

반복최적화 무감독 분광각 분류 기법을 이용한 하이퍼스펙트럴 영상 분류에 관한 연구

A Study on the Hyperspectral Image Classification with the Iterative Self-Organizing Unsupervised Spectral Angle Classification

조현기* · 김대성** · 김용일***

Jo, Hyun Gee · Kim, Dae Sung · Kim, Yong Il

* 서울대학교 대학원 지구환경시스템공학부 석사과정(E-mail: hkcho79@hotmail.com)

** 서울대학교 대학원 지구환경시스템공학부 박사과정(E-mail: mutul94@empal.com)

*** 서울대학교 공과대학 지구환경시스템공학부 부교수(E-mail: yik@plaza.snu.ac.kr)

요지 분광각(Spectral Angle)을 이용한 분류는 같은 종류의 지표 대상물의 분광 특성이 대기 및 지형적인 영향으로 인해 원점을 기준으로 선형적인 분포 모양을 가진다는 가정에 기초한 새로운 접근의 분류 방식이다. 최근 분광각을 이용한 무감독 분류에 대한 연구가 활발히 이루어지고 있으나, 원격탐사 데이터의 특성을 반영한 효과적인 무감독 분류에 대한 연구는 미진한 상태이다. 본 연구는 하이퍼스펙트럴 영상 분류에 있어서 기존 무감독 분광각 분류(USAC, Unsupervised Spectral Angle Classification) 연구에서 해결하지 못한 문제점들을 보완한 반복최적화 무감독 분광각 분류(ISOUSAC, Iterative Self-Organizing USAC) 기법을 제안하고 있다. 이를 위해, 무감독 분광각 분류에 적합한 각 분할(Angle Range Division) 기법을 적용하여 군집 초기 중심을 설정하였으며, 병합(Merge)과 분할(Split)을 통한 유동적인 군집 분석을 수행하였다. 결과를 통해, 제안된 알고리즘이 기존의 기법보다 수행 시간뿐 아니라 시각적인 면에서도 우수한 결과를 도출함을 확인할 수 있었다.

1. 서 론

모든 지표 대상물은 빛의 반사 및 흡수 등으로 인해 각 파장에 대해서 고유한 밝기값을 나타내기 때문에, 같은 종류의 지표 대상물들은 비슷한 분광반사 패턴을 가진다(Youngsinn, 2002). 하이퍼스펙트럴 데이터는 이러한 지표 대상물의 파장 특성을 높은 분광해상도를 이용하여 효과적으로 취득함으로써 다중분광 데이터보다 좀 더 세분화된 분석이 이루어질 수 있다. 하지만, 많은 밴드

수와 신뢰도가 낮은 트레이닝 데이터의 취득으로 인하여 영상을 분류함에 있어서 오히려 낮은 결과 정확도를 나타내는 현상이 발생하기도 한다. 이런 문제를 해결하기 위해 최근에는 하이퍼스펙트럴 데이터에 효과적인 새로운 알고리즘에 대한 연구가 활발히 이루어지고 있다.

분광각(Spectral Angle)을 이용한 원격탐사 데이터의 분석은 같은 종류의 지표 대상물들의 분광특성은 대기 및 지형적인 영향으로 인해 원점을 기준으로 선형적인 분포 모

양을 가진다는 특성에 기초한 새로운 접근 방식이다(Youngsinn, 2002). 하지만, SAC에 관한 연구들은 주로 감독 분류에 적용되었으며 (Kruse et al, 1993b; Youngsinn, 2002), 군집(Cluster) 중심을 계산하는데 있어 거리 개념을 사용하였기 때문에 완전한 분광각 기반 분석이 이루어졌다고 볼 수 없다 (Kai-Yi, 2002; Arel Weisberg, 1999).

