

효율적인 중요도 지도 생성 기법

김한상, 김준성, 김창수
고려대학교

{khan75, junssi153, changsukim}@korea.ac.kr

Efficient Generation of Saliency Maps

Hansang Kim, Jun-Seong Kim, Chang-Su Kim
Korea University

요 약

본 논문에서는 중요도 지도를 만드는 상향식 방법 중 대표적 기법인 Itti 의 제안 방식[1]이 지닌 문제점을 보완한다. Itti 방식의 경우 여러 특징 지도를 단순한 선형 결합을 통해 중요도 지도를 만들어 내므로 중요 영역이 명확하지 않은 단점을 가지고 있기 때문에, 모든 특징 지도를 사용하지 않고 상대적으로 비효율적인 특징 지도를 배제하는 방법과 중요 영역을 더 강조하기 위해 색의 유사도를 이용하는 방법을 제안하고 그 결과를 비교한다.

1. 서론

중요도 지도(saliency map)는 생체 공학, 전산학을 비롯해 컴퓨터 공학에 이르기 까지 다양한 학문 분야에서 다루지고 있는 주제이다. 단순 이미지 처리라기 보다는 인간의 시각적, 생물학적 특징을 실험을 통해 모델링하고 시각으로 들어오는 입력 영상이 사람에게 어떻게 인지되는가를 다루기 때문이다. 또한 중요도 지도를 통해 얻은 기초 정보를 통해 물체 감지, 콘텐츠 인식을 통한 영상 크기 변경, 관심 영역(region-of-interest)의 썸네일 확대 등 다양한 방면에서 활용이 가능하다.

중요도 지도를 만들어 내는 방법은 다양하게 제시되고 있다. 크게는 상향(bottom-up) 방식[1,2,3]과 하향(top-down) 방식[4] 그리고 이 두 가지가 결합된 하이브리드 방식[5]으로 나뉜다. 상향 방식의 경우는 시각신경으로 들어오는 자극을 생물학적 요소 별로 나누어 병렬로 특징 대비 등을 이용하며, 하향 방식의 경우는 자연 통계학, 특징 견본 정합(feature-template match)등을 이용한다. 하이브리드 방식은 위 두 방식의 특성을 결합한 것이다.

본 논문의 구성은 다음과 같다. 2 절에서는 기존 방식 중에서도 Itti 방식의 내용과 단점을 설명하고, 3 절에서는 Itti 방식의 단점을 보완하기 위해 제안하는 기법을 설명한다. 마지막으로 4 절에서는 제안한 기법을 적용하여 적용 결과를 비교 및 확인한 후, 결론을 맺는다.

2. 기존 방식

본 논문에서 기본으로 다루는 방식은 외부의 정보 없이 이미지 자체만을 분석하여 중요도 지도를 만들어 내는 상향 방식이며, Itti 의 방법[1]은 상향 방식 중 대표적인 방식이다.

Itti 는 입력 영상을 밝기(intensity), 색, 방향성 등의 성분으로 나누고 각각을 다양한 비율의 크기를 갖는 가우시안 피라미드(Gaussian pyramid)로 분해한 후 평면간 중심-주변(center-surround) 차이와 정규화를 통해 특징 지도(feature map)를 획득한다. 그리고 각 지도의 다른 크기의 평면간 결합 및 정규화를 통해 눈에 잘 띄는(conspicuous) 지도를 만들어 낸다. 최종적으로는 각 지도를 선형 결합함으로써 중요도 지도를 만들며, 후보군들 중 승자 독식(winner-take-all) 방식을 적용해 관심(attention) 영역을 찾아낸다.

Itti 방식의 단점은 다음 두 가지를 들 수 있다.

첫째, 특징지도를 만들 때 사용되는 특징 중 하나인 방향성의 경우, 0° , 45° , 90° , 135° 의 네 가지 방향에 대해 가보(Gabor)필터를 사용하기 때문에 연산 량의 증가를 유발하고, 다른 평면과의 선형 결합으로 인해 방향성 특징 지도의 영향은 제한적이다.

