

# 지형 분석을 위한 영역 기반 장면 분할 기법

전태균\*, 전문구\*

\*광주과학기술원 정보통신공학과

e-mail : tgjeon@gist.ac.kr

## Region based Scene Segmentation method for Topography Analysis

Taegyun Jeon\*, Moongu Jeon\*

\*Dept. of Information and Communications, Gwangju Institute of Science and Technology

### 요 약

본 연구에서는 일반 야외 영상 및 항공 시뮬레이션 영상에 대한 지형 분석을 위해 영역 기반 장면 분할 기법을 제시한다. 영역의 분류를 위해 MeanShift 기법을 기반으로 한 표현과 Texton, SIFT, 위치정보를 특징으로 하는 기법을 제안하고 실험을 통해 주요 대상 영역이 분할되는 결과를 보인다. Sowerby 데이터 셋과 Google Earth 데이터로부터 자체적으로 제작한 데이터 셋에 대해 실험하였으며 수풀지형, 초목지형, 도로 등에 대해 분류하였다.

### 1. 서론

장면 분할은 입력 영상에서 나타나는 각기 다른 의미적 물체 및 공간에 대한 분리하는 기술이며 최근 감시, 정찰, 자료수집, 안내 등의 다양한 지능형 영상 처리 분야에 응용되고 있다. 지형 분석은 지리정보시스템(GIS)에서 최종적인 목표로 삼고 있는 형태이며, 지형에 관한 입력 영상이 주어질 때, 각 영역의 지형적인 특징을 바탕으로 해당 지형이 속하는 클래스로 분류하고자 하는 것이 목표이다. 이러한 지형 분석은 사용하는 센서에 따라 전기광학(Electro Optical) 카메라, LiDAR 등과 같은 레이더로 나뉜다. 전기광학 카메라는 가격이 저렴하고 촬영이 용이한 편이나 기후 변화 및 야간 상황에서 영상획득에 어려움이 있다. 반면 LiDAR 등과 같은 레이더는 지상 표면의 높낮이를 바탕으로 정확한 자료 획득이 가능하나, 장비의 가격과 운용에 대한 문제가 있다. 또한 관측되는 위치에 따라 위성영상, 항공영상, 저고도 UAV 와 같은 다양한 관측시스템에서 취득된 영상정보를 이용한다. 하지만, 사전 촬영계획에 따라 특정 영역 및 일정한 촬영방향으로 공간영상정보를 취득하는 기존의 위성이나 항공관측 기술은 재난상황이나 긴급한 상황에 있어서 빠른 영상획득이 힘들다. 반면, 저고도 UAV는 좁은 지역에 대해 최신의 영상을 획득할 수 있는 장점이 있다. 본 논문의 구성은 다음과 같다. 먼저 2장에서는 장면 분할과 관련된 연구를 분석하고, 3장에서는 영역 기반 장면 분할 기법을 제안한다. 4장에서는 야외, 항공 데이터셋에 대한 실험 결과를 보이며, 마지막 5장에서는 결론으로 마무리 한다.

### 2. 관련연구

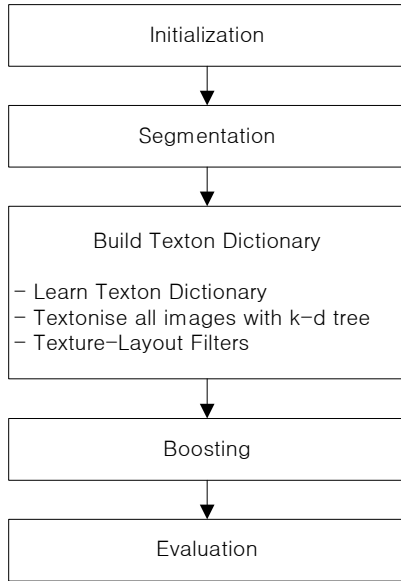
의미론적인 장면 분할에 관한 연구는 크게 외형적 특징과 개체 추적을 이용한 두 방향으로 분류된다[1].

