



장면 텍스트 영역 추출을 위한 적응적 에지 강화 기반의 기울기 검출 및 보정

백재경¹ · 장재혁² · 서영건^{1*}

¹경상대학교 컴퓨터학과

²경상대학교 산학협력단

The Slope Extraction and Compensation Based on Adaptive Edge Enhancement to Extract Scene Text Region

Jaegyung Back¹ · Jaehyuk Jang² · Yeong Geon Seo^{1*}

¹Department of Computer Science, Gyeongsang National University, Gyeongnam, 528-28, Korea

²Office of Academy and Industry Corporation, Gyeongsang National University, Gyeongnam, 528-28, Korea

[요 약]

실세계에서 텍스트가 포함된 장면은 텍스트를 추출하고 인식하여 많은 정보를 얻을 수 있으므로, 장면의 텍스트 영역을 추출하고 인식하는 기술들은 꾸준히 발전하고 있다. 장면에서 텍스트 영역을 추출하는 기술은 크게 텍스처를 기반으로 하는 방법과 연결요소방법, 그리고 이 둘을 적절히 혼합하는 방법들로 구분 할 수 있다. 텍스처를 기반으로 하는 방법은 영상의 색상, 명도 등의 정보를 이용하여 텍스트가 다른 요소와는 다른 값을 갖는다는 것을 기반으로 한다. 연결 요소 방법은 장면의 각 화소마다 인접해 있는 유사 화소를 연결 요소로 만들어 기하학적인 특성을 이용하여 판별한다. 본 논문에서는 텍스트 영역 추출의 정확도를 높이기 위해 영상의 기울기를 검출하고 보정한 후 에지를 적응적으로 변경하는 방법을 제안한다. 제안 방법은 영상의 기울기를 보정한 후 텍스트가 포함된 정확한 영역만 추출하기 때문에 MSER보다 15%, EEMSER보다 10% 더 정확하게 영역을 얻었다.

[Abstract]

In the modern real world, we can extract and recognize some texts to get a lot of information from the scene containing them, so the techniques for extracting and recognizing text areas from a scene are constantly evolving. They can be largely divided into texture-based method, connected component method, and mixture of both. Texture-based method finds and extracts text based on the fact that text and others have different values such as image color and brightness. Connected component method is determined by using the geometrical properties after making similar pixels adjacent to each pixel to the connection element. In this paper, we propose a method to adaptively change to improve the accuracy of text region extraction, detect and correct the slope of the image using edge and image segmentation. The method only extracts the exact area containing the text by correcting the slope of the image, so that the extracting rate is 15% more accurate than MSER and 10% more accurate than EEMSER.

색인어 : MSER, EEMSER, 장면 텍스트, 기울기 보정, 텍스트 영역 추출

Key word : MSER, EEMSER, Scene text, Slope compensation, Text region extraction

<http://dx.doi.org/10.9728/dcs.2017.18.4.777>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 17 July 2017; Revised 27 July 2017

Accepted 28 July 2017

*Corresponding Author; Yeong Geon Seo

Tel: +82-55-772-1392

E-mail: young@gnu.ac.kr

1. 서론

현대 사회에서 정보는 매우 중요한 요소이고, 우리 주위의 장면들은 정보를 담고 있다. 특히 문자가 포함되어있는 장면은 문자를 추출, 인식하는 것만으로도 많은 정보를 얻을 수 있다. 이러한 문자 정보를 이용해 시각장애인을 위한 시각보조 시스템이 개발 중이고 외국인들이 여행 중에 휴대용 장비를 통해서 외국어로 된 지리정보와 관광정보 등을 텍스트 검출과 변환 기술로 외국인들이 이해할 수 있는 문자로 변환 가능하도록 하는 시스템이 개발 중이다. 이런 이유로 장면의 문자 영역을 추출하고 인식하는 기술들은 꾸준히 발전하고 있으며 장면에서 문자를 추출하는 기술은 크게 텍스처를 기반으로 하는 방법과 연결요소 방법, 그리고 이 둘을 적절히 혼합하는 방법[1-3]들로 구분 할 수 있으며 이 외에도 다양한 텍스트 영역 검출 방법에 관한 연구가 진행되어 왔다[4-6].

