

지역적 연결요소 및 에지 구조 성분 특징을 이용한 자연이미지로부터 문자영역 검출

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

Text Region Detection Using Regional Connected Component and Edge Structure Component Feature From Natural Scene Images

Jong-Cheon Bak*, Dong-Guk Hwang*, Gyo-Hyeon Gwon*,
Byeong-Min Jeon*

*School of Electrical&Computer Engineering,
Chungbuk National University

요 약

최근 모바일 영상기반 응용 분야에 관한 연구가 활발히 진행되고 있으며 모바일기기로 촬영된 영상에서 문자정보를 추출하고자 하는 많은 연구도 진행되고 있다. 자연이미지로부터 문자정보를 추출을 위한 전단계로 문자영역 검출이 필수적이다. 본 연구는 문자영역의 지역적 에지 및 연결요소 특징을 고려하여 조명 및 복잡한 배경에서도 문자영역을 검출하는 방법을 제안한다. 에지 검출은 캐니-에지 검출기로 추출하고, RGB 컬러분포 패턴을 분석하여 컬러 양자화를 함으로서 연결성분을 추출한다. 각각 추출된 에지 및 연결성분으로부터 문자후보 영역을 검출하고, 각각의 결과를 결합하여 최종적인 문자 후보 영역을 검출하고, 문자 후보 영역에 대한 검증을 수행함으로써 최종적인 문자영역을 검출한다. 제안한 방법은 다양한 환경에서 얻어진 자연이미지를 대상으로 실험한 결과, 에지 및 연결성분의 두 가지 특징을 결합함으로써 자연이미지에 존재하는 다양한 형태의 문자영역을 효과적으로 검출하였다.

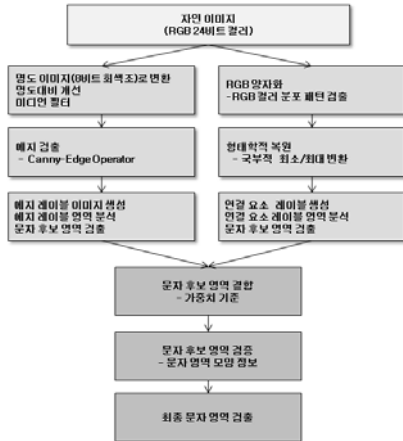
1. 서론

전통적인 문자인식에서는 스캐너로 읽어 들인 문서 영상을 대상으로 하였다. 따라서 조명이 균일하게 유지되며, 배경은 종이 색깔인 흰색 글자는 검정 색이 일반적이다. 그러므로 문서 영상에서 문자 영역을 검출하는 연구는 어느 정도 완료되어 상용화되었다. 반면에 카메라 기술의 발전과 동영상의 증가에 따라, 사진 영상에서 글자부분을 추출하는 시스템에 대한 수요가 늘고 있다. 카메라 기술과 PDA기술이 결합되면, 문자 인식에 대한 수요는 증가한다. 예를 들어, 외국인이 시내에서 관광할 경우, 그 나라의 문자를 모르면 어려움

이 많다. 길거리의 이정표와 간판들을 인식하지 못함으로 원하는 곳에서 식사하거나, 쇼핑을 하기가 힘들다. 이런 상황에서 카메라가 달린 PDA로 간판과 이정표들을 촬영하면, 영상중의 문자 영역을 찾아내어 인식을 수행한 후에, 외국인이 이해할 수 있는 언어로 번역하여 표시하는 시스템이 있다면 관광이 편리할 것이다. 또한 비디오 영상에서 자막의 추출 및 인식은 아날로그로 저장되어 있는 동영상들을 인덱싱 하는데 사용될 수 있다. 최근 연구 중에는 시각장애인을 위한 시각보조 시스템[1]이 개발 중에 있고, 외국인이 여행 중에 휴대용 장비로 외국어로 된 지리정보와 관광정보를 문자영역 검출과 변환기술로 외국어로 변환 가

능하도록 하는 시스템[2]도 개발 중이다.

본 논문은 기존의 에지기반 연구 방법[3-6]과 연결요소 분석방법을 결합하여 문자영역을 검출하는 방법을 제안한다. 본 연구는 기존의 연구 방법이 에지기반 방법이 복잡한 배경에서 문자영역을 검출하기 어려운 단점이 있고, 연결성분을 이용한 방법은 빛 또는 조명에 민감한 단점이 있으므로 이러한 단점들을 상호 보완하기 위해서 각각의 방법에서 추출한 결과를 결합함으로써 문자영역 검출 성능을 향상할 수 있었다.



