

항공 하이퍼스펙트럴 영상의 PCA기법 적용을 통한 토지 피복 분류 정확도 개선 방안에 관한 연구

A Study on the Improvement classification accuracy of Land Cover using the Aerial hyperspectral image with PCA

최병길* · 나영우** · 김승현*** · 이정일****

Choi, Byoung Gil · Na, Young Woo · Kim, Seung Hyun · Lee, Jung Il

要 旨

본 연구에서는 항공 하이퍼스펙트럴 영상에 대해 PCA를 적용하여 토지 이용 및 피복 분류 관독의 가독성을 향상시키기 위하여 고유성분이 높은 밴드를 선별적으로 조합하여 5개 유형의 PCA영상을 제작하였다. 유형별 영상은 SAM감독 분류 기법을 적용하여 영상분류를 시행하고 정확도를 평가한 결과 PCA변환 시 고유성분 포함율은 PCA변환 영상의 첫 번째 밴드에 해당하는 영상이 76.74%의 성분을 포함하며, PCA변환 영상의 두 번째 누적 밴드에 해당하는 영상이 98.40%로 대부분의 성분자료가 두 번째 영상까지에 담긴 것을 알 수 있었다. 유형별 영상의 정량적 분류정확도 평가는 전체정확도, 생산자 및 사용자 정확도를 분석한 결과 유사한 패턴을 가지며, 특이한 사항은 정성적인 분류정확도 평가는 PCA변환 영상의 네 번째 밴드 이상이 포함되어야 정확도가 확보되는 것으로 판단되나 정량적인 분류 정확도 평가에서는 PCA변환 영상의 두 번째 밴드까지를 포함하는 영상이 가장 높은 정확도를 나타내는 것을 알 수 있었다.

핵심용어 : 하이퍼스펙트럴, 센서, PCA, 감독분류, 밴드

Abstract

The researcher of this study applied PCA on aerial hyper-spectral sensor and selectively combined bands which contain high amount of information, creating five types of PCA images. By applying Spectral Angle Mapping-supervised classification technique on each type of image, classification process was carried out and accuracy was evaluated. The test result showed that the amount of information contained in the first band of PCA-transformation image was 76.74% and the second accumulated band contained 98.40%, suggesting that most of information were contained in the first and the second PCA components. Quantitative classification accuracy evaluation of each type of image showed that total accuracy, producer's accuracy and user's accuracy had similar patterns. What drew the researcher's attention was the fact that the first and the second bands of the PCA-transformation image had the highest accuracy according to the classification accuracy although it was believed that more than four bands of PCA-transformation image should be contained in order to secure accuracy when doing the qualitative classification accuracy.

Keywords : hyperspectral, sensor, PCA, Supervised classification, Band

data received: 2014.02.13, data revised: 2014.03.06, data accepted: 2014.03.18

* 정회원 · 인천대학교 도시과학대학 건설환경공학과 교수(Member, Professor, Dept. of Civil and Environmental Engineering, Incheon National University, bgchoi@incheon.ac.kr)

*** 교신저자 · 정회원 · 인천대학교 산학협력중점교수(Corresponding author, Member, Professor of Hub-Industry-Academic Cooperation, Incheon National University, survey@incheon.ac.kr)

*** 선영엔지니어링 기술영업부 부장(Section chief, Sunyoung ENG, seunghyun@nate.com)

**** 정회원 · 인천대학교 일반대학원 건설환경공학과 박사과정(Member, Doctoral Student Dept. of Civil and Environmental Engineering, Incheon National University, iji2000@hanmail.net)

1. 서론

하이퍼스펙트럴 영상(Hyperspectral image)은 수십 개에서 수백개의 연속된 분광밴드로부터 동시에 분광 정보를 취득하여, 각 화소별로 분광특성 및 공간정보로 분류가 가능한 영상이다(Park, 2006). 현재 다양한 연구기관 및 학교, 기업 등에서는 이러한 특성을 이용하여 식생, 연안, 해양환경 등에 대해 연구와 시범사업이 활발히 진행되고 있어 하이퍼스펙트럴 관련 기술이 더욱 각광받을 것으로 예상된다. 공중 하이퍼스펙트럴 영상은 대상지역이 대규모지역이거나 현장접근이 어려운 지역일수록 효율적인 자료취득방식으로 이용 가능하다. 근래까지는 광학인공위성의 저해상도 영상을 중심으로 지표의 관측과 해석에 활용해 왔으나 현재는 인공위성보다 수백 배 이상 해상도가 높은 항공기를 이용한 하이퍼스펙트럴 센서의 등장으로 상당부분 고해상도 영상으로 대체되고 있으며 산림 및 연안해역의 연구 및 사업을 중심으로 활용도가 높다. 따라서 향후 항공기를 이용한 고해상도 하이퍼스펙트럴 영상의 활용이 증대될 것으로 보인다.