텍스처를 기반으로 하는 방법은 영상의 색상, 명도 등의 정보를 이용하여 문자와 문자가 아닌 요소와는 다른 값을 갖는다는 전제하에 이러한 영상에 포함된 정보를 이용하여 문자를 찾아 추출한다. 이에 반해 연결 요소 방법은 장면의 각 화소마다 인접해 있는 유사 화소를 연결요소로 그룹을 만드는 것과 같은 기하학적인 특성을 이용하여 문자와 문자가 아닌 요소를 판별한다. 이런 연결요소 방법의 역사에 큰 획을 그은 이론이 있다면 SWT(Stroke Width Transform)[4]와 MSER(Maximally Stable Extremal Regions)[5]일 것이다. SWT는 문자가 있는 장면에서 문자의 폭을 찾는 방법으로, 이는 문자 인식에서 같은 영역에 있는 문자의 폭은 일정하다는 점에서 착안되었다. MSER은 영상에서 명도 정보를 이용한 이진화를 통해 영역을 찾아내어 모든 이진화 값에 대해 연결영역을 찾아내는데 각 이진화 값의 변화에 따라 연결영역의 변화가 적은 곳을 찾는다. 이 결과를 바탕으로 변화가 적은 곳을 하나의 영역으로 보는 방법이다. 이러한 SWT와 MSER에 대한 많은 연구도 진행되어 왔다[6]. 문자를 추출할 때에는 문자의 경계선 또한 중요하다. 기본적인 경계선 추출은 필터링을 통하여 경계가 되는 부분에서 색이 다른 경우에 그 부분을 경계로 추출하는 것이다.

대표적인 경계선 추출의 방법으로는 캐니 연산자[5]가 있다. 캐니 연산자는 잡음을 없애기 위한 블러링, 기본적인 경계선 검출, 비최대값 제거, 두 개의 임계값으로 크기를 구분, 검출된 경계선들의 연결로 다섯 단계를 거쳐 경계선을 추출한다. 또한 연결요소 방법과 경계선 추출을 혼합한 방법인 EEMSER(Edged-Enhanced MSER)이 있다[7]. 하지만 EEMSER에서 경계선을 추출할 때 사용하는 캐니 연산자는 임계값에 따라 결과가 상이하게 나오는 문제점을 가지고 있다. 또한 장면에 포함된 문자는 때때로 기울기를 가지고 있다. 이런 기울어진 문자를 포함한 장면에서 문자 영역 추출의 정확도를 높이기 위해 기울기의 추출 및 보정이 필요하다.

기울기를 추출 및 보정하는 방법으로는 경계선과 최소자승

법을 이용한 기울기 보정 방법[8], 객체의 위치와 최소자승법을 이용한 기울기 보정 방법[9], 투영 히스토그램을 이용한 방법[10]등이 있다. 이 방법들은 문서 영상 또는 문자열이 다수 포함된 장면에서는 빠르고 정확도가 높으나 문자와 배경이 혼합된 장면에서는 상대적으로 낮은 정확도를 보이고 있다. 특히 최소자승법을 이용한 방법은 정상적인 데이터 분포에 속하지 않는 데이터가 존재한다면 잘못된 근사결과를 낼 수 있는 최소자승법의 한계로 인해 문자와 배경이 혼합된 장면에서는 적용하기 힘든 방법이다.