둘째, 눈에 잘 띄는 지도는 모든 특징 지도의 결과를 평균, 분산 등의 관계를 통해 각 평면 별 가중치 값을 계산하게 된다. 이 때 평면에 따라서는 대부분의 픽셀 값이 0 혹은 255 에 가까운 극한 값을 갖지만 소수 개의 반대 극한 값(255 혹은 0)을 갖는 픽셀이 존재하면 가중치가 커져 최종 결과에 상대적으로 큰 영향을 미치게 된다.

3. 제안하는 기법

본 논문에서는 Itti 방식의 단점을 보완한다. 첫 번째로 방향성 특징을 제외하고, 두 번째로 남은 평면 중에서 다른 평면의 결과에 영향을 주지 않거나 영향이 적은 평면을 계산에서 제외시킴으로써 계산 량의 감소를 이끌어 낸다.

색의 강도(IT)와, 적녹(RG), 황청(BY) 세 개의 평면 중 제외시킬 한 개의 평면을 판단하는 기준 및 순서는 아래와 같다. 첫 번째 조건에 맞는 평면이 없다면 두 번째 조건에 맞는 평면을 제외하고, 두 번째 조건도 만족하는 평면이 없으면 세

본 논문은 2011 년도 교육과학기술부의 재원으로 한국연구재단의 지원을 받아 <실감교류 인체감응술루션> 글로벌프린티어사업 (한국연구재단-M1AXA003-2011-0031648)으로 진행되었으며, 또한 2012 년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2012011031)

번째 조건을 만족하는 평면을 제외하는 방식이다. 만약 두 번째 조건을 만족하는 평면이 2 개 이상일 경우는 세 번째 조건을 통해 배제할 평면을 선택하여 최종적으로 2 개의 평면만을 이용하여 중요도 지도를 만들게 된다.

- 1) 최대값이 상대적으로 작은 평면

$$smallest(G_{max_IT}, G_{max_RG}, G_{max_BY})$$

- 2) 극대값의 평균값이 극한인 평면

$$find_abnormal(LB, UB, L_{avg_IT}, L_{avg_RG}, L_{avg_BY})$$

- 3) 극대값의 분산 대 극대값의 평균의 비가 최소인 평면

$$smallest(L_{var_IT}/L_{avg_IT}, L_{var_RG}/L_{avg_RG}, L_{var_BY}/L_{avg_BY})$$

위에서 사용된 기호는 다음과 같이 정의된다. 입력 영상을 8x8 블록으로 나누어 각 블록의 max 값은 L_{max} (극대값), L_{max} 중 가장 큰 값은 G_{max} (최대값), L_{max} 의 평균과 분산은 각각 L_{avg} , L_{var} 이다. 그리고, *smallest* 는 세 값 중 가장 작은 값을 갖는 평면을, *find_abnormal* 는 세 값 중 상한(UB)과 하한(LB) 사이의 영역에 들어가지 않는 비 정상적인 평면을 찾는 함수이며, 상한과 하한은 각각 G_{max} 중 최대 값의 95%와 5%이며 이 값은 실험적으로 구한 값이다.

특정 지도는 흑백(gray color)의 분포가 산재되어 있는 특성이 있기 때문에 중요도 지도를 만들 때 정보를 제공하지 않고 잡음으로 작용할 수 있다. 따라서, 본 논문에서는 위 조건을 확인하기 전에 아래 수식 (1)과 같은 가중치 적용을 통해 색의 유사도를 이용해 중요 영역을 강조한다. 유사한 색 값을 갖고 있는 픽셀 간의 거리를 고려하여 밝은 색이 모여있는 곳에서는 더 밝게, 어두운 색이 모여있는 쪽은 더 어둡게 만들어준다.

$$W_p = \sum_{q \in P, q \neq p} \frac{I(p)I(q)}{d_{pq}} \quad (1)$$

위 수식에서 W_p 는 영상의 p 위치에 적용되는 가중치 계수이다. W_p 는 p 주변의 유사한 색 값을 갖는 임의의 점 q 와의 거리 (d_{pq})와는 반비례, p 와 q 위치에서의 값 $I(p)$, $I(q)$ 와는 비례한다.

