

# 문자-에지 맵의 패턴 히스토그램을 이용한 자연이미지에서의 텍스트 영역 추출

박종천, 황동국, 이우람, 권교현, 전병민  
충북대학교 컴퓨터 공학과  
e-mail: simplepjc@hanmail.net

## Text Region Extraction using Pattern Histogram of Character-Edge Map in Natural Images

Jong-Cheon Park, Dong-Guk Hwang, Woo-Ram Lee,  
Kyo-Hyun Kwon, Byoung-Min Jun  
School of Electrical & Computer Engineering,  
Chungbuk National University

### 요 약

자연이미지에 포함된 텍스트는 많은 중요한 정보를 포함하고 있다. 그러므로 자연이미지에서 텍스트를 추출할 수 있다면 다양한 분야에서 활용될 수 있다. 본 논문에서는 문자-에지 맵 패턴 히스토그램 분석함으로써 텍스트 영역을 추출하는 방법을 제안한다. 캐니-에지 검출기로 에지를 추출하여 16가지 에지 맵을 생성하고, 에지 맵을 조합하여 문자 특징을 갖는 8가지 문자-에지 맵을 생성한다. 8가지 문자-에지 맵과 16가지 에지 맵을 이용하여 텍스트 후보 영역을 추출하고, 문자-에지 맵의 패턴 히스토그램 및 텍스트 영역의 구조적 특징을 이용하여 텍스트 후보 영역에 대한 검증은 수행하였다. 제안한 방법은 다양한 종류의 자연이미지를 대상으로 실험하였고, 복잡한 배경, 다양한 글꼴, 다양한 텍스트 컬러로 구성된 자연이미지에서 텍스트 영역을 효과적으로 추출하였다.

### Abstract

The text to be included in the natural images has many important information in the natural image. Therefore, if we can extract the text in natural images, It can be applied to many important applications. In this paper, we propose a text region extraction method using pattern histogram of character-edge map. We extract the edges with the Canny edge detector and creates 16 kind of edge map from an extracted edges. And then we make a character-edge map of 8 kinds that have a character feature with a combination of an edge map. We extract text region using 8 kinds of character-edge map and 16 kind of edge map. Verification of text candidate region uses analysis of a character-edge map pattern histogram and structural feature of text region. The method to propose experimented with various kind of the natural images. The proposed approach extracted text region from a natural images to have been composed of a complex background, various letters, various text colors effectively.

### 1. 서 론

자연이미지에 포함된 텍스트는 중요한 정보를 제공하고 있으므로, 텍스트를 실시간으로 추출할 수 있는 기술개발의 필요성이 증가하고 있는 추세이다. 최근 연구 중에는 시각장애인을 위한 시각보조 시스템이

개발 중에 있고[1], 또한 외국인이 여행 중에 휴대용 장비를 통해서 외국어로 된 지리정보와 관광정보 등을 텍스트 추출과 변환기술로 외국인들이 이해할 수 있는 문자로 변환 가능하도록 하는 시스템도 개발 중이다[2].

자연이미지로부터 텍스트 영역 추출에 관한 많은 연구가 진행되어 왔으며, 다양한 연구 방법이 제시되었다. 텍스트 영역 추출 방법들은 크게 두 가지 방법으로 분류할 수 있다[3]. 첫 번째는 연결요소 분석방법(Connected component based method)[4,5]이고, 두 번째는 에지 기반 분석 방법(Edge based method)[6,7,8,9,10]이다.

연결요소 분석 방법은 텍스트 영역 추출 속도가 빠른 반면 복잡한 이미지 배경에 영향을 많이 받는다. 에지 기반 분석 방법은 텍스트 영역의 위치가 정밀하지 못한 단점이 있다.

