

Contrast map과 Salient point를 이용한 중요객체 자동추출

곽수영^o 고병철 변혜란
연세대학교 컴퓨터과학과

{ksy2177, soccer1, hrbyun}@cs.yonsei.ac.kr

Automatic salient-object extraction using the contrast map and salient point

Sooyeong Kwak^o Byoungchul Ko Hyeran Byun

Department of Computer Science, Yonsei University, Seoul, Korea, 120-749

요 약

본 논문에서는 Contrast map과 Salient point를 이용하여 영상에서 중요한 객체를 자동으로 추출하는 방법을 제안한다. 우선 인간의 시각 체계와 유사한 밝기(luminance), 색상(color) 그리고 방향성(orientation) 3가지의 특징정보를 이용하여 각각의 특징정보로부터 feature map을 생성하고, 이 3가지의 feature map을 선형 결합하여 contrast map을 생성한다. 이렇게 생성된 하나의 contrast map을 이용하여 대략적인 Attention Window (AW)의 위치를 결정 한다. 다음으로, 영상으로부터 웨이블릿 변환을 적용하여 salient point를 찾고, salient point의 분포와 contrast map의 중요도에 따라 AW의 크기를 실제 중요 객체의 크기와 가장 유사하도록 축소시킨다. 이렇게 선택되고 축소된 AW안에서 실제 중요 객체를 추출하기 위해 AW 내부에 존재하는 영상에 대해서만 영상 분할을 하고 불필요한 영역을 제거하여 자동으로 중요객체를 추출하도록 한다.

1. 서 론

MPEG4에서는 객체 기반의 독립적인 부호화 방식을 채택하고 있지만 실제적으로 영상 또는 동영상으로부터 의미 있는 객체를 추출하는 것은 어려운 작업이다. 이러한 이유로 대부분의 객체추출방법은 정확한 객체 추출을 위해 반자동추출 방법[6]이나 사전에 객체가 중심에 있다는 가정을 두는 방법[2]을 이용하고 있다. 이러한 한계점을 극복하기 위해 본 논문에서는 반자동이나 기본적인 가정 없이 영상으로부터 의미 있는 객체를 자동으로 추출하는 방법을 제안한다. 이러한 중요 객체 자동 추출 연구는 다음과 같은 여러 가지 응용분야에 적용될 수 있다.

①워터마킹: 일반적으로 워터마킹 기법은 전체영상에 공개 키를 삽입하는 방법을 사용하지만 실제로 이러한 방법은 영상에 손실이 발생할 경우 워터마크가 손상될 우려가 있다. 따라서, 일반적으로 영상을 편집한다 하더라도 중요한 객체는 그대로 보존한다는 사실을 이용하여 중요객체를 추출하고 이 객체에 워터마크를 삽입한다면 영상 편집으로 인한 워터마크의 손실을 최소화 시킬 수 있다.

②MPEG4: MPEG4는 객체기반 압축방법으로 이를 이용하여 모바일 폰이나 PDA와 같이 다양한 크기의 디스플레이를 갖는 단말기에 동일한 크기의 영상을 전송하는데 사용될 수 있다. 하지만 전송 하려고 하는 단말기 디스플레이의 크기가 상대적으로 작다면 같은 크기의 영상이라도 작게 축소시켜 전송하여야 함으로 선명한 영상을 전송할 수 없다. 이 경우 원 영상에서 중요한 객체를 추출하고 이를 중심으로 영상을 축소하여 전송한다면 디스플레이의 크기가 작은 단말기에서도 보다 선명한 영상을 전송받을 수 있다.

③디지털 카메라: 디지털 카메라를 이용하여 사진을 촬영할

경우 영상에서 중요한 객체가 무엇인가를 카메라가 자동으로 인식할 수 있다면, 사용자의 수작업 없이도 중요한 객체를 중심으로 조명 또는 색상을 보정하여 촬영할 수 있는 기능을 추가 시킬 수 있다.

④영상검색: 현재 '영상에 의한 질의(query-by-image)'는 가장 대표적인 내용기반 영상검색 기술의 하나다. 하지만, 영상검색 시스템에서 영상의 전역적인 특징만을 이용하여 영상을 검색한다면 배경색만 바뀐다 하여도 유사영상으로 검색되지 않은 오류를 범할 수 있다. 이러한 단점을 해결하기 위해 영상의 중요한 객체를 추출하여 객체 또는 특정영역에 대한 영상검색을 할 수 있다.

본 논문에서는 이러한 여러 가지 응용분야에 유용하게 적용될 수 있는 중요객체 추출을 contrast map과 salient point를 이용하여 자동으로 추출하는 방법을 제안한다.

