

## 중력모델에 기반한 하이퍼스펙트럴 영상 분류 Classification of Hyperspectral Images based on Gravity type Model

변영기<sup>1)</sup>·이정호<sup>2)</sup>·김용민<sup>3)</sup>·김용일<sup>4)</sup>

Byun, Young Gi·Lee, Jeong Ho·Kim, Yong Min·Kim, Yong Il

<sup>1)</sup> 서울대학교 대학원 지구환경시스템공학부 박사과정(E-mail: kko071@snu.ac.kr)

<sup>2)</sup> 서울대학교 대학원 지구환경시스템공학부 박사과정(E-mail: ilome79@snu.ac.kr)

<sup>3)</sup> 서울대학교 대학원 지구환경시스템공학부 석사과정(E-mail: kym0210@snu.ac.kr)

<sup>4)</sup> 서울대학교 공과대학 지구환경시스템공학부 부교수(E-mail: kiyun@snu.ac.kr)

### Abstract

Hyperspectral remote sensing data contain plenty of information about objects, which makes object classification more precise. Over the past several years, different algorithms for the classification of hyperspectral remote sensing images have been developed. In this study, we proposed method based on absorption band extraction and Gravity type model to solve hyperspectral image classification problem. In contrast to conventional methods that are based on correlation techniques, this method is simple and more effective. The proposed approach was tested to evaluate its effectiveness. The evaluation was done by comparing the results of preexisting SFF(Spectral Feature Fitting) classification method. The evaluation results showed the proposed approach has a good potential in the classification of hyperspectral images.

### 1. 서 론

하이퍼스펙트럴 영상자료는 지표물의 완전한 분광반사특성곡선(Spectral reflectance curve) 정보를 제공하고, 다중분광영상자료에 비해 높은 분광해상도를 가지므로 지표물에 대한 세밀한 분석이 가능하다. 그러나 하이퍼스펙트럴 영상의 많은 밴드수로 인하여 신뢰도가 낮은 트레이닝 데이터를 취득할 경우 영상을 분류하는데 있어서 오히려 낮은 분류정확도를 나타낼 수 있다. 따라서 기존의 다중분광영상에 적용되었던 일반적인 분류기법들을 그대로 적용하는데 어려움을 겪을 수 있다. 이러한 문제를 해결하기 위해 하이퍼스펙트럴 데이터에 적합한 분류 알고리즘 개발에 대한 연구가 활발히 진행되어왔다.

특히 물체의 분광특성을 가장 잘 표현하는 하이퍼스펙트럴 영상의 장점을 최대한 활용하여 보다 정확한 지표물의 종류를 측정하기 위한 여러 가지 스펙트럼 분석기법들이 제시되고 있다. 스펙트럼 분석기법은 기존에 알려져있는 대상물체의 기준 분광반사값을 이용하여 하이퍼스펙트럴 영상에서 얻어지는 반사값과의 분광특성 유사성을 분석하여 각 화소에 대한 대상물체의 종류를 식별하는 방법으로써, 유사도를 측정하는 방법과 기준에 따라 다양한 분류 알고리즘이 연구되어 왔다(James, 2003). 기존의 대표적인 분광특성정합 기법으로 분광흡수밴드를 이용하여 기준 스펙트럼과 대상스펙트럼과의 유사성을 제시하는 방법인 SFF(Spectral Feature Fitting)을 들 수 있다.

본 연구에서는 중력모델을 이용하여 기존의 거리요소와 스펙트럼 특성정보를 통합한 효율적인 분류방법을 개발하고 이를 하이퍼스펙트럴 영상에 적용하여 개발된 방법의 성능을 평가하는데 그 목적이 있다. 또한, 결과 비교를 위해, 하이퍼스펙트럴 영상분류에 있어 정형화된 SFF 분류기법을 이용하여 동일 지역에 대해 분류를 수행하고 두 방법의 분류 정확도 및 효율성을 평가하고자 한다.

## 2. 자료 처리 및 분석

### 2.1 사용데이터

Hyperion 센서는 가시/근적외(VNIR) 분광계와 단파적외(SWIR) 분광계로 구성되어 있다. 다중분광센서와 달리 연속적인 분과영역에 대해 242개의 고유한 분광채널을 갖는다. USGS에서 제공하는 레벨-1B 영상의 경우에 두 분광기 자료의 정렬 및 불량 픽셀 보간, 포화된 검출소자 보고서 등을 통하여 보정이 이루어졌으나, 여전히 검출소자의 민감도 차이로 인하여 많은 밴드에서 수직 방향의 보정되지 않은 줄무늬를 볼 수 있다. 본 연구에서 실험에 사용한 자료는 2002년 4월 3일에 취득한 EO-1 Hyperion 센서 영상으로 안양과 군포 일대의 250pixels × 500 pixels 영상을 대상지역으로 사용하여 알고리즘을 평가하였다.