첫 번째로, 외형적 특징(Visual feature)기반의 방식은 색상, 재질, 외곽선, 형태, 위치 등의 특징점을 중심으로 영상에 나타나는 영역을 Bag-of-Word 특징 벡터로 표현하고, 이를 학습 및 분류를 통해 장면을 분할하게 된다. 이 방법은 장면에 나타나는 각 영역 및 물체가 보기 좋게 잘 나타나도록 해상도가 뒷받침되어야 하며, 학습단계를 위해 모든 영상에 대해 픽셀 및 영역 단위 별 레이블이 주석되어야 한다. 또한 시점, 조명, 크기, 각도에 변함없는 특징점이 요구된다. 두 번째는, 추적을 기반으로 하는 장면분할 기법이다. 이 방법은 고정된 카메라 영상에 대해 객체가 이동 가능한 영역에 대한 가정을 두고 추적을 통해 정보를 획득한다. 물체에 대한 크기, 위치, 속도, 방향 등이 주요 특징이 된다. 하지만 항공영상과 같이 이동하며 촬영되는 영상에 대한 장면 분할에는 적합하지 않다.

### 3. 영역 기반 장면 분할 기법

본 연구에서는 앞서 설명한 바와 같이 외형적 특징을 사용하는 영역 기반 장면 분할 기법을 제안하고자 한다. 본 논문에서 제안한 장면 분할 처리과정은 아래 (그림 1과 같다. 초기화 단계에서는 데이터 셋의 모든 이미지에 대해서 Filterbank 를 적용한다. 여기서 사용되는 Filterbank 는 총 17 차원이며 Gaussians, Laplacians, Derivatives of Gaussians 과 같이 총 3 가지 종류의 필터를 적용한 결과를 종합한다. Gaussians 필터는 Lab 색상공간에서 각 채널 별로  $\sigma = 1, 2, 4$  를 적용한다. Laplacians 필터는 L 채널에만 한하여  $\sigma = 1, 2, 4, 8$  를 적용한다. Derivative of Gaussians 는 L 채널에만 한하여  $\sigma = 2, 4$  와 x 축/y 축으로 각각 적용하여 총 4 개의 결과가 나오게 된다.

Segmentation 단계에서는 각 영역의 최적화된 표현을 위해 추상화된 superpixel 의 표현을 위해 Meanshift segmentation 기법[2]을 적용하였다.



(그림 1) 영역 기반 장면 분할 기법 처리 과정

Textons 을 생성하는 단계에서는 크게 3 단계의 세부 과정을 거치게 된다. 우선 Textons 의 사전을 학습하는 단계에서 전체 이미지의 각 픽셀별 필터 반응 결과들을 군집화하여 대표 Textons 을 추출하게 된다. 본 연구에서는 K=400 으로 수행하였다. Textonise 과정에서는 학습 및 평가 데이터 셋 모든 이미지에 대해 각 픽셀마다 나온 필터 반응을 가장 가까운 대표 Textons 과 매핑을 하기 위해 k-d tree 를 사용하였다. 이 단계에서 Textons map 이 생성된다. Texture-Layout filter 는 각 특징점에 대해 임의의 크기를 가지는 100 개의 직사각형  $r_i$  과 Texture  $t_i$  에 대해 전 단계에서 나타난 Textons map 에 임의의 위치에 해당하는 반응들과 비교하게 된다. 이때 Textons map 에서  $r_i$  내에서  $t_i$  가 차지하는 비중을 0 과 1 사이의 확률로 나타내어 texture 와 layout 간의 반응을 얻을 수 있다.

Boosting 단계에서는 각각의 weak learners 가 분류하는 오류를 최소화하는 단계를 반복적으로 수행함으로써 분류기를 학습하게 된다. Textons 특징점 이외에도 물체 인식에서 가장 널리 쓰이는 SIFT[3] 특징과 해당 클래스가 이미지에서 보편적으로 나타나는 위치에 대한 가정을 기반으로 사용되었다.