에지를 이용하는 기존의 연구 방법들이 단순하게 기본적인 에지의 특징을 이용함으로써 에지가 배경에 많이 존재하는 복잡한 이미지에서 텍스트 영역을 정확히 추출하지 못하는 문제점이 있었다. 따라서 본 연구에서는 이러한 문제점을 해결하기 위해서 기본적인 에지들의 특징을 조합하여 문자-에지 맵 패턴 히스토그램 분석과 텍스트 영역의 구조적인 특징을 이용함으로써 복잡한 배경을 갖는 이미지에서도 텍스트 영역을 정확하게 추출할 수 있도록 하였다.

**2. 텍스트 영역 추출**

텍스트 영역 추출은 그림 1과 같은 과정으로 수행하였고, 자연이미지를 캐니 에지 검출기[11]를 이용하여 에지를 추출하고, 추출된 에지로부터 16가지의 에지 맵을 생성한다. 생성된 에지 맵을 조합하여 문자 특징을 갖는 8가지의 문자-에지 맵을 추출한다. 문자-에지 맵의 및 구조적 특징을 이용하여 텍스트 후보영역을 추출하고, 추출된 텍스트 후보영역에 대한 검증은 문자-에지 맵 패턴 히스토그램 분포를 이용하여 최종적인 텍스트 영역을 추출하였다.

**2.1 에지 추출 및 에지 맵 생성**

에지 검출을 위해 컬러 이미지를 명도이미지(Gray-level)로 바꾼다. 명도이미지를 대상으로 캐니 에지 검출기로 에지를 검출한다. 텍스트 영역은 강한 대비 특성 갖고 있으므로 텍스트 영역의 에지를 정확하게 검출할 수 있다. 또한 텍스트 영역에 존재하는 에지는 기본적인 에지들의 조합으로 구성됨을 알 수 있으므로 텍스트 영역에 존재하는 에지의 특징을 추출위해 수평, 수직, 대각선 그리고 역 대각선 방향의 에지를 조합하여 16가지의 에지 맵을 생성한다.

에지 맵은 에지 추출 결과 생성된 에지를 3×3영역에 존재하는 에지의 형태를 16가지로 분류한 것이다. 에

지 맵의 분류 기준은 김수겸, 장유정[12]이 제안한 에지의 종류를 이용하였고, 4종류의 에지 구조는 텍스트 에지 구조에 맞게 수정하였다. 16가지의 에지 맵은 중심화소를 기준으로 2개의 방향으로 연결성을 갖는 성분으로 구성된다. 그림 2는 16가지의 에지 맵을 나타낸다.

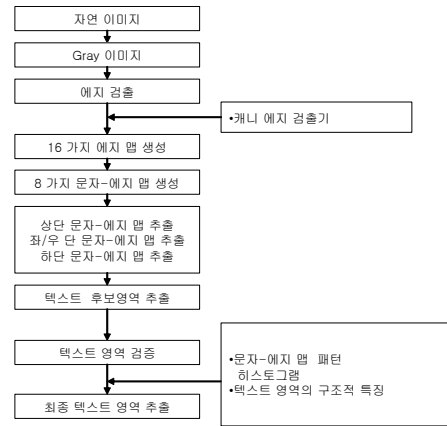


그림 1. 텍스트 영역 추출과정

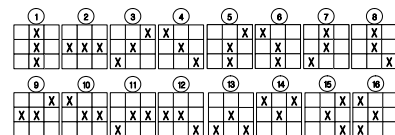


그림 2. 16가지의 에지 맵

텍스트 영역의 에지 형태를 분석한 결과, 그림 3의 텍스트 영역을 구성하는 16가지의 에지 맵이 서로 연결되어 문자를 이루고 있는 것을 알 수 있다. 따라서 텍스트 영역의 16가지 에지 맵을 조합하여 8가지의 문자-에지 맵(Character-Edge Map)을 생성하여 텍스트 영역 추출의 주된 특징으로 사용하였다.

