

자연영상에서 코너 에지 특징 분석방법을 이용한 한글 텍스트 검출기법에 관한 연구

Hangul Text Detection using Text Corner Edge Feature Analysis in Natural Scene Images

박종천, 권교현, 전병민
충북대학교

Park Jong-Cheon, Kwon Kyo-Hyun,
Jun Byung-Min
Chungbuk Univ.

요약

본 연구에서는 자연 이미지에서 한글 텍스트가 갖고 있는 에지 코너 특징을 이용한 한글 텍스트 검출방법을 제안한다. 자연영상으로부터 에지를 검출하고, 검출된 에지를 20종류의 에지 구조 성분을 갖는 에지 맵을 생성한다. 생성된 에지 맵에서 한글 텍스트 특징 갖는 특징들을 조합하여 모두 8가지의 텍스트 영역 후보 특징을 추출한다. 추출된 텍스트 영역의 특징을 수평 및 수직방향으로 검사하여 텍스트의 시작 라인과 끝 라인을 검출하여 텍스트 영역의 수평좌표를 구한다. 추출된 텍스트 후보 영역에서 최종적으로 텍스트 영역을 결정한다. 제안한 방법은 다양한 종류의 자연 이미지에서 텍스트 영역을 검출에서 좋은 성능을 나타냈다.

I. 서론

자연영상에서 문자는 중요한 정보를 포함하고 있다. 문자정보를 실시간으로 검출할 수 있는 기술개발의 필요성이 증가하고 있는 추세이다. 최근 연구 중에서는 시각장애인을 위한 시각보조 시스템이 개발 중에 있다. 이 시스템은 휴대장비에 부착된 카메라로부터 획득한 자연영상에서 문자정보를 추출하여 이를 음성으로 문자정보를 전달해주는 시스템이다. 또한 외국인이 여행 중에 휴대용 장비를 통해서 외국어로 된 지리정보와 관광정보 등을 문자 추출과 변환기술로 외국인의 이해할 수 있는 문자로 변환 가능하도록 하는 시스템도 개발 중이다.

자연영상에서 문자 추출에 관한 연구는 연결요소 성분 분석 방법(Connected Component-based method)[1,2]와 에지-기반 분석 방법(Edge-based method)[3-5]으로 분류된다. 이 방법은 텍스트 속성 중에서 텍스트 영역에서 색상의 유사성을 특징으로

하는 방법이다.

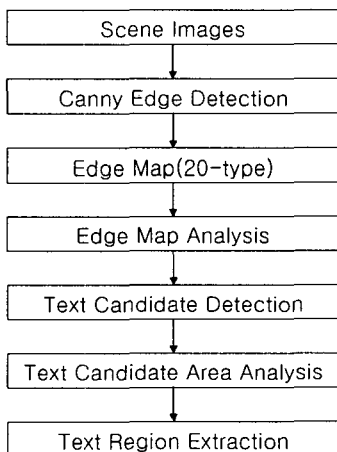
Jain et al.[1]는 다양한 이미지로부터 그래픽 텍스트를 추출하였다. 컬러 연속성과 그레이 레벨 값을 특징으로 사용하였다. 이 방법은 이진 이미지, 웹 이미지 그리고 비디오 이미지에서 좋은 성능을 나타내는 방법이다. 그러나 컬러 이미지에서는 성능이 떨어진다.

Ohya et al.[2]는 4단계 텍스트 추출 방법을 제안하였다. 1단계는 지역 임계치를 이용한 이진화, 2단계는 그레이 레벨 차이 값을 이용한 후보 텍스트 영역 검출, 3단계는 데이터베이스에 저장된 표준 패턴과 텍스트 후보 문자와의 유사도를 측정, 4단계는 데이터베이스에 저장된 표준 패턴의 유사도를 갱신하는 과정 등으로 구성된다. 이 방법은 텍스트 정렬이 되어야 하고, 컬러 텍스트를 제외하는 제한조건을 갖고 있다. 이 방법의 실험결과 100개의 이미지로부터 텍스트 위치 추출은 85.4%를 추출된 텍스트의 인식률

은 66.5%의 성능을 보였다.

에지-기반 방법은 텍스트와 배경은 대비가 강한 특징을 갖고 있다는 점에 중점을 둔 방법이다. 따라서 텍스트 경계에서 에지를 식별하고 이를 조합한다. 그리고 여러 가지 휴리스틱을 적용하여 텍스트 영역과 배경영역을 분리하고 검출하는 방법이다. 일반적으로 에지검출과정을 수행하고 이를 연결성분으로 만드는 smoothing 연산을 수행하거나, 형태학적 연산을 수행한다. 그리고 마지막으로 이들을 통합하는 단계를 수행한다.