본 연구는 기존의 분광각 분류 연구를 토대로 개선된 분광각 분류 기법을 무감독 분류에 적용하였으며, 기존의 연구와는 달리 군집 중심을 결정하는데 있어 단위벡터(Unit Vector)를 이용한 각 평균을 사용함으로써, 분광각을 이용한 무감독 분류 기법이 실제 원격탐사 데이터에 좀 더 효과적인 무감독 분류 기법임을 증명하였다. 또한 군집 초기 중심(Seed Points)을 결정하는 방법을 제시하였고, 병합(Merge)과 분할(Split)을 통한 유동적인 군집 분석을 수행하여 더욱 개선된 분광각 분류 알고리즘을 개발하였다.

2. 알고리즘

2.1 무감독 분광각 분류

무감독 분광각 분류는 같은 종류의 지표 대상물의 분광특성은 대기 및 지형적인 영향으로 인해 원점을 기준으로 선형적인 분포를 가진다는 특성에 기초하여 유사한 분광각을 지니는 화소들을 군집화한다. 분광차원에서 최소 거리 개념(Spectral Distance Rule)을 기반으로 분류를 수행하는 K-means, ISODATA, Hierarchical Clustering 등의 무감독 분류 기법과는 달리 무감독 분광각 분류는 최소 분광각 개념(Spectral Angle Rule)에 근거하여 분류를 수행한다. 무감독 분광각 분류는 다음의 식을 통해서 화소와 군집의 중심 간의 분광각($\theta_{i,c}$)을 구한다.

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

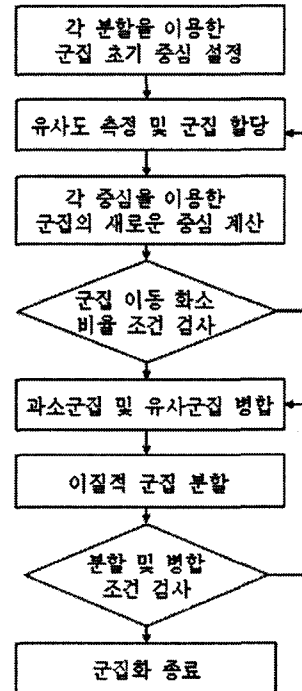
여기서,

$x_{i,k}$: 임의 화소의 k 밴드 화소값

$\mu_{c,k}$: c 군집 중심의 k밴드 화소 값

m : 밴드 수

2.2 반복최적화 무감독 분광각 분류



<그림 1> 반복 최적화 무감독 분류 과정

반복최적화 무감독 분광각 분류는 최소 분광각 법칙을 기반으로 분류를 수행한다는 점에서 무감독 분광각 분류와 유사하다고 할 수 있다. 그러나 군집화 과정 중 군집의 중심을 계산하는 단계에서 무감독 분광각 분류는 유클리드 거리(Euclidean Distance)를 이용하여 클러스터의 중심을 구함으로써, 완전한 분광각 분석이 이루어졌다고 볼 수 없다. 반면 반복최적화 무감독 분광각 분류는 각을 이용하여 군집 중심을 구하므로, 완전한 분광각 분석을 수행할 수 있다(김대성, 2005). 또한 군집 초기 중심을 설정 시, 각 분할 방법을 적용하고, 군집화 종료 후, 병합과 분리를 통한 유동적인 군집 분석을 수행함으로써 기존의 무감독 분광각 분류를 보완한 새로운 알고리즘이다. 반복최적화 무감독 분광각 분

류 과정은 <그림 1>에서 설명하고 있다.

2.2.1 각 분할 초기 군집 중심 설정

각 분할 방법은 모든 화소에 대해서 각각의 밴드와 화소가 이루는 각 중 최대각과 최소각을 검색하여 최대각과 최소각을 사용자가 지정한 군집의 수만큼 등각으로 분할하여 군집의 초기 중심(Seed Point)을 지정하는 기법이다. 먼저 각 화소가 밴드와 이루는 각($\theta_{i,k}$)을 계산하여, 각 밴드에 대해서 최대각과 최소각을 등각으로 나누어 군집의 초기 중심각($\theta_{c,k}$)을 구한다(식2). 각각의 중심각 벡터에 해당하는 단위 벡터의 밴드별 좌표값을 군집 초기 중심 화소값으로 지정한다. 여기서, n 은 사용자가 지정한 군집의 개수이며, C 는 C 번째 군집을 말한다.