본 논문에서는 기울기를 가진 문자와 배경이 혼합된 장면에서 문자의 기울기를 검출하기 위해 모폴로지를 이용하여 전처리한다. 전처리 방법으로는 회색조 영상을 두고 닫기 후 열기 연산을 한다. 모폴로지 연산 후 결과영상에 캐니 연산자를 사용하여 경계선을 검출한다. 또한 영상의 최적 기울기를 구하기 위해 검출된 경계선의 영상을 3X3으로 분할한다. 그리고 경계선이 없는 영역의 경우에는 연산 속도와 기울기의 정확도를 위해 배제한다. 다음으로는 분할된 각 영역마다 최대 경계선 집중도 갖는 기울기를 구한다. 최대 경계선 집중도를 갖는 기울기를 구하는 방법으로는 영상의 분할된 영역을 -45°부터 45°까지 1°간격으로 투영하여 경계선 투영 히스토그램을 생성하고 이를 이용하여 분할된 각 영역의 최대 경계선 집중도를 갖는 기울기를 구한다. 이 때 분할된 각 영역의 최대 경계선 집중도를 갖는 기울기를 최빈값과 중앙값을 이용하여 영상의 최적 기울기를 구하고 보정한 후 EEMSER의 캐니 연산자 임계값을 명도 히스토그램을 이용해 적응적으로 적용하는 방법을 제안한다. 제안된 방법을 이용하여 기울기를 가진 문자가 포함된 장면에서 문자 추출 결과 MSER과 EEMSER보다 높은 추출 정확도를 보이는 것을 알 수 있다.

II. 관련 연구

2-1 기울기 검출

[10]은 문서 영상에서 고속 기울기 검출을 위해 3단 coarse-to-fine 탐색 알고리즘을 사용한 기울기 검출을 제안하였다. 이 방법은 문자 영역에서 캐니 연산자를 사용하여 경계선 영상을 구한다. 이 후 경계선 영상을 다양한 방향으로 투영시켜 경계선의 집중도를 계산하고 최대 경계선 집중도가 생성되는 최적의 기울기를 구한다. 최적 기울기를 검출하는 3단 coarse-to-fine 탐색 알고리즘은 다음과 같다.

단계1) -45° ~ +45°에서 3° 간격으로 최대 경계선 집중도를 갖는 θ_{max1} 을 검출.

단계2) θ_{max1} 을 중심으로 -2° ~ +2°에서 1° 간격으로 최대 경계선 집중도를 갖는 θ_{max2} 를 검출.

단계3) θ_{max2} 을 중심으로 $-0.5^\circ \sim +0.5^\circ$ 에서 0.1° 간격으로 최대 경계선 집중도를 갖는 θ_{max3} 을 검출.

그림 1은 경계선 투영의 예로 기울기에 따른 경계선 투영과 투영 히스토그램을 보여주고 있다. 그림 1은 문서의 기울기와 투영 방향이 다른 경우와 문서의 기울기와 투영 방향이 일치한 경우에 따른 경계선이 집중된 결과를 보여주고 있다. 그림 2는 문서의 기울기에 따른 경계선의 집중도를 그래프로 나타낸 것이다. coarse-to-fine 탐색 알고리즘을 사용한 기울기 검출은 기울기가 일정한 문서 영상에서 아주 빠르고 강력하다. 하지만 문자와 배경이 혼합된 장면 영상에서는 상대적으로 최적의 기울기가 추출되지 않을 수 있다.