Smith and Kanade [3]은 입력 이미지를 3×3 수평 성분 추출 필터를 적용하고 임계치를 사용하여 수직 에지 성분을 검출하였다. 그리고 smoothing 연산을 수행하여 작은 에지성분은 제거 및 끊어진 에지는 연결하였고, 개개의 영역을 클러스터에 의해서 생성되고, 개개의 영역에 대한 사각형 영역을 계산하고, 사각형 영역에 대한 3가지의 휴리스틱을 적용한다. 사각형 영역의 비율과 사각형 영역의 채움 속성, 사각형 영역의 크기 등이다. 이 방법은 비디오 영상에서 수평방향으로 정렬된 텍스트에서 좋은 성능을 나타낸다.



▶▶ 그림 1. 텍스트 영역 검출 과정

었다. 계산의 복잡도를 줄이기 위해서 작은 윈도우 영역에서 오직 한 개의 에지 포인트만 크기와 위치가 측정되었다. 텍스트의 에지는 이 크기 정보를 이용하여 개선되었고 에지들을 연결하여 클러스터를 만들기 위해서 형태학적 팽창 방법이 수행되었다. 그리고 수평과 수직 비율과 높이 등과 같은 여러 가지 휴리스틱을 적용하여 텍스트 영역과 배경영역을 분리하였다. 이러한 특징들을 평가하기 위해서 두 개의 가보형태의 필터를 사용하였다. 그리고 이러한 필터의 결과 값을 평가하기 위해서 신경망을 사용하였다.

Zong et al.[4] 컬러 연속성과 그레이 레벨 값들의 분포 변화도를 특징으로 사용하였다. 두 개의 특징을 조합하여 하였다. 이 방법은 다양하게 스캔된 이미지에서 좋은 성능을 보이지만, 텍스트와 배경의 대비가 약하거나, 텍스트 컬러와 배경이 비슷한 경우 그리고 텍스트가 수평으로 정렬되지 않은 경우에는 성능이 저하되는 단점이 있다.

본 연구에서는 한글 텍스트가 갖는 에지 코너 특징을 기반으로 텍스트 영역을 검출한다. [그림 1]에 텍스트 영역 검출 방법에 대한 전체적인 구조를 보여준다. 자연 이미지로부터 에지를 검출한다. 에지 검출은 Canny-에지 검출 연산자[6]를 사용한다.

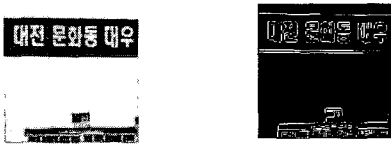
II. 한글 텍스트 영역 검출

1. 에지 검출 및 에지 맵 생성

본 일반적으로 텍스트 영역은 배경영역과 대비가 강하다는 특징을 갖고 있다. 따라서 이러한 특징을 추출하기 위해서 에지검출 연산자를 사용한다. Canny-에지 검출기를 사용하여 에지를 검출하였다. Canny-에지 검출기는 대비가 약한 이미지에서도 에지 성분을 가장 잘 검출한다. [그림 2]는 자연영상으로부터 Canny-에지 검출기를 이용하여 에지 성분을 추출한 결과이다.

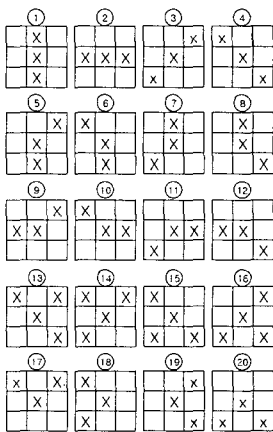
Canny-에지 검출기를 사용한 결과 텍스트 영역에 존재하는 에지를 잘 검출하는 것을 알 수 있다. 그러

나 [그림 2]에서 텍스트 영역이 아닌 건물영역에서도 에지가 검출되는 것을 볼 수 있다. 따라서 텍스트에 존재하는 에지와 건물등과 같은 텍스트가 아닌 영역에 존재하는 에지의 특징을 분석하여 텍스트와 배경을 분리해야 한다.



▶▶ 그림 9. Canny-에지 검출 결과

에지 맵은 에지 추출 결과 생성된 에지를 3x3영역에서 존재하는 에지의 형태를 20가지로 분류하였다. 이러한 분류 기준은 김수겸, 장유정[7]의 에지추출 방법에 따라 선정하였고, 4가지 구조는 텍스트 에지 코너 구조에 맞게 수정하였다. 20가지의 에지 맵은 2-연결 에지 성분과 12개와 8-연결 에지 성분 8개로



▶▶ 그림 3. 20가지 에지 구조

구성된다. [그림 3]은 20가지 에지 맵을 보여준다. [그림 3]에서 ①-⑫는 2-연결 성분을 갖는 에지들이고, ⑬-⑯는 3연결 성분을 갖는 에지 코너 성분들

이다. 그리고 ⑰-⑳은 [7]에서 3-연결성분으로 제안되었지만 텍스트 영역 검출을 위해서 에지 형태를 2연결 성분으로 수정하였다. 실험에서는 20가지 에지 맵 중에서 텍스트 영역 많은 비중을 차지하고 있는 수직과 수평에지(1,2) 그리고 수직과 수평을 연결하는 에지 코너 성분(5-12), 그리고 문자의 시작과 끝에서 나타나는 에지 코너(17-20)를 사용하였다.