$$\theta_{c,k} = \theta_{\min,k} + \frac{\theta_{\max,k} - \theta_{\min,k}}{n-1} \times (C-1) \quad (2)$$

2.2.2 유사도 평가

유사도 평가 단계에서는 최소 분광각 범칙에 근거하여, 먼저 임의의 화소와 군집 중심과의 분광각을 계산한 후, 임의의 화소를 최소 분광각을 가지는 군집에 할당한다. 임의의 화소와 군집 중심과의 분광각($\theta_{i,c}$) 계산은 앞에서 언급한 식 1을 통해서 구할 수 있다.

2.2.3 군집 중심 계산

군집 중심을 계산하기 위해 군집에 할당된 화소에 대하여 각 밴드축과 이루는 각을 계산한 후(식 3), 군집에 할당된 모든 화소의 각 밴드축과 이루는 각을 평균한다(식 4). 끝으로 단위 벡터와의 곱을 통해 군집의 중심을 구한다(식 5).

$$\theta_{i,k} = \cos^{-1}\left(\frac{x_{i,k}}{\sqrt{\sum_{k=1}^m x_{i,k}^2}}\right) \quad (3)$$

$$\mu_{\theta_{i,k}} = \frac{1}{N_c} \sum_{i=1}^{N_c} \theta_{i,k} \quad (4)$$

$$\mu_{c,k} = \text{unitvector} \times \cos(\mu_{\theta_{i,k}}) \quad (5)$$

여기서,

N_c : 군집에 할당된 화소수

$$\text{unitvector} = \sqrt{\sum_{k=1}^m p_{i,k}^2} = 1$$

2.2.4 병합

군집에 포함된 화소 수가 사용자가 지정한 임계값보다 작거나(식 6), 두 군집이 매우 유사한 경우 즉, 두 군집의 중심 간 분광각이 사용자가 지정한 임계값보다 작을 경우(식 7) 두 군집을 병합하도록 설정하였다.

$$\text{if } N_c < \text{Threshold} \rightarrow \text{Merge} \quad (6)$$

$$\text{if } \theta_{\mu_i, \mu_j} < \text{Thereshold} \rightarrow \text{Merge} \quad (7)$$

2.2.5 분할

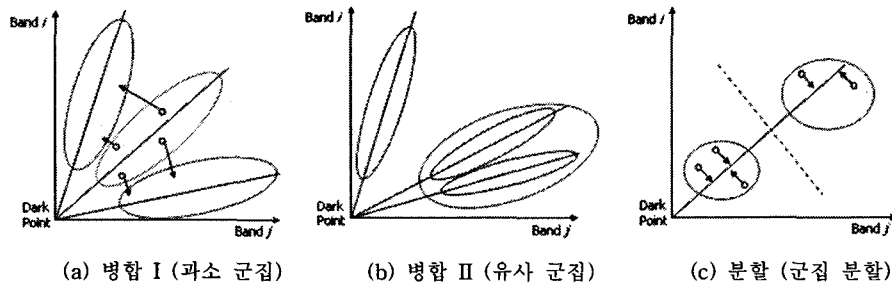
한 군집 내에 이질적인 화소 집단이 존재할 경우 군집 분할을 수행한다. 먼저 군집 내에 포함된 각각의 화소를 군집 중심 벡터에 투영 시킨 후(식 8), 투영된 각 화소에 대해서 원점으로부터의 거리를 계산하여 평균 및 표준 편차를 구한다(식 9). 이때 표준편차가 사용자가 지정한 임계값보다 클 경우 군집의 중심을 기준으로 군집을 분할한다.