PCA(Principal Component Analysis)는 주성분 분석으로 불리며, 국내에서는 PCA기법을 적용하여 하이퍼스펙트럴 영상융합 알고리즘을 2단계 분광혼합기법 기반하에 연구한 사례가 있으며(Choi et al., 2006), 다중분광자료의 중첩을 제거하거나 노이즈를 감소시키기 위하여 사용되는 기법이다. 본 연구의 목적은 하이퍼스펙트럴 영상에 대해 PCA기법을 적용하여 영상의 분류 정확도를 평가하고 개선의 정도를 정량적으로 확인하고자 한다.

2. 이론적 고찰

2.1 하이퍼스펙트럴 시스템

항공 하이퍼스펙트럴 카메라의 촬영시스템은 크게 카메라 본체, IMU, 제어시스템, 촬영 지원 모듈로 구성된다. 국외에서는 VNIR(가시광), SWIR(중적외), LWIR(열적외)파장대와 같은 제품을 사용하고 있다. 현재 국내의 운영 사례는 초분광카메라로는 VNIR파장대로서 필란드 SPECIM사의 AISA Eagle 제품과 캐나다 ITRES사의 CASI-1500이 사용되고 있으며, 이번 실험에는 SPECIM사의 AISA Eagle 성과를 이용하여 실험 및 분석에 사용하였다.

다수의 연속된 분광밴드를 통해 데이터의 특성을 분석하고 화소별로 완전한 분광정보 및 공간정보를 획득하여 지표의 식생과 정보를 얻는 하이퍼스펙트럴 원격

Hardware



Figure 1. System of Hyperspectral



Figure 2. Specifications of Aisa Eagle and inner System.

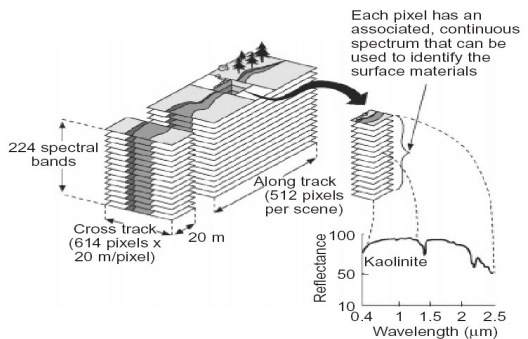


Figure 3. Hyperspectral Remote Sensing

탐사는 대규모 지역 혹은 접근불능지역에 효과적으로 적용할 수 있는 기술로 “Imaging Spectrometry”나 “Imaging Spectroscopy”라는 용어로 사용하기도 한다. 일반적으로 많은 수의 밴드(통상 40개 이상), 좁은 폭

(0.01~0.02μm), 연속적인 분광 밴드를 가지고 있으며 분광해상도가 매우 높은 것이 특징이다.

Fig. 3은 지상의 정보를 하이퍼스펙트럴 센서로 취득한 후, 단일 화소에 대한 각 파장대별 분광정보들을 나타낸 것이다.

2.2 SAM 분류 기법

SAM(Spectral Angle Mapping) 분류 기법은 분광 공간상에서의 거리 개념이 아닌 각도 개념을 적용하는 분류 방법으로 화소와 분류 항목간의 유사도를 판별하는데 있어 n차원에서의 각도를 이용한다. 입력 조건으로써 최대 유효 각도를 입력할 수 있어, 화소와 가장 작은 각도를 나타내는 트레이닝 자료까지의 각도가 이 각도보다 큰 경우 미분류 자료로 남는다.