2. 중요객체 자동추출

본 논문에서 제안한 방법은 contrast map의 중요도(saliency)를 이용하여 AW의 대략적인 위치를 결정하고, 영상에 포함된 salient point의 분포와 contrast map의 정도에 따라 AW의 크기를 줄여가며 객체와 가장 근사한 AW의 크기와 위치를 결정한다. AW에 대한 최적의 크기와 위치가 결정되면, AW내부에 존재하는 영역에 대해서 영상분할을 실시하여 배경이나 중요하지 않은 영역들을 자동으로 제거 하고 중요하다고 생각되는 객체만을 자동으로 추출한다. AW는 중요하지 않은 객체 또는 배경을 제거하는 역할을 하기 때문에 중요객체를 추출하는 시간을 줄여줄 뿐만 아니라 보다 정확한 위치를 파악하게 해주는 중요한 역할을 한다.

2.1 Contrast map 생성

대략적인 AW위치를 파악하기 위해서는 먼저 contrast map을 만들어야 한다. contrast map을 이용하는 목적은 중요한 객체의 대략적인 위치정보를 얻을 수 있기 때문이다. 하나의 contrast map은 밝기(luminance), 색상(color) 그리고 방향성(orientation)의 3가지 특징으로 구성된 feature map을 결합하여 생성한다. 본 논문에서는 contrast map을 생성하기 위해 Itti [3]가 제안한 방법을 변형하여 적용하였다. 우선, 밝기에 대해서 6개, 색상에 대해서는 12개, 그리고 방향성에 대해서는 18개의 feature map을 생성하였다. 따라서, 각각의 특징에 대해서 luminance map, color map 그리고 orientation map이 완성되고 이를 같은 비율로 결합하여 하나의 contrast map을 생성하게 된다.

2.1.1 Luminance map

Luminance map을 생성하기 위해서 우선 gray영상에 가우시안 필터를 적용하여 잡음을 제거한다. 다음에 영상의 크기와 필터 크기를 달리 하여 모두 6개의 feature map을 생성하여 결합시켜 준다. 영상의 크기와 필터의 크기를 달리 해 주는 이유는 각각의 feature map을 생성할 때 사용되는 필터의 크기에 매우 민감하기 때문이다. 객체의 크기가 필터의 크기보다 클 경우에는 saliency가 낮게 나올 수 있고, 배경이 객체보다 오히려 높은 saliency값을 가질 수도 있다. 이러한 단점을 보완하기 위해서 각 영상의 크기를 1/2배, 1/3배, 1/4배로 다운 샘플링하고 각각의 다운 샘플링된 영상마다 필터의 크기를 달리하여 6개의 feature map을 완성하게 된다. 이 6개의 map을 모두 결합하여 luminance map을 생성하게 된다. 이때 다운 샘플링된 영상의 크기가 모두 다름으로 각각 원 영상의 1/4배로 줄여 더해준다. 수식(1)에서 $G(c, s)$ 는 가우시안 필터를 적용한 gray영상을 1/c 배의 크기로 다운 샘플링 하고, s크기의 필터는 적용한 영상을 나타낸다. 함수 D)는 괄호안의 이미지를 다운 샘플링 한다는 것을 의미한다.

$$\bar{L} = \frac{1}{6} \left(\sum_{c=2}^4 D \left(\sum_{s=c+3}^{c+4} G(c, s) \right) \right) \quad (1)$$

2.1.2 Color map

Color map을 생성하기 위해서 Lab 색상모델을 사용한다. L은 명도(luminosity), a와 b는 각각 빨강/초록, 노랑/파랑의 보색 축을 의미한다. color map에서는 Lab 색상모델에서 색상정보를 가지고 있는 a 와 b를 이용하였다. a, b의 색상 모델에 대해서 luminance map을 생성할 경우와 같이 영상의 크기와 필터의 크기를 달리하여 a 색상체계에 대해 6가지의 feature map, b 색상체계에 대해 6개의 feature map을 생성한다. 결국 12가지의 feature map이 생성된다. 12가지의 feature map의 경우도 영상의 크기가 각기 틀리므로 영상의 크기 중 가장 작은 영상크기인 원영상의 1/4배로 줄여서 모두 더해 하나의 color map을 완성한다. 이 map을 생성할 때에도 전처리 단계로 영상처리 기법 중 가우시안 필터링