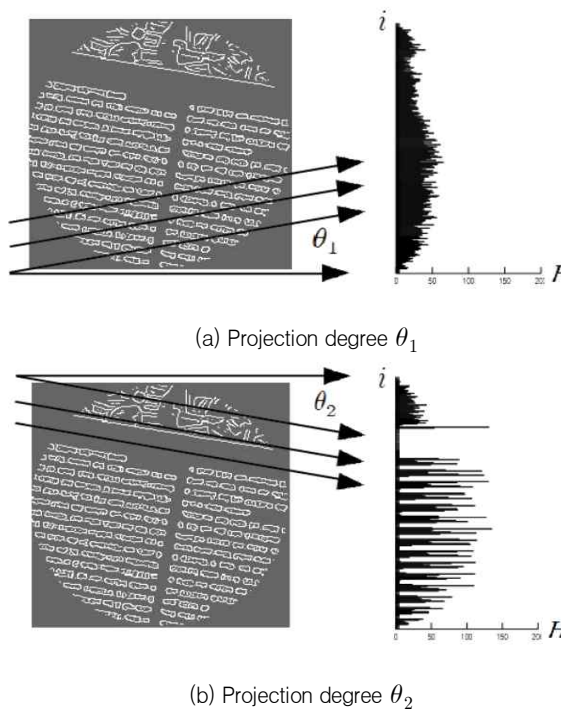


그림 1. 에지 주사의 예
Fig. 1. Examples of edge projection

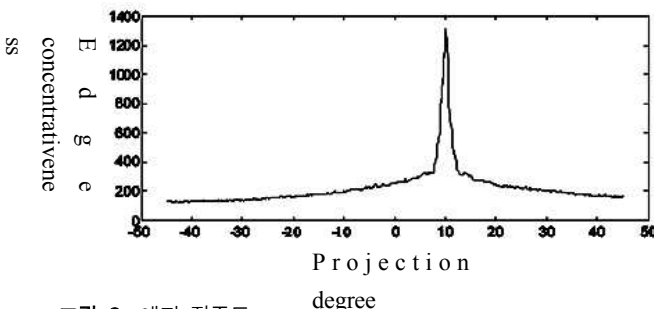


그림 2. 에지 집중도
Fig. 2. Edge Concentrativeness

[11]은 흑백 문서 영상에서 투영 히스토그램을 사용한 기울기 검출 방법을 제안하였다. 이 방법은 흑백 문서 영상을 각기 다른 기울기로 투영하여 흑백 또는 백흑의 전이점의 개수를 구한 후 이를 식 (1)에 적용하여 최대 전이점 집중도를 구하여 문서 영상이 기울기를 검출한다.

$$V(\theta) = \frac{1}{n} \sum_{i=1}^{i=n} (N_i - M)^2, M = \frac{\sum_{i=1}^{i=n} N_i}{n} \quad (1)$$

최대 전이점 집중도를 구하기 위하여 각각의 기울기마다 전이점 집중도를 구해야 한다. 전이점 집중도는 전이점 투영 히스토그램의 분산(V)이고 n은 투영한 행의 합이다.

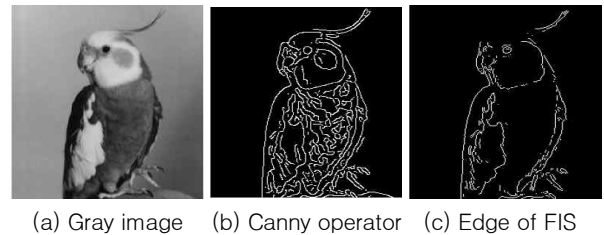


그림 3. 흑백영상으로부터 FIS의 캐니연산자와 에지
Fig. 3. Canny Operator and edge of FIS from gray image

2-2 경계선 추출

[12]는 FIS를 이용한 경계선 추출을 제안하였다. 영상의 명도 히스토그램을 이용하여 각 화소들을 명도에 따른 그룹으로 나눈다. 각 그룹마다 경계선을 추출할 때 사용할 임계값을 추론하는데 3가지 통계치를 이용한다. 이후에 18개의 규칙이 정의된 FIS에 입력되어 멤버십 함수에 의해 정도가 표현되고 이 정도를 이용하여 경계선을 추출한다.

그림 3은 새의 영상을 두고 캐니 연산자를 이용한 영상과 FIS를 이용해 경계를 추출한 영상을 비교하기 위해 사용한 것이다. 캐니 연산자는 강력한 성능을 자랑하기 때문에 윤곽선뿐만 아니라 내부에서 경계로 삼을 만한 것까지 경계로 추출하는 반면 FIS에 의한 경계 추출은 윤곽선을 중심으로 경계를 추출한 것을 알 수 있다. 하지만 장면 텍스트 추출에서 상대적으로 크기가 작은 텍스트들의 경계는 규칙에 의해 표현되지 않을 수가 있다.