SAM 분류기법은 분광값을 표현하는데 있어서 구하고자 하는 화소값이 정확하게 보정이 되었다는 전제하에 참조 데이터 벡터와 각도상으로 유사성을 찾고자 하는 방식이다. 밴드 수와 공간의 차원상 벡터가 같다는 것을 기본 영역에서 각각 트레이닝된 화소값들에 대한 그룹이 형성되고 이것이 최적의 벡터로 표시된다. 알지 못하는 하나의 화소를 분류 할 경우 앞서 생긴 클래스의 벡터와 가장 가까운 각(0에서 π/2까지의 라디안 범위)을 가지는 클래스로 Fig. 4와 같이 분류 한다.

분광차원 상의 각을 구하기 위해 벡터의 내적을 이용하면 다음 식(1)과 같다.

$$\alpha = \cos^{-1} \frac{\sum XY}{\sqrt{\sum(X)^2 \sum(Y)^2}} \quad (1)$$

여기서, α: 참조 스펙트럼과 영상 스펙트럼과의 각

X : 영상 스펙트럼

Y : 참조 스펙트럼

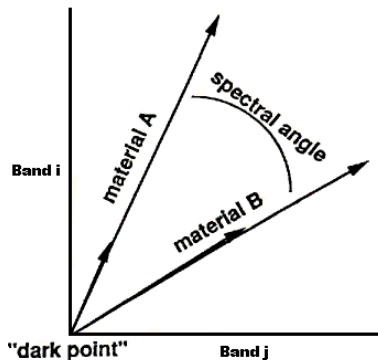


Figure 4. Spectral Angle Mapping

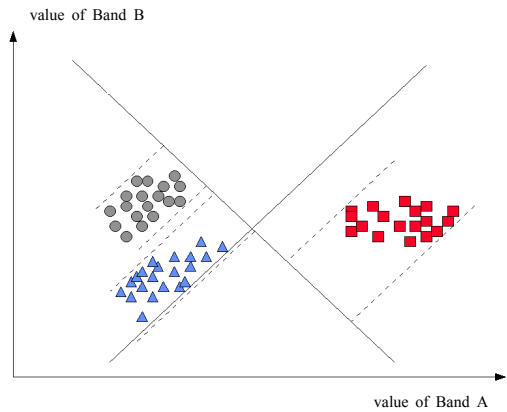


Figure 5. Conceptual diagram of PCA

2.3 PCA 분석

PCA는 주성분 분석으로 불리며, 다중 분광자료의 중복을 제거하거나 노이즈를 감소시키기 위하여 사용되는 기법이다. 변량 사이의 상관관계에 주목해서 최대한 정보를 상실하지 않고 많은 변량 측정값을 적은 개수의 종합지표로 집약하여 나타내는 방법이다.

n변량(n차원) 측정값을 m차원의 주성분으로 집약한다는 의미에서 차원을 줄이는 방법이라고도 말할 수 있다. 주성분분석을 위해서는 서로 상관된 다중파장대 데이터를 변환하여야 한다. 다중파장대 영상에서는 각 밴드 자료간에 상관이 있는 경우가 많기 때문에 주성분분석을 이용하여 현재의 영상에 포함되어 있는 대부분의 정보를 가상적 소수 밴드로 나타낼 수 있다.

3. 수치 해석 실험 및 분석

3.1 실험대상 지역 선정

실험 대상지역은 충청남도 태안군 소재의 의항리 일원으로 실험대상 지역의 면적은 약 250,000m²로 항공 하이퍼스펙트럴 영상과 공간해상도 1m 급, 128밴드의 분광영상을 획득하였다.

Fig. 6의 상단 왼쪽은 50cm급 참조용 항공사진이며, 오른쪽은 항공 하이퍼스펙트럴 영상의 RGB밴드 영상을 나타낸 것이다. 그림의 하단은 실험지역의 전경사진을 나타내고 있다.

3.2 트레이닝 자료 수집

감독분류를 위한 트레이닝 자료 수집을 위해 고해상도 참조사진을 이용하여 정확한 대상을 선정하였다.