을 취해준다. 수식(2)에서 $A(c, s)$ 와 $B(c, s)$ 는 Lab 색상 모델에서 a색상모델과 b색상모델을 1/c배의 크기로 다운 샘플링 하고, s크기의 필터는 적용한 영상을 나타낸다

$$\bar{C} = \frac{1}{12} \left(\sum_{c=2}^4 D \left(\sum_{s=c+3}^{c+4} A(c, s) \right) + \sum_{c=2}^4 D \left(\sum_{s=c+3}^{c+4} B(c, s) \right) \right) \quad (2)$$

2.1.3 Orientation map

중요한 영역은 주변 지역에 비해 눈에 띄는 질감정보를 가지고 있기 때문에 본 논문에서는 orientation map을 생성하기 위해 질감정보를 가장 잘 나타내어 주는 웨이블릿 변환방법을 이용하였다. 이를 위해 1단계 웨이블릿 변환만을 취하여 수평, 수직, 대각선 방향성분을 추출하고 이 특징들을 이용하여 18개의 feature map을 생성하게 된다. 3가지의 방향성 성분에 대해서 각각 2.1.1과 2.1.2의 luminance map과 color map을 생성할 때와 같이 영상의 다운 샘플링 정도와 필터의 크기를 달리하여 모든 feature map을 완성하여 더해준다. 수식(3)에서 $\alpha(c, s)$ 는 웨이블릿 변환을 하여 얻은 수직, 수평 그리고 대각선 성분을 나타낸 영상을 1/c배의 크기로 다운 샘플링 하고, s크기의 필터는 적용한 영상을 나타낸다.

$$\bar{O} = \frac{1}{18} \left(\sum_{HL, HL, LH} \left(\sum_{c=2}^4 D \left(\sum_{s=c+3}^{c+4} \alpha(c, s) \right) \right) \right) \quad (3)$$

위의 수식에 의하여 luminance map, color map 그리고 orientation map이 완성된다. 이 3가지 feature map은 수식(4)와 같이 모두 같은 비율로 선형 결합하여 하나의 contrast map을 완성한다.

$$S = \frac{1}{3} (\bar{L} + \bar{C} + \bar{O}) \quad (4)$$

그림1은 원 영상을 수식(1),(2),(3)을 이용하여 생성된 3가지의 feature map, 즉 luminance map, color map 그리고 orientation map의 결과의 결과와 이 3가지의 영상을 결합하여 생성된 contrast map을 보여 주고 있다.

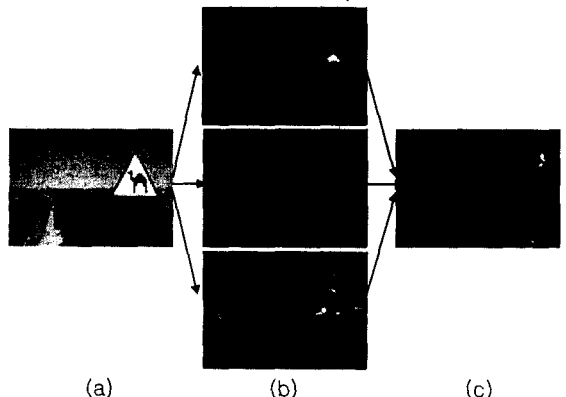


그림 1. Contrast map 추출 과정. (a) 원 영상, (b) luminance, color, orientation map, (c) 생성된 contrast map

2.2 Salient point를 이용한 AW 추출

Contrast map으로 대략적인 AW의 위치를 찾게 되면 영상에서 salient point를 찾아야 한다. 본 논문에서는 웨이블릿 변환을 이용한 salient point 추출법[5]을 이용하여 salient point를 추출한다. 이 포인트는 주파수의 변화가 심한 영역에서 많이 발생하므로 객체와 같이 배경에서 두드러지게 분리되는 경계부분에 많은 salient point를 가지게 된다. 따라서 중요한 객체는 항상 다수의 salient point들을 가지게 되고 이러한 사실을 이용하여 salient point분포가 가장 큰 영역을 찾아 잠정적인 AW의 위치를 결정한다. 위치가 결정되고 나면 AW를 객체에 가장 근사하게 축소하게 되는데 이때는 contrast map의 saliency값이 특정 임계값 이상이고 해당 saliency에 salient point가 동시에 존재할 경우 해당 영역으로 AW를 축소시켜 준다. 이렇게 AW의 최적의 위치와 크기가 결정된 결과가 그림 2이다. AW의 정확한 위치가 결정되면 AW내부에 존재하는 영역에 대해서 영상 분할[4]을 적용하고 하고 불필요한 영역을 제거하여 중요한 객체만을 자동으로 추출한다.



