

중요도 맵과 최댓값 필터를 이용한 영상 자막 영역 추출

김영진 김만배

강원대학교 IT대학 컴퓨터정보통신공학과

{kimyj09, manbae}@kangwon.ac.kr

Image Caption Area extraction using Saliency Map and Max Filter

Youngjin Kim and Manbae Kim

Dept. of Computer and Communications Eng., IT College, Kangwon National University

요약

본 논문에서는 Saliency map과 Max Filter를 이용한 영상의 자막영역을 추출한다. Saliency map은 눈에 띄는 영역, 즉 영상에서 주변영역에 비해 밝기 차이가 심한 영역과 윤곽선에 대한 특징이 강한 영역을 돌출하는 것을 말하며, MaxFilter는 중심 픽셀을 최대 윈도우 값을 사용하는 것으로 극단적인 Impulse Noise를 제거하는데 효과적이며 특히 어두운 스파이크를 제거하는데 유용하게 사용된다. 이 두 가지의 특징들을 이용하여 영상의 자막 영역을 추출한다.

1. 서론

멀티미디어 서비스 기술이 다방면으로 광범위하게 발달하고 지능적으로 변하면서 최근 IT분야가 핫이슈로 떠오르고 있다. 멀티미디어 서비스 영역 중 영상분야에서 영상 내 자막은 TV를 보는 시청자들에게 TV내용에 대한 이해와 콘텐츠에 대한 다양한 정보를 제공되고 있어 시각적으로 직접 내용을 전달해주는 효율적인 부가 기능의 역할을 한다.

또한 자막검출은 비디오 정보 검색이나 영상 검색 기술 등에서 중요하게 사용되어지고 있다. 기존의 자막 추출 방법으로는 DCT기반 히스토그램을 이용한 밝기대비로 자막 검출알고리즘을 제시한 방법[1]과, Otsu방법을 변형하여 체인코드를 이용한 자막검출[2], 그리고 헤리스 코너 검출기를 이용한 자막검출 기법[3], PCT를 이용하여 입력영상의 Saliency map을 구하여 자막후보영역을 찾은 뒤 labeling하여 자막영역을 추출[4] 등이 있다.

본 논문에서는 자막이 포함된 다양한 영상들을 이용하여 자막영역 추출 방법을 제안하고자 한다. 영상 내 자막의 특성은 구조적 및 시각적으로 에지 성분을 포함하고 있어 에지 기반의 방법으로 알고리즘을 만들면 기존의 자막영역 검출알고리즘보다 효율적인 결과를 보여준다.

2. 제안하는 알고리즘

첫 번째로 진행할 작업은 원본영상을 입력영상으로 받아온 뒤 전체 영상을 GrayScale로 변환시킨다. 변환시키기 전 자막영역은 대부분 Top 부분과 Bottom부분에 위치하고 있어 GrayScale영상에서의 가운데 영역은 중요하지 않기 때문에 별도작업을 통해 Black으로 처리를 한다. 그 후 5x5 크기의 Max Filter를 만든 후 앞서 변환시킨 GrayScale영상과 접목시키면 전체영상이 두툼해지는 영상이 출력되는데 이때의 자막영역은 이전 GrayScale영상에 비해 더욱 돋보이게 나타난다.

두 번째로 진행할 작업은 앞서 출력된 Max Filter Image에서 R영상을 구하는 것이다. R 영상은 식 (1)과 같이 입력영상의 전체 밝기 값에서 입력영상의 밝기평균값을 뺀 뒤 입력영상의 표준편차를 나눈 영상을 말한다.

$$R(x,y) = \frac{[I(x,y) - AVG(I(x,y))]}{\sigma^2(I(x,y))} \quad (1)$$

하지만 우리는 조금 변형시켜 (2)와 같이 입력영상의 전체영역 밝기평균값을 구하는 것이 아닌 그림 1의 자막영역 (A(x,y))을 제외한 나머지 픽셀들의 밝기평균값을 구하여 입력영상에서 뺀 뒤 표준편차로 나눈 값을 이용한다. 그 후 MaxFilter Image로부터 출력된 R영상을 이진화하면 편차 차이가 큰 후보 영역만을 보여주게 된다.

$$R(x,y) = \frac{[I(x,y) - AVG(I(x,y) - A(x,y))]}{\sigma^2(I(x,y))} \quad (2)$$



그림 1. 자막 영역

마지막으로 진행할 작업은 이진화 된 Max Filter Image를 다시 한 번 MaxFilter를 접목시켜 자막이외의 불필요한 잡음들을 제거시키면서 남아있는 잔상을 더욱 부각시켜준다. 그 다음 Sobel Edge의 수직영역을 접목시켜 영상의 수직영역을 검출한다. 이는 자막이 같은 높이에 수평으로 일정한 크기로 나열되어 있다는 자막의 특징을 이용한 것이다. 전체 제안 방법은 그림 2에서 보여진다.

3. 실험 결과

본 논문에서는 앞서 제안한 알고리즘을 가지고 성능을 평가하기 위해 640 x 480 크기의 영상 콘텐츠에서 자막이 포함된 다양한 영상을 수집하였다.

