

교통 표지판에서 텍스트 추출 및 기울기 추출

Text Extraction and Skew Detection in Traffic Sign

최규담
(광운대학교, 석사과정)

김성동
(계원조형예술대학, 교수)

최기호
(광운대학교, 교수)

Key Words : 표지판, 텍스트, 기울기, 인식,

목 차

- | | |
|------------------------|----------------|
| I. 서론 | 3. 텍스트 영역 분류 |
| II. 교통 표지판에서 텍스트 추출 설계 | 4. 후처리 |
| 1. 전처리 | III. 실험결과 및 분석 |
| 2. 텍스트 후보 영역화 | IV. 결론 |

I. 서론

최근 컴퓨터의 발전과 인간의 시각기관과 유사하게 컴퓨터와 로봇에 부착된 카메라를 통해 자연을 받아들여 적절히 분석하고 이해하고자 많은 연구가 진행되었다. 그 중 특히 교통 표지판의 텍스트 정보는 지역 위치나 방향 지시 등에 중요한 정보를 주지만 영상이나 그래픽과 같은 그림 형태로 구성되어 있기 때문에 당장 직접 활용에는 많은 제약점들이 따르게 된다[1][2].

자연 영상 특히 교통 표지판에서 텍스트 추출을 위한 연구는 이제껏 많은 방법들이 제안되고 있다[1-7]. 기존 연구로써 텍스트를 구성하는 연결요소의 크기와 밀도는 그래픽의 경우와 다르다는 특징을 이용한 Page Segmentation 방법[3]과 텍스트 부분의 흑화소 분포의 복잡도를 수치화하는 방법으로 Cross correlation approach[4], Kolmogorov complexity measure[5]등이 제안 되었다. 그러나 이러한 연구들은 대부분 배경이 단순한 이미지를 대상으로 실험 하였으며, 텍스트의 크기나 위치, 기울기 등에 심한 제약을 받는 단점이 있다.

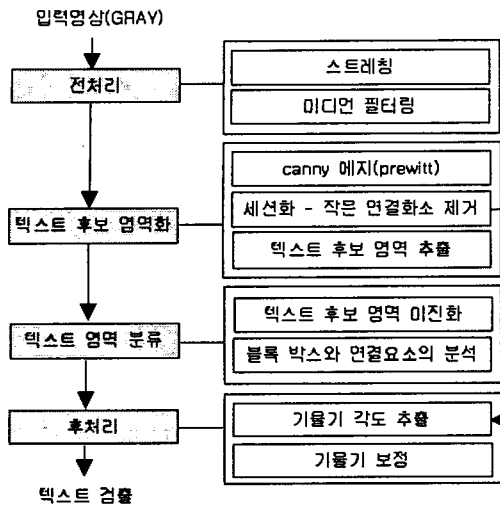
또한 텍스트의 기울기 추출을 위한 기존 연구로는 허프변환(Hough transform)[6], 투영 윤곽(profile)분석에 의한 방법[7] 등을 사용하고 있으나 처리 시간이 늦고 알고리즘의 구조상 텍스트를 추출하기에는 다소 어려운 점이 있다.

본 논문에서는 위에서 살펴 본 기존의 연구들의 단점들을 보완하여 교통 표지판에 포함되어 있는 텍스트 정보를 정확하게 추출하고 기울어진 텍스트를 보정한다. 여기서 제안한 방법은 명도 이미지를 가지고 밝기 차를 이용한 에지영상을 사용한다. 에지영상은 세선화(Thinning)를 거쳐 잡영과 선 정보를 제거하고 세선화를 거친 영상은 기울기 각도를 추출할 때 다시 사용된다. 잡영과 선이 제거된 영상은 텍스트들의 특

징들을 분석함으로써 텍스트 후보영역을 추출하게 되고 여기서 구해진 후보영역 위치정보를 가지고 스트래칭 된 영상으로부터 각 영역별로 이진화 과정을 수행하게 된다. 다음으로 후보영역의 외각 사각형과 연결요소들의 관계를 분석하여 불필요한 연결요소를 제거함으로써 텍스트를 검출하게 된다. 또한, 기울어진 텍스트의 기울기 각도를 추정하기 위해 전에 세선화를 거친 영상에서 5x5 윈도우를 사용하여 가로·세로축으로 각각 $-45^\circ \sim 45^\circ$ 인 직선을 찾아낸다. 찾아진 직선들은 각각의 각도를 구해 확률적으로 가장 많이 나온 각도를 가지고 회전벡터(rotation vector)를 이용하여 회전 함으로써 바르게 보정한다