2. 에지 맵 분석

텍스트 영역 검출을 위해서 20가지 에지 맵을 분석함으로써 텍스트 후보 영역을 검출할 수 있다. [그림 4]는 텍스트 영역의 에지 코너 구조를 보여주고 있다.

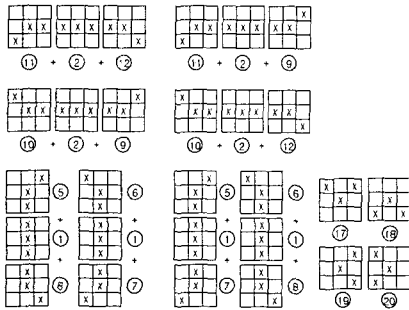
[그림 4]에서 텍스트 영역에 존재하는 여러 가지 에지 코너 구조를 볼 수 있다. 따라서 텍스트 영역을 검출하기 위해서 이러한 텍스트 영역에 존재하는 8가지 에지 맵의 조합을 추출하였고, 추가로 4가지의 에지 맵을 사용하였다. [그림 5]은 8가지 에지 맵의 조합과 4가지 에지 맵을 보여준다.

[그림 5]에서 ⑪,⑫,⑬의 조합은 텍스트 상단 부분에 주로 존재하고, ⑩,⑭,⑮의 조합은 텍스트 하단 부분에 주로 존재한다. ⑤,①,⑧의 조합은 텍스트 왼쪽



▶▶ 그림 4. 텍스트 영역 에지 구조

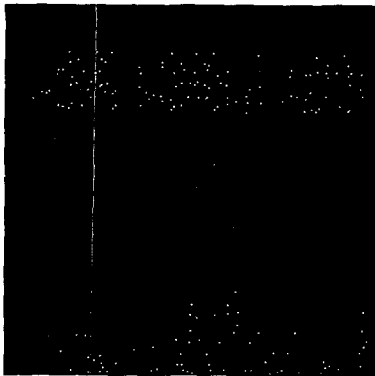
외곽부분에 주로 존재하고, ⑥,①,⑦의 조합은 텍스트 오른쪽 외곽부분에 주로 존재한다. 그리고 텍스트 내부의 수직부분에는 ⑤①⑦조합과 ⑥,①,⑧조합이 주로 존재하고, 수평부분에는 ⑪,⑫,⑨의 조합과 ⑩, ②,⑬의 조합이 주로 존재한다.



▶▶ 그림 5. 8가지 에지 맵 조합

3. 한글 텍스트 후보 영역 검출 및 텍스트 영역 검출

텍스트 후보 영역을 검출하기 위해서 텍스트 영역의 시작 부분과 끝 부분을 찾는다. 텍스트 영역의 시작 부분을 주로 [그림 5]에서 ⑪, ⑫, ⑬의 에지 맵



▶▶ 그림 6. 8가지 에지 맵의 분포

패턴을 갖는다. 그리고 텍스트 영역의 끝 부분은 대부분 ⑭, ⑮, ⑯의 에지 맵 패턴을 갖는다. 따라서 위와 같은 두 가지 형태의 텍스트 시작 영역의 에지 맵 패턴과 끝 영역의 에지 맵 패턴을 구조적 분석에 의해서 검출하여 텍스트 영역의 수평시작과 끝 좌표를 얻는다. [그림 6]은 텍스트에 존재하는 8가지 에지 맵 분포를 각각의 픽셀 포인트로 보여준다.

텍스트 시작 부분은 텍스트 시작 상단부분에 텍스트 배경의 특성이 존재하므로 에지가 존재하지 않는 특성

을 이용하여 효과적으로 검출가능하다. 텍스트의 끝 부분의 경우에도 마찬가지로 바로 하단의 에지 패턴은 배경의 특성으로 인해서 에지가 존재하지 않는다.

III. 실험 및 결과분석

1. 실험 환경

본 연구에 사용된 자연영상은 올림푸스 CW-5060 디지털 카메라 로 촬영된 이미지를 사용하였고, 실험에 사용된 이미지는 총 100개의 다양한 종류의 텍스트가 존재하는 자연영상을 대상으로 하였다. 실험 이미지의 해상도는 256x256을 사용하였다. 사용된 컴퓨터 환경은 펜티엄4-2.0GHz이고, 주기억 용량은 512MB이다.