### 2.2 전처리 과정

하이퍼스펙트럴 센서의 특성상 매우 많은 영상촬영소자를 이용하게 되므로 측정기기 내의 열로 인한 잡음 현상과 검출기의 불량이나 민감도 등의 이유로 영상 내에 노이즈가 발생한다. 그러므로 사용자의 Band selection이나 노이즈제거와 같은 전처리 과정이 필요하다. 본 연구에서는 초기 Hyperion 영상에 수 많은 수직선 형태의 노이즈들이 많이 존재하므로 이를 고려하여 Band selection 과정과 원영상에 MNF 변환을 실시하여 신호대잡음비(SNR)가 매우 우수한 MNF 성분들만을 이용하는 방법을 적용하여 영상 전처리를 수행하였다.

#### 2.2.1 MNF(Minimum Noise Fraction) 변환

MNF 변환은 주성분분석(PCA) 변환과 유사한 변환이다. 하지만 자료들의 산포도를 최대화하는 주성분분석과는 달리, MNF변환은 신호대잡음비를 최대화하는 성분을 선정하는 선형변환이다. MNF변환 후에 생성되는 낮은 순서의 성분들에는 증가된 잡음이 존재하기 때문에 필터링 등의 후처리를 실시하고 이를 MNF 역변환을 실시하여 노이즈를 감소시킨다.

#### 2.2.2 Band Selection

MNF를 사용한 후에도 영상 전체가 의미 없는 정보를 담고 있는 밴드를 처리를 할 수 없다. 특히, Hyperion 영상에는 이러한 밴드들이 많이 존재하는데, 정보가 전혀 없는 밴드와 노이즈로만 채워져 있는 밴드로 구분할 수 있다. 이를 제거하기 위해서는 다양한 알고리즘을 적용할 수 있지만, 본 연구에서는 USGS에서 공식적으로 보정되지 않은 밴드로 명명한 밴드 44개와 수작업을 통하여 노이즈라고 판단되는 밴드 63개를 제거하여 242개의 밴드로부터 135개의 밴드만을 추출하였다.

#### 2.2.3 Reflectance Transformation

하이퍼스펙트럴 영상에서 얻을 수 있는 가장 중요한 정보는 각 화소에 해당하는 지표물의 연속적인 분광반사값(spectral reflectance)이라 할 수 있다. 현재 초분광영상은 대부분 센서에서 감지된 복사량(radiance)으로 표현된 자료로 제공되고 있다. 추출된 135개의 밴드 자료는 모두 16bit의 Radiance값이다. 영상의 Radiance 값은 각 파장에 대하여 지표에서 반사된 에너지의 양이기 때문에 밴드별로 스케일이 모두 다르다. 따라서 이들 값들은 지표의 분광반사 특성 값들로 변화시켜주어야 분류의 성능을 향상시킬 수 있다. 또한, 변환된 분광반사값들은 0~1의 값을 가지기 때문에 스케일 문제 또한 해결할 수 있다. 분광반사로의 변환식은 식 1과 같다.

$$\rho_p = \frac{\pi \times L_\lambda \times d^2}{ESUN_\lambda \times \cos \theta_s} \quad (1)$$

여기서,  $\rho_p$  : 영상의 반사율,  $L_\lambda$  : 영상의 분광 복사휘도,  $d$  : 지구와 태양사이의 거리,  $ESUN_\lambda$  : 센서의 태양 방사도,  $\theta_s$  : 태양 천정 각이다.

