

# Mean Shift 분석을 이용한 그래프 컷 기반의 자동 칼라 영상 분할\*

김정환<sup>o</sup>, 박안진, 정기철  
 송실대학교 IT대학 미디어학과  
 kjw598@hanmail.net, {anjin, kcjung}@ssu.ac.kr

## Graph Cut-based Automatic Color Image Segmentation using Mean Shift Analysis

Jungwhan Kim<sup>o</sup>, Anjin Pak, Keechul Jung

Department of Digital Media, College of Information Technology, Soongsil University

### 1. 서 론

그래프 컷은 데이터 항(data term)과 스무드 항(smoothness term)으로 구성된 에너지 함수를 전역적으로 최소화하는 방법으로, 데이터 항은 특징공간에서 주어진 사전정보와 각 픽셀간의 유사도를 나타내며, 스무드 항은 영상공간에서 이웃하는 픽셀간의 유사도를 나타낸다. 여기서 데이터 항은 특징공간에서 스무드 항은 영상공간에서 유사도를 계산하기 때문에 두 공간을 동시에 고려할 수 있는 장점이 있지만 데이터 항을 위한 사전정보가 필요하다.

데이터 항의 사전정보를 자동으로 설정하기 위해 GMM(Gaussian mixture models)이 주로 이용되고 있으며, EM 알고리즘에 의해 계산된 평균과 공분산(covariance)이 각 클래스를 위한 사전정보로 이용되었다[1]. 이 방법은 특징공간상에 표현된 특징값의 밀도함수를 추정함으로써 쉽게 사전정보를 구할 수 있지만, 평균과 공분산으로 클래스를 표현하기 때문에 클래스의 모양이 초구(hyper-sphere) 또는 초타원(hyper-ellipsoid)일 때만 좋은 성능을 보이며 클래스의 수 또한 미리 알아야 한다. 그러므로 GMM을 이용한 방법들은 실제 특징공간 상의 데이터를 정확하게 분석, 추정하기에 적당하지 않다[2].

다양한 클래스의 모양에서 좋은 성능을 유지하기 위해, 본 논문에서는 mean shift 분석 방법을 이용한 그래프 컷 기반의 자동 영상 분할 방법을 제안한다. mean shift 기반의 방법은 특징공간에서 다양한 클래스의 모양을 보다 정확하게 분석할 수 있는 장점 때문에 GMM 기반 방법의 단점들을 해결할 수 있다[2]. 본 논문에서는 임의로 선택된 샘플로부터 밀도가 높은 지역인 모드를 찾아가는 경로의 집합을 각 클래스를 위한 사전정보로 이용하며, 이 사전정보를 기반으로 데이터 항을 자동으로 설정한다. 물체 분할을 위해 이용된 최초의 그래프 컷 방법은 단지 두 개의 터미널만을 가지기 때문에 이진 레이블링 문제에만 적용이 가능하므로, 본 논문에서 다중(multi) 레이블링을 위해  $\alpha$ -expansion 방법을 이용하며, 스무드 항은  $\alpha$ -expansion 방법에 최적화된 Potts 모델을 이용한다[3]. 실험에서  $L^*U^*V^*$  색상이 영상 분할을 위한 특징값으로 이용되었고, 최근에 많이 이용되고 있는 mean shift와 normalized cut기반의 영상 분할방법에 대한 문제점을 분석하였으며, 제안된 방법이 기존의 문제점을 해결함으로써 Berkeley Segmentation Dataset을 기반으로 하여 이전의 두 가지 방법과 GMM을 이용한 그래프 컷 기반의 방법보다 더 좋은 성능을 보였다.