육안으로 식별이 명확한 식생, 수역, 암석자갈, 모래구역의 트레이닝 자료를 수집하였다. 트레이닝 자료의

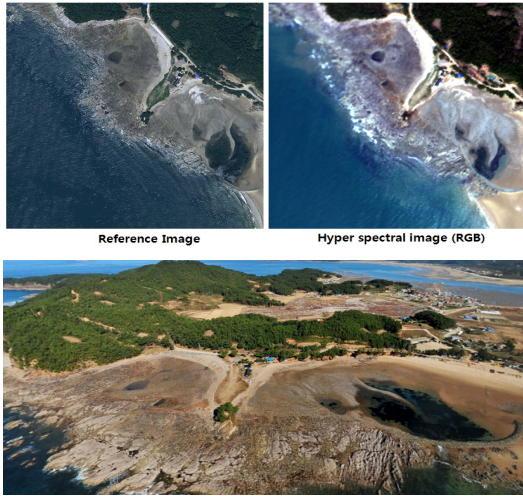


Figure 6. Study area

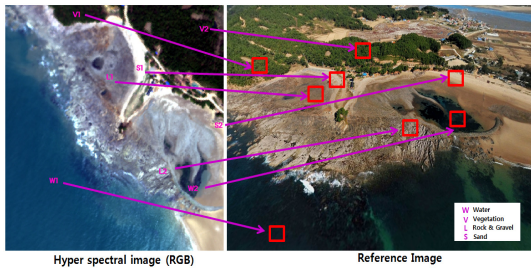
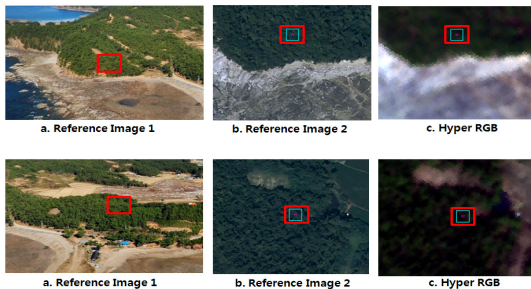


Figure 7. Location of training data

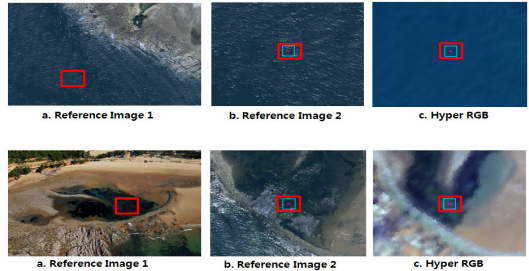
신뢰성과 분류 정확성을 높이기 위해 분류 대상 항목 당 2개소에 해당하는 트레이닝 자료를 수집하였다.

Fig. 7에서 좌측 항공 하이퍼스펙트럴 영상에서의 트레이닝 사이트에 대해 우측 참조 이미지의 해당 위치를 나타낸다.

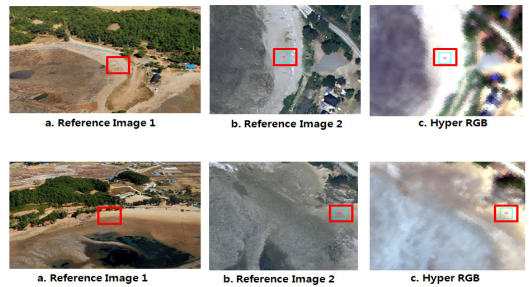
- Vegetable area of Training



- Water area of Training



- Sand area of Training



- Rock area of Training

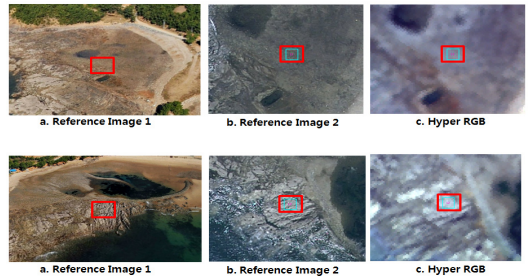


Figure 8. Different Training Data for study

3.3 PCA변환 결과 및 분석

128밴드의 원영상에 대한 PCA변환 결과는 Fig. 9와 같으며, 고유성분 포함률은 PCA변환 영상의 첫 번째 밴드에 해당하는 영상이 76.74%의 성분을 포함하며, PCA변환 영상의 첫 번째와 두 번째 누적 밴드에 해당하는 영상이 98.40%로 대부분의 성분자료가 두 번째 영상에 담긴 것을 알 수 있었다. 또한 두 번째 영상 이후부터 성분의 증가량은 완만하게 증가하여 여덟 번째 PCA밴드와 원본 영상(128밴드, PCA 미적용)과는 미소한 차이를 나타냈다.