그림 2. Contrast map과 salient map을 이용한 AW 생성 및 축소 결과

3. 실험결과

본 실험은 웹과 Corel Photo CD에서 수집한 자연영상, 그래픽영상, 그림 영상 등을 포함하는 3000개의 영상을 실험데이터로 하였다. 이 시스템은 AW의 크기가 객체의 크기에 따라 유동적이므로 AW의 크기가 고정되어 있을 때 보다 중요객체의 좀 더 정확한 위치정보를 얻을 수 있다. 또한 영상에서 중요하다고 생각되는 객체나 눈에 띄는 부분이 뚜렷이 나타나 있을 경우에는 자동으로 중요객체를 추출한 결과가 대체적으로 좋다고 평가할 수 있다. 이 시스템으로 자동으로 중요객체를 추출한 결과영상들을 그림3에서 볼 수 있다. 하지만 영상에서 특정 객체가 존재한다고 말하기 힘든 자연영상의 경우나, 배경이 아주 복잡한 영상에서는 자동으로 객체를 추출하는 것에는 한계점이 있다.

4. 결론 및 향후 과제

본 논문에서는 영상에서 중요객체를 자동으로 추출하는 방법을 제안하였다. 이 시스템은 객체가 영상에 중심에 존재한다는 가정을 이용하여 실험한 Kim[2]등의 중요객체를 추출 방법과는 달리 아무런 사전 가정 없이 자

동으로 전체 영상에서 중요객체를 추출한다. 또한 [7]과 달리 AW의 크기가 고정되어 있지 않고 객체의 크기에 따라 유동적으로 변하기 때문에 좀 더 정확한 객체의 위치를 추출할 수 있다는 장점이 있다. 하지만 다수의 객체가 존재할 경우 본 논문에서는 하나의 AW만을 생성하기 때문에 여러 개의 객체를 정확히 추출한다는 것에는 한계점이 있다. 또한 영상의 배경이 아주 복잡한 경우 contrast map과 salient point가 비 중요영역에서도 크게 발생함으로 잘못된 AW를 생성할 수 있다. 최종적으로 지금까지의 연구에서는 객체추출방법에 대한 명확히 정의된 성능평가 방법이 없기 때문에 객체를 추출 했을 때, 사람의 주관적인 견해 없이 객체의 추출정도를 평가하는 방법에 대해서도 연구해야 할 필요성이 있다.

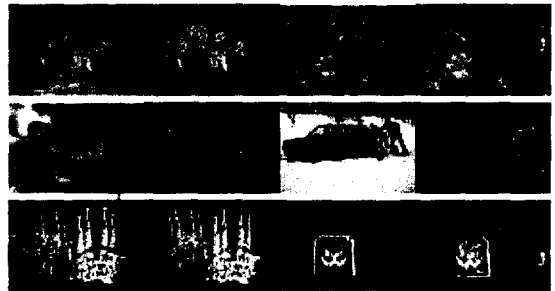


그림 3. 중요 객체 추출 결과

감사의 글

본 연구는 학술진흥재단(KRF-2002-005-H20002)의 지원을 받아 수행하였습니다.

5. 참고문헌

- [1] Y.F. Ma and H.J. Zhang, "Contrast-based Image Attention Analysis by Using Fuzzy Growing," Int. Conf. on Multimedia, pp.355-358,2003.
- [2] S. Kim, S. Park and M. Kim, "Central Object Extraction for Object-Based Image Retrieval," Int. Conf. on Image and Video Retrieval, pp.39-49,2003.
- [3] L. Itti, C. Koch and E. Niebur, "A Model of Saliency-based Visual Attention for Rapid Scene Analysis," IEEE Trans. on PAMI, Vol.20, pp.1254-1259,1998
- [4] B.C. Ko and H.R. Byun, "Region-based image retrieval: A new method for extraction of salient regions and learning of importance scores," Int. Journal of Patt. Recog. and Arti. Intelli. Vol.17.No.8 Dec. 2003.
- [5] E. Loupias and N. Sebe, "Wavelet-based Salient Points for Image Retrieval," Research Report RR 99.11, RFV-INSA Lyon, 1999.
- [6] Q. Tian, Y. Wu, and T.S. Huang. "Combine user defined region-of-interest and spatial layout for image retrieval," IEEE Int. Conf. Image Processing, Vol.3. pp.746-749, 2000.
- [7]곽수영, 고병철, 변해란, "영상검색을 위한 SVM-기반 중요 영역 추출 알고리즘," 제16회 영상처리 및 이해에 관한 워크샵, pp. 337-342, 2004.