$$pd_{x_{i,c}} = \sum_{k=1}^m x_{i,k} \mu_{c,k} / \sum_{k=1}^m \mu_{c,k}^2 = \sum_{k=1}^m x_{i,k} \mu_{c,k} \quad (8)$$

다음 군집에 포함된 화소들에 대해서 원점으로부터 군집 중심에 투영된 위치까지 거리에 관한 평균과 표준 편차를 구하여, 사용자가 지정한 임계치보다 큰 경우 군집의 중심에서 분할을 수행한다.

$$\text{if } \text{Std}(pd_{x_{i,c}}) > \text{Tereshold} \rightarrow \text{Split at mean}(pd_{x_{i,c}}) \quad (9)$$

위의 방법으로 군집 분할을 수행함으로써 유사한 분광각을 가지나 원점으로부터의 거리가 매우 다른 화소를 동일한 군집으로 구분한다는 기존 기법의 단점을 개선할 수 있다. 그림 2는 병합과 분할을 수행하는 원리를 나타낸 것이다.



<그림 2> 병합 및 분할 원리

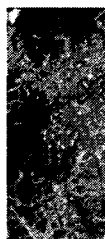
3. 알고리즘의 적용 및 결과

3.1 사용 데이터

본 연구에서는 EO-1에 탑재된 Hyperion 센서에 의해 취득된 영상을 데이터로 사용하였다. Hyperion 센서는 pushbroom 방식 센서로 약 7.7km 관측폭을 가지고, 30m 공간 해상도와 0.4 μ m~2.5 μ m의 분광해상도를 가지며, 242개의 밴드로 구성되어 있다. 대상 지역은 서울, 경기 지역으로 2004년 4월 3일에 취득된 영상이다. 그림 3은 연구 대상 지역과 사용된 Hyperion 영상이다.

3.2 연구 결과

제안된 알고리즘의 유효성을 검증하기 위해 두 가지 방법으로 그 결과를 분석하였다. 우선 군집 초기 중심 설정에 있어 각분할 방법이 더 효과적인지에 대해 결과의 시각적인 분석과 반복 연산 횟수를 통해 단일 패스 방법(Single Pass) 및 범위 분할 방법(Range Division)과 비교해 보았다. 다음으로 반복최적화 무감독 분광각 분류가 K-means 및 기존 무감독 분광각 분류보다 우수한 결과를 도출할 수 있는지 확인하였다.



<그림 3>. 대상 지역의 Hyperion 영상

3.2.1 각분할 군집 초기 중심 설정 결과

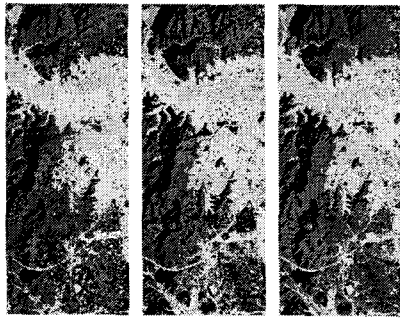
군집 초기 중심 설정에 있어서 군집 11개, 12개에 대해 각분할 방법, 단일 패스 방법, 범위 분할 방법의 반복 연산 횟수를 비교하였으며, 군집 개수 11개 및 12개에 대해서 수계 지역 분류 결과를 비교해 본 결과는 표 1과 같다. 단일 패스 방법을 적용한 결과 도시 지역 어두운 부분이 수계 지역으로 오분류가 발생하였고, 범위 분할 방식과 각분할 방식은 수계지역을 잘 분류하였으나, 연산 횟수에 있어서 각분할 방식이 보다 효율적임을 확인할 수 있었다.