각 PCA변환 후 각 밴드별 고유성분 누적포함비율은 Table 1과 같다.

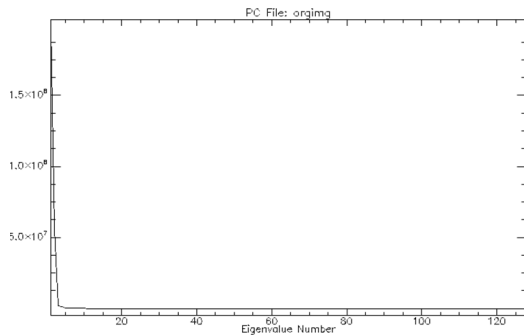


Figure 9. Value of each PCA band

Table 1. Rate of specific component for each PCA

Bnad	EV(%)
PCA 1B	76.74
PCA 2B	98.40
PCA 4B	99.62
PCA 6B	99.78
PCA 8B	99.82
128B(Org)	100.00

Table 1에서 보는 바와 같이 첫번째 PCA영상의 경우 가장 많은 정보를 포함하게 되며, 두 번째 PCA밴드를 포함할 경우 98.40%로 늘어나게 된다. Fig. 10을 보면 PCA 여덟 번째 영상에서 부터는 노이즈가 명확하게 나타남을 알 수 있었다.

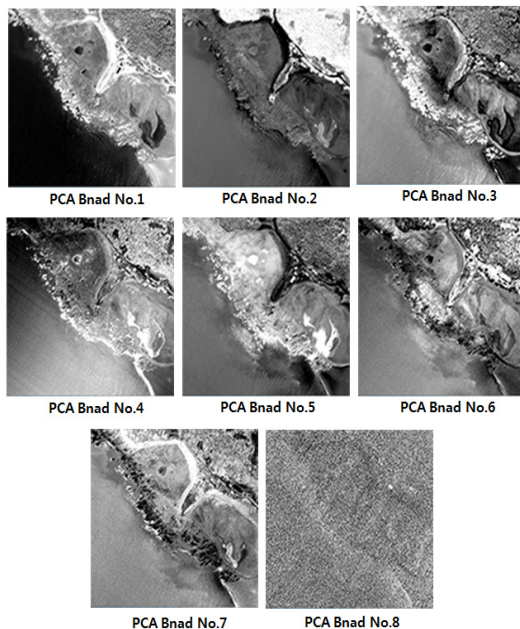


Figure 10. Image of each PCA band

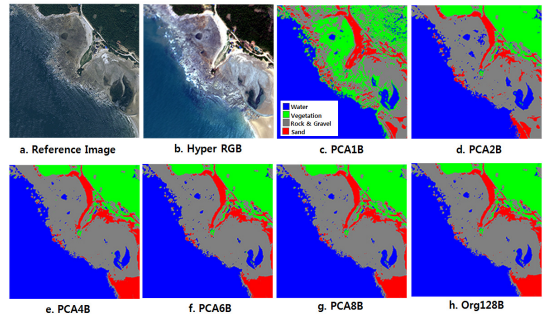


Figure 11. Result of classification PCA image

실험에 이용되는 SAM 감독 분류의 경우는 분광특성을 나타낸 곡선을 이용하게 된다. SAM 분류 기법은 전술한 바와 같이 공간상에서의 각도 개념을 적용하는 분류 방법이다. 따라서 화소와 분류 항목간의 유사도를 판별하는데 있어 n차원에서의 각도를 이용한다. 트레이닝 데이터를 이용하기 위해서 PCA밴드를 역 변환하여 128밴드의 노이즈가 제거된 유형별 분광 영상으로 변환하였다.

3.4 영상 분류 결과 및 분석

고유성분이 높은 밴드를 선별적으로 조합하여 5개의 유형별 PCA영상을 주성분 축에 따른 증가된 변환영상으로 제작 후 SAM감독분류 기법을 적용하여 영상을 분류한 결과이다.

