모바일폰 영상에서 연결요소 특징을 이용한 텍스트 영역 검출

권교현*, 박종천, 전병민
안동과학대학교*, 충북대학교 전기전자컴퓨터공학부
e-mail:khgamja@naver.com, simpleego@hanmail.net,
bmjun@chungbuk.ac.kr

Text Region Detection Using Connected Component Feature in Mobile Phone Images

Gyo-Hyeon Gwon, Jong-Cheon Park, Byoung-Min Jun
Andong Science College, School of Electrical&Computer Engineering,
Chungbuk National University

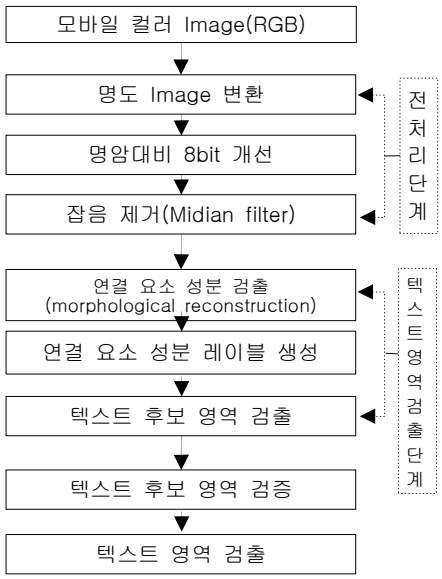
요 약

본 논문에서는 모바일 폰으로 획득한 영상의 텍스트영역 검출을 제안한다. 최근 모바일 폰을 이용한 영상기반 응용 분야의 연구가 활발히 진행되고 있으며, 특히 영상에서 텍스트를 인식하기 위한 전단계로 텍스트 영역 검출은 중요하다. 본 논문은 텍스트 영역 검출을 위해 먼저, 컬러 영상을 입력 받아 그레이 이미지로 변환하여 영상내에 내포된 잡음을 제거하고 열림/닫힘 연산의 특징을 이용해 각 연결요소를 검출하고 검출된 요소들을 레이블링 한다. 레이블링 된 영상은 텍스트가 갖는 특정 조건에 의해 텍스트 영역인지 텍스트 영역이 아닌지를 검출하고 검출된 텍스트 영역은 검증을 통해 최종 텍스트 영역을 검출한다. 제안한 방법은 기존의 택스트 영역 겁출보다 정확도가 향상할 수 있다.

1. 서론

최근 급진하는 정보화 발전에 따라 컴퓨터 지원기능을 추가한 지능형 휴대폰으로 편리한 인터페이스를 갖춘 스마트폰은 휴대폰 기능과 PDA 기능, 무선 인터넷 기능 등 다양한 최첨단 멀티미디어와 컨텐츠를 제공한다. 최근 이러한 폰에 필수적으로 장착된 카메라의 역할은 카메라로 획득한 영상내에 포함된 텍스트를 검출하여 다양한 응용분야에 활용하기 위한 연구가 꾸준히 이루어지고 있다. 휴대 장비로 위치정보와 관광정보 등을 실시간으로 외국어로 변환 가능하도록 하여 외국인도 쉽게 정보를 인지할 수 있는 시스템[1]도 있다.

본 논문은 모바일 폰으로 획득한 영상에서 기존의 모폴로지 재구성 연산[2][3]의 연결요소 특징을 이용하여 복잡한 배경과 빛과 조명 등에 의한 저 대비를 갖는 등 다양한 영상 내에 존재하는 텍스트 영역을 검출한다. 실험 결과 텍스트 영역의 정확률과 저 대비를 갖는 영상에서 효과적으로 텍스트 영역을 검출할 수 있었다.



[그림 1] 제안한 텍스트영역 검출 흐름도

2. 제안한 텍스트 영역 검출

모폴로지 재구성 연산의 연결요소 성분 특징을 이용하여 컬러와 조명 변화가 있는 영상에 내포된 텍스트 영역을 검출한다. 텍스트 영역 검출 과정은 그림 1과 같은 흐름도를 수행한다.

