

## 화소간 유사도 측정 기법을 이용한 하이퍼스펙트럴 데이터의 무감독 변화탐지에 관한 연구

### A Study on the Unsupervised Change Detection for Hyperspectral Data Using Similarity Measure Techniques

김대성<sup>1)</sup> · 김용일<sup>2)</sup>

Kim, Dae Sung · Kim, Yong Il

<sup>1)</sup> 서울대학교 대학원 지구환경시스템공학부 박사과정(E-mail: mutul94@empal.com)

<sup>2)</sup> 서울대학교 공과대학 지구환경시스템공학부 부교수(E-mail: yik@plaza.snu.ac.kr)

#### Abstract

In this paper, we propose the unsupervised change detection algorithm that apply the similarity measure techniques to the hyperspectral image. The general similarity measures including euclidean distance and spectral angle were compared. The spectral similarity scale algorithm for reducing the problems of those techniques was studied and tested with Hyperion data. The thresholds for detecting the change area were estimated through EM(Expectation-Maximization) algorithm. The experimental result shows that the similarity measure techniques and EM algorithm can be applied effectively for the unsupervised change detection of the hyperspectral data.

#### 1. 서론

지상의 대상물(Objects)이나 그 속에서 발생하는 현상(Phenomenon)들은 계속 변화하고, 그 변화에 대한 모니터링의 중요성 때문에 변화탐지는 원격탐사에 있어 중요한 주제이며 현재까지도 새로운 기법들이 계속 제안되고 있다(D. Lu 등, 2004). 변화탐지는 크게 감독 변화탐지(Supervised Change Detection)와 무감독 변화탐지(Unsupervised Change Detection) 기법으로 나눌 수 있다<sup>1)</sup>(Lorenzo B. 등, 2000). 감독 변화탐지는 변화속성에 대한 정보를 얻기 위한 방법으로 일반적으로 분류 또는 군집화 기법을 이용한다. 무감독 변화탐지는 영상의 화소값 연산에 의해 변화지역을 탐지하는 기법이다. 이 연구는 무감독 변화탐지에 한정하여 연구의 초점을 맞추고 있으며, 최근 원격탐사 데이터로 관심이 높아지고 있는 하이퍼스펙트럴 데이터를 이용하고자 한다.

하이퍼스펙트럴 데이터는 지상의 정보를 연속적이며 좁은 파장 폭(약 10~20nm)의 수백 개 밴드로 이루어져 있어 높은 분광 및 방사해상도 특성을 지닌다(Freek D. V. 등, 2001). 최근 EO-1(Earth Observing-1) 위성의 발사와 함께 위성 하이퍼스펙트럴 데이터의 연구가 활발히 진행되고 있으며, 토양, 지질, 산림 분야에 있어 그 응용 영역을 확대해 가고 있다. 하지만, 낮은 SNR(Signal-to-Noise)과 좁은 촬영 폭은 데이터 분석, 특히 변화탐지 적용에 있어 그 한계가 발생한다. 따라서 하이퍼스펙트럴 데이터가 가지는 많은 밴드의 정보를 최대한 활용할 수 있는 변화탐지 알고리즘의 적용이 필요하며, 본 연구를 통해 두 시기 데이터간 스펙트럼 정보의 유사도를 측정함으로써 변화를 추정할 수 있는 기법을 제안하자 한다. 또한 기대최대화(Expectation-Maximization) 기법을 통해 무감독 변화탐지의 단점으로 알려진 한계값을 보다 객관적으로 도출하였으며, 자동화된 무감독 변화탐지 기법을 수행해 보았다.