영상분류 결과를 간단하게 확인하면, PCA 1B영상의 경우 분류의 정도가 상당히 떨어짐을 알 수 있다. 이것은 PCA 1B가 고유성분포함율 76.74%인 것에 기인한다고 판단된다. 정보 포함률이 98.40%인 PCA 2B영상 부터는 단계적으로 평가 결과가 양호한 값을 얻게 되며, PCA 4B 영상부터는 원본 영상의 분류결과까지 유사한 결과를 나타냈다. PCA 2B영상의 경우 PCA 1B 영상보다 양호하지만 하단부 모래지역이 암석자갈구역으로 오분류된 것을 육안으로 판단할 수 있다. 육안평가 결과 PCA8B영상과 원본 영상의 차이가 미소하므로 유효한 정보는 PCA 8B에 포함되었다고 추정할 수 있다.

4. 정확도 분석 및 평가

4.1 정량적 정확도 분석 및 평가

유형별 PCA영상의 고유성분포함율 결과를 살펴보면, PCA 1B영상의 경우는 원본 영상을 PCA변환 후, Table 2에서 보는 바와 같이 첫 번째 밴드가 70%이상의 성분을 포함하고 있으며, 밴드가 누적되어 갈수록 성분이 증가하는 것을 알 수 있다.

Table 2. Classification result of each band accuracy
(unit : %)

Bnad	EV 고유성분	ALLA 전체 정확 도	PMA 생산자	UMA 사용자
PCA1B	76.74	64.30	63.01	65.18
PCA2B	98.40	90.97	91.25	92.17
PCA4B	99.62	89.68	90.00	91.98
PCA6B	99.78	89.89	90.21	92.09
PCA8B	99.82	90.11	90.42	92.21
128B(Org)	100.00	90.11	90.42	92.21

여기서 특이한 사실은 첫 번째 밴드와 두 번째 밴드가 차지하는 성분 포함률이 98%를 육박하고 있다는 점이다. 따라서 대부분의 성분자료가 PCA 2B 영상에 담긴 것으로 나타났다. PCA 네 번째 변환 영상을 포함하는 경우인 PCA 4B 영상은 고유성분이 99.62%이며, 이후부터 128번째 원본 영상까지의 고유성분 포함률은 완만하고 균등한 기울기로 증가하는 것으로 나타났다.

유형별 PCA 영상의 전체 분류정확도 결과를 살펴보면, 전체 분류정확도는 PCA 1B 영상이 64.30%로 낮은 정확도를 나타냈으며, PCA2B 영상이 90.97%로 가장 높은 정확도를 나타냈다.

그러나 PCA4B 영상에서는 정확도가 1.29% 감소한 89.68%로 나타났다.

PCA6B 영상부터는 정확도가 완만하게 증가하였다. 유형별 PCA 영상의 생산자 분류정확도 결과를 살펴보면, 생산자 분류정확도는 PCA1B 영상이 63.01%로 가장 낮은 정확도를 나타냈으며, PCA2B 영상이 91.25%로 가장 높은 정확도를 나타냈다. 그러나 PCA4B 영상에서는 정확도가 1.25% 감소한 90.00%로 나타났다. PCA6B 영상부터는 원본 영상(PCA미변환, 128밴드)까지는 정확도가 완만하게 증가하였다. 유형별 PCA 영상의 사용자 분류정확도 결과를 살펴보면, 사용자 분류정확도는 PCA1B 영상이 65.18%로써, 가장 낮은 정확도를 나타냈으며, PCA2B 영상이 92.17%로 나타났다. PCA2B 영상부터 완만한 정확도 증가를 나타냈으며, PCA8B 영상과 원본 영상(PCA미변환, 128밴드)과는 92.21%로 가장 높은 정확도를 나타냈다.