2.1 전처리 단계

모바일 기기로 획득한 컬러 영상을 Gray 영상으로 변환하고 영상내에 내포된 빛이나 조명에 의해 대비가 낮은 영상은 영상의 히스토그램을 식 1에 의해 구하고 그 결과에 따라 영상의 질을 개선하기 위해 프로그램 내에 자동 명도 대비 개선 과정을 수행한다. 영상에 포함되어있는 잡음들은 3-by-3 마스크의 크기를 이용하여 미디언 필터(Midian filter)를 통해 제거한다. 그림 2는 원 영상과 전처리 단계의 결과 영상이다.

$$Histogram \quad n = \sum_{i=0}^{255} m_i \qquad \text{식 1}$$

[그림 2] 전처리 결과 영상

2.2 제안한 텍스트 영역 검출

영상에서 연결요소 성분 특징을 이용하기위해 수리 형태학(Mathematical Morphology) 연산은 영상에서 객체에 대한 모양, 크기 등의 여러 특징을 검출할 수 있다. 특히 경계선과 영상의 형태를 유지하면서 영역의 내부 및 외부를 평탄하게 하는 형태학적 재구성(morphology Reconstruction)[4] 연산은 영상 분할에서 영상을 단순화하여 분할하는데 효율적으로 이용된다. 형태학적 재구성(Morphology Reconstruction)은 두 개의 영상과 하나의 구조 요소가 관계된 형태학적 변환으로 재구성에 의한 열림(opening by reconstruction) 연산과 재구성에 의한 닫힘(closeing by reconstruction) 연산으로 분류되고, 이것은 영상의 형태를 유지하면서 잡음을 제거한다. 형태학적 재구성 연산의 한 영상은 영상의 재구성을 위한 변환의 시작점인 마커(marker)이고, 다른 하나는 영상의 재구성을 제한하는 마스크(mask)이다. 사용된 구조 요소는 연결성(connectivity)을 정의하고 재구성 연산은 식 2와 같이 표현된다.

$$h_{k+1} = (h_k \oplus B) \cap g \qquad \text{식 2}$$

여기서, h_k는 원 영상에 침식 연산을 한 마커 영상 f 의 초기값이고, g 는 마스크 영상이다. B는 3×3 크기의 구성요소이고 ⊕는 팽창(dilation) 연산이다. 그림 3은 형태학적 재구성 연산에 의한 연결요소 검출 결과 영상이다. 즉, 마커와 마스크 이미지를 설정하고 형태학적 팽창(Dilation)연산과 침식(Erosion)연산을 반복적으로 수행하여 지역적 극소/극대 값을 검출하는 방법이다.

[그림 3] 연결 요소 검출 결과 영상

2.3 후보 텍스트 영역 검출 및 검증

후보 텍스트 영역 검출은 극소/극대 변환으로 검출된 연결요소 특징을 바탕으로 텍스트 영역 특징을 만족하는 후보의 개별텍스트를 검출하는 것이다. 본 논문에서는 이 과정을 수행하기 위해서 앞에서 검출한 성분들 즉, 지역적 극소/극대 연결요소 성분은 일정 픽셀 미만의 영역을 제거하고 얻어진 영상의 값에 다시 단일 화소인 구멍(전경 화소들에 의해 포위된 배경 화소들)을 채우고 8-연결성(connected)에 의해 레이블화 한다. 최소/최대(Local minima/maxima) 변환을 이용한 연결요소 성분으로 후보 텍스트 영역을 검출한다. 또한 텍스트 영역 수평/수직 프로파일 특징을 이용하여 텍스트 영역이 맞는지를 검증한다. 그리고 검출된 후보 텍스트 영역 검증은 검출된 후보 텍스트 영역에서 단일 텍스트에 대한 지역적 인접성을 검사함으로서 해당하는 단일 텍스트의 텍스트 영역이 텍스트 영역인가, 비텍스트 영역인가의 여부를 결정한다. 수평방향으로 텍스트 높이의 2배 만큼 인접한 영역 안에 유사한 특징을 갖는 단일 텍스트가 있는지의 존재 여부를 체크하는 것이다. 다시 말해 후보 텍스트 영역 검증은 텍스트영역은 일반적으로 일정한 방향으로 군집을 형성한다. 따라서 이러한 특징이 존재하면 최종적으로 텍스트영역으로 확정한다. 군집화 방법은 검출된 텍스트 후보 영역의 레이블을 기준으로 주변영역에 다른 텍스트 후보 영역 레이블이 존재하는지를

