

특징 지도를 이용한 자동적인 중심 객체 추출

박기태⁰ 김종혁 문영식
 한양대학교 컴퓨터공학과
 {parkkt⁰, jhkim, ysmoon}@cse.hanyang.ac.kr

Automatic Attention Object Extraction Using Feature Maps

Ki Tae Park⁰, Jong Hyeok Kim, Young Shik Moon
 Department of Computer Science and Engineering, Hanyang University

요 약

본 논문에서 제안하는 방법은 영상에서 중심 객체를 추출하기 위해 에지와 색상 정보에서 추출한 특징 지도와 배경의 영향을 줄이기 위해 참조 지도(reference map)를 제안한 것이 특징이다. 특징 지도는 다른 영역과 현저하게 구분되는 영역을 검출하기 위해서 영상의 특징 값(feature)들을 이용해서 구성한 영상이라고 할 수 있다. 그리고 참조 지도는 배경의 영향을 최소화 하면서, 객체가 존재할 확률이 높은 부분을 나타내는 지도이다. 제안하는 방법은 밝기 차 정보를 가지고 있는 에지와 YCbCr 컬러모델과 HSV 컬러모델의 색상 성분을 특징 값으로 사용한다. 이들 특징 값을 이용해서 특징 지도를 구성하는 방법으로 영상 내 색상 차에 의해서 나타나는 경계부분을 구하는 방법을 사용한다. 이 방법을 사용하여 에지 지도와 두 개의 색상 지도의 3가지 특징 지도를 생성한다. 다음으로, 영상 배경의 영향을 줄이기 위해 참조 지도를 구한다. 구해진 참조 지도와 특징 지도들을 이용해서 결합 지도(combination map)를 생성한다. 결합 지도로부터 다각형의 객체 후보 영역을 구하고, 객체 후보 영역에 영상분할을 적용하여 중심 객체를 추출한다. 실험에 사용된 영상들은 Corel DB를 사용하였으며, 실험결과로서 precision은 84.3%, recall은 81.3%의 성능을 보인다.

1. 서 론

최근 컴퓨터 및 통신기술이 빠르게 발달하면서 사진이나 동영상 같은 멀티미디어 정보 제공에 대한 사용자의 욕구가 증가하고 있다. 따라서, 네트워크상에서 원하는 데이터를 빠르고 효율적으로 검색하고자 하는 욕구가 증대되고 있다. 이에 따라 내용 기반 영상 검색이 활발히 연구되어 왔다[1,2]. 영상들은 색상, 질감, 형태 등의 특징값 등에 의해서 검색된다. 그러나, 저수준의 특징값들은 호환이, 자동차와 같은 특정한 의미가 내포되어 있는 영상을 검색할 경우에는 문제점이 있다. 이러한 문제점을 해결하기 위하여 영상으로부터 의미 있는 영역이나 객체를 추출하려고 많은 노력을 하고 있다.

2. 제안하는 알고리즘

본 논문에서는 에지와 색상 정보를 이용한 특징 지도를 이용하여 영상에서 중심 객체를 추출하는 방법을 제안한다. 그림 1은 제안하는 알고리즘의 순서를 보여주고 있다. 제안한 알고리즘은 4 단계로 구성되어 있다. 첫 번째 단계는 에지와 색상 정보를 이용하여 특징 지도를 생성한다. 두 번째 단계는 에지 지도와 색상 지도를 결합하여 대표 지도를 생성한다. 세 번째 단계는 대표 지도에서 객체가 존재할 수 있는 객체 후보 영역을 검출하고, 마지막 네 번째 단계에서는 영상 분할 알고리즘을 이용하여 중심 객체를 추출한다.

2.1. 특징 지도(Feature map)

본 논문에서 제안하는 특징 지도(feature map) 생성방법은

Itti가 제안한 방법과 동일하다[3]. 하지만 기존의 방법이 영상 내의 다른 영역과 현저하게 차이가 나는 곳에 사선이 무의식적으로 집중되는 인간의 시각 특성을 반영하여 중요도가 높은 부분을 구해내는 것과 달리 제안하는 방법에서는 영상내의 색상 차에 의해서 나타나는 경계 부분을 구해내는 것에 중점을 둔다. 왜냐하면 영상의 배경이 복잡해 질수록 배경의 영향을 많이 받아서 객체를 제대로 표현해 주지 못하기 때문이다. 특징 지도 구성에 사용하는 특징값은 기존의 밝기(Intensity), 색상(Color), 질감(Texture) 정보가 아닌 에지와 두 개의 색상 정보를 사용하여 생성한다. 그림 2는 제안하는 방법에 의해 생성된 특징 지도를 보여주고 있다.