4.2 시각적 정확도 분석 및 평가

육안을 통해 시각적인 분류 정확도 평가를 실시하였다. 분류대상이 모두 포함된 혼합구역과 분류 대상항목별로 대상이 확실한 구역에 대해 분류 대상별로 구역을 설정하여 시각적 평가를 실시하였다. PCA 1B 영상이 다른 유형별 영상의 결과보다 정확도가 현저히 떨어지

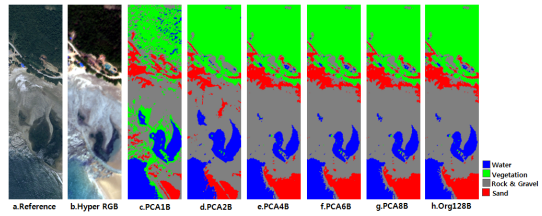


Figure 12. Analysis of entire classification result

을 확인할 수 있었다. PCA 1B 영상의 경우 고유성분 포함률이 76.74%인 것에 기인한다고 판단된다. 고유성분 포함률이 98.40%인 PCA 2B 영상부터는 단계적으로 분류의 정도가 양호한 결과를 나타냈으며, PCA 4B 영상부터는 원본 영상(PCA미변환, 128밴드)의 분류결과와 유사한 결과가 나타났다. PCA 2B 영상의 경우 PCA 1B 영상보다 분류결과가 양호하지만 그림에서 아래지역인 모래구역은 암석자갈 구역으로 오분류된 것을 육안을 통해 확인할 수 있었다. 육안평가결과 PCA8B 영상과 원본 영상과의 차이가 미소하므로 유효한 대부분의 고유성분은 PCA8B라고 추정할 수 있다.

4.2.1 분류항목별 분석

- 모래구역

모래지역의 PCA 밴드별 영상분류 결과 정보 포함률이 98.40%인 PCA2B 영상까지 오분류가 되었으나 PCA4B 영상부터는 정확하게 분류된 것으로 나타났다.

- 수계구역

수계지역의 PCA 밴드별 영상분류 결과 정보 포함률이 98.40%인 모든 유형의 PCA 밴드에서 정확하게 분류된 것으로 나타났다.

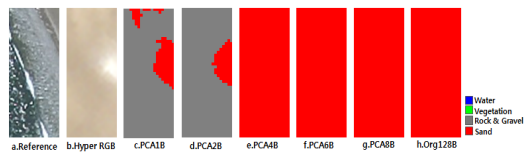


Figure 13. Classification result of PCA band for sand area

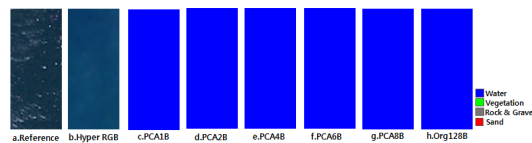


Figure 14. Classification result of PCA band for water area

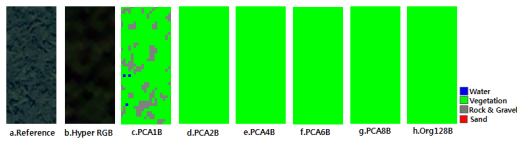


Figure 15. Classification result of PCA band for vegetation area

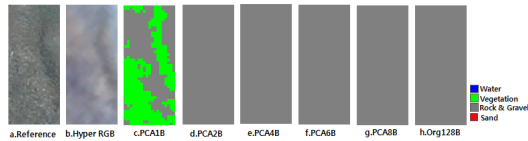


Figure 16. Classification result of PCA band for rock & gravel area

- 식생구역

식생지역의 PCA 밴드별 영상분류 결과 정보 포함률이 76.74%인 PCA 1B 영상까지 오분류가 되었으나 정보 포함률이 98.40%가 PCA2B 영상부터는 정확하게 분류된 것으로 나타났다.

- 암석자갈구역

모래암석지역의 PCA 밴드별 영상분류 결과 정보 포함률이 76.74%인 PCA1B 영상까지 오분류가 되었으나 정보 포함률이 98.40%이 PCA2B 영상부터는 정확하게 분류된 것으로 나타났다.

5. 결론

본 연구에서는 항공 하이퍼스펙트럴 영상에 대해 PCA를 적용 후 고유성분이 높은 밴드를 선별적으로 조합하여 5개의 유형별 PCA 영상을 제작하였다. SAM 감독 분류 기법을 적용하여 분류 정확도를 평가한 결과 다음과 같은 결론을 도출하였다.

첫째. PCA 변환 결과 고유성분 포함률을 살펴 본 결과, PCA 변환 영상의 첫 번째 밴드에 정보포함률이 76.74%를 포함하며, 두 번째 밴드에서의 누적 성분은 98.40%로써, 대부분의 성분정보가 PCA 변환 후 두 번째 영상까지에 담긴 것으로 나타났다.

