 다른 지역에서도 수행해 본 결과 거의 비슷한 역치가 채택되었으므로 일반적으로 대부분의 위성영상분류에 적용될 수 있으리라 판단된다. 다만 이에 대한 전제조건은 2가지 피복 상태만 혼합된 경우로 한정된다. 예측변수군의 개수가 달라지면 정준벡터의 요소수도 마찬가지로 개수이며, 이럴 경우 각 피복별로 정준가중치의 비율 기준이 달라지므로 역치도 크게 변화한다.

4. 結 論

본 연구에서는 정준상관분류과정에서 산출되는 정준벡터를 분석하여 혼합화소를 해석하는 방법을 제시하였고, 그 효용성을 검토하였다. 結論적으로 本研究의 遂行結果에서 나타나는 特徴 및 意味를 整理하면 다음과 같다.

첫째, 각 토지피복항목의 正準加重值 中에서 가장 큰 값과 두 번째 큰 값과의 比에 의해 混合畫素 地域을 抽出할 수 있으며, 실험대상영역에 대해 최적의 혼합화소를 추출할 때의 比를 역치로 결정할 수 있다. 본 연구에서와 같이 9월 1일을 전후한 TM 데이터에서 다리와 물 항목의 2가지 토지피복만 혼합된 경우에는 대략 4.0이 적합한 것으로 생각된다.

둘째, 정준벡터요소의 총합은 대략 1에 근사하므로 피복별 정준가중치를 통해 혼합화소의 대략적인 분포 비율을 추정할 수 있을 것이다.

셋째, 本 研究에서 試圖된 혼합화소 해석방법은 지상 실제데이터 자체의 오차 때문에 100%까지 신뢰할 수는 없지만, 실험영역에 대해서 혼합화소 추출의 정확도가 90%인 것은 절대적으로 높은 것이며 遠隔探査의 영상분류기법 분야에서 유용하게 사용될 수도 있을 것으로 생각된다.

參考文獻

1. 朴珉虎, "遠隔探査 Data의 正確度 向上을 위한 Bi-temporal Classification 技法의 適用", 서울大學校 大學院 工學碩士學位論文, 1988, 28-29.
2. 朴珉虎 外 2人, "An Analysis of Mixed Pixel in the Remote Sensing Image Data", 韓國 地形空間情報學會誌, 第3卷 第2號, 1995. 12, 91-100.
3. 朴珉虎, "正準相關分析을 이용한 映像分類技法의 開發", 서울大學校 大學院 工學博士學位論文, 1996.
4. 朴珉虎, "군집분석을 활용한 정준상관분류의 LANDSAT TM영상 분류정확도 향상에 관한 연구", 大韓土木學會論文集, 第18卷 第III-3號, 1998. 5, 301-315.
5. 李喜演, "地理 統計學", 法文社, 1991, pp. 485-518.
6. M. S. Bartlett, "The Statistical Significance of Canonical Correlations", *Biometrika*, **32**, 1941, 29-37.
7. Fangju Wang, "Improving Remote Sensing Image Analysis through Fuzzy Information Representation", *PE & RS*, **56**(8), 1990, 1163-1169.
8. Mark A. Karaska, Robert L. Huguenin, Donald Van Blaricom and Basil Savitsky, "Subpixel Classification of Cypress and Tupelo Trees in TM Imagery", *ACSM/ASPRS*, 1995, 856-865.
9. Patricia G. Foschi, "Estimating Land Cover Proportions Within Mixed Pixels Using a Moving-Window Algorithm", *ACSM/ASPRS*, 1993, 102-107.
10. Richard A. Johnson and Dean W. Wichern, "Applied Multivariate Statistical Analysis", Third Edition, Prentice Hall, 1992, 459-466, 484.
11. Thomas M. Lillesand and Ralph W. Kiefer, "Remote Sensing and Image Interpretation", First Edition, John Wiley & Sons, 1979, 457-461.
12. Thomas M. Lillesand and Ralph W. Kiefer, "Remote Sensing and Image Interpretation", Third Edition, John Wiley & Sons, 1994, 468.
13. R. S. Chhikara, "Error analysis of crop acreage estimation using satellite data", unpublished manuscript, University of Houston, 1985.
14. G. M. Foody and D. P. Cox, "Estimation of sub-pixel land cover composition from spectral mixture models", spatial data 2000, proceedings of a joint conference of the photogrammetric society, Remote Sensing society, and American society for photogrammetry and Remote Sensing, Oxford, 1991, 186-195.
15. H. M. Horwitz, R. F. Nalepka, P. D. Hyde and J. P. Morgenstern, "Estimating the proportions of objects within a single resolution element of a multispectral scanner", proceedings of 7th international Symposium on Remote Sensing of Environment, 1971, 1307-1320.
16. R. D. Jackson, R. J. Reginto, P. J. Pinter, Jr. and S. B. Idso, "Plant canopy information extraction from composite scene reflectance of row crops", *Applied Optics*, **18**, 1979, 3775-3782.
17. S. E. Marsh, P. Switzer, W. S. Kowalik and R. J. P. Lyon, "Resolving the percentage of component terrains within single resolution elements", *Photogrammetric Engineering and Remote Sensing*, **46**, 1980, 1079-1086.
18. A. J. Richardson, C. L. Wiegand, H. W. Gausman, J. A. Cuellar and A. H. Gerbermann, "Plant, soil, and shadow reflectance components of row crops", *Photogrammetric Engineering and Remote Sensing*, **41**, 1975, 1401-1407.
19. E. A. Work and D. S. Gilmer, "Utilization of satellite data for inventorying prairie ponds and lakes", *Photogrammetric Engineering and Remote Sensing*, **42**, 1976, 685-694.