### 2. 본 론

본 논문에서는 영상 분할 문제를 레이블링 문제로 고려한다. 레이블링 문제는 주어진 영상  $\mathbf{P}$ 의 각 픽셀  $p$ 에 레이블 집합 중 하나를 할당하는 것으로  $\mathbf{F}=\{f_1, f_2, \dots, f_{D_n}\}$ 로 표시하며, 영상분할에서 레이블은 분할된 각 영역의 인덱스(index)로 이용된다.  $\mathbf{F}$ 의 각 요소는 레이블의 집합 중 하나가 할당되고  $D_n$ 은 픽셀의 수를 나타낸다. 입력영상의 각 픽셀에 적합한 레이블을 전역적으로 할당하기 위해, 입력영상은 먼저 에너지 함수(식 1)로 표현하며 표현된 에너지 함수를 그래프 컷 방법을 이용하여 전역적으로 최소화한다.

$$E(F) = \sum_{p \in P} D_p(f_p) + \lambda \sum_{p, q \in N} \delta(f_p, f_q) \cdot V_{p, q} \quad (1)$$

여기서  $D_p(f_p)$ 은 픽셀  $p$ 가  $f_p$ 에 얼마나 적합한지를 나타내는 데이터 항이다. 다르게 표현해서 픽셀  $p$ 가 레이블  $f_p$ 에 대한 사전정보와 유사하면  $D_p(f_p)$ 에 낮은 비용(cost)을 할당하며, 유사하지 않으면 높은 비용을 할당한다.  $\delta(f_p, f_q) \cdot V_{p, q}$ 는 스무드 항으로  $V_{p, q}$ 가 두 개의 이웃하는 픽셀  $p, q$ 가 유사한지를 나타내며,  $\delta(f_p, f_q)$ 는 델타 함수로

\* 이 논문은 한국과학재단(KOSEF)의 특정기초연구 No.(R01-2006-000-11214-0)의 지원에 의하여 연구되었음

$f_p$ 와  $f_q$ 가 같은 레이블을 가지면 0, 아니면 1의 값을 가진다.  $\lambda$ 는 두 항의 상대적인 중요성을 나타내며,  $N$ 은 입력영상의 근접한 이웃을 나타내는 픽셀 쌍의 집합으로  $n-links$ (neighborhood links)라 불린다.

그래프 컷을 이용하기 위해 먼저 픽셀과 동일한 노드를 가진 그래프  $G=\langle U, \varepsilon \rangle$ 를 생성해야 한다. 레이블 정보를 위해 두 개의 추가적인 노드 source( $S$ )와 sink( $T$ )가 필요하며 각 노드는  $t-links$ (terminal links)라 불리는 에지를 추가하게 된다. 그래프의 가중치는  $t-links$ 와  $n-links$ 에 할당되며, 각각의 터미널과 노드를 연결하는  $t-links$ 는 픽셀과 레이블 사이의 적합도를 표현하는 데이터 항이, 근접한 이웃 픽셀을 연결하는  $n-links$ 에는 두 근접한 노드 사이의 불연속성을 나타내는 스무드 항이 그래프의 가중치로 각각 할당된다. 데이터 항의 비용을 할당하기 위해 사전정보가 필요하고 이 사전정보를 위해 본 논문에서 특징공간에서 임의로 선택된 샘플로부터 클러스터의 중심으로 이용되는 모드를 찾아가는 mean의 경로를 각 클러스터의 사전정보로 이용한다. 이는 하나의 모드를 향하는 mean의 경로들은 명시적으로 하나의 클러스터에 속하기 때문이다.

그래프 컷은 그래프  $G$ 가 설정된 후에 가중치의 합이 최소가 되는 컷(cut)을 찾는 방법으로 컷된 두 터미널  $S$ 와  $T$ 에 각각 연결된 각 픽셀에 레이블을 할당한다. 최소 컷은  $S$ 에서  $T$ 로 흐르는 max flow를 찾는 것으로 해결된다. 일반적인 그래프 컷 방법은 배경에서 물체를 추출하는 것과 같이 이진화 문제에서 좋은 성능을 보이지만, 대부분의 레이블링 문제는 세 개 이상의 클래스를 포함하기 때문에 결론적으로 다중 레이블링을 위한 확장된 그래프 컷이 요구된다. 본 논문에서 다중 레이블링을 위해 스무드 항이 Potts 모델일 때 에너지 함수를 전역적으로 최적할 수 있는 장점을 가지고 있는  $\alpha$ -expansion 알고리즘을 이용한다. 레이블링  $F$ 가 주어졌을 때 expansion의 움직임은 굉장히 많은 수가 존재하기 때문에 그래프 컷 방법이 최적의 expansion의 움직임을 찾기 위해 이용되며,  $\alpha$ -expansion은 그래프  $G$ 에 최소 컷과 일치하는 레이블링을 계산하는 것으로, 그래프는  $\alpha$ 에 새로운 레이블을 할당할 때마다 동적으로 바뀐다.

