

에지 및 국부 최소/최대 변환을 이용한 자연이미지로부터 텍스트 영역검출

박중천*, 황동국*, 전병민*
*충북대학교 전기전자컴퓨터공학부
e-mail: simplepjc@hanmail.net

Text Region Detection using Edge and Local Minima/Maxima Transformation From Natural Scene Images

Jong-Cheon Park*, Dong-Guk Hwang*, Byoung-Min Jun*
*School of Electrical&Computer Engineering,
Chungbuk National University

요 약

자연이미지에 내포된 텍스트는 많은 정보를 제공함으로써 이를 효과적으로 검출하여 다양한 응용분야에 활용될 수 있다. 본 논문에서는 텍스트 영역의 에지 특징과 국부 최소/최대 변환을 이용하여 자연이미지로부터 텍스트 영역 검출 방법을 제안한다. 에지 검출은 캐니-에지 검출기로 추출하고, 국부 최소/최대 변환을 이용하여 텍스트 영역의 연결성분을 추출한다. 각각 추출된 에지 및 연결성분으로부터 텍스트 영역 후보를 검출하고, 각각의 결과를 결합하여 최종적인 텍스트 후보 영역을 검출하고, 후보 텍스트 영역에 대한 검증은 수행함으로써 최종적인 텍스트 영역을 검출한다. 제안한 방법은 다양한 종류의 자연이미지를 대상으로 실험한 결과, 에지 및 연결성분의 두 가지 특징을 결합함으로써 자연이미지에 존재하는 다양한 형태의 텍스트 영역을 효과적으로 검출하였다.

1. 서 론

자연이미지에 포함된 텍스트는 중요한 정보를 제공함으로써, 이러한 텍스트 영역을 검출함으로써 문자기반 응용분야에서 활용될 수 있다. 최근 연구 중에는 시각장애인을 위한 시각보조 시스템[1]이 개발 중에 있고, 외국인이 여행 중에 휴대용 장비로 외국어로 된 지리 정보와 관광정보를 텍스트 영역 검출과 변환기술로 외국어로 변환 가능하도록 하는 시스템[2]도 개발 중이다.

본 논문은 기존의 에지기반 연구방법[3-6]과 연결요소 분석방법을 결합하여 텍스트 영역을 검출하는 방법을 제안한다. 제안한 방법은 에지기반 방법이 복잡한 배경에서 텍스트 영역을 검출하기 어려운 단점이 있고, 연결성분을 이용한 방법은 빛 또는 조명에 민감

한 단점이 있으므로 이러한 단점들을 상호 보완하기 위해서 각각의 방법에서 추출한 결과를 결합함으로써 텍스트 영역 검출 성능을 향상할 수 있었다.

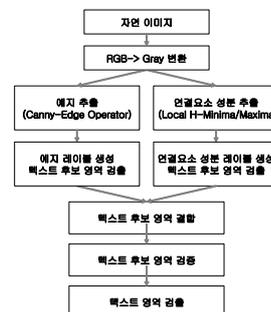


그림 1. 텍스트 영역 검출과정

2. 텍스트 영역 검출

텍스트 영역 검출은 그림 1과 같은 과정을 수행한다. 명도정보를 이용한 에지추출은 캐니에지 검출기[7]를 이용하였고, 실험결과 평균적인 임계값은 하한값은 0.1, 상한값은 0.2, 그리고 가우시안 필터 처리를 위한 시그마(sigma)값은 1로 설정하였다. 색 정보를 이용한 연결요소 성분 추출은 local h-minima와 h-maxima 변환을 수행하여 추출한다. local h-minima/maxima 변환[8]은 gray-scale 모폴로지를 수행하여 국부 최소값을 갖는 영역과 최대값을 갖는 영역을 검출한다. Local h-minima/maxima 변환을 위한 임계값은 텍스트 영역과 배경과의 편차가 큰 특징 있으므로 이미지 전체의 표준편차를 이용하였다. 그림 2는 Local h-minima/maxima를 수행하여 텍스트 영역의 연결요소 성분으로 추출한 결과를 보여준다. 그림 2의 왼쪽은 h-maxima의 변환으로 추출된 연결요소 성분을 것으로 gray-scale 이미지로부터 텍스트 영역이 배경에 비해 상대적으로 밝은 경우 효과적으로 추출됨을 알 수 있다.