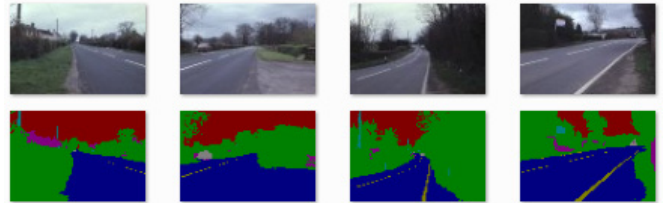
평가 단계에서는 평가 데이터에서 나타나는 각 클래스의 픽셀 대비 예측한 클래스의 픽셀의 일치 여부를 누적하여 평가하였다. 실험은 픽셀 레벨에서 위 처리과정을 거친 결과와 Segmentation 단계에서 세부적으로 나눈 영역에 나타난 모든 픽셀에 대해 가장 많은 빈도수를 보인 클래스로 종합하여 분류하는 영역 레벨 분류로 2 가지 방법으로 수행하였다.

#### 4. 실험결과

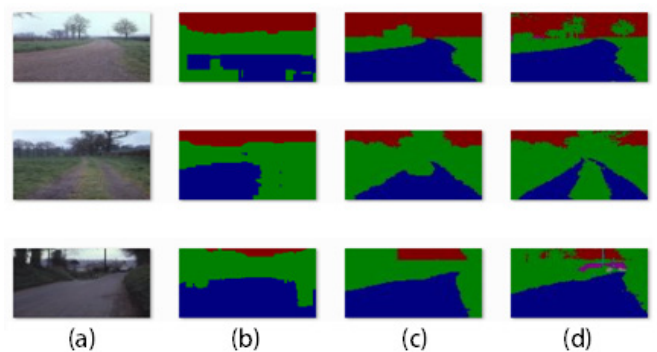
##### 4.1 실험 데이터셋

본 논문에서는 실험한 데이터 셋은 야외 장면을 포함하고 두 종류이다. 첫 번째는, Sowerby 데이터 셋[4]이며 실외 도심지역의 장면을 담고 있으며, 검증자료

데이터를 제공하고 있다. 하늘, 초목지, 도로, 도로표지, 건물, 자동차, 전신주로 구성된 7 개의 카테고리들을 포함하고 있으며, 각 카테고리는 검증자료에서 서로 다른 색상으로 픽셀 단위로 표현 되어있다. 주요 영상 및 검증자료는 아래 (그림 2와 같다.