II. 교통 표지판에서 텍스트 추출 설계

본 논문에서 제안하는 텍스트 추출과정은 네 단계를 거친다. 첫 번째 에지추출을 위한 전처리 과정단계를 이루고 두 번째는 텍스트 후보영역화(segmentation) 단계 세 번째는 텍스트 영역과 비텍스트 영역을 분류(Block classification)하는 단계 마지막으로 후처리로써 추출된 텍스트의 기울기를 추정·보정하는 단계로 이루게 있다. [그림 1]은 이 과정을 요약하여 나타낸 그림이다.



<그림 1> 텍스트 추출 과정

1. 전처리

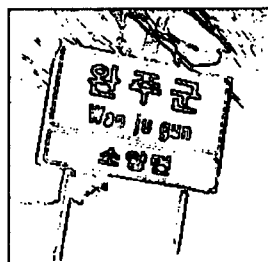
교통 표지판의 텍스트는 중요한 정보를 알려줘야 하기 때문에 사람의 눈으로 잘 구별할 수 있도록 그 주변 배경과 보색관계를 가지고 있다. 따라서 본 연구는 텍스트 영역과 주변 배경과의 명도차를 이용한 에지를 사용한다. 우선 에지 추출을 위한 전처리로 명도의 밝기 차를 더 부각 시키기 위해 어떤 특정부분에 집중되어 있는 화소값을 명암대비가 끌고루분포 되도록 화소 밝기를 스트레칭(Stretching) 시키고 잡영의 일부를 제거 하기 위해 3x3 마스크 미디언(Median)필터를 적용한다.

2. 텍스트 후보영역화

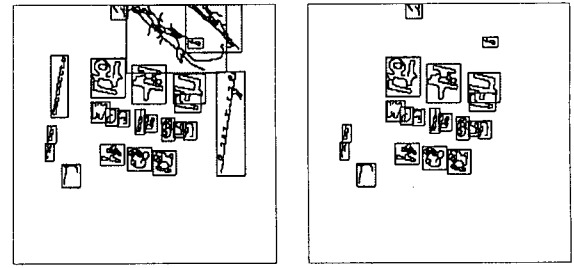
텍스트 후보영역을 구하기 위해 우선 canny 에지검지를 사용하여 에지 영상을 검출하고 각 연결화소를 세션화 한 후 작은 연결요소로 이루어진 잡영들을 제거한다. 다음으로 남아 있는 연결화소를 둘러싸여 있는 외각사각형의 가로 세로의 길이중 큰 값을 이용하여 분포도를 구하고, 어떤 길이의 값들 보다 확연히 다른 값을 가지는 영역을 제거함으로써 텍스트 후보영역을 결정한다.



(a) 스트레칭 영상



(b) Canny 에지 영상



(c) 세션화 후 잡영 제거 영상

(d) 텍스트 후보영역

<그림 2> 텍스트 후보영역 추출 과정

1) 에지검출

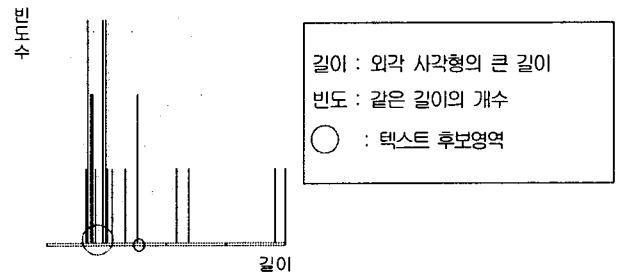
대부분의 에지 추출 검지기는 잡음에 대해 민감하므로, 작은 잡음이라도 그것을 에지로 간주하여 추출하는 경우가 많다. 텍스트를 추출하기 위해서는 정확한 경계선의 검출을 위해 끊어짐 현상이 방지되어야 하는데 많은 테스트 결과 prewitt 마스크, 가우시안의 표준편차 $\sigma=0.6$, 이중 경계값, high값=70, low값=40을 이용한 Canny 에지 검지기를 사용하였을 때, [그림 2]의 (b)와 같이 텍스트를 검출을 하기 위한 가장 좋은 영상을 얻는다.

