

자연영상에서 저 대비 영역의 텍스트 영역 검출을 위한 연결요소 기반 방법에 관한 연구

박종천*, 권교현*, 전병민*

*충북대학교 전자정보대학 컴퓨터공학부
e-mail:simplejic@hanmail.net

Study on Text Detection of Low-contrast Region in Natural Images Using Connected-Component Based Method

Jong-Cheon Bak*, Gyo-Hyeon Gwon*,
Byeong-Min Jeon*

*School of Electrical&Computer Engineering,
Chungbuk National University

요 약

최근 모바일 기기 응용 분야에 관한 연구가 활발히 진행되고 있으며 모바일기기로 촬영된 영상에서 텍스트 정보를 추출하고자 하는 많은 연구도 진행되고 있다. 자연영상으로부터 텍스트 인식을 위한 전 단계로 텍스트 영역 검출은 필수적이므로 본 연구는 텍스트 영역의 지역적 연결요소 성분 특징을 고려하여 저 대비 영역에 존재하는 텍스트 영역을 검출하는 방법을 제안한다. 연결요소 성분은 전역 및 지역적 표준편차를 이용하여 연결성분을 추출하고 추출된 연결성분으로부터 후보 텍스트 영역을 검출하고, 후보 텍스트 영역에 대한 검증을 수행함으로써 최종적인 텍스트 영역을 검출한다. 제안한 방법은 낮은 대비를 갖는 자연영상에서 텍스트 영역 검출 성능을 개선하는 결과를 얻었다.

1. 서 론

모바일 기기의 발전과 더불어 모바일 응용 소프트웨어의 중요성이 부각되고 있다. 모바일 기기에 장착된 카메라, GPS, 중력감지 센서, Wi-Fi, 터치 센서 등을 기반으로 다양한 응용 프로그램 개발이 급속도로 개발되고 있고 이들 응용프로그램의 활용함으로써 보다 더 편리한 생활을 영위하고 있다. 영상처리 및 컴퓨터 비전 분야에서는 카메라를 이용한 다양한 응용 프로그램을 개발 및 연구 중에 있다.

시각장애인을 위한 시각보조 시스템이 개발 중에 있고[1], 외국인이 여행 중에 휴대용 장비로 외국어로 된 지리정보와 관광정보를 텍스트 영역 검출과 변환 기술로 외국어로 변환 가능하도록 하는 시스템도 개발 중이다[2].

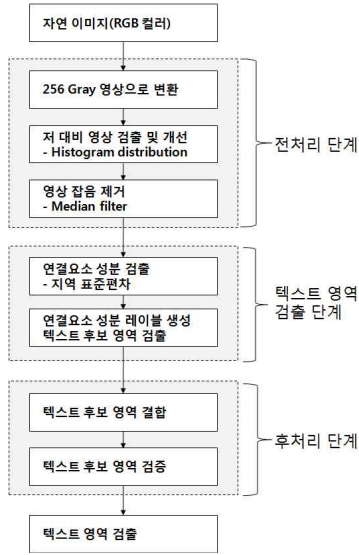
본 연구는 자연영상에서 저 대비(Low contrast) 영역에 존재하는 텍스트 영역 검출을 위해서 연결요소-기반 방법을 제안한다. 기존의 연구 방법은 낮은 대비

를 갖는 자연영상에서 텍스트 영역을 효과적으로 검출하지 못하는 문제점을 갖고 있다. 따라서 전처리 과정에서 저 대비 영상을 검출하고 영상의 대비를 개선을 함으로서 대비가 낮은 자연영상에서도 텍스트 영역을 검출 하는 것이다.

2. 저 대비 영상 검출 및 대비 개선

텍스트 영역 검출은 그림 1과 같은 과정을 수행한다. RGB 24컬러 영상을 명암도(Gray-scale) 영상으로 변환하고, 전처리 단계로 저 대비 영상을 검출하기 위해서 영상의 히스토그램을 식 1을 이용하여 구하고, Low_Level 와 High_Level 의 밝기 비율을 각각 식 2-3으로 구한다. 그 결과 Low_Level 와 High_Level 의 값이 모두 10% 미만인 경우 저 대비 영상으로 결정하고 영상 개선 과정을 수행하도록 한다.

$$Histogram \quad n = \sum_{i=0}^{255} m_i \quad (1)$$



[그림 1] 텍스트 영역 검출 과정

$$Low-Level = \sum_{i=0}^{50} Histogram(i) \quad (2)$$

$$High-Level = \sum_{i=200}^{255} Histogram(i) \quad (3)$$

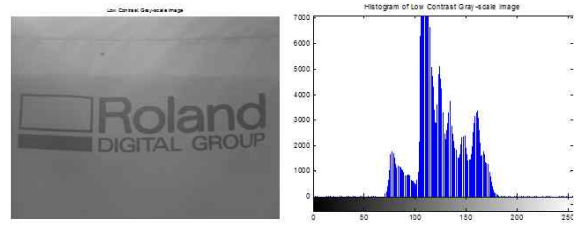
영상 대비 개선은 매트랩 함수(imadjust(I,[low_in, high_in],[low_out, high_out], gamma))를 이용하였다.