영상의 에지 지도를 구하기 위해서 먼저 원영상의 크기를 1/2배, 1/3배, 1/4배로 축소하고 각각의 영상마다 소벨(Sobel) 연산자를 사용하여 에지 지도 영상을 생성한다. 그런 다음에 각각의 에지 지도를 하나의 영상으로 결합시켜준다. 영상을 결합할 때는 서로 다른 크기의 에지 지도를 모두 1/4배로 축소시켜서 결합한 후 원영상의 크기로 복원한다. 에지 지도를 생성하는 방법은 수식 1과 수식 2로 정의한다.

$$F_s(x, y) = \frac{1}{24} \left(\sum_{i=1}^{24} |I_s(x, y) - N_i| \right) \quad (1)$$

$$RFM_{Edge} = \frac{1}{3} \sum_{s=2}^4 N \left(F_{\frac{1}{s}}^{Edge} \right) \quad (2)$$

영상에서 색상지도를 생성하기 위해서 에지 지도를 생성하는 방법과 동일하게 영상을 축소한 후 각각의 영상들에 대해서 색상 차를 이용하여 색상지도를 생성하고, 모든 영상을 1/4로 축소시켜서 하나의 영상으로 결합한다. 마지막으로,

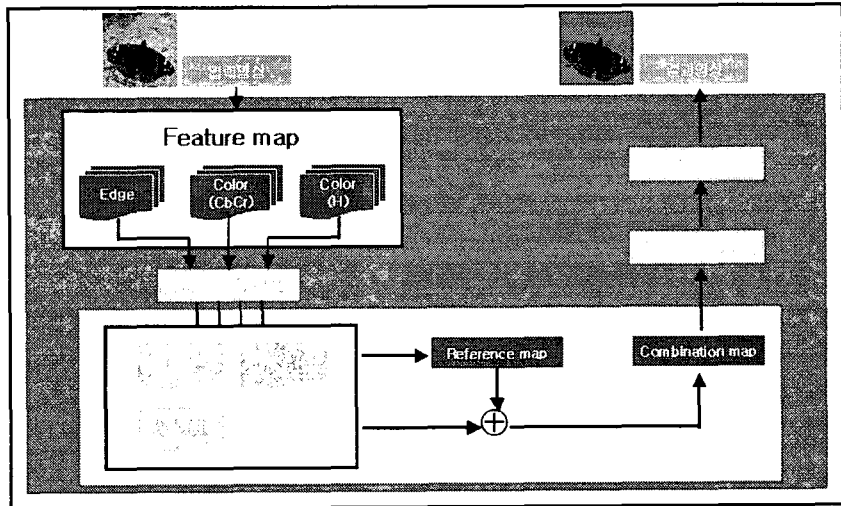


그림 1. 전반적인 제안한 방법의 순서도



그림 2. 특징 지도 결과 영상

원영상의 크기로 복원한다. 본 논문에서는 YCbCr 색상 모델에서의 Cb, Cr 요소와 HSI 색상 모델에서 H 요소를 결합해서 사용한다. 색상지도 생성 방법은 에지 지도를 생성하는 방법과 동일하다.

2.2. 결합 지도(Combination map) 생성

에지 지도 및 색상 지도가 완성한 후에 특징 지도들을 결합해야 한다. 기존 방법에 의해 구해진 결합 지도들은 선형 결합을 통한 지도를 생성한다. 하지만, 이 방법은 각각의 특징 지도 영상에서 중심 객체 이외의 배경의 영향을 많이 받을 수 있게 된다. 따라서 본 논문에서는 배경의 영향을 최대한 줄이면서 객체를 표현할 수 있는 참조 지도를 제안한다.