2) 세션화를 통한 잡영제거

자연 이미지의 텍스트 주변에는 텍스트 후보영역을 추출하는데 불필요한 잡영들이 많이 발생하게 된다. 비록 이것이 텍스트 영역이라 하더라도 인식이 어렵기 때문에 제거해 주어야 한다. 이런 잡영들을 제거하기 위해서 각 영역들의 연결요소들을 세션화를 거친 후 전체 이미지의 세로, 가로길이 중 큰 값의 10분의 1을 임계값으로 정해 영역의 연결요소의 개수가 임계값 이하이면 그 화소들을 제거 시킨다. 추가적으로 각 외각사각형 내의 연결요소의 개수가 그 외각사각형 넓이에 비해 차지하는 비율이 작으면 제거한다. 이것은 선 같은 요소를 제거하는 효과를 얻을 수 있으며 이를 적용한 영상은 [그림 2]의 (c)에서 보여주고 있다.

3) 텍스트 후보영역 추출

텍스트의 특징은 각 연결요소를 둘러싼 외각사각형에서 세로와 가로의 길이 중 큰 값에 비슷한 길이로 분포되어 있다. 이것을 히스토그램으로 나타내면 [그림 3]과 같이 그래프로 나타낼 수 있다.

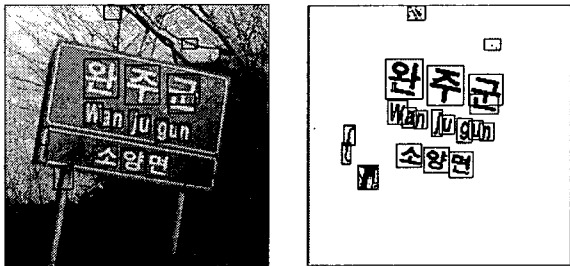


<그림 3> 외각사각형의 가로·세로 중 큰 값을 가지고 있는 길이 히스토그램

[그림 3]의 히스토그램을 보면 텍스트는 비슷한 길이에 밀집된 부분과 같은 길이의 빈도 수가 많은 부분에 있을 가능성을 가지고 있다. 이와 같이 길이와 빈도의 상관 관계를 따짐으로써 [그림 2]의 (d)와 같은 텍스트 후보영역을 결정한다.

3. 텍스트 영역 분류

텍스트 영역 분류 단계는 텍스트 후보영역에 대해서 텍스트-비텍스트 판별을 수행한다. 우선 2.1장에서 스트레칭된 영상에 각 후보영역의 위치 정보를 가지고 있는 외각 사각형을 매칭시켜 각 영역별로 이진화를 시킨다. 그리고 나서 각 영역별로 라벨링 방법을 취해 외각사각형 선에 붙어 있는 비텍스트 연결요소는 제거시킨다. 마지막으로 작은 잡영을 제거함으로써 완전한 텍스트를 얻는다.



(a) 스트레칭 영상 (b) 이진화 영상
(c) 텍스트 추출 영상 (d) 기울기가 보정된 결과 영상
〈그림 4〉 텍스트 영역 분류 과정

1) 텍스트 후보영역 이진화

2.1장 전처리에서 스트레칭 한 영상에서 후보영역내의 텍스트를 추출하기 위해서 이진화 작업을 수행한다. 우선 외각사각형의 크기를 한 픽셀씩 상·하·좌·우 확장 시키고 각 텍스트 후보영역내의 화소의 평균값을 임계값으로 정해 영역내의 화소가 임계값보다 작으면 흑화소로 크면 백화소로 변환시킨다. 다음 외각사각형 선의 흑화소의 존재를 조사하여 0.9%이상을 가지고 있으면 흑과 백화소를 반전시킨다.