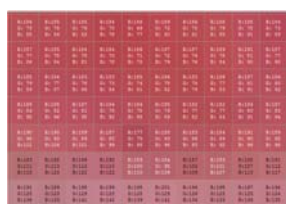
[그림 1] 제안한 문자영역 검출과정

2. 문자영역 검출

문자영역 검출은 그림 1과 같은 과정을 수행한다. 명도정보를 이용한 에지추출은 캐니에지 검출기[7]를 이용하였고, 실험결과 평균적인 임계값은 하한값은 0.1, 상한값은 0.3, 그리고 가우시안 필터 처리를 위한 시그마(sigma)값은 0.7로서 3-by-3 마스크 영역에 가우시안 계수를 설정하여 처리한다.



[그림 2] 단일 문자 색상 변화



[그림 3] RGB 컬러 분포 패턴

연결요소는 RGB 컬러 분포 패턴을 분석함으로써 24 비트 컬러 대표색상 컬러로 양자화 할 수 있다. 대표색상은 16컬러, 24컬러 등 다양하게 설정 가능하다. 문자 영역의 컬러는 일반적으로 배경과 높은 대비 특성을 갖고 있지만 다양한 자연환경에 노출된 물체에 입혀진 색상은 그 색상의 순도가 떨어지며 조명 및 그림

자 등의 영향으로 단일 문자 영역의 순도가 다르고 배경과의 대비가 낮아지는 특징을 갖고 있다. 그러나 일반적으로 인간의 시각은 같은 색상으로 인지됨으로써 이를 반영하여 RGB 색상 분포 패턴 특징을 이용함으로써 이와 같은 문제를 해결할 수 있다. 그림 2는 자연 이미지의 문자 영역의 문자의 컬러가 조명의 영향으로 색상의 대비가 서로 다른 경우를 보여준다.

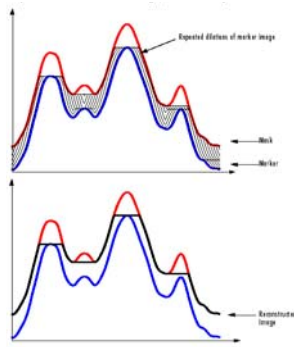
그림 3은 그림 2의 “Times” 문자의 “T”자의 RGB 컬러 분포를 보여준다. 조명의 영향으로 하단의 RGB 컬러 분포는 상단의 분포보다 R 채널 성분의 값은 낮아지고 G와 B 채널은 반대로 높아졌다. 그러나 전체적인 패턴은 여전히 R채널의 값은 RGB 채널 중에서 가장 높은 값을 갖고, 나머지 G와 B채널은 서로 비슷한 값을 갖는 패턴임을 알 수 있다. 따라서 시각적으로 여전히 빨간색으로 인지되는 것을 알 수 있다. 이러한 RGB 컬러 분포 패턴을 기반으로 RGB 컬러 대표 컬러 값으로 양자화하여 문자 영역의 연결요소 성분을 검출한다. 그림 4는 컬러 양자화 결과를 비교하기 위해서 RGB 컬러를 인덱스 6컬러로 양자화한 결과와 RGB 24비트 컬러를 R 채널을 기준으로 컬러 양자화를 수행한 결과를 제시하였다. 인덱스 6 컬러로 양자화한 결과에서는 “Bus Times”의 문자영역이 조명의 영향으로 거의 검출되지 않음을 보여준다. 반면 RGB 컬러 분포 패턴을 이용한 결과는 이미지 전체의 빨간색 영역을 정확히 검출함으로써 “Bus Times”의 문자영역이 컬러성분을 정확히 검출함을 알 수 있다. 그러나 RGB 컬러 분포 패턴을 이용한 방법은 수행 시간이 많이 필요하고, 분포 패턴을 결정하는 방법이 휴리스틱한 알고리즘을 이용한 대표 컬러값을 이용함으로써 문자 영역 배경과 비슷한 컬러를 사용한 경우 문자 영역의 컬러를 정확히 검출할 수 없는 경우도 있다.



[그림 4] 인덱스 6컬러(왼쪽) 이미지와 RGB 컬러 R 채널 양자화 이미지(오른쪽)

RGB 컬러 양자화 결과 추출된 이미지를 명도 이미지로 변화하고 명도 이미지 형태학적 복원 (Morphological Gray-scale Reconstruction)[8]과 RGB 컬러 양자화 방법을 이용하여 추출한다. 그림 5

는 명도 이미지에서 형태학적 복원 과정을 개념적으로 보여준다. 즉, 마커와 마스크 이미지를 설정하고 형태학적 팽창(Dilation)연산과 침식(Erosion)연산을 반복적으로 수행하여 지역적 최대/최소 값을 검출하는 방법이다. 형태학적 복원과정에서 중요한 파라미터는 마스크 이미지의 설정으로 본 연구에서는 이미지 전체의 표준편차를 이용하였다.