## 2.3 방법론

일반적으로 하이퍼스펙트럴 영상을 이용한 감독분류(Supervised classification)의 과정은 사용자가 각 분류항목의 특성을 지정하는 트레이닝(Training)단계와 트레이닝 단계에서 추출된 특징을 기반으로 하는 분석단계로 이루어진다. 트레이닝 단계에서는 영상에서 각 분류항목별로 확실한 영역(Training Field)을 지도나 현장조사를 통하여 설정하고, 이 트레이닝 데이터로부터 분류항목별 특징을 추출하게 된다. 분석 단계에서는 선정된 특징값을 이용하여 일정한 기준에 의하여 모든 화소들을 순차적으로 분류하게 된다. 본 연구에서는 분류항목별 특징으로서 평균스펙트럼벡터를 이용하였고, 중력모형을 이용하여 기존의 거리요소와 스펙트럼특성정보를 통합한 새로운 유사도측정기법을 사용하여 영상분류를 수행하였다.

### 2.3.1 중력모형

중력모형은 사회과학에 응용되어 지리적 공간에서 서로 떨어진 지역의 공간상호작용(spatial interaction effect)를 이해하는데 많이 이용되어져 왔다. 즉 공간상에서의 상호작용을 중력 계산에 이용되는 질량에 해당하는 유인 요소와 거리요소를 주요변수로 하여 적용목적에 따른 대상 지점간의 상대적 정량적 지수정보로 표현하는 방법이다. 이러한 중력모형은 도시계획이나 교통 분야에서 통행수단 비용, 시간, 거리등을 주요 변수로 하여 위치선정 및 교통수요예측 등을 연구하는데 많이 이용되어 왔다. 본 연구에서는 이러한 중력모형을 스펙트럼 공간상에 적용하여 식 2 와 같은 유사도 측정지수를 개발하여 하이퍼스펙트럴 영상분류의 유사도측정도구로 사용하였다.

$$SIV = \frac{[SID(\vec{x}, \vec{y})]^{-1}}{d(\vec{x}, \vec{y})} \quad (2)$$

$$d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\| = \left[ \sum_{i=1}^N (x_i - y_i)^2 \right]^{1/2}$$

$$SID(\vec{x}, \vec{y}) = D(\vec{x} \parallel \vec{y}) + D(\vec{y} \parallel \vec{x})$$

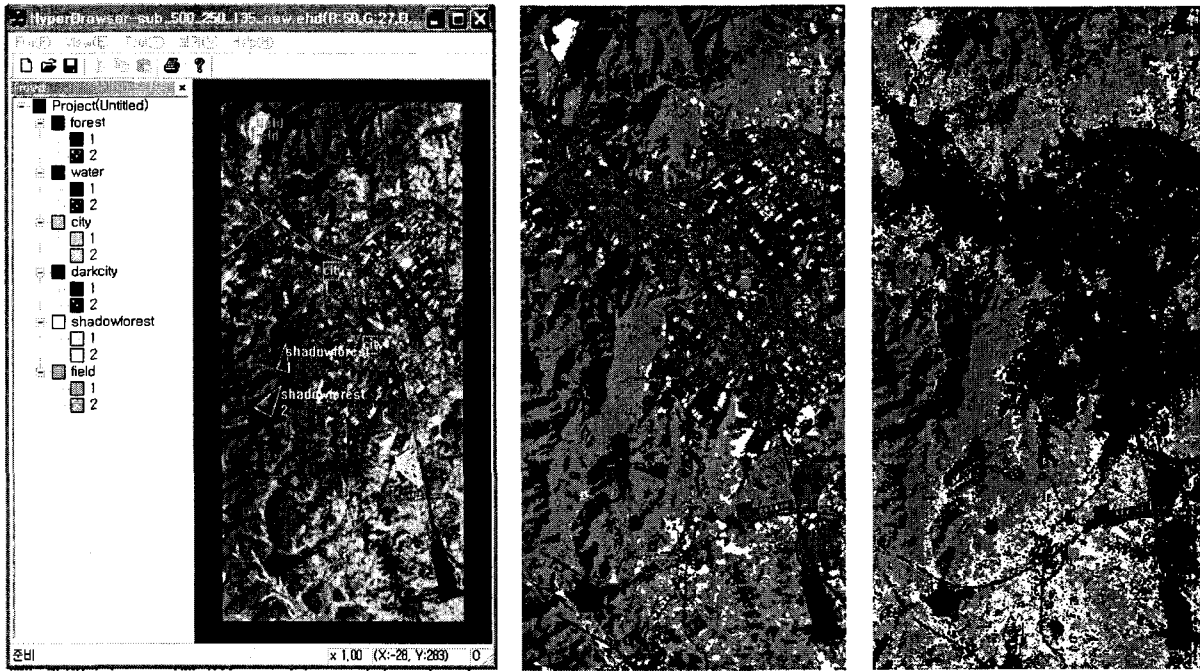
$$D(\vec{x} \parallel \vec{y}) = \sum_{i=1}^N p_i \log(p_i/q_i), \quad D(\vec{y} \parallel \vec{x}) = \sum_{i=1}^N q_i \log(q_i/p_i)$$