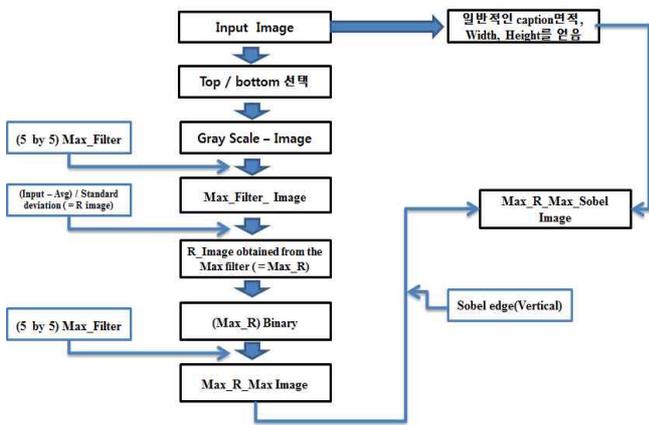


그림 2. 제안하는 자막 추출 알고리즘

3에 비하면 조금 잡음이 있지만 자막영역은 선명하게 출력되었다. 마지막 그림 5를 입력영상으로 설정하고 출력하였을 때의 결과는 잡음이 많이 잡히는 영상이 출력되었다.

실험결과에서 그림 3와 같이 뉴스영상에서는 입력영상의 밝기 값 차이가 적고 대체적으로 안정적이기 때문에 잡음을 거의 제거한 상태에서 자막을 얻을 수 있었다. 그림 4과 같은 다큐멘터리의 경우에도 뉴스영상과 비슷하게 자막영역이 추출되었다는 것을 확인할 수 있었다. 하지만 예능, 음악방송의 경우 무대의 빛들이 워낙 다양하고 입력영상들이 대체적으로 화려하기 때문에 자막영역 이외에다 각종 잡음들이 함께 검출된 것을 확인할 수 있다. 다시 말해 자막은 밝기 값에 민감하게 반응하기 때문에 이진화를 한 경우 임계치보다 높은 밝기 값들의 배경영상들도 추출되는 것이다. 그림 5의 경우에는 완벽하게 자막이 추출되지 않기 때문에 다양한 접근 방법을 제시하여 더 세밀하고 항상 된 검출 기법이 요구된다.

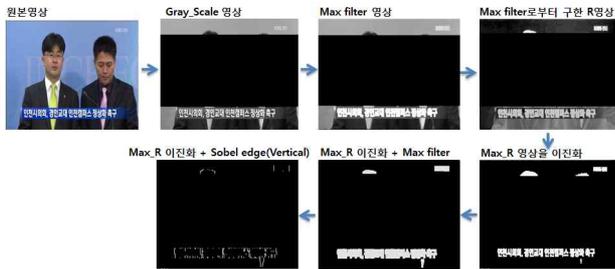


그림 3. 뉴스영상의 알고리즘에 따른 출력 영상과정

4. 결론

본 논문에서는 자막이 포함된 다양한 TV콘텐츠 영상을 가지고 제안한 알고리즘으로 실험하였다. 제안 방법은 실시간 처리가 가능하도록 설계되었다. 자막은 위 아래에 존재한다는 가정하에 국부 영역에 적용된 saliency는 전체 영상으로부터 얻어진 것보다는 보다 성능이 우수하다. 따라서 자막 검출 정확률이 향상되게 된다.

자막은 일반적으로 다른 객체와 분리되지만, 또한 혼합되어 보여지기도 한다. 이러한 경우에 자막 추출은 매우 어려운 문제이고 극복되어야 하는 주제이다.

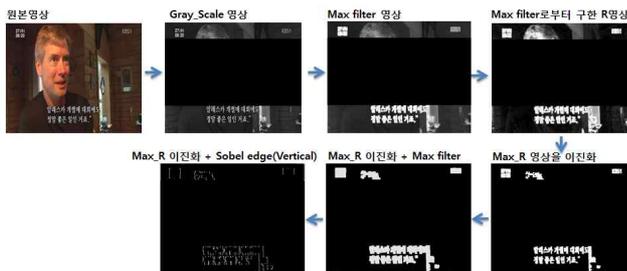


그림 4. 다큐멘터리 영상의 알고리즘에 따른 출력 영상과정

참고 문헌

- [1] 장현성, 강경욱, "DCT기반 압축영상에서 자막 영역을 검출하기 위한 효율적인 방법", 2002년도 한국방송공학회 정기총회 및 학술대회, 2002.11, 127-130
- [2] 이영주, "내용 기반 비디오 검색을 위한 자막 영역 추출 및 프레임 통합에 관한 연구", 2005년 인하대학교 학위논문(석사)
- [3] 김원준, 김창익, "헤리스 코너 검출기를 이용한 비디오 자막 영역 추출, 정보과학회논문지 : 소프트웨어 및 응용 제34권 제7호, 2007.7, 646-654
- [4] 박민수, 박기태, 문영식, "헤리스 코너 검출과 Visual Saliency Map을 이용한 자막 영역 검출 기법", 한국정보기술학회논문지 제11권 제2호(JKIIT, Vol.11, No.2), 2013.2, 53-60

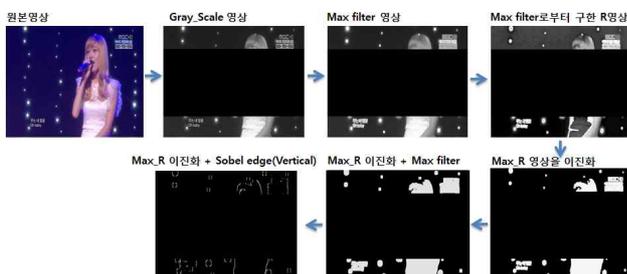


그림 4. 예능, 음악 영상의 알고리즘에 따른 출력 영상과정

앞의 그림들은 입력영상의 종류에 따라 각기 다른 결과영상들이다. 먼저 그림 3의 원본영상을 입력영상으로 설정하고 출력하였을 때의 자막영역은 선명하고 뚜렷하게 나오는 것을 볼 수 있었다. 다음 그림 4의 원본영상을 입력영상으로 설정하고 출력하였을 때의 자막영역은 그림