1) 변화강조(Change Enhancement) 기법과 변화속성탐지(Change Nature Detection)으로 나누기도 함

## 2. 변화탐지 알고리즘

### 2.1 감독 변화탐지와 무감독 변화탐지

변화탐지란 다른 시간 데이터를 통해 대상물이나 현상의 관점에서 그 차이를 인식하는 과정으로 정의할 수 있으며(Singh A., 1989), 원격탐사의 다양한 해상도 중 시간해상도(Temporal Resolution)를 효과적으로 이용할 수 있는 기법이다. D. Lu 등(2004)은 변화탐지 연구를 1) 대수적 방법(Algebra) 2) 변환 기법(Transformation) 3) 분류 방법(Classification) 4) 발전된 형태의 변화탐지 기법(Advanced Model) 5) 지리정보시스템 접근(GIS Approach) 6) 시각적 분석(Visual Analysis) 7) 기타 등 7개 카테고리로 나누고 있다. 이 중 감독 변화탐지 기법은 분류 방법과 발전된 형태의 변화탐지 기법, 지리정보시스템 접근, 시각적 분석을 포함한다. 감독 변화탐지가 각 영상의 다른 대기 상태에 대한 영향을 최소화할 수 있는 효과적인 장점을 가지는 반면, 분류를 위한 훈련자료(Training Data)를 취득하는 것이 어렵고, 분류 정확도에 따라 변화탐지 신뢰도가 결정되는 단점이 발생한다(Lorenzo B. 등, 1997). 무감독 변화탐지 기법은 감독 변화탐지에 비해 추출할 수 있는 변화의 정보가 적고 두 시기 데이터의 전처리 과정이 필요한 반면 알고리즘 적용이 쉽고, 자동화된 변화탐지를 수행할 수 있다. 대수 기법으로 차연산, CVA(Change Vector Analysis) 등이 많이 사용된다. PCA 또는 Tasseled Cap 알고리즘과 같은 변환 기법은 밴드간 데이터 중복을 줄이고 추출된 성분의 다른 정보들을 강조할 수 있다(D. Lu 등, 2004).

### 2.2 유사도 측정 기법을 이용한 무감독 변화탐지

유사도 측정 기법은 많은 밴드를 가진 하이퍼스펙트럴 데이터의 특성을 반영하기 위해 스펙트럼 정보를 이용하는 분류 기법(Image Classification)과 유사하다. 본 연구는 이러한 원리를 이용하여 두 시기 스펙트럼의 차이가 변화 정도를 나타낼 수 있다는 가정에서 시작하였다. 이는 화소간 “유사도 측정 기법”이라는 용어로 사용할 수 있으며, 관련 알고리즘으로 유클리드 거리(Euclidean Distance), 분광각(Spectral Angle), 상관계수, SSS(Spectral Similarity Scale), SID(Spectral Information Divergence) 등이 있다(Emmanuel Christopher 등, 2005).

분광각은 다차원으로 확장된 공간에서 측정되는 두 벡터간 각을 말하며, 스펙트럼 모양의 유사도를 측정하는데 효과적이다(식 1). 하지만 분광각 알고리즘은 벡터의 게인(Gain) 값이 민감하지 않고 원점 부근의 화소들을 잘 분리시키지 못하는 단점이 발생한다.

$$\theta_{i,c} = \cos^{-1} \left[ \frac{\sum_{k=1}^m x_{i,k} x_{c,k}}{\sqrt{\sum_{k=1}^m x_{i,k}^2 \sum_{k=1}^m x_{c,k}^2}} \right] \quad (\text{식 1})$$

여기서,  $x_{i,k}$ :  $i$  영상의  $k$  밴드의 화소값,  $x_{c,k}$ :  $c$  영상의  $k$  밴드의 화소값,  $m$ : 밴드 수

SSS는 거리와 각이 가지는 문제점을 해결하기 위해 제안된 것으로(Sweet J. N., 2004) 두 가지 가정을 통해 성립된다. 동일한 벡터는 그 방향(Direction)과 크기(Magnitude)가 같아야 한다는 것(벡터의 일치성)과 벡터의 크기와 방향은 서로 독립적이어야 한다는 것(벡터의 독립성)이다. DAED와 DiCorr 값은 각각 1로 정규화되어 계산되며 SSS의 최대값은  $\sqrt{2}$ 가 된다.(식 3 참고).

$$SSS = \sqrt{DAED^2 + DiCorr^2} \quad (\text{식 2})$$

$$\text{여기서, } DAED_{i,c} = \sqrt{\frac{1}{m} \times \sum_{k=1}^m (x_{i,k} - x_{c,k})^2}, \quad DiCorr_{i,c} = \frac{\sum_{k=1}^m (x_{i,k} - \mu_i) \times (x_{c,k} - \mu_c)}{(m-1) \times \sigma_i \times \sigma_c}$$