<표 1> 반복 연산 횟수 비교 결과

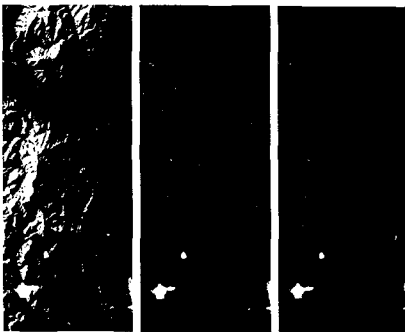
	군집개수 11	군집개수 12
단일 패스	28	33
범위 분할	38	42
각 분할	25	27

3.2.2 반복최적화 무감독 분광각 분류 결과

반복최적화 무감독 분광각 분류 결과의 유효성을 확인하기 위해 K-means 및 기존 무감독 분광각 분류 결과를 시각적으로 평가하였다. 군집의 구분이 명확한 수계 지역에 대해 그 결과를 비교하였으며, K-Means 분류는 어두운 산림 지역과 수계 지역을 구분하지 못하는 데 반해 기존의 무감독 분광각 분류와 반복최적화 무감독 분광각 분류는 수계 지역을 잘 구분한 것을 확인할 수 있었다. <그림 4>는 각 알고리즘의 결과 영상을 나타내며, <그림 5>는 수계 지역에 대한 분류 결과이다.



(a) K-Means (b) USAC (c) ISOMUSAC
 <그림 4> 분류 결과



(a) K-Means (b) USAC (c) ISOMUSAC
 <그림 5> 수계 지역 분류 결과

4. 결 론

본 연구에서는 기존 무감독 분광각 분류 기법의 한계점인 거리 중심을 사용함으로써 완전한 분광각 분석이 이루어지지 못한 점을 보완하여 각을 이용한 군집 초기 중심 설정 및 군집 중심 계산을 수행하였으며, 보다 유동적인 군집 분석을 수행하기 위해 ISODATA 분류에 기반한 병합과 분할과정을 수행한 반복최적화 무감독 분광각 분류를 제안하였다. 각 분할을 이용한 군집 초기 중심 설정은 수계 지역 분류 및 효율성 면에서 보다 개선된 결과를 보였다. 최종 분류 결과는 K-means 보다 수계 지역의 분류에 있어 보다 좋은 결과가 나타났으며, 기존의 무감독 분광각 분류 기법을 보완한 개선된 결과를 산출하는 것을 확인할 수 있었다. 향후 정량적 정확도 평가 및 거리 개념을 접목한 무감독 분광각 분류 알고리즘에 대한 연구가 수행되어야 할 것으로 판단된다.

< 참고 문헌 >

1. 김대성, 김용일, "The Modified Unsupervised Spectral Angle Classification (MUSAC) of Hyperion, Hyperion-FLAASH and ETM+ Data Using Unit Vector", ISRS, pp 134-137, 2005.
2. Arel Weisberg, Michelle Najarian, Brett Borowski, Jim Lisowski, Bill Miller, "Spectral Angle Automatic cLuster routine(SAALT) : An Unsupervised Multispectral Clustering Algorithm", Aerospace Conference, IEEE, Vol.4, pp. 307-317, 1999.
3. Kai-Yi Huang, "A Synergistic Automatic Clustering Technique (SYNERACT) for Multispectral Image Analysis", PE & RS, ASPRS, Vol. 68, No.1, pp. 33-49, 2002.
4. P.K. Agarwal, N.H. Mustafa, "K-Means Projective Clustering", PODS, pp. 155-165, 2004.
5. S.Phillips, "Reducing the Computation time of the ISODATA and K-means unsupervised classification algorithms", in Proc. 22nd IEEE Intl. Ceosci. and Remote Sensing Sump., vol. 3, pp. 1627-1629, 2002.
6. T. Kanungo, D. M. Dount. N. S. Netanyahu, C. Piatko, R. Silverman, A. Y. Wu, "An Efficient K-means Clustering Algorithm : Analysis and Implementation", IEEE Trans, Pattern Anal. Mach, Intell., vol. 24, pp. 881-892, 2002.
7. Youngsinn Sohn and N. Sanjay Rebello, "Supervised and Unsupervised Spectral angle Classification", PE & RS, ASPRS, Vol. 68, No. 12, pp. 1271-1280, 2002.