imadjust함수는 영상의 대비를 개선하는 함수로 영상의 입력 명암도를 그 에 해당하는 출력 명암도와 매핑함으로서 영상의 명암도를 개선하는 함수이다. 인자 값 low_in, high_in는 입력 영상의 최소값과 최대값을 기술하는 것으로 stretchlim(G)를 이용하여 구한 결과를 이용하였다. 인자 값 low_out, high_out값은 각각 0, 1로 설정하였다. 즉, 0~255 명암도 매핑하도록 하였다. 그림 2는 저 대비 영상과 히스토그램 분포이고, 그림 3은 대비 개선된 영상과 히스토그램 분포를 보여 준다. 그리고 영상의 잡음을 제거하기 위해 5×5 크기를 갖는 미디언 필터 처리를 한다.

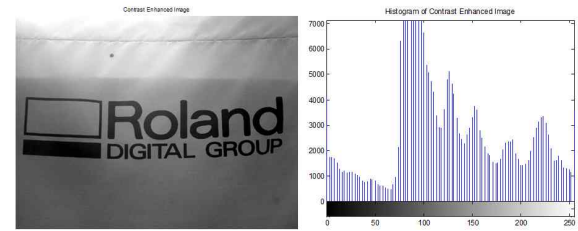
3. 텍스트 영역 검출

기존의 연결요소 성분 검출기반 연구 방법은[3-6] 주로 텍스트 영역 내부의 특징을 이용하였으나 본 연구에서는 텍스트 영역 경계부분의 표준편차 특징을 이용하였다. 영상의 전역 표준편차 값을 식 4로 구하고 지역적 표준편차 값은 매트랩 함수 stdfilt(I)를 이용하여 지역(3×3 이웃픽셀) 영역의 표준편차를 갖는 영상을 구한다. 실험 결과 영상 전체의 표준편차가 70보다 크면 지역적 표준편차의 임계값을 5로 설정하고

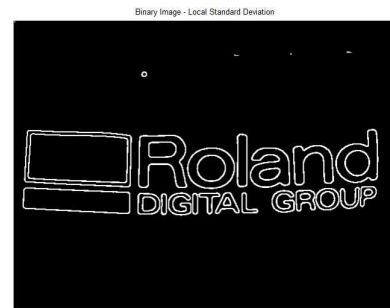
그렇지 않으면 원 영상의 편차평균 값의 2배로 설정하여 연결요소 성분을 검출한다.



[그림 2] 저 대비 영상과 히스토그램 분포도



[그림 3] 대비 개선된 영상과 히스토그램 분포도



[그림 4] 연결요소 추출 결과

지역적 표준편차 및 전역 표준편차를 이용하여 연결요소 성분을 검출한 결과는 그림 4와 같다.

텍스트와 배경의 밝기 차이로 인해서 표준편차가 텍스트 외곽선에서 상대적으로 텍스트 내부보다 큰 값을 갖게 된다.

$$S = \left(\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2 \quad \text{단,} \quad \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (4)$$

검출된 연결요소 성분을 레이블하고 외곽사각형 영역을 구한다. 레이블 영역의 크기가 30 픽셀 미만이면 텍스트를 구성할 수 없으므로 제거하였다. 레이블 영역의 구조적 특징은 표 2에 제시하였다. 표 2에 제시한 조건은 텍스트 영역 에지는 텍스트 에지 구조의 특징으로 최소 2번 이상 존재해야만 한다는 것을 반영한 것이다. 레이블 결과로 얻은 이진 영상으로 텍스트 영역이 갖는 구조적 및 통계적 특징을 얻기 위해서 매트

랩 함수 regionprops을 수행한다. regionprops는 이진 영상의 다양한 정보 픽셀 분포 및 모양정보를 제공한다.

[표 1] 텍스트 레이블 영역의 구조적 조건

```
nr는 영상의 높이, nc는 영상의 너비
width는 레이블 영역의 너비
height는 레이블 영역의 높이
Number_Nested_Label은 레이블 영역에 포함된 다른 레이블의 수
1) Max_Width_Height=max(width,height);
   Max_Width_Height < 10
2) width > height*3 or (width <= 4 and
   height >= width*5) or height >= width*10
3) Number_Nested_Label >= 10
4) width < (nc*0.9) and height < (nr*0.9)
```

위의 조건 1)은 너비와 높이 중에서 가장 긴 것을 기준으로 최소 10픽셀 미만의 작은 텍스트 영역을 제외하는 것이고 조건 2)는 텍스트의 문자가 갖는 가로와 세로의 비율이고, 조건 3)은 잡음 영역을 제외하는 것이고, 조건 4)는 너무 크거나 작은 텍스트 영역을 제외하는 것이다.