### 2.3 무감독 변화탐지의 한계값(threshold) 선택 기법

무감독 변화탐지 기법들을 통해 추출한 결과값을 이용하여 변화 유무를 판단하기 위한 작업은 두 개의 클래스를 나누는 작업과 유사하다(Farid M. 등, 2002). 일반적 한계값 결정 방법은 크게 1) 수동적인

시행착오를 통한 방법(Trial-and-Error Procedure)과 2) 통계적인 분석을 통한 접근(Statistical Measures)의 두 가지로 나누어진다(D. Lu 등, 2004). 하지만, 두 방법 모두 한계값을 결정하기 위해사용자가 개입되어야 하는 문제가 발생한다. 본 연구에서 이러한 문제를 기대최대화 기법을 이용하여 해결해보고자 하였다. 기대최대화 기법은 입력 자료로부터 반복적인 연산을 통해 확률 밀도 함수의 파라미터(평균, 분산, 사전확률)를 결정하는 방법이다(박노욱 등, 2003). 가우시안 확률 분포를 가정하며, 파라미터들의 기대값을 계산하는 기대화 과정(E-step)과 최도우도를 가지는 새로운 추정치를 계산하는 최대화 과정(M-step)으로 나누어진다. 이전 단계의 추정치와 새로운 추정치의 값이 차이가 거의 없을 때 반복 계산은 멈추게 되고 두 밀도 함수가 가지는 확률이 동일한 지점이 한계값이 된다. 한계값은 식 3에서 나열된 2차 방정식을 통해 계산할 수 있다.

$$(\sigma_n^2 - \sigma_c^2)\widehat{T}_0^2 + 2(\mu_n\sigma_c^2 - \mu_c\sigma_n^2)\widehat{T}_0 + \mu_c^2\sigma_n^2 - \mu_n^2\sigma_c^2 + 2\sigma_n^2\sigma_c^2 \ln\left[\frac{\sigma_c P(w_n)}{\sigma_n P(w_c)}\right] = 0 \quad (\text{식 3})$$

여기서,  $\widehat{T}_0$ 는 한계값,  $\sigma_n, \sigma_c$ 는 표준편차,  $\mu_n, \mu_c$ 는 평균,  $P(w_n), P(w_c)$ 는 사전확률

### 3. 알고리즘 적용 및 결과

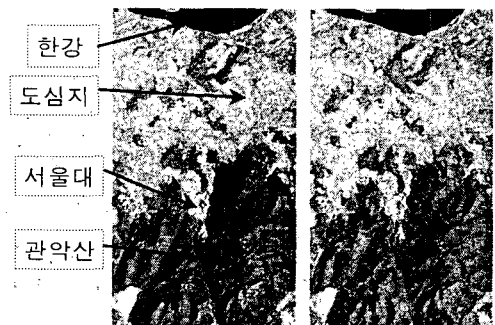
#### 3.1 연구 방법

알고리즘 적용을 위해 사용된 데이터는 Hyperion 영상이다. Hyperion은 Pushbroom 방식으로 촬영하며 30m 공간해상도, 242개의 밴드로 구성되어 있다. 242개의 밴드 중 보정되지 않은 밴드와 보정은 되었으나 파장이 중복되는 밴드를 1차적으로 제거하여 196개의 밴드를 생성하였다. 시각적 판단에 의해 노이즈 밴드를 제거한 최종 131개 밴드를 선택하여 350×200화소 크기의 데이터로 구성하였다.

이 실험은 두 가지 경우에 대해 알고리즘의 성능을 비교하고 있다. 첫 번째 경우는 유사도 측정 기법이 노이즈에 효과적임을 보이기 위해 수행한 실험으로, 안산 부근을 대상지역으로 선정하여 세 군데 임의의 지역에 대해 화소의 변화를 주었으며<sup>2)</sup>, 일정량의 노이즈( $\sigma=293$ )를 첨가하여 실험을 수행하였다(그림 1 참고). 노이즈가 발생한 단밴드 영상의 영상차 결과와 시각 및 정량적 평가를 수행하여 제안된 기법의 우수성을 알아보았다. 두 번째 경우에 사용된 영상은 서울 남부지역을 2002년 4월 3일과 4월 19일 촬영한 것으로 봄 시기의 활엽수 성장 변화를 관측할 수 있을 것으로 예상하고 실험을 진행하였다. 4월 3일 영상을 기준으로 기하보정을 실시하였고, 두 데이터간 화소차를 최소화하기 위해 히스토그램 매칭을 실시하였다. 사용된 데이터는 그림 2에서 확인할 수 있다.