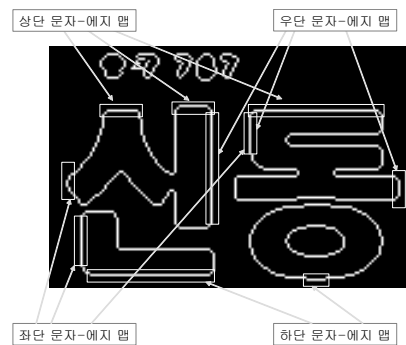


그림 3. 텍스트 영역 에지 구조

**2.2 문자-에지 맵 생성**

그림 3에서 텍스트 영역의 16가지 에지 맵 간에 상호 연결되어 문자를 형성하고 있는 구조적 특징을 잘 나타내고 있다. 따라서 이러한 텍스트 영역에 존재하

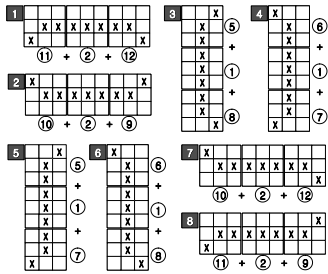


그림 4. 문자-에지 맵

는 에지 구조의 특징을 이용하기 위해 문자의 구조적 특징에 맞도록 16가지 에지 맵을 조합하여 8가지 문자-에지 맵을 정의한다. 그림 4는 8가지 문자-에지 맵을 보여준다.

문자-에지 맵은 위치에 따라 1부터 8까지의 번호를 부여하였다. 8가지 문자-에지 맵 중에서 주로 1,2,3,4번의 에지 맵을 사용하게 된다. 그 이유는 문자의 외각사각형 영역을 결정하는데 중요한 특징을 갖기 때문이다.

### 2.3 텍스트 후보영역 추출

텍스트 후보 영역을 추출은 텍스트 영역 너비를 설정하고 설정된 텍스트 영역의 너비를 시작으로 텍스트 높이를 검출하여 텍스트 영역 상/하, 좌/우 좌표를 설정한다. 설정된 텍스트 너비 영역을 바탕으로 텍스트 후보 영역 높이를 설정함으로써 텍스트 후보 영역을 검출한다.

텍스트 영역은 그림 4에서 제시한 8가지의 문자-에지 맵으로 구성된다. 그러므로 텍스트 영역에 존재하는 문자-에지 맵을 검출함으로써 텍스트 영역을 추출한다. 텍스트 영역 추출을 위한 구조적 특징과 경험적인 정보를 조건으로 제시한다.

- 조건 1. 텍스트 영역의 높이는 최소 문자 높이 값(10 픽셀) 이상이어야 한다.
- 조건 2. 텍스트 하단 문자-에지 맵(그림 5의 2번) 영역 아래에는 에지가 존재하지 않는다.
- 조건 3. 텍스트 상단 및 하단에지 사이에는 좌/우 에지(문자-에지 맵 “3,4”번)가 존재해야 한다.
- 조건 4. 텍스트 영역의 가로/세로 비율이 아래와 수식 1을 만족해야 한다.

$$Ratio_{wh} = \frac{Width\ of\ Region}{Height\ of\ Region} \geq 0.2 \quad (1)$$

일반적으로 텍스트가 존재하지 않는 영역도 텍스트 영역과 비슷한 문자-에지 맵을 갖는 경우가 존재한다.

그러나 텍스트 영역은 일정한 패턴으로 문자-에지 맵이 존재하는 것을 알 수 있다. 그림 5는 텍스트 후보 영역 추출결과를 보여준다. 그림 5에서 텍스트가 아닌 부분이 텍스트 영역으로 추출되는 것을 볼 수 있다. 따라서 텍스트 영역 검증 과정에서 문자-에지 맵 패턴 히스토그램을 이용하여 잘못 추출된 텍스트 영역을 제거 한다.