제안된 방법을 평가하기 위해 Normalized cut(NCut) 알고리즘[4], Mean Shift 기반의 분할(MS) 알고리즘[5], GMM을 이용한 그래프 컷 기반의 영상 분할 방법[1]과 비교하였으며, 제안된 방법은 Berkeley image segmentation dataset으로부터 선택된 60개 컬러 영상을 이용하였고 NCut 알고리즘과 GMM기반의 알고리즘에서 각 영상의 클러스터 개수는 사람이 나눈 결과의 평균적인 개수와 같게 설정하였다. Berkeley image segmentation dataset은 각 영상마다 5~7개의 사람이 분할한 결과를 가지고 있으며, 우리는 이 결과를 ground truth로 이용하였다. 영상 분할의 평가는 precision과 recall로부터 계산된 F-measure로 평가하여 제안된 방법은 앞선 3가지 방법보다 좋은 성능을 보여주었고 다른 알고리즘보다 사람이 분할한 값에 보다 가까웠다.

### 3. 결 론

그래프 컷 방법은 데이터 항과 스무드 항으로 구성된 에너지 함수를 전역적으로 최소화하는 방법으로 컴퓨터 비전과 이미지 처리에 최근까지 좋은 결과를 보여주었지만 자동적으로 데이터 항을 위한 사전정보를 획득하기 어렵다. 앞서 GMM을 이용한 그래프 컷 기반의 영상 분할 방법이 일반적으로 사용되고 있지만 특징 공간상에서 클래스의 모양이 초구나 초타원일 때만 좋은 성능을 보이는 단점이 있다. 본 논문에서 다양한 클래스의 모양에서 좋은 성능을 보이기 위해, Mean shift 기반으로  $L^*U^*V^*$  색상 공간에서 mean을 찾아가는 모드들의 경로들을 각 클래스의 사전정보로 사용하는 그래프 컷 기반의 영상 분할방법을 제안했고 제안된 방법은 앞선 세 가지 방법들의 문제점을 극복하였으며 Berkeley image segmentation dataset을 이용하여 앞선 세 가지 방법보다 좋은 결과를 보여주었다. 그러나 데이터 항 설정시간과 멀티 레이블 링을 위해 사용하는  $\alpha$ -expansion은 약 23초 정도의 많은 계산시간이 필요하다. 따라서 앞으로 우리는 이러한 계산시간을 줄이기 위한 연구를 계속할 것이다.

### 참고문헌

- [1] J-S. Kim and K-S. Hong, "A New Graph Cut-based Multiple Active Contour Algorithm without Initial Contours and Seed Points," Machine Vision and Applications, to be published.
- [2] Y. Li, J. Sun, C-K. Tang, and H-Y. Shum, "Lazy Snapping," ACM Transactions on Graphics, Vol. 23, Issue 3, pp. 303-308, 2004.
- [3] Y. Boykov, O. Veksler, and R. Zabih, "Fast Approximate Energy Minimization via Graph Cuts," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 11, pp. 1222-1239, 2001
- [4] J. Shi and J. Malik, "Normalized Cuts and Image Segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 8, pp. 888-905, 2000.
- [5] D. Comaniciu and P. Meer, "Mean Shift: A Robust Approach toward Feature Space Analysis," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 5, pp. 603-619, 2002.
- [6] D. Martin, C. Fowlkes, and J. Malik, "Learning to Detect Natural Image Boundaries using Local Brightness, Color, and Texture Cues," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 26, No. 5, pp. 530-549, 2004.