

# 자연영상에서 적응적 문자-에지 맵 특징을 이용한 텍스트 영역 검출

박종천, 황동국, 이우람, 전병민  
충북대학교 컴퓨터공학과  
e-mail: simplepjc@hanmail.net

## Text Region Detection using Feature of Adaptive Character-Edge Map in Natural Images

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

### 요 약

자연영상에 포함된 텍스트는 많은 중요한 정보를 포함하고 있으므로 자연영상에서 텍스트 정보를 검출하는 연구가 활발히 진행되고 있다. 본 논문에서는 문자 영역의 구조적인 특징을 배열문법으로 정의한 적응적 문자-에지 맵을 제안하여 텍스트 영역을 검출한다. 캐니-에지 검출기로 에지를 추출하고, 생성된 에지 이미지를 레이블링하고 그 영역의 문자구조 특징을 분석하기 위해서 적응적 문자-에지 맵을 분석한다. 적응적 문자-에지 맵의 분포 상태를 분석함으로써 텍스트 후보 영역을 검출하고, 텍스트 영역의 에지 히스토그램 프로파일을 분석함으로써 텍스트 후보 영역에 대한 검증을 수행하여 최종적인 텍스트 영역을 검출한다. 제안한 방법은 다양한 종류의 자연영상을 대상으로 실험하였고, 기울어진 텍스트와 다양한 크기의 텍스트 구성된 자연영상에서 텍스트 영역을 효과적으로 검출하였다.

### 1. 서 론

자연이미지에 존재하는 텍스트는 중요한 정보를 제공하고 있으므로, 이러한 텍스트를 검출함으로써 텍스트 정보를 이용한 응용분야의 활용도가 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 검출과 변환기술로 외국인들이 이해할 수 있는 문자로 변환 가능하도록 하는 시스템도 개발 중이다[2]. 텍스트 영역 검출 방법에 관련된 많은 연구가 진행되어 왔다[3-10].

본 논문에서 이용하는 에지 기반 방법은 단순하게 기본적인 수평, 수직, 대각선 방향의 에지 특징을 이용함으로써 에지가 배경에 많이 존재하는 복잡한 영상에서 에지가 많이 검출되어 텍스트 영역을 정확히 검

출하지 못하는 문제점이 있다. 따라서 본 연구에서는 이러한 문제점을 해결하고자 문자 영역에 존재하는 에지의 구조를 분석한 결과 문자에 존재하는 문자-에지 맵을 배열문법으로 정의하여 문자의 특징을 추출하고, 추출된 문자들의 클러스터를 분석하여 텍스트 영역을 검출한다.

### 2. 텍스트 영역 검출

텍스트 영역 검출은 그림 1과 같은 과정으로 수행함으로써 이루어진다. 캐니에지 검출기[11]로 추출된 에지 이미지를 레이블링 과정을 수행하여 레이블 영역을 생성한다. 레이블 영역에서 긴 수평 및 수직 형태의 영역은 제거된다. 또한 레이블 영역의 픽셀의 수가 10 픽셀 이하로 구성된 작은 영역은 제거된다.