[4]는 SWT를 이용하여 글자 영역을 추출하는 방법을 제안했다. 이미지에서 글자의 테두리를 추출하는데 이는 캐니 연산자를 이용한다. 테두리의 한 화소에서 기울기가 큰 방향으로 광선을 쏜다. 이는 영상의 x, y방향에 대한 각각 편미분값으로 알 수 있다. 이 값을 통하여 반대편 글자 화소에서도 광선을 쬐었을 때 서로 마주보게 될 시에 이는 글자를 가로지르는 획으로 간주하고 선을 거리만큼 채우면 된다. 이러한 방식으로 모든 글자 영역을 채우게 되면 한 글자는 비슷한 폭으로 채워지게 된다.

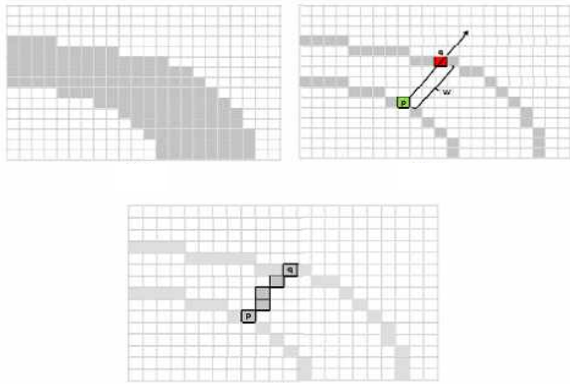


그림 4. SWT의 원리
Fig. 4. Principles of SWT

그림 4는 SWT의 원리를 이해하는 것을 돕기 위해 제시한 그림이다. SW 값들이 있는 영상에서 CCs(Connected Components)를 사용하여 글자 객체를 추출한다. 글자 필터 기준은 다음과 같다. 각각의 객체는 SW값의 분산이 낮은 경향이 있다. 중형비를 0.1과 10 사이의 값으로 제한하여 같은 영역의 글자는 자간이 글자의 폭의 3배를 넘을 수 없고 글자는 같은 색을 가지고 있다고 가정하여 글자의 영역을 판단한다. 하지만 폰트에 따라 글자가 항상 같은 폭을 갖지는 않기 때문에 글자가 다른 폭을 갖는 형태라면 글자로 판단하지 못할 수 있다. [7]은 MSER과 캐니 연산자를 혼합한 경계선을 강화하여 결과영상을 추출하는 방법인 EEMSER을 제안했다. 기존의 MSER 결과영상에서 캐니 연산자의 결과영상을 교차 연산하여 중간 영상을 구할 수 있고 기존 MSER 결과를 반전시켜 중간 영상과 다시 교차 연산한다.

그림 5는 MSER을 통해 검출된 영역인 블러링 영역이고 빨간색 선은 캐니 연산자를 통해 검출된 경계 영역으로 번짐 효과를 제거하여 영역의 경계를 강화시키는 것을 알 수 있다. 경계를 늘린 후에는 연결요소를 이용하여 기하학 처리와 SW 정보를 가지고 비문자 영역을 제거한다. 캐니 연산자의 효과는 이미 입증된 바 있으나 사용되는 임계값의 설정에 따라 다른 결과를 얻을 수 있다.