$$p_i = x_i / \sum_{i=1}^N x_i, \quad q_i = y_i / \sum_{i=1}^N y_i$$

여기서  $\vec{x}_i$ 는 기준 스펙트럼벡터,  $\vec{y}_i$ : 대상스펙트럼, N:사용된 밴드 수를 나타낸다. 식 2에서 볼 수 있듯이 거리요소에 직선거리(Euclidean distance)를 유인요소에 Spectral Information Divergence를 적용함으로써, 두 스펙트럼에서 나타나는 반사값의 차이와 모양의 유사함을 동시에 고려할 수 있다. 거리를 측정하는 방법에는 여러 거리척도들이 제시되었으나 일반적으로 직선거리 측정 방법이 가장 많이 사용되는 방법이다. SID는 Chang(1999)이 제안한 유사도 측정방법으로 기본적으로 스펙트럼의 특성정보로, 두 스펙트럼의 상대적 엔트로피(relative entropy)를 이용한 방법이라 할 수 있다.

## 2.4 분석 및 평가

본 연구에서 영상 분류에 적용된 클래스 선정은 실험 대상 영상 내에서 트레이닝이 가능하고, 충분히 식별가능한 클래스를 검토한 결과 그림 1과 같이 forest, shadow-forest, water, city, dark city, field 등의 6 개 클래스를 선정하였다. 또한 결과 분석을 위해 데이터 셋을 샘플링하여 각 분류항목별 특징(feature)을 추출하기 위한 트레이닝 데이터 1202개와 정량적인 검증을 위한 Validation 데이터 708개를 생성하였다.



(a) 트레이닝, 검증데이터 (b) SIV 분류결과 (c) SFF 분류결과  
 그림 1. 클래스별 트레이닝, 검증 데이터 및 분류결과 영상

본 연구에서는 결과 비교를 위하여 동일데이터에 대한 SFF 기법을 적용하여 두 방법의 정확도 평가를 수행 하였다. SIV와 SFF기법의 분류결과와 검증 데이터의 클래스 일치여부를 비교한 분류정확도는 각각 96.2%, 92.7%의 정확도를 보였다. 그림 1의 (b)영상은 본 연구에서 제안된 방법으로 분류한 결과로서, 시각적으로 볼 수 있듯이 도심지의 고층 밀집 건물 지대와 도로의 선형특성을 잘 나타내고 있으나 일부 나대지 지역과 도심지역이 약간 혼합되어진 것을 볼 수 있다. SFF 분류기법의 경우 시각적 평가 및 정확도 평가 결과에서 알 수 있듯이 제안된 분류방법에 비해 좋지 못한 결과를 보이고 있다.

### 3. 결 론

본 연구에서는 중력모델을 이용하여 기존의 거리요소와 스펙트럼 특성정보를 통합한 효율적인 분류방법을 개발하고 이를 하이퍼스펙트럴 영상에 적용하여 보다 효율적인 분류방법을 제시하였다. 또한, 기존의 정형화된 SFF 분류기법과 제안된 분류기법을 수행하여 분류정확도 평가한 결과 다음과 같은 결론을 얻었다. 제안된 분류방법의 경우 기존의 유사도 측정치들을 통합함으로써 보다 효과적으로 하이퍼스펙트럴 영상을 분류 할 수 있었다. 정확도 평가 결과 제안된 알고리즘을 통하여 생성된 분류영상이 SFF기법을 이용한 분류영상보다 전체 정확도 측면에서 3.5% 정도 증가함으로써 전반적으로 정확도가 향상되었다. 향후 더 다양한 항목의 분류 및 지형도를 활용한 더욱 정밀한 정확도 평가를 수행해야 할 것이다.

### 참고문헌

Chein-I Chang, 1999, Spectral information divergence for hyperspectral image analysis, Goescience and Remote Sensing Symposium, 1999. IGARSS '99 Proceedings. IEEE 1999 International, Vol 1, pp. 509-511.  
 N. James, 2003, The Spectral Similarity Scale and its Application to the Classification of Hyperspectral Remote Sensing Data, IEEE Work Shop, pp. 92-99.