(a) 원영상 (b) 시기 I(밴드 10) (c) 시기 II(밴드 10)  
그림 1. 사용 영상(경우 I)



(a) 시기 I (b) 시기 II  
그림 2. 사용 영상(경우 II)

#### 3.2 알고리즘 적용 결과 - 경우 I

알고리즘을 통해 추출되는 결과는 변화의 정도를 나타내는 단밴드 영상이며, 높은 값을 가지는 화소

2) 화소변화를 준 각각의 지역은 도심지에 산림지 화소를 부여한 경우(20×20화소), 산림지에 내대지 화소를 부여한 경우(12×12화소), 산림지에 도심지 화소를 부여한 경우(20×20화소)로서 총 944화소를 변화시켰음(그림 1 (c) 참고).

들은 변화 정도가 높게 나타남을 의미한다. 그림 3은 유클리드 거리, 분광각, SSS, 단밴드 영상차를 적용하여 도출된 결과 영상이며 결과값의 최대, 최소를 표 1에서 정리하였다.

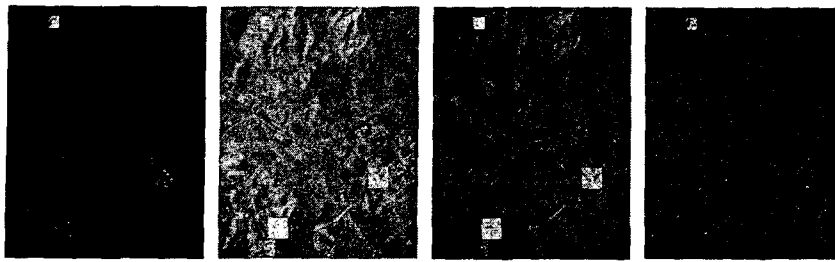


표 1 알고리즘 결과 최대/최소

	최대	최소
유클리드거리	22106	3415.7
SSS	1.1092	0.0481
분광각	28.908	3.5706
단밴드 영상차	2386	0

그림 3. 알고리즘 적용 결과 영상(좌로부터 유클리드 거리, SSS, 분광각, 단밴드 영상차)

변화의 여부를 판단하기 위해 기대최대화 기법을 적용하여 한계값을 설정하였다. 그림 3의 결과값을 히스토그램 화하여 그림 4로 표현하였으며, 히스토그램 분포를 기준으로 기대최대화 기법을 적용하게 된다. 그림 4에서 하늘색(왼쪽)의 가우시안 곡선은 변화되지 않은 지역을 나타내고, 빨간색(오른쪽) 가우시안 곡선은 변화 지역을 나타낸다. 한계값 결정을 방법을 좀 더 자세히 설명하기 위해 그림 4 (d)에서 한계값 부분을 확대하여 보여주고 있으며, 두 곡선이 만나는 지점에서 한계값이 결정된다. 그림 5 (a), (b), (c), (d)는 한계값 설정을 통해 추출된 변화지역의 2진 영상이다.