2.2.1. 참조 지도 생성

입력영상의 배경이 복잡해질수록 객체를 제대로 추출할 수 없기 때문에 배경의 영향을 줄이기 위해서 참조 지도(Reference map, RM)를 제안한다. 참조 지도는 배경의 영향을 최소화하면서 객체가 존재할 확률이 가장 큰 영역을 나타내는 지도라고 정의한다. 참조 지도를 구하는 방법은 에지 지도를 기준으로 나머지 두 개의 색상 지도와 비교해서 공통적인 부분을 추출한다. 에지 지도는 배경의 영향을 덜 받으면서 객체를 잘 표현해 주고, 두 개의 색상 지도들과는 다른 밝기 차 정보를 가지고 있다. 그리고, 경계 부분은 명확하게 표현해 주기 때문에 이 에지 지도를 기준으로 정한다. 참조 지도를 생성하는 방법은 수식 3으로 정의한다.

$$CRFM_{E_CbCr} = \frac{1}{2}(RFM_{Edge} \oplus RFM_{CbCr})$$

$$CRFM_{E_H} = \frac{1}{2}(RFM_{Edge} \oplus RFM_H) \quad (3)$$

$$RM = \frac{1}{2}(CRFM_{E_CbCr} \oplus CRFM_{E_H})$$

식에서, \oplus 는 두 지도의 공통적인 부분들을 결합해주는 연산자이고, $CRFM$ 은 두 지도들의 공통적인 부분들을 더해서 만들어진 중간 단계의 지도이다. 두 개의 $CRFM$ 을 더해서 최종적으로 참조 지도(RM)가 생성된다.

2.2.2. 결합 지도 생성

중심 객체를 추출하기 위해 특징 지도들과 참조 지도를 선형 결합을 통하여 결합 지도를 생성한다. 생성방법은 수식 4와 같이 참조 지도와 각각의 특징 지도들을 비교해서 공통적으로 존재하는 부분의 값은 그대로 유지해 주고, 공통적이지 않은 부분의 값은 낮춰주어 공통적인 부분을 상대적으로 높여주는 방법을 사용한다.

$$RFM(x, y) = \begin{cases} FM(x, y) & \text{if } FM(x, y) \cap RM(x, y) \neq 0 \\ \omega \times FM(x, y) & \text{otherwise} \end{cases} \quad (4)$$

식에서, RFM 은 참조 지도에 기반한 수정된 특징 지도를 나타내고, ω 는 [0,1] 사이의 값을 갖는 가중치를 표현한다. 에지 지도와 색상 지도들에 대해서 각각 수정된 특징 지도들을 생성한 후에, 수식 5를 이용하여 최종 결합 지도를 생성한다. 그림 3은 결합 지도 결과 영상을 보여주고 있다.

$$CM = \frac{1}{3}(RFM_{Edge} + RFM_{CbCr} + RFM_H) \quad (5)$$

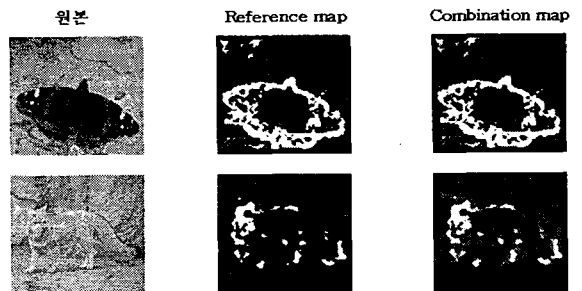


그림 3. 결합 지도 결과 영상

2.3. 객체 후보 영역 검출

본 논문에서는 영상 내에서 중심 객체를 찾기 위해서 먼저 객

체가 존재할 수 있는 후보 영역을 검출한다. 결합 지도를 생성한 후 배경의 영향으로 인한 노이즈 제거를 하기 위해서 겹치지 않도록 영상을 8*8 크기의 서브 블록으로 나눠준다. 서브 블록 단위로 객체로 판단되는 화소의 수가 문턱치보다 높은 블록만 고려하고, 그렇지 못한 블록은 잡음으로 간주되어 제거한다. 그런 다음, 객체 블록들에 대해서 레이블링 알고리즘을 수행하여 각 레이블들에 대해서 일정 개수의 이하인 블록을 다시 한번 제거한다. 그림 4는 서브 블록들에 대해서 레이블링을 통한 노이즈를 제거한 결과 영상을 보여주고 있다.



그림 4. 레이블링을 통한 노이즈를 제거한 결과 영상

그런데, 그림 4를 보면 객체를 표현하고 있는 서브 블록들이 객체의 윤곽선을 중심으로 표현되고, 객체 전체를 표현하고 있지 않는 것을 확인할 수 있다. 노이즈 블록이 제거된 영상에서 객체를 포함할 수 있는 다각형 영역을 객체 후보 영역으로 결정한다. 본 논문에서는 Andrew가 제안한 monotone chain 알고리즘[4]을 사용하여 다각형의 객체 후보 영역을 결정한다. 그림 5는 최종적인 객체 후보 영역을 검출하기 위한 방법들을 보여주고 있다.

