

해리스 코너 검출기를 이용한 배경 영상에서의 문자 검출

김민하* · 김미경* · 차의영*

*부산대학교

Character Detection in Complex Scene Image using Harris Corner Detector

Min-ha Kim* · Mi-kyung Kim* · Eui-young Cha*

*Pusan National University

E-mail : minhakim@pusan.ac.kr

요 약

본 논문은 복잡한 배경 영상에서 필기체가 아닌 수평, 수직 성분이 많이 포함된 문자 검출 방법을 제안한다. 본 논문에서 검출하고자 하는 문자는 코너 성분이 많이 밀집되어 있으며 배경 영상은 그에 비해 코너 성분이 적고 드문드문하다는 특징을 이용하여 먼저 해리스 코너 검출기를 이용하여 전체 영상에서 코너를 검출한다. 검출된 코너들의 위치 정보를 이용해 밀집되어 있는 코너들을 클러스터링 함으로써 문자 영역을 검출한다. 검출된 문자 영역간의 위치 정보와 히스토그램 분포를 비교하여 비슷한 특징을 갖는 영역들을 합치고 문자 성분의 특징을 갖지 않는 영역은 필터링 하여 문자 영역을 개선한다. 문자 영역에서 R채널, G채널, B채널 각각의 채널에 대한 히스토그램 분포를 분석하여 문자를 검출한다.

ABSTRACT

In this paper, we propose a detection method of the character rather than cursive, containing many components of the vertical and horizontal direction in complex background image. The characters have many dense corners but the background has few sparse corners. So we use harris corner detector and cluster the corners by using the position of the detected corners for detecting character regions. To merge or filter character regions, we analysis a histogram of gray image of character regions. In each improved region, we compare histograms of R, G, B channels to detect characters.

키워드

Character Detection, Complex Image Binarization, Harris Corner, Histogram Analysis

I. 서 론

본래의 영상만으로는 표현하고자하는 의미를 확실하게 전달하지 못할 때 영상에 문자를 삽입하여 명확하게 의미를 전달한다. 따라서 영상의 배경보다는 삽입된 문자를 인식하여 그 의미를 파악해야 하는 경우가 있다.

영상에 삽입된 문자를 인식하기 위해서는 먼저 영상에서 배경을 제외한 문자 영역을 검출하고 그 문자 영역 내에서 문자를 검출해야 한다.

본 논문에서는 영상에 삽입된 문자들은 필기체와는 다르게 수평, 수직 성분이 많이 포함되어 있고 주위 배경과 뚜렷한 색상 차이를 갖는 특

징을 이용하여 해리스 코너 검출기로 문자 후보 영역을 검출하고 R, G, B 각 채널별 히스토그램 분석을 통해 문자를 검출하는 방법을 제안한다.

II. 관련 연구

엣지 기반의 문자 검출 방법은 엣지 정보가 색상과 밝기 등에 관계없이 문자 검출에 사용하기 좋은 특징이라는 사실을 기반으로 엣지의 크기, 밀도, 방향을 주요 특징으로 사용한다[1].

모폴로지 연산을 이용한 문자 검출 방법은 문

자가 배경보다 밝은 색 일 때에는 열림 연산을 반복한 영상에서 닫힘 연산을 반복한 영상을 차영상하고 문자가 배경보다 어두운 색 일 때에는 닫힘 연산을 반복한 영상에서 열림 연산을 반복한 영상을 차영상하여 문자를 검출한다[2,3,4]. 이는 모폴로지 연산의 반복 횟수에 따라 검출되는 문자의 크기가 다양하다는 장점이 있다. 하지만 문자의 밝기를 사전에 알고 있어야 하고 모폴로지 연산 횟수를 결정하는 것에 어려움이 있다.

다평면 분할(Multi-plane Segmentation)을 이용한 문자 검출 방법은 원 영상을 여러개의 작은 영역으로 분할하고 각 영역 내에서 지역적, 전역적으로 특징을 추출하고 유사한 특징을 갖는 성분끼리 하나의 평면에 표현하여 각 평면마다 문자를 검출하여 합하는 방법이다[5]. 이는 다른 방법들 보다 적응적으로 문자 검출이 가능한 장점이 있다.

III. 본 론

3.1 문자 영역 검출

영상에 어떠한 정보를 글로써 나타내기 위해 추가적으로 삽입한 문자들은 수직 성분과 수평 성분을 많이 함유하고 두 성분이 교차하는 지점에서 코너 성분이 나타난다. 문자들은 수평 방향으로 정렬되어 있다는 가정 하에 한 단어 이상의 글자들이 수평 방향으로 정렬되어 코너 성분이 수평 방향으로 밀도가 높고 그에 비해 문자가 아닌 배경 영상은 수직 성분과 수평 성분을 적게 함유하고 있고 이들은 드문드문하게 나타나는 특징을 이용하여 해리스 코너 검출기를 사용하여 영상에 삽입된 문자의 코너 성분을 검출한다.