둘째. PCA 밴드가 조합된 유형별 영상의 정량적인 분류정확도를 평가한 결과, 전체정확도와 생산자 정확도는 유사한 그래프 패턴을 가지고 있었다. PCA 변환 영상의 첫 번째 밴드에 해당하는 영상이 가장 낮은 분류 정확도를 나타냈으며, 두 번째 밴드까지 조합된 영상이 가장 높은 정확도를 나타냈다. 네 번째 밴드가 조합된 영상부터는 전체적으로 정확도가 완만하게 증가하였다.

셋째. PCA 밴드가 조합된 유형별 영상의 시각적인 평가를 실시한 결과, PCA 변환 영상의 첫 번째 밴드에 해당하는 영상이 제일 낮은 정확도를 나타냈다. 고유성분 포함률이 급격하게 높아진 두 번째 밴드까지 누적된 영상부터는 점진적으로 증가하였다. 이후의 영상들은 양호한 분류결과를 나타냈다. PCA 변환 영상의 여덟 번째 까지 누적된 밴드의 영상은 원본 영상과의 차이가 미소하여 고유성분의 유효정보는 여덟 번째 밴드라고 추정할 수 있었다.

넷째. 고유성분 포함률, 분류정확도를 평가한 결과, 정량적 분류 정확도 수치상으로는 두 번째 밴드까지 누적된 영상이 가장 높았으나 시각적인 평가결과는 여덟 번째 밴드를 포함해야 실용적으로 이용이 가능한 것으로 평가된다.

연구결과 PCA 영상의 정량적인 정확도와 시각적 평가의 차이가 존재하므로 실용적인 용도를 고려하여 적절한 밴드를 취사하여 사용하여야 한다.

References

1. Cho, D. Y. and Kim, E. M., 2010, Extraction of spatial information of tree using LIDAR data in urban area, Journal of The Korean Society for Geospatial Information System, Vol. 18 No. 4, pp.11-20.
2. Choi, J. W., Kim, D. S., Lee, B. K., Kim, Y. I. and Yu, K. Y., 2006, Hyperspectral image fusion algorithm based on two-stage spectral unmixing method, Korean Journal of Remote Sensing, Vol. 22 No.4, pp.295-304.
3. Du, Q., 2012, A new sequential algorithm for hyperspectral endmember extraction, IEEE Geoscience and Remote Sensing Letters, Vol.9 No.4, pp.695-699.
4. Han, D. Y., Cho, Y. W., Kim, Y. I. and Lee, Y. W., 2003, Feature selection for image classification of hyperion data, Korean Journal of Remote Sensing, Vol. 19 No.2, pp.171-179.
5. Han, Y. K., Byun, Y. K, and Kim, Y. I., 2012, Automatic estimation of geometric translations between high-resolution optical and SAR images, Journal of the Korean Society for Geospatial Information System, Vol. 20 No. 3, pp.41-48.
6. Lee, C. K., Kim, B. O. and Kim, N. Y., 2007, Investigation of shoreline change by photogrammetric method, Journal of Korean Society of Geospatial Information System, Vol. 15 No. 2,

- pp.15-23.
7. Liangrocapt, S. and Petrou, M., 1998, Mixed pixels classification, Proc. SPIE 3500, Image and Signal Processing for Remote Sensing IV, pp.72-83.
 8. Mars, J. C., Garrity, C. P., Houseknecht, D. W., Amoroso, L. and Mears, D. C., 2005, Digital-elevation and surface-classification maps of the Fish Creek Area, Harrison Bay Quadrangle, Northern Alaska, U.S. Geological Survey.
 9. Park, M. H., 2006, Usefulness of canonical correlation classification technique in hyper-spectral image classification, Korea Society of Civil Engineers, Vol. 26 No. 5D, pp.885-894.
 10. Shin, J. I., Kim, S. H., Yoon, J. S., Kim, T. G. and Lee, K. S., 2006, Spectral mixture analysis using hyperspectral image for hydrological land cover classification in urban area, Korean Journal of Remote Sensing, Vol. 22 No. 6, pp.565-574.