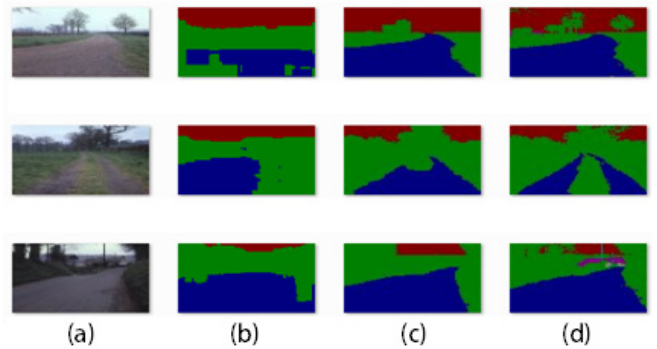


(그림 2) Sowerby Image Dataset



(그림 3) Sowerby Image Dataset 처리 결과

위

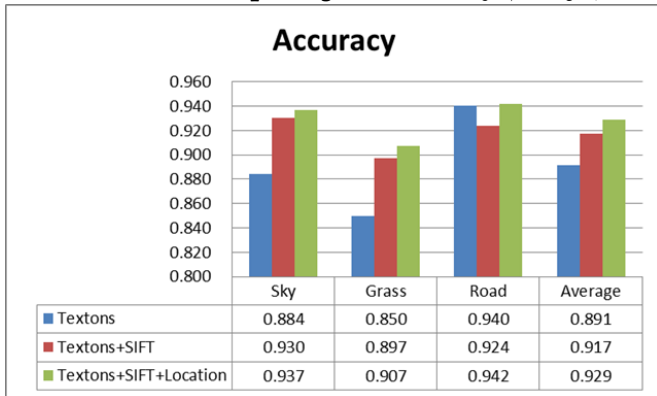


(그림 3) Sowerby Image Dataset 처리 결과

은 Sowerby 데이터 셋에 대한 실험 결과를 나타낸다. (a)열은 원본 이미지, (b)열은 픽셀 레벨에서 수행한 결과, (c)열은 영역 레벨에서 픽셀 레벨 수행결과를 종합한 결과이다. (d)열은 검증자료이다.

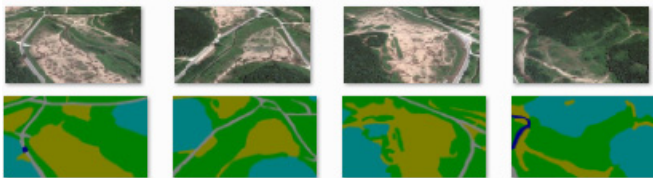
픽셀 레벨에서는 영역의 위치나 어느 정도의 형태는 분류가 가능하나 정확한 영역의 윤곽이나 대부분의 미세함이 많이 떨어짐을 보인다. 하지만, 영역 기반의 실험 결과는 검증자료와 비교했을 때 넓은 영역에 나타나는 클래스에 대해서는 거의 흡사한 결과를 나타낸다. 하지만 미세한 도로나 나무 부분은 영상 해상도에서 일정 크기에 미치지 못하여 정확히 분류가 어려움을 알 수 있다.

<표 1> Sowerby Image Dataset 정확도 평가



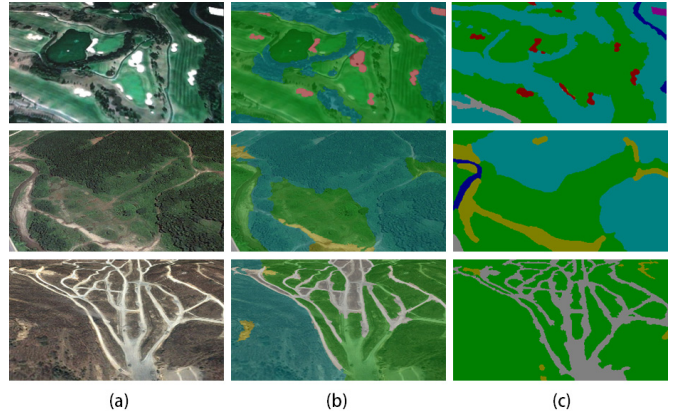
위 <표 1>에서는 Sowerby 데이터 셋에서 주요 관심 대상인 하늘, 수풀지형, 도로에 대한 인식 정확도에 대해 나타내었다. Texton 특징점만 사용하는 것보다 SIFT를 같이 사용함으로써 성능에 향상이 있었고, 위치 특징 또한 소폭적이지만 성능 향상에 도움이 됨을 알 수 있다.

두 번째 데이터 셋은 본 연구에서 자체 제작한 시뮬레이션 영상이다. Google Earth[5]를 통해 지상과의 고도는 평균 300m, 지표면과의 각도는 30~40도를 유지하는 경사영상으로 이루어져있다. Sowerby 데이터셋과 같은 방식으로 본 연구팀이 직접 손으로 해당 레이블에 대해 색상으로 검증자료를 표현하였다. 주요 영상 및 검증자료는 아래 (그림 4과 같다.