그림 5. EEMSER의 기본 원리
Fig. 5. Basic principle of SWT

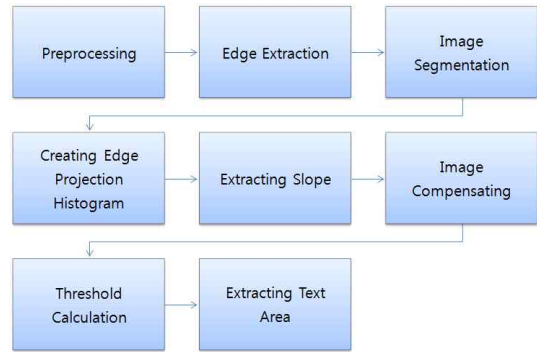


그림 6. 제안 방법의 흐름도
Fig. 6. Flowchart of the proposed method

III. 기울기 계산 및 보정과 임계값

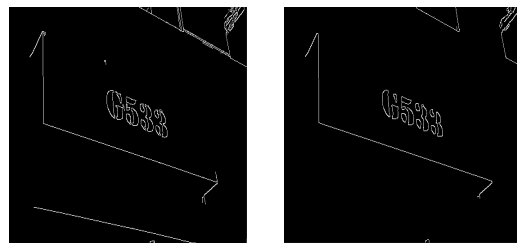
본 장에서는 장면 영상에서 문자열 영역 추출의 정확도를 높이기 위해 장면 영상의 기울기를 추출하고 보정하여 EEMSER에서 경계선 임계값을 계산하기 위한 캐니 연산자 과정을 적응적 방법을 적용한 방법을 제안한다.

3-1 모폴로지 연산을 이용한 전처리

그림 6은 제안 방법의 전체 흐름을 나타내고, 전처리 과정에서는 영상에서 경계선을 효과적으로 추출하기 위하여 모폴로지 연산을 사용한다. 그림 7은 모폴로지 연산 전 회색조 영상을 두고 캐니 연산자를 이용하여 경계선을 추출한 영상과 닫기 후 열기 모폴로지 연산을 한 후 캐니 연산자를 이용하여 경계선을 추출한 영상을 비교하였다.



(a) Gray image (b) Morphology operator



(c) Canny operator (d) Canny operator after based on gray image morphology operator

그림 7. 캐니 연산자와 모폴로지 연산 후의 캐니 연산자
Fig. 7. Canny operator and the one after morphology operator

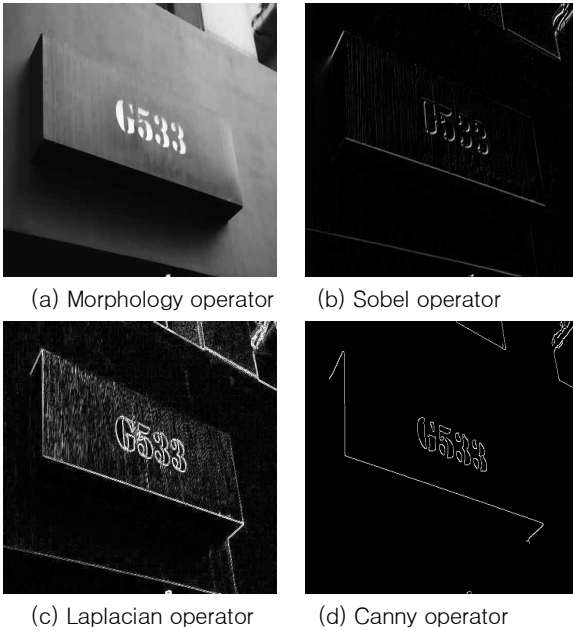


그림 8. 모폴로지 연산 후의 연산자의 비교
Fig. 8. Comparison of the operators after morphology operation

3-2 경계선 추출 및 영상 분할

제안 방법에서는 영상에서 경계선을 추출하기 위해 캐니 연산자를 사용한다. 캐니 연산자는 소벨 연산자와 라플라시안 연산자에 비해 잡음에 민감하지 않으며 강한 경계선을 추출할 수 있다. 그림 9는 모폴로지 연산 후 영상을 두고 소벨 연산자, 라플라시안 연산자, 캐니 연산자를 이용하여 경계선을 추출한 영상을 비교하였다. 소벨 연산자 영상은 잡음에는 민감하지 않으나 경계선이 뚜렷하지 않고 라플라시안 연산자 영상은 경계선은 강하나 잡음에 민감하다. 캐니 연산자 영상은 잡음에도 민감하지 않으며 경계선도 뚜렷하여 본 연구에서는 캐니 연산자를 사용하여 경계선을 추출한다. 경계선 추출 영상을 이용하여 경계선이 포함되지 않는 영역을 배제하기 위해 그림 9와 같이 경계선 추출 영상을 분할한다.