검출된 코너 간의 수평 거리와 수직 거리를 비교하여 일정 거리 이내에 있는 코너들을 하나의 문자 후보 영역으로 검출한다.

3.2 문자 영역 병합 및 문자 검출

문자 후보 영역간의 병합은 그림 1의 순서에 의해 진행된다.

먼저 문자 영역 후보군을 R_{all} 이라 하고 $R_a, R_b \in R_{all}$, $a \neq b$ 인 두 영역 R_a, R_b 가 식 1을 만족할 때 이 두 영역을 이웃 영역으로 판단한다.

$$\begin{aligned}
 & (R_y(a) > R_y(b) \text{ and} \\
 & R_y(a) < R_y(b) + R_h(b)) \\
 & \text{or} \\
 & (R_y(a) + R_h(a) > R_y(b) \text{ and} \\
 & R_y(a) + R_h(a) < R_y(b) + R_h(b))
 \end{aligned} \tag{1}$$

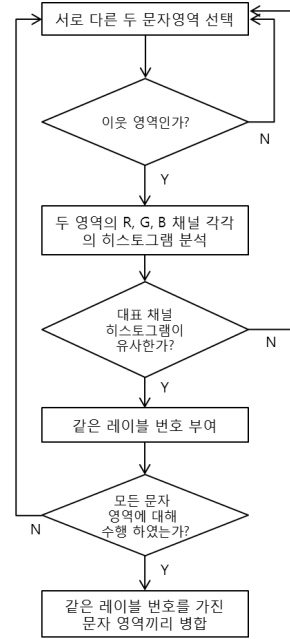


그림 1. 문자 후보 영역 병합 순서도

식 1에서 $R_y(k)$ 는 문자 후보 영역 k의 y좌표, $R_h(k)$ 는 문자 후보 영역 k의 높이를 의미하고 문자 후보 영역의 좌표는 그림 2와 같다.

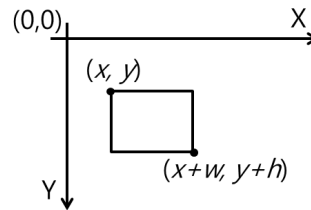


그림 2. 영상의 픽셀 좌표계를 기준으로한 문자 경계 사각형의 좌표

이웃 영역 각각에서 R, G, B 3채널의 히스토그램 분포를 구한다. 그리고 각 히스토그램에서 그림 3과 같이 peak에 해당하는 두 개의 색상 값 $I_1, I_2 (I_1 < I_2)$ 를 구하고 I_1 과 I_2 의 차를 구한다. 그림 3에서 I 는 각 채널의 0~255사이의 색상 값이고 N 은 해당 색상 값을 가지는 픽셀 수를 의미한다. 여기서 N 은 0~1사이의 정규화된 값이다. $peak1$ 과 $peak2$ 이는 두 peak값을 나타낸다. I_1 과 I_2 는 각 채널에서 배경과 문자의 대표 색상 값을 나타내고 I_1 과 I_2 의 차는 배경 색상과 문자 색상의 차이를 나타낸다. R, G, B 채널의 히스토그램 중에 I_1 과 I_2 의 차가 가장 큰 채널을 해당 후보 영역의 대표 히스토그램으로 선택한다.

이웃한 두 문자 후보 영역의 대표 히스토그램 간의 분포 차이를 식 2로 측정하고 D_{hist} 가 20 이

하일 때 두 문자 후보 영역을 같은 문자 영역으로 판단하고 같은 레이블(label)을 부여한다. 임계치 20은 실험에 의해 결정되었다.

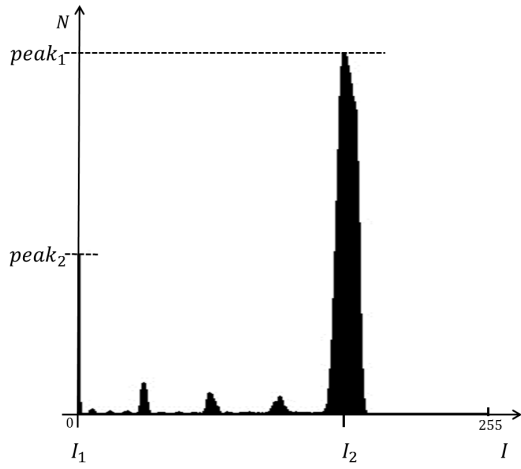


그림 3. 정규화된 히스토그램에서 peak값을 가지는 두 색상 값 I_1, I_2

모든 문자 영역에 대해 레이블링(labeling)을 마치고 나면 같은 레이블을 가진 문자 영역들을 하나로 병합한다.

$$D_{hist} = |I_k(a) - I_k(b)|, \quad k=1,2 \quad (2)$$

문자 후보 영역들을 레이블링하고 병합한 후 수평 성분과 수직 성분이 포함되어 있지 않은 영역은 제거한다.