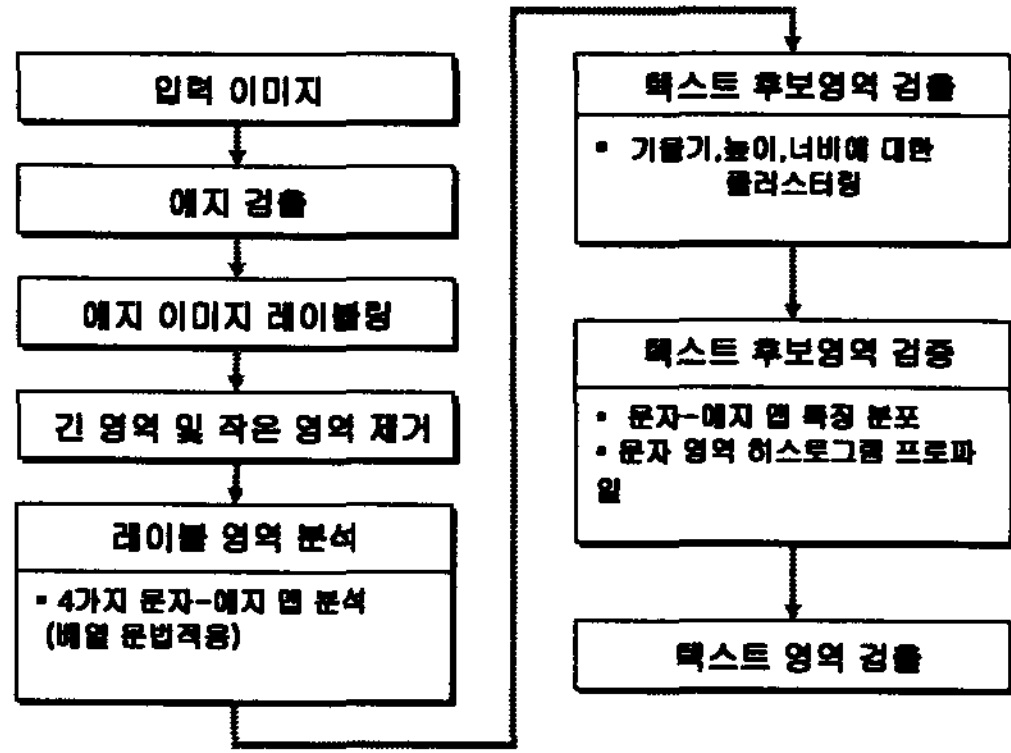


그림 1. 텍스트 영역 검출과정

2.1 레이블 영역 분석 및 적응적 문자-에지 맵

레이블 영역에 존재하는 에지 구조의 특징을 분석함으로써 레이블 영역의 에지 구조의 특징 값이 문자인지 아닌지를 결정한다. 본 논문에서 제안한 적응적 문자-에지 맵의 특징을 이용하여 문자 여부를 판단하였다. 적응적 문자-에지 맵은 문자의 크기와 기울기 등의 특성을 고려하여 정의된 것이다. 따라서 문자의 크기와 기울기 등에 적응적인 특징을 갖는다.

적응적 문자-에지 맵은 Freeman의 8 방향코드[12]의 기본 생성원을 바탕으로 에지 경계를 표현하고 배열 문법을 적용하여 문자-에지 맵의 구조를 정의하였다. 그림 2는 기울어진 텍스트 이미지로부터 4가지의 문자-에지 맵의 형태가 존재함을 보여준다. 따라서 1차 텍스트 후보 영역 검출과정에서 이러한 문자-에지 맵의 특징을 갖는 레이블 영역을 추출한다.

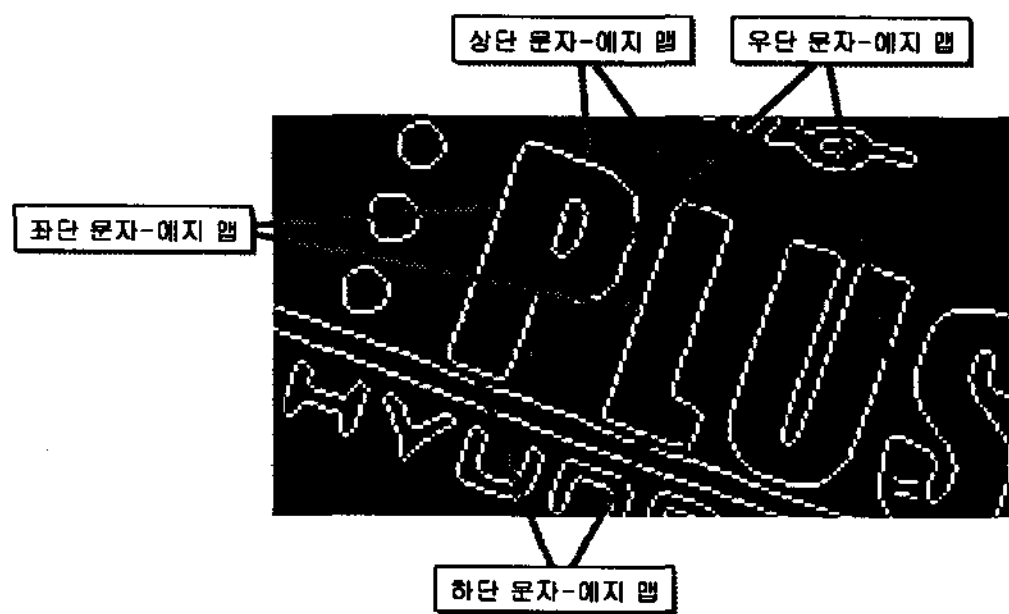


그림 2. 문자 영역의 문자-에지 맵 분포

2.2 적응적 문자-에지 맵 과 배열문법

적응적 문자-에지 맵의 기본적인 구조는 그림 3과 같다. 적응적 문자-에지 맵을 정의하기 위해서 체인코드로 에지의 경계를 추적하고, 에지 추적결과 4가지 형태의 문자-에지 맵을 검출하게 된다. 문자-에지 맵의 검출 알고리즘은 배열 문법을 적용하였다