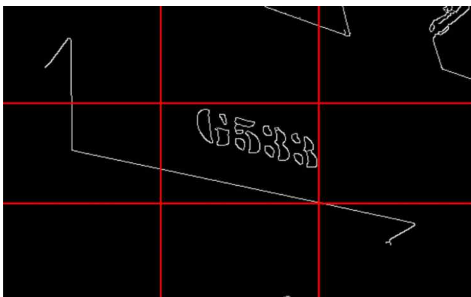


그림 9. 3x3 영상 분할
Fig. 9. 3x3 image segmentation

3-3 경계선 투영 히스토그램을 이용한 최적 기울기 검출

영상에서 문자열은 일반적으로 가로 방향으로 정렬되어 있다. 이런 문자열의 특성은 가로 방향으로 경계선을 많이 발생시킨다. 이를 이용하여 경계선 누적 히스토그램을 생성하고 최적 기울기를 검출한다. 최적 기울기를 검출하기 위해 투영 히스토그램을 이용하여 3단계 검출 알고리즘을 사용하는데, 다음은 세 번째 단계인 단계 3만 보이고 있다.

-단계 3 :

```

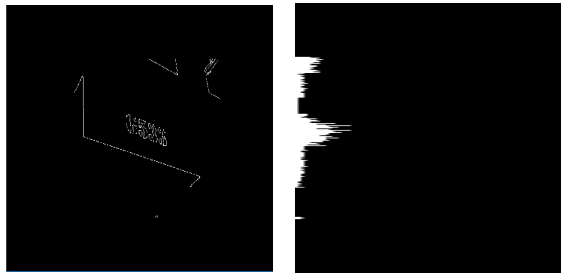
입력: 중복된 최빈값을 갖는 기울기의 집합(mode_array)
출력: 중앙값과 가장 근접한 기울기(max3)
For i = 1 to mode_array.length
    If mean < mode_array[i]
        temp <- mode_array[i] - mean
    else
        temp <- mean - mode_array[i]
    End
    If max3 > temp then
        max3 <- temp
    End
End
    
```

단계 1은 분할된 영역의 영상을 $-45^\circ \sim +45^\circ$ 의 범위에서 1° 간격으로 투영하여 경계선 투영 히스토그램을 생성한다. 여기서 생성된 경계선 투영 히스토그램을 사용하여 아래의 (2)를 계산 후 최대 경계선 집중도를 갖는 $\max1^\circ$ 을 검출한다.

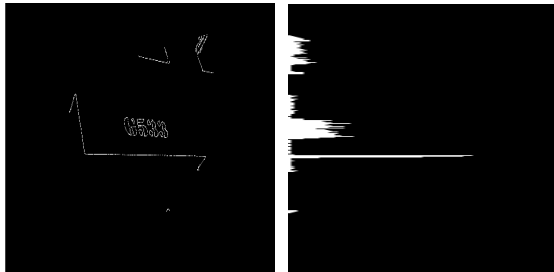
$$V(\theta) = \frac{1}{n} \sum_{i=0}^n (A_i - M)^2, M = \frac{1}{n} \sum_{i=0}^n A_i \quad (2)$$

