모바일 영상에서 한글 문자의 기하학적 특징을 이용한 에지 기반 텍스트 검출

박종천*, 오명관**,전병민*
*충북대학교 컴퓨터공학과
**혜전대학교 디지털서비스과
E-mail:simplepjc@hanmail.net

Edge-based Text Localization Using Geometrical Features of Hangul Character in Mobile Images

Jong-Cheon Park*, Myoung-Kwan Oh**, Byeong-Min Jeon*
*Dept. of Computer Engineering, Chungbuk National University
**Dept. of Digital Service, Hyejeon College

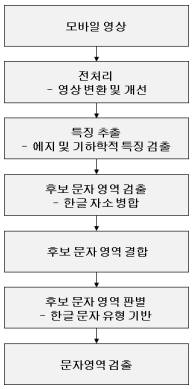
요 약

최근 모바일 기기의 보급이 일반화됨으로서 모바일 영상을 다루는 많은 응용프로그램이 개발되고 있다. 모바일 영상을 분석하여 정보를 추출한 결과를 인터넷 검색의 키워드로 연계함으로서 직관적인 멀티미디어 검색을 가능하도록 한다. 본 연구는 모바일 영상에 포함된 한글 문자영역을 검출하는 방법을 제안하였다. 한글 문자의 기하학적인 특징을 추출하고 이를 분석함으로서 후보 한글문자 영역을 검출하고, 검출된 후보 문자영역을 한글 자소 병합 알고리즘을 이용하여 병합한다. 그리고 후보 문자 영역을 한글 6가지 한글 문자 유형 특징을 이용하여 한글 문자 영역을 판별함으로서 최종적인 한글 문자 영역을 검출한다. 실험결과, 문자영역 검출률의 성능 평가 요소로서 재현률이 향상됨을 알 수 있었다.

1. 서 론

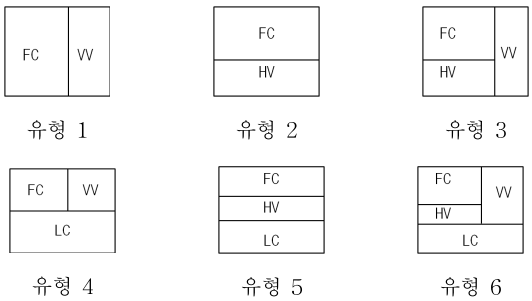
모바일 기기의 사용이 보편화되고 이에 발맞춰서 모바일 응용 소프트웨어의 중요성이 부각되고 있다. 모바일 기기 중에서 가장 일반적인 모바일 폰에 장착된 카메라, GPS, 중력감지 센서, Wi-Fi, 그리고 터치 센서 등을 기반으로 다양한 응용 프로그램 개발이 급속도로 이루어지고 있고, 이를 활용함으로서 보다 더 편리한 생활을 영위하고 있다. 영상을 기반으로 하는 분야에서는 카메라를 이용한 보안, 식별, 인식, 그리고 로봇 등의 다양한 응용 분야에서 연구 개발에 박차를 가하고 있다.

시각장애인을 위한 시각보조 시스템이 개발 중에 있고[1], 외국인이 여행 중에 휴대용 장비로 외국어로 된 지리정보와 관광정보를 한글문자 영역 검출과 변환기술로 외국어로 변환 가능하도록 하는 시스템도 개발 중이다[2]. 최근에는 스마트폰의 보급 확대로 증강현실의 한 분야로 모바일 카메라로 획득된 영상을 분석하여 문자영역 검출하고 인식함으로서 얻은 정보를 검색엔진으로 연계하여 멀티미디어 검색을 가능하도록 하는 시스템도 연구되고 있다.



[그림 1] 한글문자 영역 검출 순서도

본 연구는 모바일 영상에서 한글문자의 기하학적인 특징과 한글 고유의 6가지 문자 유형을 이용하여 한글 문자영역을 병합하고 그 결과를 판별함으로서 한글 문자영역을 검출하는 방법을 제안하였다. 문자의 기하하적인 특징을 이용한 연구사례는 주로 문자 인식에서 많이 이용되는 방법이다[7-8]. 제안 방법의 전체적인 순서도는 그림 1에 제시하였다.



[그림 2] 한글문자 유형

2. 한글 문자 영역 검출

RGB 24컬러 영상을 256 밝기(Gray-scale) 영상으로 변환하고, 에지 특징을 검출하기 위해서 캐니-에지 연산자[3]를 사용하였다. 검출된 에지 성분을 분석하여 한글 문자가 갖는 특징이 있는지 조사하여 이에 해당하지 않는 것은 제거함으로서 후보 문자 영역을 검출한다. 검출된 후보 문자영역은 한글 자소 병합 과정과 판별 과정을 수행함으로서 문자영역을 검출한다.

2.1 한글 자소 병합

추출된 후보 문자영역은 한글 자소 병합과정을 수행함으로서 하나의 문자영역으로 병합된다. 한글은 초성, 중성, 종성으로 구성됨으로 이러한 특징으로 초성을 기준으로 오른쪽과 아래방향으로 병합을 진행함으로서 정확한 한글 자소 병합을 수행한다. 그림 2는 한글의 6가지 문자 유형을 제시하였다. 그림 2의 FC(First Consonant)는 초성, LC(Last Consonant)는 종성, VV(Vertical Vowel)는 수직모음, HV(Horizontal Vowel)는 수평모음을 의미한다. 그림 2의 한글 문자 유형은 한글 문자가 갖는 고유한 특징으로서 기하학적인 특징을 갖고 있음을 보여준다.