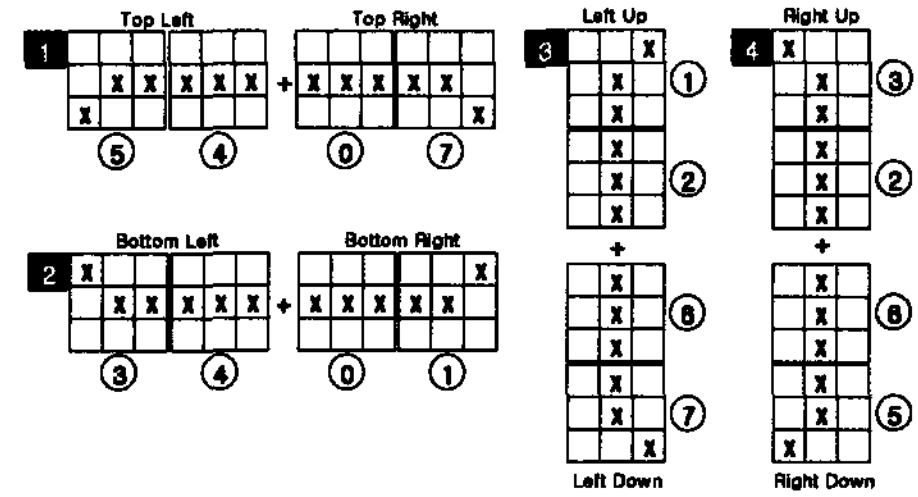


그림 3. 적응적 문자-에지 맵

그림 3에서 보여주는 문자-에지 맵은 문자 영역 상/하, 좌/우 경계에 존재하는 에지 형태를 의미한다. 그림 2의 문자들이 기울어져 있어도 이러한 특징은 존재하는 것을 볼 수 있다. 그러나 기울어진 특성으로 인해서 그림 3에서 제시한 문자-에지 맵과 똑같은 형태는 아닌 것을 알 수 있다. 따라서 기울어진 경우에도 문자-에지 맵의 형태를 검출할 수 있도록 적응적 문자-에지 맵을 제안하여 이러한 문제를 해결하였다.

적응적 문자-에지 맵은 배열문법을 이용하여 기술되는 것으로 배열문법은 패턴인식 방법 중에서 구조적 방법이고, 구조적 방법 중에서 형식언어 이론을 이용하는 구문론적 방법을 말한다[13-14]. 본 연구에서는 문자에지 패턴을 2차원 이미지에서 직접 파싱을 하는 방법 중의 하나인 동형배열(isometric array) 파서(parser)를 이용하여 문자에지 패턴을 추출한다.

2.3 동형배열문법

동형배열문법  $G$ 는  $G=(V_N, V_T, P, S, \#)$ 으로 정의된다. 여기서,  $V_N \rightarrow$  논터미널기호의 유한집합이고,  $V_T \rightarrow$  터미널기호의 집합이며,  $V_N \cup V_T = \emptyset, V_N \neq \lambda \neq V_T$ 이다.  $\#$ 은  $\#(V_N \cap V_T)$ 인 공백기호이고,  $P$ 는  $\alpha \rightarrow \beta$ 에 의해 표현되는 생성의 유한집합이며,  $S \in V_N$ 은 시작부호이다. 생성  $A \rightarrow B$ 와 같이 생성의 오른쪽이 단지 한 개의 논터미널을 갖고, 모든 기호의 이미지는  $V = V_N \cup V_T$ 이다. 즉,  $A \in V_N, \beta \in V^*$ 인 문법  $G$ 는 Context Free Grammar 이다.

- (  $A \rightarrow Bx, 0$  ) (  $A \rightarrow BC, 1$  ) (  $A \rightarrow Bx, 2$  )
- (  $A \rightarrow xB, 3$  ) (  $A \rightarrow Bx, 4$  ) (  $A \rightarrow xB, 5$  )
- (  $A \rightarrow Bx, 6$  ) (  $A \rightarrow xB, 7$  ) (  $A \rightarrow Bx, 8$  )
- (  $A \rightarrow BC, 9$  ) (  $A \rightarrow x, \quad$  )

여기에서,  $A, B, C$ 는 논터미널기호 이고,  $x$ 는 터미널기호이다.

위의 Context Free Grammar를 이용하여 그림 3에서 제시한 적응적 문자-에지 맵 중에서 1번의 Top-Right 패턴을 정의하면 아래와 같다.

동형배열문법  $G$ 가  $G = (V_N, V_T, P, A, \#)$  여기서,

$V_N = \{A, B, C\}$ ,  $V_T = \{x\}$ , A는 시작기호, 그리고  $\# \in \{0, 1, 2, 3, 4, 5, 6, 7\}$ 인 Freeman의 8-방향기호이다. 생성규칙 P는 다음과 같다.

