

# 사용자 입력 기반의 전경 추출 영상 분할 알고리즘 분석

김규목, 이존하, 박지열, 박진원, 정승원  
동국대학교 멀티미디어공학과  
e-mail : kyoomok@gmail.com

## An analysis on Object Segmentation Based User Interactive Algorithm

Kyumok Kim, Jon-Ha Lee, Ji Yeol Park, Jinwon Park, and Seung-Won Jung  
Dept. of Multimedia Engineering, Dongguk University

### 요 약

본 논문에서는 컴퓨터 비전 기술에서 중요하게 인식되고 있는 영상 분할 기법에 대해 분석한다. 사용자가 입력하는 표시 (Seed) 정보를 이용하여 영상에서 중요한 객체를 추출하는 대표적인 알고리즘의 접근방식 및 차이점을 비교한다.

### 1. 서론

최근 사용자 입력 방식의 영상 분할 기법은 활발한 연구 분야로 알려지고 있으며, 새로운 방식의 알고리즘이 자주 등장하고 있다. 영상에서 중요한 객체로 여겨지는 부분이 추출 될 때, 사용자의 입력 정보 없이 자동으로 추출이 완벽하게 이루어지면 아주 만족스러운 알고리즘이 될 것이다. 하지만 현재 사용자의 도움이 없이 영상을 추출할 때에는 매우 불만족스러운 결과물이 나타나고 있고 효율적이지 못한 경우가 대부분이다. 따라서 좀 더 만족스러운 영상 추출을 위해 사용자의 입력을 기반으로 많은 알고리즘들이 제시되었다.



(그림 1) 영상에서 비중이 높은 영역을 사각형으로 표시하는 GrabCut 알고리즘[2]을 이용하여 영역을 추출한 입력 영상과 결과 영상

본 논문에서는 사용자 입력 기반의 영상 분할 알고리즘의 대표적인 세 가지의 알고리즘들의 접근 방식 및 적용할 수 있는 응용을 분석한다. 첫 번째 알고리즘 [1]은 사용자의 입력 정보를 입력 점의 연속적인 형태로 얻고 전경(foreground)과 배경(background) 영역을 구분하여 각 영역의 경계(boundary)와 관련된

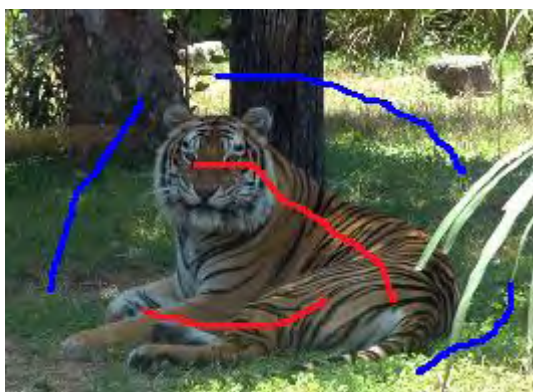
에너지 값과 지역적(regional)인 에너지의 값이 최소화 되도록 하는 분할하는 방식을 이용한다. 두 번째 알고리즘은 'GrabCut'[2] 방식으로 첫 번째 알고리즘의 개선된 방식이다. 사용자의 입력 정보를 점의 연속된 형태가 아닌 비교적 간단한 사각형의 형태로 입력하여 전경을 추출하는 알고리즘을 이용한다. 세 번째 알고리즘[3]은 입력 영상을 먼저 'Mean-shift' 분할 방식을 이용하여[4] Superpixel 형태의 영상 분할 결과물을 먼저 생성한 후 사용자의 입력 정보를 이용하여 여러 Superpixel 영역을 반복적인 연산 과정을 통해 통합하는 단계를 거쳐 전경을 추출하는 방식을 이용한다.

### 2. 영상 분할 알고리즘 분석

#### (1) Graph-Cut[1] 방식

Graph-Cut[1] 기반의 영상 분할 방식은 사용자의 입력 정보를 기반으로 영역과 관련된 변수와 테두리 부분과 관련된 변수의 합을 최소화 하는 에너지 식을 사용한다[1]. 이를 최적화된 방식으로 계산하기 위해서 graph 형태의 데이터 구조를 이용하고 이 것을 자르는 형태의 계산 방식을 이용한다.[5]

Graph-cut 형태의 영상 분할은 입력 정보를 상세하게 표시했을 경우 영상 분할 결과로 영상의 전경에 해당하는 부분을 만족스럽게 추출할 수 있다. 하지만 경우에 따라 상세하게 입력 정보를 표시하는 만큼 시간 소요가 크고 사용자의 표시 정보를 입력할 때 정확하게 표시해야 하기 때문에 사용자의 시간적 노력이 필요하다.