한글 자소 병합과정에서 오른쪽 방향과 아래쪽 방향의 자소의 기하학적인 특징을 고려하여 병합이 이루어진다. 오른쪽 방향은 한글 자소가 존재한다면 항상 수직모음 성분이므로 이를 검사함으로서 한글 자소를 병합한다. 한글 자소의 수직 모음 성분 특징 값을 표 1에 제시하였다.

2.2 한글문자 후보영역 판별

한글 자소 병합과정으로 검출된 후보 문자영역은 6가지 한글 문자 유형[4] 알고리즘을 이용하여 한글 문자영역 여부를 판별한다. 표 3에 한글 6가지 유형을 제시하였다. 표 2는 6가지 한글 유형별 특징을 보여준다. 수직/수평 방향으로 Run의 수와 수직모음의 기하학적 위치를 갖는 것을 알 수 있다. 이러한 기하학적인 유형별 특징으로 한글 문자영역 여부를 판별한다.

[표 2] 6가지 한글 유형별 특징

6 가지 한글 유형	Run의 수		모음 위치	
	수평	수직	수평	수직
I	[2, 3]	[2, 4]	없음	오른쪽
II	[2, 4]	[2, 4]	중간 아래	왼쪽
III	[2, 3]	[2, 5]	없음	오른쪽 위
IV	[3, 3]	[2, 6]	중간	없음
V	[2, 3]	[2, 4]	중간	오른쪽
VI	[2, 4]	[4, 5]	중간	오른쪽

[표 1] 한글 자소의 수직 모음 성분 특징 값

수직 모음	Run의 수		위치	
	수평 방향	수직 방향	수직 획	수평 획
ㅣ	1	1	왼쪽	없음
ㅏ	1	1	왼쪽	중앙
ㅑ	1	2	왼쪽	중앙 위 중앙 아래
ㅓ	1	1	오른쪽	중간
ㅕ	2	1	오른쪽	중앙 위 중앙 아래
ㅐ,ㅔ	[1, 2]	2	왼쪽 오른쪽	중앙 위 중앙 아래

3. 실험 결과 및 결론

본 논문에서는 모바일 폰 카메라로 획득한 자연 영상에 존재하는 한글 문자영역을 검출하기 위해 한글 문자의 기하학적인 특징 및 한글 유형 특징을 이용한 방법을 제안하였다. 실험의 성능을 평가하기 위해서 실험 영상은 KAIST 인공지능 연구실 사이트[5]에서 제공하는 자연영상으로 하였다. 실험 대상 영상은 모바일 폰 카메라를 사용하여 촬영되었고, 표 4과 같은 다양한 환경의 영상을 대상으로 하였다. 성능 평가를 위한 방법은 식 1의 정확률(Precision)과 식 2의 재현율(Recall)로 평가하였다[6].

$$p = \frac{c}{|E|} \quad (1)$$

$$r = \frac{c}{|T|} \quad (2)$$

식 1과 2에서 c는 문자영역의 수를 나타내고 식1의 |E|는 검출된 문자영역의 수를 의미하고 식2의 |T|는 영상에 존재하는 모든 문자영역의 수를 의미한다.

표 3은 제안한 방법으로 자연영상에서 문자 영역을 검출률을 보여준다. 실외 영상과 그림자 영상은 글자영역 주변에 복잡한 영상이 존재함으로서 전체적으로 정확한 에지 검출이 어렵고, 글자 영역이 아닌 영역에서 글자와 비슷한 에지를 검출하는 오류를 초래함으로서 낮은 정확률을 보인다. 이러한 문제점은 한글 문자 유형에 대한 추가적인 기하학적 특징을 추출하고 이를 문자영역 판별 기준으로 설정함으로서 해결될 수 있도록 해야 할 것이다. 또한 기울어진 문자영역의 경우에는 Run의 특징을 찾는데 문제가 있으므로 영상의 기울기를 검출하고 이를 기준으로 각각의 문자 영역에 대한 기울기 보정함으로서 기울어지 영상에서도 그 특징을 검출할 수 있도록 하는 것이다.

[표 3] 문자영역 검출 성능 평가 결과

영상종류	Recall	Precision
Outdoor	0.88	0.83
Indoor	0.93	0.92
Shadow	0.88	0.85
Light	0.94	0.83
Book Cover	0.96	0.93
Others	0.92	0.87
전체 평균	**0.92**	**0.87**

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] Canny, J(1986), “A Computational Approach to Edge Detection”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-8, No. 6, pp.679-698.

[4] 신봉기, 김진형, “자소 탐색에 기반한 온라인 한글 인식”, 한국정보과학회 논문지(B), Vol.23 No.11, pp. 1135-1144, 1996.

[5] http://ai.kaist.ac.kr/home/node/85

[6] 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.

[7] Jinsu Jo, Jihyun Lee, Yillbyung Lee , “Stroke-Based Online Hangul/Korean Character Recognition”, Pattern Recognition, 2009. CCPR 2009. Chinese Conference on Date of Conference: 4-6 Nov. 2009 pp. 1-5, 2009.

[8] Kumar V., Singh, A.K., Negi, T., Miglani, J. “Character Recognition Using Geometrical Features of Alphabet: A Novel Technique”, Communication Software and Networks, 2009. ICCSN '09. International Conference on Date of Conference: 27-28 Feb., pp. 119 - 125, 2009.