1. ( A  $\rightarrow$  Bx, 0)    5. ( C  $\rightarrow$  xC, 6)
2. ( B  $\rightarrow$  xB, 0)    6. ( C  $\rightarrow$  xC, 5)
3. ( B  $\rightarrow$  xC, 7)    7. ( C  $\rightarrow$  x    )
4. ( C  $\rightarrow$  xC, 7)

이면, 문법 G에 의해 생성될 수 있는 이미지 L(G)는

$$L(G) = \begin{pmatrix} xxx & xxx & xx \\ & x & x & x \\ & x & & x \dots \\ x, & x, & x, \end{pmatrix}$$

= { 두 개의 에지 성분이 임의의 길이인 Top-Right }

이미지 L(G)는 가능한 형태의 Top-Right 문자-에지 맵을 보여준다. 그러나 기울어진 경우와 문자 크기가 다양한 경우는 문법 G에 의해서 모든 형태의 Top-Right 문자-에지 맵을 검출할 수 없으므로 기울기 정보를 이용하며, 시작 터미널로부터 기울기가 1보다 큰 에지가 존재하면 검출과정을 종료하게 된다. 다른 모든 형태의 문자-에지 맵도 위와 같은 배열 문법에 의해서 정의되고, 검출한다. 그림 4는 적응적 문자-에지 맵을 이용하여 검출된 문자 영역의 이미지이다.



그림 4. 적응적 문자-에지 맵으로 검출된 문자 영역

### 3. 텍스트 후보영역 검증

그림 4는 적응적 문자-에지 맵으로 검출된 이미지로 문자가 아닌 영역을 검출하는 것을 볼 수 있다. 따라서 문자가 아닌 영역을 제거하기 위해서 문자 영역에 대한 수평/수직 에지 히스토그램 프로파일을 분석함으로써 문자가 아닌 영역을 제거하였다. 수평/수직 에지 히스토그램 프로파일 수행결과 수평/수직 에지

의 개수가 2 미만인 것의 분포가 최소한 각각의 방향에 10% 미만이어야 한다. 그리고 식1과 같은 조건을 만족하는 것을 문자 영역으로 설정한다.

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

텍스트 영역에 대한 검증 결과 최종적인 텍스트 영역을 추출한다.

## 4. 실험 및 결과분석

### 4.1 실험 및 성능 평가

실험대상 이미지는 ICDAR 2003 표준 이미지와 실험실에서 디지털 카메라로 획득한 640×480 크기의 24비트 컬러 이미지를 사용하였다. 성능평가 결과는 표 1과 같은 텍스트 영역 검출률을 제시하였다. 적중률(Hit rate), 실패율(Miss rate), 오 검출(False alarm)은 아래와 같이 정의한다.

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

$$실패율 = 100 - 적중률 \quad (3)$$

$$오검출 = \frac{\text{잘못검출된 텍스트 영역의 수}}{\text{이미지에 존재하는 텍스트의 영역의 수}} \times 100 \quad (4)$$

표 1. 텍스트 영역 검출률

이미지 종류	실험 이미지 수	텍스트 영역 수	적중률	실패율	오검출
ICDAR 이미지	50	270	85%	15%	20%
실험실 이미지	50	250	90%	10%	15%

## 5. 결론

본 연구에서는 자연영상에서 적응적 문자-에지 맵을 이용하여 텍스트 영역을 검출하는 방법을 제안하였다. 텍스트 영역에 존재하는 문자-에지 맵을 이용하여 텍스트 후보 영역을 검출하고, 추출된 문자 영역의 수평/수직 방향 에지 히스토그램을 분석하여 최종적인 텍스트 영역을 추출하였다. 따라서 본 논문에서 제안한 텍스트 영역 검출 기법은 에지의 구조적인 특징만 사용하여 텍스트 영역을 검출하였고, 제안한 방법은 다양한 크기를 갖는 텍스트와 기울어진 텍스트를 효과적으로 검출하였다. 향후 연구 과제는 텍스트 영역 검출을 위한 전처리 과정에서 보다 더 정확한 에지 검출 방법과 텍스트 영역 검증을 위한 방법이 연구되

어야하며, 컬러정보와 연결요소 기반 방법의 장점을 적용하는 것이다.

#### 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] H. Freeman, "On the Encoding of arbitrary Geometric Configurations,". IEEE Trans. Electron Comput, vol. EC-10, pp. 260 - 268, June 1961.
- [13] A. Rosenfield, Picture Languages, Academic Press, NewYork, 1979.
- [14] Parktrick Shen-Pei Wang, "Hierarchical Structure and Complexities of Parallel Isometric Languages", IEEE Trans, on Pattern Recognition and Machine Intelligence, VoL. PAMI-5, No. 1. pp. 92-99, 1983.