검사하여, 존재하면 이들을 하나의 군집으로 간주하여 텍스트영역을 검증하게 된다. 주변 텍스트 영역의 설정 기준은 두 가지로 설정하였다. 첫째 조건은 기준 텍스트영역 레이블과 유사한 크기의 너비와 높이를 갖고 있어야 하며, 둘째 조건은 기준 텍스트영역의 너비/높이의 크기 중 가장 큰 값을 기준으로 2배 이내의 거리에 이웃 텍스트영역 레이블이 존재해야 한다.

[그림 4] 후보 텍스트 영역 검출 결과 영상

3. 실험결과

제안한 텍스트 영역 검출을 위한 실험대상 영상은 SCH-W770 모바일기기의 카메라로 획득한 카이스트(http://ai.kaist.ac.kr/home/node/85)에서 제공되는 표준 이미지를 대상으로 실험하였다. 그림 4는 후보 텍스트 영역 검출 검증을 완료한 후 최종 후보 텍스트 영역 검출 결과 영상이다. 점선 사각형이 검출한 텍스트 영역 위치이다. 평가방법은 식(3), 식(4), 식(5)을 이용한 정확률과 재현율로 평가하였다[5].

$$Sum = (True + Part + Missing) \quad \text{식 3}$$

$$Recall = \frac{True}{Sum} \quad \text{식 4}$$

$$Sum = (True + Part + Missing) \quad \text{식 5}$$

True : 정확히 찾은 텍스트 영역의 개수
Part : 텍스트의 일부분만을 찾은 개수
Error : 비텍스트 영역을 텍스트로 검출한 개수
Missing : 텍스트를 찾지 못한 개수
Sum : True, Part, Missing의 각 검출수를 합한 개수

4. 결론

본 연구에서는 모바일 폰으로 획득한 영상에서 텍스트영역 검출하는 방법으로 연결요소 성분의 특징을 이용하였다. 본 논문에서 제안한 텍스트영역 검출 기법은 복잡한 이미지나 빛이나 조명에 의한 저대비 영상에 대해 비텍스트 영역을 텍스트 영역으로 잘못 추출하는 경우가 많았으나, 연결성분 특징을 이용한 방법으로 이러한 단점을 개선하였다.

향후 연구 과제는 텍스트 영역 검출을 위한 전처리 과정에서 정확한 임계값 설정을 할 수 있는 방안을 연구하며, 텍스트영역 검증을 위한 보다 일반적인 방법에 대한 연구가 필요하다.

참고문헌

[1] J. Yang, X. Chen, J. Zhang, Y. Zhang, A. Waibel, “Automatic detection and translation of text from natural scenes”, Acoustics, Speech, and Signal Processing, IEEE International Conference on Volume 2, pp.2101-2104, 2002.

[2] 장인영·고병철·변혜란, “모폴로지를 이용한 비디오 영상에서의 자동 문자 추출,” 정보과학회 추계학술대회 논문집, 28(10), 418-420, 2001.

[3] Wolf, C., J.M. Jolion and F. Chassaing, “Text Localization, Enhancement and Binarization in Multimedia Documents,” In Proceedings of the International Conference on Pattern Recognition, Vol. 4, pp. 1037-1040, 2002.

[4] Luc Vincent, “Morphological Grayscale Reconstruction in Image Analysis:Applications and Efficient Algorithms”, Image Processing, IEEE Transactions on Volume 2, pp. 176-201, 1993.

[5] Junker, M. and R. Hoch, "On the Evaluation of Document analysis components by recall, precision, and accuracy, Proceedings of ICDAR, pp.713-716, 1999.