2) 비텍스트 제거

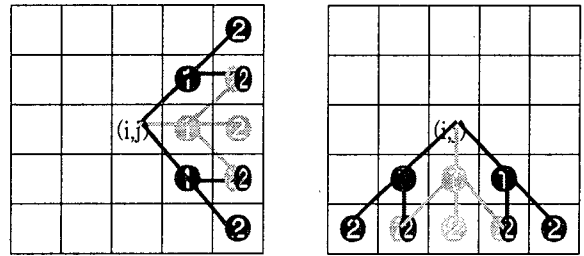
텍스트는 후보영역 외각사각형의 넓이를 확장했기 때문에 외각사각형 선상에 닿지 않고 안쪽에 놓이게 된다. 따라서 외각사각형 선에 흑화소가 존재 한다면 그 연결화소는 비텍스트라고 할 수 있으므로 그 연결요소를 제거하고 마지막으로 작은 잡영들을 제거 시켜줌으로써 [그림 4]의 (c)와 같은 텍스트만을 추출하게 된다.

4. 후처리

추출된 텍스트 영역을 정확히 인식하기 위해서는 기울어진 텍스트를 보정해 주어야 할 필요가 있다. 기울어진 텍스트의 각도를 알기 위해서는 먼저 2.2.2절에서 세선화 후 잡영처리를 한 영상을 가지고 가로·세로 선을 추출하여 각도를 추정한다. 추정된 기울기 각도는 양선영 보간법을 이용한 회전(rotation)을 적용하여 보정함으로써 최종 텍스트 영상 결과를 얻는다.

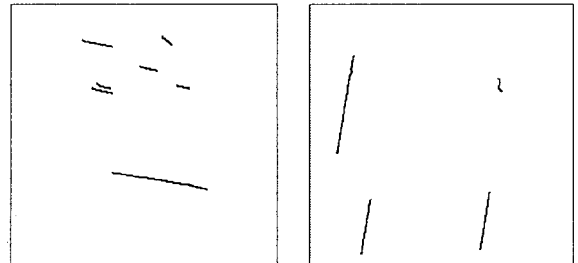
1) 기울기 추정

기울기 각도를 추출하기 위해서 2.2.2절의 세선화 된 영상을 가지고 [그림 5]의 원도우를 사용하여 [그림 6]과 같은 가로·세로 직선 요소만을 추출한다.



가로(-45° ~ +45°) 추출 원도우 가로(-45° ~ +45°) 추출 원도우

〈그림 5〉 직선 추출 원도우



가로(-45° ~ +45°) 직선 추출 세로(-45° ~ +45°) 직선 추출

〈그림 6〉 직선 추출 원도우

[그림 5]는 각각 가로·세로 -45° ~ +45°의 직선을 추출하는 5x5 원도우로 (i,j)에 있는 흑화소에 대해 색상이 같은 ①과 ②가 연속되어 있는 화소만 남기고 나머지 화소는 제거한다. 추가적으로 작은 길이의 선은 영상의 기울기 각도를 추출하는데 방해가 되므로 같이 제거한다. 이렇게 추출된 선 요소의 양쪽 끝 좌표를 다음 식(1)을 이용하여 각도를 추출한다.

$$\angle = \frac{\tan^{-1} \left(\frac{y_2 - y_1}{x_2 - x_1} \right)}{\pi \times 180} \dots\dots\dots \text{식(1)}$$

$(x_1, y_1), (x_2, y_2)$: 각 직선의 양쪽 끝 좌표 값

각 직선에 대해 구해진 각도는 빈도수를 조사하여 확률적으로 가장 많이 나온 각도를 대표값으로 기울어진 텍스트 영상

에 적용한다.

2) 기울기 보정

최종적인 기울기 보정은 양선형 보간기법을 이용한 회전 방법으로 현재위치(실수값) 근처의 이웃하는 네 점(정수값)의 밝기값을 이용하여 현재 위치와 각 지점과의 거리에 따라 반 비례하여 가중치를 달리하는 기법을 사용한다.

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x + c_x \\ y + c_y \end{pmatrix} + \begin{pmatrix} c_x \\ c_y \end{pmatrix} \dots\dots\dots\text{식(2)}$$

(x, y) : 회전 전의 원 영상 좌표값

(x', y') : 회전 후의 새로운 영상 좌표값

(C_x, C_y) : 원 영상 회전 중심의 좌표값

θ : 회전각

[그림 4]의 (d)는 회전각 -10° 을 식(2)에 적용시켜 최종적인 텍스트를 보정한 결과 영상이다.