[표 1] 텍스트 검출 성능 비교표

		Total	True	Part	Error	False
텍스트 이미지	영역 개수	100	75	10	15	13
	Precision = 77%					

2. 실험 및 성능 평가

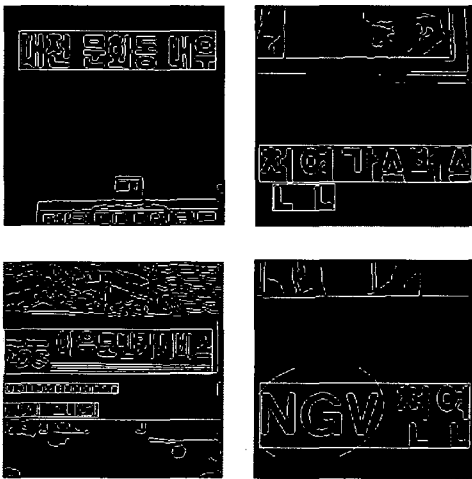
성능평가를 위해서 [표 1]과 같은 검출 성능 비교표를 사용하였다. 총 100개의 자연영상으로부터 텍스트 영역의 검출률을 측정하였다. Total은 실험 대상 이미지에 존재하는 전체 텍스트의 개수이며, True는 정확하게 검출한 텍스트의 개수, Part는 텍스트의 일부분만을 검출한 개수, Error은 텍스트를 검출하지 못한 개수, False는 텍스트 영역이 아닌 영역을 텍스트로 잘못 찾은 개수를 나타낸다. 텍스트는 실험자의 눈으로 인식 가능한 크기의 단어 단위 텍스트를 대상으로 설정하였다. Precision은 True/(True+Part+False)*100으로 계산하였다. 즉, Precision은 텍스트 영역 검출의 정밀도를 의미한다. [그림 7]에서 텍스트 영역 검출 결과를 보여준다.

텍스트 영역 검출의 정확성을 떨어뜨리는 경우는 한글 텍스트의 특성상 초성과 중성 그리고 종성으로

구성됨으로 중성이 받침으로 존재하는 경우 다른 텍스트 영역으로 분리되는 경우가 발생한다. 그리고 텍스트 영역과 비슷한 특징을 갖는 영역이 텍스트 영역으로 검출되는 경우도 발생한다. 또한 동일 수평 영역에 글자의 높이가 다른 문자가 배열하는 경우 텍스트 높이가 큰 텍스트를 기준으로 텍스트 영역이 검출된다.

IV. 결론

본 연구에서는 자연영상으로부터 텍스트 영역을 검출하는 방법을 제안하였다. 텍스트 영역에 존재하는 에지 코너의 형태를 분석하여 텍스트 영역에 존재하는 에지의 조합을 생성하여 에지 코너 특징만을 사용하여 텍스트 영역과 배경영역을 분리하여 텍스트 영역을 검출 할 수 있었다.



▶▶ 그림 7. 텍스트 영역 검출 결과

향후 연구과제는 텍스트 영역이 정렬되어있지 않아도 텍스트 영역을 검출 할 수 있는 방법과 비 문자 영역에 존재하는 에지 특성을 파악하여 텍스트 영역과 비 문자 영역을 효과적으로 분리하는 것이다. 그리고 색상정보를 이용한 텍스트 영역 검출 방법을 같이 적용하는 방법도 연구할 과제이다.

■ 참고 문헌 ■

- [1] A. K. Jain, B. Yu, "Automatic Text Location in Images and Video Frames," *Pattern Recognition*, Vol.31, No. 12, pp.2055-2076, 1998.
- [2] J. Ohya, A. Shio, S. Akamatsu, "Recognizing Characters in Scene images," *IEEE Trans. on PAMI*, Vol.16, No.2, pp.67-82, 1995.
- [3] M.A. Smith and T. Kanade, *Video Skimming for Quick Browsing Based on Audio and Image Characterization*, Technical Report CMU-CS-95-186, Carnegie Mellon University, July 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, 2001, pp.192-197.
- [5] Yassin M. Y. Hasan and Lina J. Karam, *Morphological Text Extraction from Images*, *IEEE Transactions on Image Processing*, 9 (11) (2000) 1978-1983.
- [6] Canny, J. "A Computational approach to edge detection", *IEEE Trans. Pattern Anal. and Machine Intelligence*, 8(6) pp.679-698, 1986.
- [7] 김수겸, 장유정, "에지의 구조적정보를 이용한 에지추출", *정보처리학회논문지*, Vol.3 No.5, pp.1337-1345, 1996.