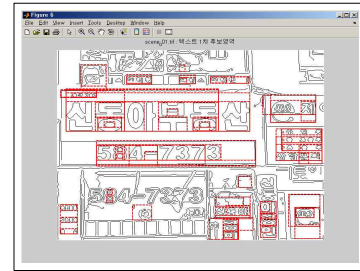


그림 5. 텍스트 후보 영역

### 2.5 텍스트 영역 검증

텍스트 영역에 대한 검증은 텍스트 후보 영역에 대한 문자-에지 맵 패턴 히스토그램 분석으로 이루어진다. 텍스트 영역에 대한 문자-에지 맵 패턴 분포는 그림 6과 같다.

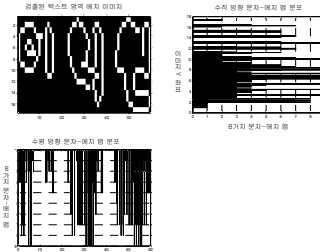


그림 6. 텍스트 영역 문자-에지 맵 패턴 히스토그램

그림 6은 정확히 검출된 텍스트 영역에 대한 수평 및 수직 방향에 대한 8가지 문자-에지 맵 패턴을 히스토그램으로 표현한 것이다. 수평방향의 문자-에지 맵의 패턴은 좌단 문자-에지 맵으로 시작해서 우단 에지 맵으로 끝난다. 그리고 그 사이에 상/하단 에지가 존재하고, 기타 5, 6, 7, 8번 문자-에지 맵이 일정한 거리를 갖고 분포한다.

수평방향의 문자-에지 맵 패턴은 상단 문자-에지 맵으로 시작해서 하단 문자-에지 맵으로 끝나고, 사이에 좌/우단 에지가 존재하며, 기타 5, 6, 7, 8번 문자-에지 맵이 일정한 거리를 갖고 분포하게 된다. 그림 7는 잘못 검출된 텍스트 영역의 문자-에지 맵 패턴 히스토그램 분포를 보여준다. 텍스트 영역이

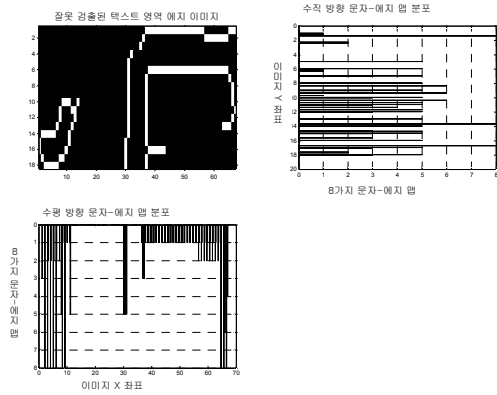


그림 7. 잘못된 텍스트 영역 문자-에지 맵 패턴 히스토그램

아닌 경우 문자-에지 맵 패턴의 분포가 규칙적이지 않음을 알 수 있다.

텍스트 영역 검증은 8가지 문자-에지 맵 패턴을 수평 및 수직 방향으로 히스토그램 분석함으로써 수행될 수 있었다.

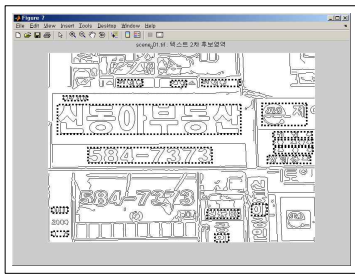


그림 8. 텍스트 영역 추출 결과

그림 8은 텍스트 영역 검증 결과 추출된 이미지로 대부분의 겹침 영역을 제거하였고, 텍스트 영역에 두 개의 텍스트 라인이 존재하는 것을 제거한다. 텍스트 영역의 최소 문자수를 2로 설정함에 따라 독립된 한 개의 문자와 세로로 배열된 문자는 텍스트 영역으로 검출하지 못한다.