(그림 4) Google Earth Image Dataset

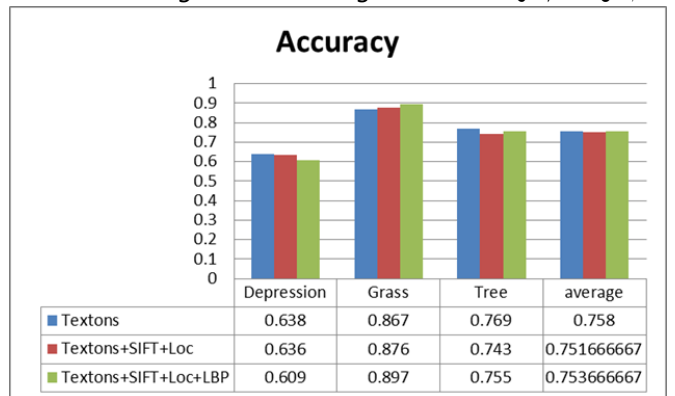
아래 (그림 5에서는 Google Earth 데이터 셋에 대한 실험 결과를 나타낸다. (a)열은 원본 이미지, (b)열은 영역 레벨에서 수행결과이며, (c)열은 검증자료다. 실제 지형에 대한 항공 영상 시뮬레이션 데이터이기 때문에 Sowerby 데이터 셋보다 영상에 나타나는 지형이 복잡하며 각 클래스 간에 배치된 형태가 일정하지 않으며, 일정한 색상이나 위치에 나타나지 않음을 알 수 있다. 검증자료와 비교해볼 때 넓은 영역을 나타내는 지역은 대부분 그 형태를 유지하며 인식함을 볼 수 있다. 그러나 자연발생적으로 생긴 길과 같은 경우에는 그 색상이 토양의 종류와 지형조건에 따라 색상, 모양, 크기 등이 매우 다양하게 나타나 매우 정확한 인식 결과를 도출하지 못함을 알 수 있다.



(그림 5) Google Earth Image Dataset 처리 결과

아래 <표 2>에서는 Google Earth 데이터 셋에 대한 주요 관심 대상인 함몰지, 수풀지형, 초목지형에 대한 인식 정확도를 보인다. 함몰지의 경우 영상에서 일정 크기 이하의 대상인 경우 특징이 충분히 확보되지 않아 주위 영역에 흡수되어 인식이 되지 않는 경우가 발생 하게 된다. 수풀지형은 영상 대부분에서 충분한 성능을 보였으며, 초목지형은 수풀지형과 인접한 영역에 있어서 인식 성능이 떨어짐을 알 수 있다.

<표 2> Google Earth Image Dataset 정확도 평가



### 5. 결론

본 연구에서는 항공 영상을 대상으로 한 영상 분석을 목적으로 영역 기반 장면 분할 기법을 제안하였다. 벤치마킹을 위한 일반적인 야외 데이터 셋에 대해서는 89~92% 가량의 우수한 성능을 보였으며, 항공 시뮬레이션 데이터 셋에 대해서는 넓은 영역에 해당하는 지역에 대해서는 우수한 성능을 보였으나 좁은 지역이나 형태가 불분명한 영역에 대해서는 주위 영역에 흡수되는 결과를 나타내었다. 이는 실질적인 UAV 및 항공영상에 대한 실험을 통해 발전시켜 나갈 여지가 있음을 드러냈다. 이를 위하여 차후 도로 및 함몰지형에 대한 형태를 반영한 특징점 개발을 통해 영역 분할 기법을 더욱 발전시켜 나갈 예정이다.

### 감사의 글

본 연구는 한국과학기술원 국방무인화기술 특화연구 센터를 통한 방위사업청과 국방과학연구소 연구비 지원으로 수행되었습니다.

### 참고문헌

- [1] L. Li, R. Socher and L. Fei-Fei. "Towards Total Scene Understanding: Classification, Annotation and Segmentation in an Automatic Framework". Computer Vision and Pattern Recognition (CVPR), 2009.
- [2] D. Comaniciu and P. Meer. "Meanshift: a robust approach toward feature space analysis", IEEE Trans. On Pattern Analysis and Machine Intelligence (TPAMI), 2002.
- [3] D. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, 2004.
- [4] D. Collings, W. A. Wright and P. Greenway, "The Sowerby Image Database", International Conference on Image Processing and Its Applications, 1999.
- [5] Google, "GoogleEarth", <http://earth.google.com>