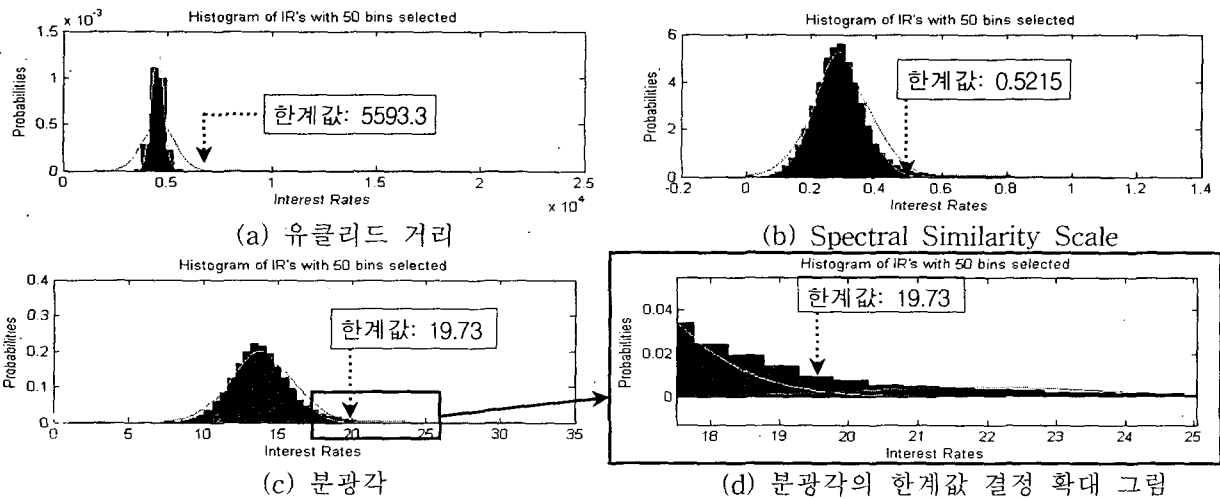


그림 4. 기대최대화 기법 적용



그림 5. 한계값을 적용하여 추출된 변화지역 2진 영상(흰색 부분이 변화된 지역임)

경우 I은 가상의 영상을 이용해 변화지역을 임의로 첨가한 것이므로 레퍼런스 영상을 통해 정확도 평가를 수행할 수 있다(그림 5 (e)). 정확도 평가는 일반적으로 분류와 마찬가지로 오차행렬을 이용한다. 하지만, 변화탐지 결과에서 변화된 지역의 면적이 변화되지 않는 지역보다 적게 분포하는 것이 일반적이다. 이 경우 전체 정확도가 높게 평가되어 잘못된 정보를 보여줄 수도 있다. 경우 I에 사용된 변화지역의 비율은  $944/50,000=1.888\%$ 로서 만약 변화지역이 모두 변화되지 않는 지역으로 잘못 탐지하더라도 전체정확도는  $98.112\%$ 로 나타나며, 변화되지 않은 지역까지 잘못 탐지된다 하더라도  $96.224\%$

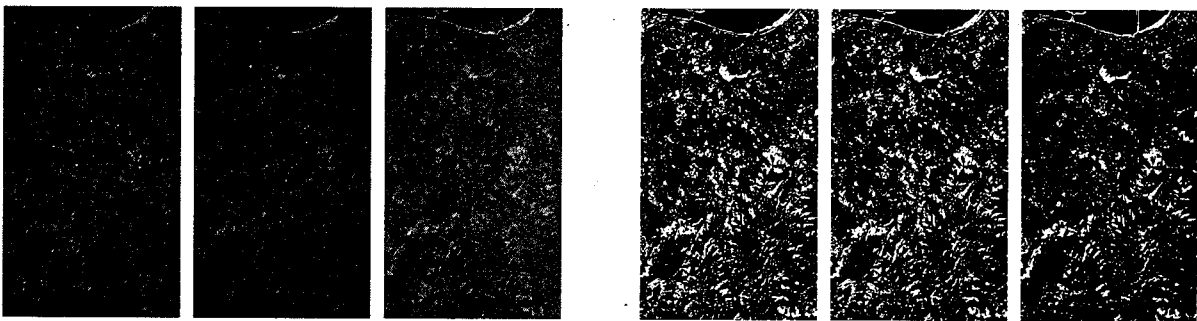
(=100-1.888×2) 이상이 된다. 이런 맹점을 해결하기 위해 Richard J. R. 등(2005)이 제안한 PCC(Percentage Correct Classification), JC(Jaccard Coefficient), YC(Yule Coefficient) 값을 이용하여 정확도를 평가하였다. 값이 높을수록 변화탐지 정확도가 높다는 것을 의미한다.