4. 실험 결과 및 결론

실험에서 사용된 원본 영상은 가로 방향 및 세로 방향이 모두 사용되었으며, 가로 방향의 입력 영상은 400x300 이다. Itti 방식의 출력 영상 [6]의 크기가 50x38 로 약 1/8 로 축소되어 상대적으로 작으며, 제안 방식은 512x512 로 일정하다.

2 절에서 제안한 기법을 적용한 결과로 생성된 중요도 지도를 Itti 의 알고리즘을 적용한 중요도 지도와 비교한 결과를 그림 1 에 도시한다. 왼쪽은 원본 영상, 가운데는 Itti 방식의 결과이며, 오른쪽은 제안 방식 결과이다. 입출력 영상의 가로 세로 길이와 화상비가 동일하지 않아 용이한 비교를 위해 가로 대 세로의 비율을 모두 1:1 로 표현했다. Itti 방식이 제안 방식의 결과 영상에 비해 정밀도가 떨어지고, 여러 특징의 통계적 특성을 일괄적으로 반영하여 국부적인 부분까지 강조가 되기 때문에 중요 영역을 제대로 표현하지 못하는 단점을 재차 확인 할 수 있다. 반면에 제안 방식이 상대적으로 중요 영역을 명확하게 표시해주는 것을 확인 할 수 있다.

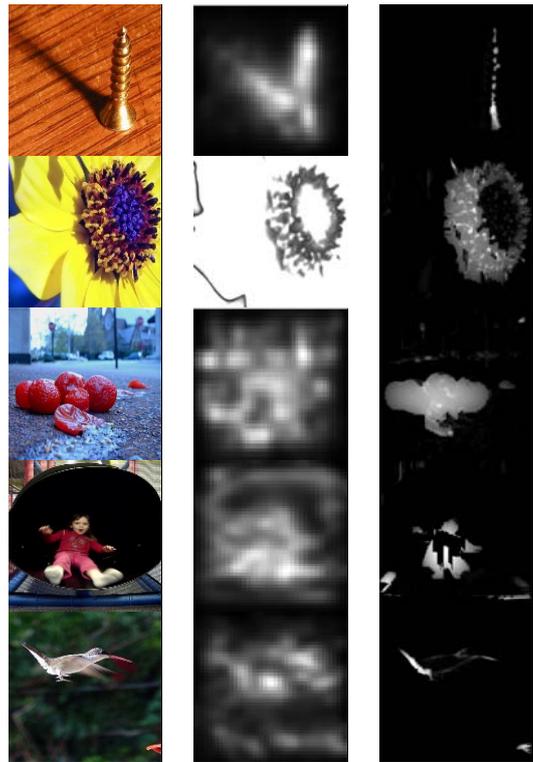


그림 1. 중요도 지도 비교
(좌) 원본 영상, (중) Itti 방식 [1], (우) 제안 방식

본 논문에서는 상향 방식 중요도 지도를 효율적으로 생성하는 방법을 제안하였다. 제안하는 기법에서는 특징지도의 개수를 줄여 연산 량을 줄이고, 주변 픽셀과의 거리를 이용해 색의 유사도를 구하여 이 값을 이용해 가중치를 줌으로써 Itti 방식에 비해 중요 영역이 더 명확하게 드러나는 것을 확인할 수 있었다.

5. 참고문헌

- [1] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20 no. 11, pp. 1254-1259, 1998
- [2] X. Hou, L. Zhang, "Saliency Detection: A Spectral Residual Approach," Proc. IEEE CVPR, pp. 1-8, 2007
- [3] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," Proc. Advances in Neural Information Processing Systems, pp.681-688, 2007
- [4] A. Oliva, A. Torralba, M. S. Castelhana, J. M. Henderson, "Top-down control of visual attention in object detection," ICIIP, pp.253-256, 2003
- [5] M. Milanova, S. Rubin, R. Kountchev, V. Todorov, R. Kountcheva, "Combined visual attention model for video sequences," ICPR, pp.1-4, 2008
- [6] <http://www.klab.caltech.edu/~harel/share/gbvs.php>