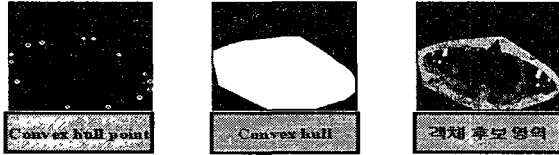


그림 5. 객체 후보 영역 검출 과정

2.4. 중심 객체 추출

객체 후보 영역은 객체 영역뿐만 아니라, 배경 영역을 포함하고 있는 것을 확인할 수 있다. 따라서, 본 논문에서는 객체 영역만을 추출하기 위해서 영상분할 알고리즘을 사용한다. 본 논문에서 사용된 영상분할 알고리즘은 속도와 성능이 좋은 것으로 알려진 Wang의 영상 분할 기법을 사용하였다[5]. 중심 객체 추출 방법은 영상분할 적용 후 비슷한 색상을 가지는 영역들을 병합한 후에 객체 후보 영역에서 경계선 부분에 접해 있는 영역들을 배경 영역이라 가정하고 객체 후보 영역에서 제외한다. 이 방법은 객체 후보 영역이 객체를 포함하는 최소의 영역이므로 수행이 가능하다. 그림 6은 실제 영상들에 대한 객체 추출 결과를 보여주고 있다.



그림 6. 중심 객체 추출 결과

3. 실험 결과

본 논문에서는 Corel 이미지 데이터베이스에서 10개의 클래스의 160장의 영상에 대해서 실험하였다. 성능평가 방법은 원 영상에서 수작업으로 추출한 정답(Ground truth) 객체 영역과 제안한 알고리즘에 의해서 뽑혀진 객체 영역에 대해서 정확율과

재현율을 계산하여 평가한다. 식 6은 성능평가를 나타내고 있다.

$$precision = \frac{(S_M \cap S_A)}{S_M} \times 100$$

$$recall = \frac{(S_M \cap S_A)}{S_A} \times 100 \tag{6}$$

위 식에서 S_M 은 실제 객체의 픽셀 수를 나타내고, S_A 는 알고리즘에 의해서 뽑혀진 객체 영역의 픽셀 수를 나타낸다. 그림 7은 실험에 사용된 영상들 중에서 각 클래스들에 대해서 대표 영상을 보여주고 있다. 그림 8은 중심 객체 추출 결과를 보여주고 있다.



그림 7. 실험에 사용된 영상들



그림 8. 중심 객체 추출 결과 영상

표 1은 실험에서 사용한 10개 클래스 영상들에 대한 중심 객체 추출 결과를 보여주고 있다.

표 1. 중심 객체 추출 결과

클래스	정확율	재현율	클래스	정확율	재현율
개	73.7	87.9	비행기	86.2	96.1
여우	80.1	80.8	나비	91.1	86.3
매	87.2	81.2	호랑이	80	74.1
말	88.9	67.8	울새	87.6	75.7
영양	80.3	76.2	독수리	87.6	75.7
			평균	84.3	81.3

4. 결론

본 논문에서는 영상의 배경의 복잡도나 객체의 위치에 관계없이 영상 내에 존재하는 중요객체를 자동적으로 추출하는 방법을 제안하였다. 배경의 영향을 줄이기 위해서 참조 지도와 결합 지도를 생성하는 방법을 제안하였다.

참고문헌

[1] A.W.M Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, " Content-based image retrieval at the end of the early years," IEEE Trans. on PAMI, pp.1349-1380, 2000.

[2] S. Michael, " Next-generation web searches for visual content," IEEE Computer, pp.46-52, Nov. 2000.

[3] L. Itti, C. Koch, and E. Niebur, "A Model of Saliency-based Visual Attention for Rapid Scene Analysis," IEEE Trans. on PAMI, pp.1254 - 1259, 1998

[4] http://softsurfer.com/Archive/algorithm_0109/algorithm_0109.htm

[5] D. Wang, " Unsupervised Video Segmentation based on Watersheds and Temporal Tracking," IEEE Trans. on Circuits and System for Video Technology, vol. 8, no. 5, pp.539-546, 1998