검출된 문자 영역의 대표 히스토그램에서 그림 4와 같이 I_1, I_2 사이에서 색상 분포가 가장 적은 색상 값 중 I_1, I_2 의 평균값에 가장 근접한 값을 임계치로 해당 채널 영상을 이진화하여 문자를 검출한다.

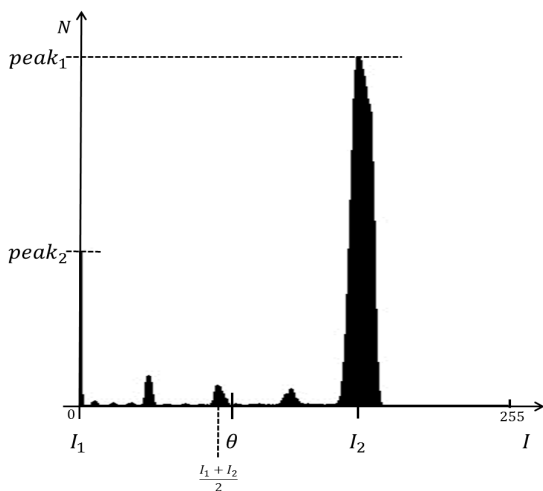


그림 4. 그림 3의 의 평균값에 가장 가까운 임계치 θ

IV. 실험 결과

실험에서는 상품 광고 영상들을 사용하여 Y. L. Chen[5]의 방법과 비교하였다. 그림 5의 (a)는 제안하는 방법으로 검출된 문자 영역이고 (b)는 Y. L. CHEN의 방법으로 문자를 검출한 결과이고 (c)는 제안하는 방법으로 문자를 검출한 결과이다.

그림 5의 (1) ~ (4)는 문자의 배경 색상이 복잡한 경우에도 문자 검출이 잘 되는 것을 보여준다. 특히 (1), (3), (4)은 Y. L. CHEN의 방법보다 제안하는 방법을 사용했을 때 우수한 성능을 보인다. (5), (6)은 문자 주위 배경에 코너 성분이 많이 분포되어 배경 영역까지 문자 영역으로 잘못 검출된 경우이다. 이러한 경우에는 문자 검출 시 배경 영역을 제거함으로써 문자를 검출할 수 있음을 보여준다. 특히 (6)의 경우에는 Y. L. CHEN의 방법보다 제안하는 방법을 사용했을 때 우수한 성능을 보인다.



그림 5. (a) 제안하는 방법으로 검출된 문자 영역, (b) Y. L. Chen의 방법으로 문자를 검출한 결과, (c) 제안하는 방법으로 문자를 검출한 결과

V. 결론

본 논문은 배경 영상에 의미를 강조하기 위하여 추가적으로 삽입된 문자 검출을 목적으로 한다. 수평, 수직 성분이 많이 포함된 문자의 코너 성분을 주요 특징으로 하여 문자 영역을 검출하고 R, G, B 각 채널의 히스토그램 분석 및 비교

를 통하여 문자를 검출한다.

제안하는 문자 영역 검출 방법은 문자 영역 검출에 있어서는 문자 주위 배경의 색상과는 무관하게 문자 영역 검출이 가능하나 배경에 수평, 수직 성분에 의한 코너가 다수 존재하는 경우에는 배경 영역까지 문자 영역으로 검출되는 문제가 있다. 하지만 이런 문제는 이후 문자 검출을 통해 배경을 제외한 문자만 검출함으로써 해결 가능하다.

참고문헌

- [1] X. Liu and J. Samarabandu, "Multiscale edge-based text extraction from complex images," pp. 1721-1724, 2006.
- [2] Y. M. Y. Hasan and L. J. Karam, "Morphological text extraction from images," *IEEE Trans. on Image Process.*, vol. 9, no. 11, pp. 1978-1983, 2000.
- [3] M. S. Uddin, T. Rahman, U. S. Busra, and M. Sultana, "Automated Extraction of Text from Images using Morphology Based Approach," 2012.
- [4] L. Gu, N. Tanaka, R. M. Haralick, and T. Kaneko, "A framework for the assessment of text extraction algorithms on complex colour images," pp. 470-473, 2010.
- [5] Y. L. Chen, "Automatic Text Extraction, Removal and Inpainting of Complex Document Images," in *International Journal of Innovative Computing*, 2011.