표 2 결과 정확도

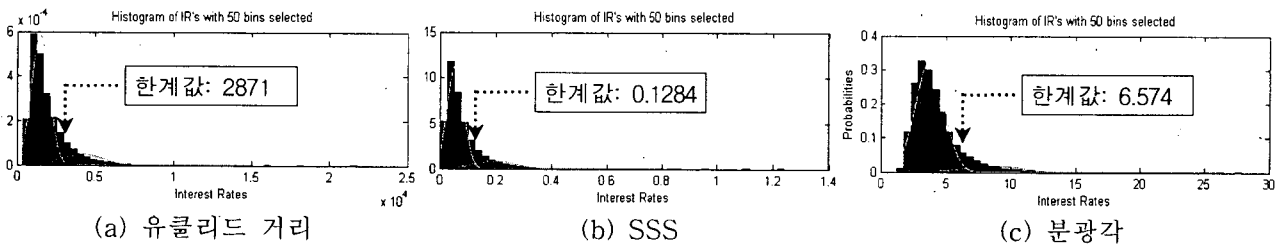
	PCC(%)	JC(%)	YC(%)	변화지역 개수	변화지역 비율(%)
유클리드 거리	99.910	95.293	96.480	923	1.846
SSS	99.030	63.285	87.791	1064	2.128
분광각	99.680	50.525	70.603	1213	2.426
단밴드 영상차	96.708	6.848	11.140	944	1.888

표 2에서 확인할 수 있듯, 결과 정확도는 유클리드 거리 > SSS > 분광각 > 단밴드 영상차 순으로 나타난다. 이 결과를 통해 하이퍼스펙트럴 영상과 같은 많은 밴드의 정보는 변화탐지에 있어 효과적으로 사용될 수 있을 것으로 판단되며, 제안된 알고리즘이 단밴드 영상차 결과보다 우수한 변화탐지 능력을 발휘할 수 있음을 확인하였다. 유사도 측정 기법 간의 결과에서는 유클리드 거리가 높은 정확도를 나타냈는데, 이는 부여된 노이즈의 형태가 크기(Magnitude)와 관련 있기 때문으로 판단된다. 실제 노이즈의 형태가 실험을 위해 부가된 노이즈와는 다른 특성을 지닐 수 있으므로 유클리드 거리 기법이 SSS, 분광각 기법보다 성능이 우수하다고 볼 수는 없다. 임의로 부여한 변화 지역의 화소개수는 944개였고, 기대최대화 기법을 통해 추출된 화소 개수는 각각 923개, 1064개, 1213개로서 유사하게 추출하는 것을 볼 수 있다. 이는 유사도 측정 기법을 통한 무감독 변화탐지에 기대최대화 기법이 효과적으로 사용될 수 있음을 의미한다.

### 3.3 알고리즘 적용 결과 - 경우 II



(a) 유클리드 거리 (b) SSS (c) 분광각 (d) 유클리드 거리 (e) SSS (f) 분광각  
그림 6. 알고리즘 적용 결과(a, b, c) 및 한계값을 적용하여 추출된 변화지역 2진 영상(d, e, f)



(a) 유클리드 거리 (b) SSS (c) 분광각  
그림 7. 기대최대화 기법 적용

경우 II는 16일 차이의 Hyperion 영상에 유사도 측정 기법을 적용한 것이다. 그림 6과 표 3은 두 시기 데이터에 각각 유클리드 거리, SSS, 분광각을 적용한 결과 영상과 값이다. 실제 레퍼런스 데이터가 존재하지 않아 정량적인 정확도 평가를 수행할 수 없었음을 밝히며, 4월이라는 시기적인 특성과 16일차 영상이라는 데이터 특성을 이용하여 시각적인 평가만을 실시하였다. 세 기법 모두 변화지역을 비슷한 양상으로

추출하였고, 특히 산림지역과 강의 경계부에서 변화탐지는 거의 동일한 결과를 보여주고 있다. 이는 4월의 시기적인 특성상 산림지역의 활엽수 생장으로 인한 변화로 판단되며, 향후 임상도를 이용하여 정확도 평가를 수행할 예정이다. 유클리드 거리와 SSS 기법에서 도심지가 과다하게 변화지역으로 추출된 것을 확인할 수 있는데, 16일의 시기동안 도심지의 변화가 심하게 나타나지는 않을 것이므로 변화지역을 잘못 추출한 것으로 볼 수 있다. 이러한 변화지역의 과다 추출은 한계값 설정이 틀린 것으로 추론할 수 있으며, 그림 7에서 볼 수 있듯 기대최대화 기법 사용 시 가정한 데이터의 가우시안 분포에 문제가 있는 것으로 사료된다. 따라서 가정한 분포 확률 모델을 결과에 맞춰 부분적인 수정이 이루어져야 할 필요가 있다.