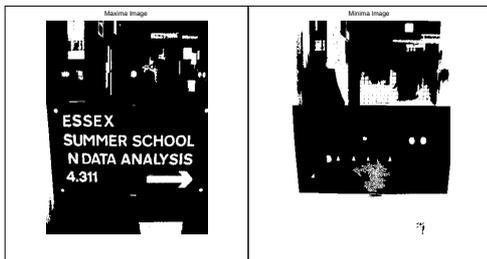


그림 2. h-minima/maxima를 이용한 연결요소 성분 추출결과

2.1 레이블 영역 분석 및 텍스트 후보영역 검출

명도정보를 이용하여 추출된 에지 성분을 레이블과정을 수행하여 에지성분의 구조정보를 이용하여 텍스트 영역의 특징을 만족하는 텍스트 후보 영역을 검출한다. 또한 색 정보를 이용하여 추출된 연결성분을 대상으로 레이블과정을 수행하여 텍스트 후보 영역을 추출한다. 에지 및 연결성분으로부터 텍스트 후보영역을 추출하는데 사용되는 주된 특징은 문자가 갖는 특징을 이용하였다. 즉 문자의 가로와 세로의 비율을 이용하였고, 연결요소 성분이 너무 큰 경우와 작은 경우는 잡음성분으로 제거된다. 텍스트 영역 에지 레이블은 다음과 같은 조건을 만족하는 것으로 한다.

첫 번째 조건은 레이블 영역이 10픽셀 미만의 크기를 갖는 것은 문자를 구성하기 불가능한 경우로 제외하였고, 또한 레이블 영역 안에 다른 레이블이 5개 이

상 있는 경우는 텍스트 영역 주변을 둘러싸고 있는 배경에서 검출된 에지로서 제외된다. 두 번째 조건은 첫 번째 조건을 만족하면 문자 에지의 일반적인 특징으로서, 오일러수가 0또는 -1이면 후보 텍스트 영역으로 선정한다. 세 번째 조건은 위의 조건을 만족하지 않는 에지로서 에지가 끊어진 형태로서 긴 성분을 갖지 않는 특징을 갖고 있으면 후보문자 영역으로 선정한다. 긴 에지 성분의 특징값은 단축길이(Minor Axis Length)를 이용하였고, 임계치는 4.0으로 하였다. 각각의 레이블 성분을 분석한 결과 추출된 에지와 연결요소 성분 텍스트 후보 영역은 논리적 OR 연산을 이용하여 최종적인 텍스트 후보 영역으로 설정된다.

3. 후보 텍스트 영역 검증

텍스트 후보 영역을 검증은 텍스트 영역의 지역적 군집화 특징을 이용하였다. 즉, 텍스트 영역은 일반적으로 일정한 방향으로 군집을 형성한다. 따라서 이러한 특징이 존재하면 최종적으로 텍스트 영역으로 확정한다. 군집화 방법은 검출된 텍스트 후보 영역의 레이블을 기준으로 주변영역에 다른 텍스트 후보 레이블이 존재하는지를 검사하여 존재하면 이들을 하나의 군집으로 간주하여 텍스트 영역을 검증하게 된다. 주변 텍스트 영역의 설정 기준은 두 가지로 설정하였다. 첫째 조건은로 기준 텍스트 영역 레이블과 유사한 크기의 너비와 높이를 갖고 있어야 하며, 둘째 조건은 기준 텍스트의 너비/높이의 크기 중 가장 큰 값을 기준으로 2배 이내의 거리에 이웃 텍스트 레이블이 존재해야 한다.