### 3. 실험 및 결과분석

#### 3.1 실험 및 성능 평가

성능평가 결과는 표 1과 같은 텍스트 영역 추출률을 나타내었다. 적중률(Hit rate), 실패율(Miss rate), 오검출(False alarm)은 아래와 같이 정의한다.

적중률 = 
$$\frac{\text{검출된 텍스트 박스의 수}}{\text{이미지에 존재하는 텍스트의 박스의 수}} \times 100 \quad (2)$$

실패율 = 
$$100 - \text{hitrate} \quad (3)$$

오 검출 =

$$\frac{\text{잘못 검출된 텍스트 박스의 수}}{\text{이미지에 존재하는 텍스트의 박스의 수}} \times 100 \quad (4)$$

그림 9에서 텍스트 영역 추출 결과 이미지를 보여준다. 텍스트 영역 추출률은 안내판 이미지가 가장 높은 결과를 나타냈다. 그 이유는 안내판 이미지 배경의 단순함으로 정확하게 에지를 추출할 수 있기 때문이다. 가장 낮은 추출률은 간판 이미지로서 다양한 형태의 텍스트가 존재하고, 또한 배경이 복잡한 특징을 갖고 있기 때문에 텍스트 영역과 비슷한 영역이 많이 존재하기 때문이다.

표 1. 텍스트 영역 추출률

이미지 종류	실험 이미지 수	텍스트 영역 수	적중률	실패율	오검출
안내판	20	70	91%	9%	13%
간판	30	91	65%	35%	44%
교통표지판	10	99	84%	16%	12%
번호판	10	20	80%	20%	27%
교통안내	20	40	73%	27%	33%
광고	10	25	75%	25%	15%

### 4. 결론

본 연구에서는 자연이미지로부터 문자-에지 맵을 이용하여 텍스트 영역을 추출하는 방법을 제안하였다. 텍스트 영역에 존재하는 에지-맵과 문자-에지 맵을 이용하여 텍스트 후보 영역을 추출하고, 문자-에지 맵 패턴 히스토그램과 텍스트의 구조적 특징으로 텍스트 영역을 검증할 수 수행하여 텍스트 영역을 추출할 수 있었다. 따라서 본 논문에서 제안한 텍스트 영역 추출 기법은 에지의 구조적인 특징만 사용하여 텍스트 영역을 추출할 수 있다는 장점을 갖고 있다.

향후 연구과제는 텍스트 영역에 존재하는 에지를 정확히 추출하는 방법과 텍스트 영역이 아닌 영역에서 텍스트 영역과 비슷한 문자-에지 맵 패턴이 존재하는 문제를 해결하는 것이다. 이러한 문제점을 해결하기 위해서 전처리 과정에서 에지 추출을 위한 적절한 파라미터 값을 구하고, 연결요소 성분 분석 방법과 컬러 정보를 병합하는 방법에 대한 연구가 필요하다.



그림 9. 텍스트 영역 추출 결과 이미지

**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] K. Jung, K. I. Kim and A. K. Jain, "Text Information Extraction in Images and Video: A Survey", Pattern Recognition, Volume 37, Issue 5, pp.977-997, 2004.

[4] A. K. Jain, B. Yu, "Automatic Text Location in Images and Video Frames," Pattern Recognition, Vol. 31, No. 12, pp.2055-2076, 1998.

[5] J. Ohya, A. Shio, S. Akamatsu, "Recognizing Characters in Scene images," IEEE Transactions on PAMI, Vol. 16, No. 2, pp.67-82, 1995.

[6] 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, July 1995.

[7] 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.

[8] Y.M.Y. Hasan, L.J. Karam, "Morphological Text Extraction from Images", IEEE Transactions on Image Processing, Volume 9 issue(11), pp.1978-1983, 2000.

[9] 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.

[10] 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.

[11] 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.

[12] 김수겸, 장유정, "에지의 구조적 정보를 이용한 에지추출", 정보처리학회 논문지, Vol.3 No.5, pp. 1337-1345, 1996.