[그림 5] 형태학적 복원

2.1 레이블 영역 분석 및 문자 후보 영역 검출

명도정보를 이용하여 추출된 에지 성분을 레이블과정을 수행하여 에지성분의 구조정보를 이용하여 문자 영역의 특징을 만족하는 문자 후보 영역을 검출한다. 또한 색 정보를 이용하여 추출된 연결성분을 대상으로 레이블과정을 수행하여 문자 후보 영역을 추출한다. 에지 및 연결성분으로부터 문자 후보 영역을 추출하는데 사용되는 주된 특징은 문자가 갖는 특징을 이용하였다. 즉 문자의 가로와 세로의 비율을 이용하였고, 연결요소 성분이 너무 큰 경우와 작은 경우는 잡음 성분으로 제거된다.

각각의 레이블 성분을 분석한 결과 추출된 에지와 연결요소 성분 문자 후보 영역은 논리적 OR 연산을 이용하여 최종적인 문자 후보 영역으로 설정된다.

3. 문자 후보 영역 검증

문자 후보 영역 검증은 문자영역의 지역적 군집화 특징을 이용하였다. 즉, 문자영역은 일반적으로 일정한 방향으로 군집을 형성한다. 따라서 이러한 특징이 존재하면 최종적으로 문자영역으로 확정한다. 군집화 방법은 검출된 문자 후보 영역의 레이블을 기준으로 주변영역에 다른 문자 후보 영역 레이블이 존재하는지를 검사하여, 존재하면 이들을 하나의 군집으로 간주하여 문자영역을 검증하게 된다. 주변 문자 영역의 설정 기준은 두 가지로 설정하였다. 첫째 조건은 기준

문자영역 레이블과 유사한 크기의 너비와 높이를 갖고 있어야 하며, 둘째 조건은 기준 문자영역의 너비/높이의 크기 중 가장 큰 값을 기준으로 2배 이내의 거리에 이웃 문자영역 레이블이 존재해야 한다.

4. 실험 결과

실험대상 이미지는 ICDAR 2003 표준 이미지를 대상으로 실험하였다. 문자영역 검출 결과에 대한 성능평가[9]는 재현률(Recall), 정확률(Precision)를 평가요소를 사용하여 표 1과 같이 제시하였다.

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

[표 1] 문자영역 검출률

실험 이미지 종류	Precision	Recall
안내판	87.7	72.8
간판	90.1	70.3
교통표지	84.1	67.4
번호판	80.0	58.9
교통안내	81.9	73.1
광고표지	88.0	94.8
전체	86.3	70.8

그림 6은 문자 후보 영역으로부터 군집화 특징을 이용하여 최종적으로 문자영역을 검출한 결과를 보여준다.



[그림 6] 문자 영역 검출 결과

그림 6의 문자 영역 검출 결과에서 효과적으로 문자영역을 검출한 것을 알 수 있다.

5. 결론

본 연구에서는 자연이미지에서 문자영역을 검출하는 방법으로 에지 및 연결요소 성분을 결합함으로써 상호간의 단점을 보완하는 방법을 제안하였다. 본 논문에서 제안한 문자영역 검출 기법은 에지 정보만을 사용하여 문자영역을 검출하는 경우, 배경이 복잡하거나 명도 대비가 낮은 이미지로부터 문자영역을 잘못 검출하거나 검출하지 못하는 경우가 있었으나 연결성분을 이용한 방법을 결합함으로써 이러한 단점을 개선하였다.

향후 연구 과제는 문자영역 검출을 위한 전처리 과정에서 이미지의 밝기 대비에 따른 에지 파라미터를 자동으로 설정하는 방법을 연구하고, RGB 컬러 양자화 방법의 대표 컬러 선정 방법의 임계값의 결정을 위한 일반화된 알고리즘을 도출하는 것이다. 또한 다른 물체와 겹쳐짐으로서 에지가 끊어짐으로서 발생하는 문제점을 해결하는 기법에 대한 연구가 필요하며, 마지막으로 문자영역 검증을 위한 보다 일반적인 방법에 대한 연구가 필요하다.

참고문헌

- [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, pp.679-698, 1986.
- [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.