표 3 결과 정확도

	알고리즘 적용 결과값		한계값 적용 후	
	최대	최소	변화지역 개수	변화지역 비율(%)
유클리드 거리(화소값)	21779	380.66	11522	23.044
SSS(단위없음)	1.2416	0.0010	11137	22.274
분광각(°)	27.890	1.2293	6926	13.848

#### 4. 토의 및 향후 연구

본 연구는 SNR이 낮은 하이퍼스펙트럴 특성에 맞는 변화탐지 알고리즘을 유사도 측정 기법을 이용하여 그 결과를 평가해 보았다. 많은 밴드의 정보를 모두 사용할 수 있다는 점에서 제안된 알고리즘은 그 사용 가능성을 확인할 수 있었으며, 임의적으로 부가된 노이즈에 대해서 양호한 변화탐지 결과를 보여주었다. 기대최대화 기법 또한 유사도 측정 기법을 통해 추출된 값의 특성에 맞는 자동화된 한계값을 결정하는 것으로 보이며 무감독 변화탐지에 효과적으로 적용 될 것으로 기대된다. 경우 I의 결과와는 달리 경우 II의 실험에서는 정확도 평가를 수행하지 못해 그 성능을 정량적으로 평가하지 못한 한계가 있으나, 4월이라는 시기적인 특성과 16일 차이 영상이라는 데이터의 특성을 고려하여 활엽수 및 도심지의 변화와 관련된 결과를 추정할 수 있었다. 향후 활엽수 지역의 변화탐지 성능을 임상도를 통해 정확도 평가를 수행해 보고자 한다. 또한 더욱더 세밀한 한계값 결정을 위해 기대최대화 기법에서 가정한 확률 모델에 대한 고려가 필요할 것으로 보인다.

#### Reference

- 박노옥, 지광훈, 이광재, 권병두 (2003), 다중시기 원격탐사 화상의 변화탐지를 위한 임계치 자동 추정, *Korean Journal of Remote Sensing*, Vol. 19, No. 6, pp. 465-478.
- D. Lu, P. Mausel, E. Brondizio and E. Moran (2004), Change Detection techniques, *International Journal of Remote Sensing*, Vol. 25, No. 12, pp. 2365-2407.
- Emmanuel Christophe, Dominique Leger and Corinne Mailhes (2005), Comparison and Evaluation of Quality Criteria for Hyperspectral Imagery, *Proceedings of SPIE*, Vol. 5668, pp. 204-213.
- Farid M., Gabriele M., Sebastiano B. S. (2002), Unsupervised Change-Detection Method for Remote-Sensing Image, *Proceedings of SPIE, Image and Signal Processing for Remote Sensing VII*, Vol. 4541, pp. 211-222.
- Freek D. Van Der Meer and Steven M. Dejong (2001), *Imaging Spectrometry - Basic Principles and Prospective Applications*, Kluwer Academic Publishers, Dordrecht, pp. XXI-XXIII
- Lorenzo B. and Diego F. P. (2000), Automatic Analysis of the Difference Image for Unsupervised Change Detection, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 38, No. 2, pp. 1171-1182.
- Richard J. R., Srinivas A., Omar A. and Radrinath R. (2005), Image Change Detection Algorithms: A Systematic Survey, *IEEE Transactions on Image Processing*, Vol. 14, No. 3, pp. 294-307.
- Singh A. (1989), Digital Change Detection Techniques Using Remotely Sensed Data, *International Journal of Remote Sensing*, Vol. 10, pp. 989-1003.
- Sweet J. N. (2003), The Spectral Similarity Scale and its Application to the Classification of Hyperspectral Remote Sensing Data, *IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, pp. 92-99.