위와 같은 4가지 조건을 텍스트 영역의 후보를 검출하는 1차적인 조건이고, 표 2에 제시한 2차 조건으로 지역성과 구조적 유사도를 적용하여 만족하지 않는 것은 후보 텍스트 영역에서 제외한다. 그림 5는 후보 텍스트 영역 검출 결과를 보여준다.

[표 2] 텍스트 지역성과 구조적 유사도

```
NB_Width = 인접한 후보 텍스트 영역의 너비
NB_Height = 인접한 후보 텍스트 영역의 높이
Similarity_Width = NB_Width/width;
Similarity_Height = NB_Height/height;
NB_Similarity =
max(Similarity_Width, Similarity_Height);
If height >= width*5 and
Similarity_Height >= 0.7 and
Similarity_Height <= 1.5
else if Similarity_Height >= 0.7 and
Similarity_Height <= 1.5 and
Similarity_Width <= 2
```

4. 텍스트 후보 영역 검증

검출한 지역적 표준편차를 이용한 연결요소 성분으로 검출된 후보 텍스트 영역에 가중치를 부여함으로써 가중치가 임계값 이상인 것을 후보 텍스트 영역을 선정한다. 그림 6은 텍스트 영역 검증 결과로서 검출된 텍스트 영역을 보여준다.



[그림 5] 후보 텍스트 영역



[그림 6] 검출된 텍스트 영역

5. 실험 결과 및 결론

본 논문에서는 모바일 폰 카메라로부터 획득한 저 대비 자연 영상에 존재하는 텍스트 영역을 효율적으로 검출하기 위해 지역적 표준편차 연결요소 성분을 특징으로 텍스트 영역을 검출하는 방법을 제안하였다. 저 대비 영상을 검출하고 이에 따른 영상 대비 개선을 수행함으로써 저 대비 영상에 존재하는 텍스트 영역을 검출할 수 있었다. 텍스트 영역 검출 성능을 평가하기 위해 다양한 환경에서 모바일 폰 카메라로 촬영된 영상을 사용하였고, 평가방법은 아래와 같은 식 5~7을 이용하여 정확률(Precision)과 재현율(Recall)로 평가하였다[6].

$$\text{Precision} = \frac{\text{True}}{\text{True} + \text{Part} + \text{Error}} \quad (5)$$

$$\text{Recall} = \frac{\text{True}}{\text{sum}} \quad (6)$$

$$\text{Sum} = (\text{True} + \text{Part} + \text{Missing}) \quad (7)$$

True는 정확히 검출한 텍스트 영역의 수, Part는 텍스트의 일부분만을 검출한 수, Error는 텍스트가 아닌 영역을 텍스트로 검출한 수, Missing은 텍스트를 검출하지 못한 수를 의미한다. 표 3은 실험영상 중에서 저 대비 영상을 추출하여 실험한 결과이다.

그림 6의 실험결과를 보면 텍스트 영역이 아닌 배경

영역에서도 텍스트 영역으로 검출되는 것을 볼 수 있고 박스형태의 텍스트를 텍스트 영역으로 검출하는 것을 알 수 있다. 이와 같은 문제점은 향후 연구 되어야 할 부분으로 텍스트 영역에 대한 검증 방법을 연구하는 것이다. 또한 전처리 과정에서 지역적으로 저 대비를 갖는 영상에 존재하는 텍스트 영역을 검출하기 위한 지역적 영상 대비 개선을 위한 방법을 연구하는 것이다.

[표 3] 저 대비 영상 텍스트 영역 검출률

| 종류 | Precision | Recalling |
|------|-----------|-----------|
| 안내표지 | 83.9 | 82.8 |
| 간판 | 74.8 | 80.8 |
| 교통표지 | 74.2 | 76.7 |
| 번호판 | 71.0 | 78.6 |
| 교통안내 | 74.2 | 74.2 |
| 광고 | 75.0 | 78.3 |

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] 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, 1995.

[4] 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.

[5] W. Jiang , Q. Shao-Lin , Q. Zhuo, W. WenYuan , "Automatic text detection in com-

plex color image", Machine Learning and Cybernetics, 2002. Proceedings. 2002 International Conference on Volume 3, pp. 1167-1171, 2002.

[6] 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.

[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.