4. 실험 결과

실험대상 이미지는 ICDAR 2003 표준 이미지 중에서 주로 텍스트 이미지를 대상으로 실험하였다. 문자 영역 검출결과에 대한 성능평가는 재현률(Recall), 정확률(Precision)를 평가요소를 사용하여 표 1과 같이 제시하였다.

- Sum : 실험 대상 이미지 내에 존재하는 전체 텍스트 영역의 수 (True+Part+Error)
- True : 정확히 검출한 텍스트 영역의 수
- Part : 텍스트 영역의 일부분만을 검출한 수
- Error : 검출하지 못한 텍스트 영역의 수
- False : 텍스트 영역이 아닌 영역을 텍스트 영역으로 잘못 검출한 수
- Precision : $True / (True+Part+False)$
- Recall : $True / Sum$

표 1. 텍스트 영역 검출률

이미지 종류	정확률(Precision)	재현률(Recall)
ICDAR 2003(251)	65%	86%



그림 3. 최종 텍스트 영역 검출 결과

그림 3은 텍스트 후보 영역으로부터 군집화 특징을 이용하여 텍스트 영역을 최종적으로 검출한 결과를 보여준다.

5. 결론

본 연구에서는 자연이미지에서 텍스트 영역을 검출하는 방법으로 에지 및 연결요소 성분을 결합함으로써 상호간의 단점을 보완하는 방법을 제안하였다. 본 논문에서 제안한 텍스트 영역 검출 기법은 에지 정보만을 사용하여 텍스트 영역을 검출하는 경우 배경이 복잡한 이미지로부터 텍스트 영역을 잘못추출하는 경우가 많았으나 연결성분을 이용한 방법을 결합함으로써 이러한 단점을 개선하였다.

향후 연구 과제는 텍스트 영역 검출을 위한 전처리 과정에서 이미지의 밝기 대비에 따른 에지 파라미터를 자동으로 설정하는 것이고, 조명의 영향과 다른 물체와 겹쳐짐으로서 에지가 끊어짐으로서 발생하는 문제점을 해결하는 기법에 대한 연구가 필요하며, 마지막으로 텍스트 영역 검출을 위한 보다 일반적인 방법에 대한 연구가 필요하다.

참 고 문 헌

[1] N. Ezaki, M. Bulacu, L. Schomaker, "Text detection from natural scene images: towards a system for visually impaired persons", Pattern Recognition, ICPR 2004. Proceedings of the 17th International Conference on

Volume 2, pp.683-686, 2004.
 [2] 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.
 [3] M.A. Smith and T. Kanade, "Video Skimming for Quick Browsing Based on Audio and Image Characterization", Carnegie Mellon University, Technical Report CMU-CS-95-186, 1995.
 [4] D. Chen, K. Shearer, and H. Bourlard, "Text Enhancement with Asymmetric Filter for Video OCR", Proc. of International Conference on Image Analysis and Processing, pp.192-197, 2001.
 [5] W. Jiang, Q. Shao-Lin, Q. Zhuo, W. WenYuan, "Automatic text detection in complex color image", Machine Learning and Cybernetics, 2002. Proceedings. 2002 International Conference on Volume 3, pp.1167-1171, 2002.
 [6] H. Xian-Sheng, Z. Liu Wenyin, Hong-Jiang, "An Automatic Performance Evaluation Protocol for Video Text Detection Algorithms", IEEE Transactions on Circuits and Systems for Video Technology, Vol 14, No. 4, pp.498-507, 2004.
 [7] J. Canny, "A Computational Approach to Edge Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-8, No. 6, 1986, pp.679-698.
 [8] Luc Vincent, "Morphological Grayscale Reconstruction in Image Analysis: Applications and Efficient Algorithms", Image Processing, IEEE Transactions on Volume 2, pp. 176-201, 1993
 [9] Vijay Raghavan, Peter Bollmann, Gwang S. Jung, "A critical investigation of recall and precision as measures of retrieval system performance", ACM Transactions on Information Systems (TOIS), Vol.7, No.3, pp.205-229, 1989