III. 실험결과 및 분석

본 논문에서 제안한 방법을 이용하여 여러 교통 표지판을 샘플로 실험을 하였다. 실험대상은 디지털 카메라로 자연 영상에서 텍스트가 배경과의 구분이 명확한 것과 원근감이 적은 교통 표지판을 대상으로 직접 획득하였고 크기에 제한을 두지 않은 100개의 영상을 가지고 Visual C++.NET 환경에서 수행하였다. 실험결과 텍스트 추출률은 93.4%를 보였으며, 텍스트 추출 정확도는 89.2%를 보였다. 이는 텍스트의 문자들 사이가 너무 많이 붙어 있는 경우와 텍스트와 배경사이의 명암 대비의 차가 현저히 적을 때 텍스트를 추출하지 못하는 경향을 보였다.

기울기 추출에 대해서는 전체 이미지에 대해서 82.5%의 추출률을 보였으며, 이는 단순한 이미지에 대해서는 정확하게 추출할 수 있었으나, 복잡한 영상은 선 요소들이 너무나 다양하고, 산만하게 분포되어 있어 정확한 기울기 추출은 어려웠다.

IV. 결 론

본 논문은 텍스트가 포함되어 있는 교통 표지판에서 텍스트는 주변 배경과 명도의 대비차가 크다는 것을 착안하여 명도 이미지에서 에지를 사용하여 텍스트들이 가지고 있는 특

정들로 텍스트만을 추출하였다. 또한 기울기를 추출하기 위해 기존 연구들에선 Hough 변환을 사용하였으나 처리시간이 길고, 텍스트 구조상 기울기 각도를 추출하기 어렵기 때문에, 본 논문에서는 5x5 윈도우를 사용하여 가로, 세로 $-45^\circ \sim +45^\circ$ 의 직선을 추출하는 방법을 제안하였다.

현재까지의 실험 결과 단순한 이미지에 대해서는 비교적 정확한 텍스트를 추출하였으나 복잡한 이미지에 대해서는 좀 더 정확한 결과 추출을 위해 텍스트 검증 단계를 추가하여 성능을 개선할 필요가 있다.

또한 에지에서 각 찾아낸 영역에 대해 좀 더 종합적이고, 상호보완적인 관계를 고려하여 텍스트 후보영역을 더 정확하게 추출하기 위하여 보다 성능개선이 뛰어난 알고리즘 개발의 연구가 필요하다.

참고문헌

1. Jiangying Zhou, Daniel Lopresti, "Extracting Text from WWW Image", Proceedings of the 4th International Conference Document Analysis and Recognition (ICDAR '97), pp. 248-252, 1997
2. Anil K. Jain, Fellow, Bin Yu, "Document Representation and Its Application to Page Decomposition", IEEE Transactions On Pattern Analysis and Machine Intelligence, Vol. 20, No.3, pp. 294-305, 1997
3. X. Li, W.Gao, S. Y. Chi, K. A. Moon and H. J. Kim, "An Efficient Method for Page Segmentation," Proc, ICICS, Vol.2, pp. 957-961, 1997
4. S. K. Yip and Z. Chi, "Page Segmentation and Content Classification for Automatic Document Image Processing," Proc. Int. Symp. Intelligent Multimedia, Video and Speech Processing, pp. 279-282, 2001
5. J. Kong and Z. chi, "Image Classification Using Kolmogorov Complexity Measure with Extracted Blocks," IEICE Trans. Inf. & Syst., Vol. 1, E81-D, pp. 1239-1246, 1998
6. Y. Nakano, Y. Shima, H. Fujisawam J. Higashino and M. Fujinawa, "An Algorithm for the Skew Normalization of documnet Image" Proc. 10th Int. Conf. on Pattern Recognition, Atlantic City, USA, June pp.8-13, 1990.
7. J. L. Fisher, "Logical Structure Descriptions of Segmented Document Images," Proc. 1st Int. Conf. on Document Analysis and Recognition, Saint-Malo, France, pp. 302-310, Sep. 1991