(그림 2) Graph Cut 방식[1]을 이용한 영상 분할의 사용자 입력 표시 영상

## (2) Grabcut[2] 방식

Grabcut[2] 방식은 Graph-Cut[1] 방식에서 좀 더 알고리즘이 보완된 방식이다. 전경과 배경 부분을 가우시안 혼합 모델(Gaussian mixture model) 형태로 구분하여 전경이 추출되도록 한다. 해당 알고리즘에서는 Graph-cut 방식에서 사용자 정보를 선의 형태로 입력한 것과는 다르게 입력 정보를 표시할 때 그림 (3)과 같이 배경에 대한 정보는 표시하지 않고 전경 정보에 대한 부분만 사각형의 형태로 표시한다. 그러므로 사용자는 선의 형태로 입력할 때와는 다르게 비교적 입력 정보를 시간적인 소요 없이 쉽게 입력할 수 있다. 따라서 해당 알고리즘은 사용자의 입력이 불안정하고 어려운 환경을 가지고 있는 응용 프로그램이나 실행 장치에서 전경을 추출하는 영상 분할 알고리즘이 사용될 때에는 해당 알고리즘 방식이 매우 적합하다.



(그림 3) Grab-Cut[2] 방식을 이용한 영상 분할의 사용자 입력 표시 영상

## (3) Superpixel 기반의 영상 분할[3]

Superpixel 화 된 영상을 이용하여 전경을 추출하는 알고리즘[3]은 그림(4)에서 나타내고 있듯이 사용자의 입력 정보를 이용하기 전에 영상의 내부에서 ‘mean-shift’ 알고리즘[4]을 통해 비슷한 픽셀들로 이루어진 영역을 크거나 작은 영역으로 나눈 형태인 Superpixel로 영상을 나누는 작업을 먼저 수행한다. 그 후 사용자의 입력 정보를 선의 형태로 입력하였을 때, 해당

입력 정보들이 나누어진 크고 작은 영역들에 포함이 되면 해당 영역 전체가 입력되는 영역으로 등록이 된다.



(그림 4) Superpixel 기반의 영상분할[3]을 이용한 영상 분할의 사용자 입력 표시 영상

해당 방식은 미리 영상을 분할 하는 작업을 포함하고 있기 때문에 다른 알고리즘 보다 사용자 표시 정보의 정확성에 덜 의존적이다. 따라서 영상 분할 시스템이 사용되는 응용에서 사용자가 정보를 입력하기에 까다로운 응용일 경우 해당 방식이 효과적으로 영상 분할을 할 수 있다. 또한 영역을 미리 나누는 단계에서 영역의 나누어지는 정도를 조절하여 이에 따라 전경을 추출할 때 결과물의 정확한 정도를 조절할 수 있다.

## 3. 결론

본 논문에서는 영상 분할에서 이용되는 알고리즘의 방식의 차이에 대해서 알아보았다. 사용자의 표시 입력 정보에 따라 알고리즘의 결과 영상이 상이할 수 있기 때문에 알고리즘의 성능을 비교하는데 어려움이 있지만 입력 방식의 난이도나 알고리즘의 소요 시간을 비교하여 알고리즘의 효율성을 비교하였다. 영상 분할 기술은 현재 차세대 분야로써 주목 받고 있는 증강 현실 및 가상 현실 기술에서 매우 큰 비중을 차지할 것으로 보인다. 각 알고리즘의 실험은 대부분 데스크탑 환경에서 이루어졌으나 최근 많은 응용들의 사용 환경인 모바일 기기나 웨어러블 기기에서는 데스크탑과는 다르게 사용자가 입력할 수 있는 방식이 제한적이다. 그렇기 때문에 제한된 환경에서도 전경 추출을 목표로 하는 알고리즘 개발이 필요할 것으로 보이고 환경에 의존적이지 않게 기술을 적용할 수 있고 효율적으로 결과물을 만들 수 있는 알고리즘 개발에 대한 노력이 필요할 것으로 보인다.

## 참고문헌

- [1] Y. Boykov and M. Jolly. Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D images. In International Conference on Computer Vision (ICCV), 2001. Proceedings. Eighth IEEE

- International Conference on Vol. 1. IEEE, 105-122.
- [2] Carsten Rother, Vladimir Kolmogorov, and Andrew Blake. 2004. GrabCut: interactive foreground extraction using iterated graph cuts. In ACM SIGGRAPH 2004 Papers (SIGGRAPH' 04). ACM, New York, NY, USA, 309-314.
  - [3] J. Ning, L. Zhang, D. Zhang, C. Wu, Interactive image segmentation by maximal similarity based region merging, Pattern Recognition 43 (2010) 445-456, Feb.
  - [4] D. Comaniciu, P. Meer. Mean shift: a robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(5) (2002) 603-619.
  - [5] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. In 3rd. Intl. Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR). Springer-Verlag, September 2001, to appear.