-45°	22°	20°
13°	12°	-11°
0	-12°	12°

그림 10. 각 블록의 최적 기울기
Fig. 10. Best slope of each area



(a) Before slope compensation



(b) After slope compensation

그림 11. 기울기에 따른 에지 투영 히스토그램

Fig. 11. Edge projection histogram with the slope

경계선 집중도는 경계선 투영 히스토그램의 분산(V)으로 n은 경계선 투영 히스토그램의 값이 0이 아닌 행의 개수이다. 투영 기울기에 따른 경계선 집중도에서 가장 큰 값인 $max1^\circ$ 을 검출한다. 단계 2는 분할된 각 영상의 최대 경계선 집중도 집합인 max_array 에서 최빈값을 갖는 $max2^\circ$ 를 검출한다. 이때, 그림 10의 분할된 영상에서 정확도를 높이기 위해 경계선이 없는 영역7을 제외한다. 단계 3은 중복된 최빈값이 있을 경우 중앙값에 가장 가까운 $max3^\circ$ 을 검출하여 정확성을 높인다. 여기서 검출된 최적 기울기를 이용하여 영상 전체에 적용, 기울기를 보정한다. 그림 11은 기울기 보정 전 영상의 경계선 누적 히스토그램과 기울기 보정 후 영상의 경계선 누적 히스토그램의 비교이다.



그림 12. 명도 히스토그램

Fig. 12. Histogram for brightness

3-4 명도 히스토그램을 이용한 임계값 계산

[7]에서 제안한 방법인 EEMSER에서 캐니 연산자를 사용하여 결과 영상을 구하는 방법을 수정했다. 캐니 연산자 과정의 네 번째 단계인 두 개의 임계 값으로 크기를 구분하는 부

분에서 일반적으로 두 개의 임계값을 사용자 임의로 입력하는 경우가 많다. 하지만 이 두 개의 임계값을 계산할 수 있는 방법들이 있는데 이 방법들 중에서 히스토그램을 이용하는 방법이 존재한다. 영상의 히스토그램은 특정 요소에 관한 분포를 보이는데 본 연구에서는 명도에 대한 히스토그램을 사용한다. 이는 최솟값, 최댓값, 평균값, 그리고 중앙값 등을 쉽게 구할 수 있기 때문이다. 임계값 계산을 위해 사용될 것은 중앙값 혹은 평균값인데 본 연구는 중앙값을 이용하였다. 히스토그램의 중앙값을 계산하는 과정은 다음과 같다.

입력: 영상의 명도 히스토그램(hist)

출력: 히스토그램의 중앙값(medianValue)

For i=0 to 255

 If $max < hist[i]$ Then

$max < -hist[i]$

$maxIndex < -i$

 End

 If $min > hist[i]$ Then

$min < -hist[i]$

$minIndex < -i$

 End

End

If $maxIndex > minIndex$ Then

$medianValue < -(maxIndex - minIndex) / 2$

Else

$medianValue < -(minIndex - maxIndex) / 2$

End

알고리즘에서 for문의 제한횟수 255는 일반적으로 명도의 값이 0~255로 표현되기 때문에 히스토그램의 x축의 범위까지만 탐색하도록 하는 의미이다. 탐색을 하면서 최솟값의 인덱스와 최댓값의 인덱스를 구한다. 출력되는 중앙값은 최댓값 인덱스와 최솟값 인덱스의 차를 2로 나눈 것이다. 이와 비슷한 방법으로 평균값 또한 구할 수 있다. 중앙값을 구한 다음에는 다음의 식을 통하여 두 개의 임계값을 결정한다. 식 (3)은 [3]에서 소개한 것으로 각각에 곱해지는 상수는 경험적 법칙에 의해 결정됐다.

$$\begin{aligned} low\ Threshold &= median \times 0.66 \\ high\ Threshold &= median \times 1.33 \end{aligned} \quad (3)$$

3-5 텍스트 영역 추출

그림 13은 식 (3)에서 계산된 두 개의 임계값을 캐니 연산자에 적용시켜 나온 영상과 기존의 MSER을 거친 영상이다.

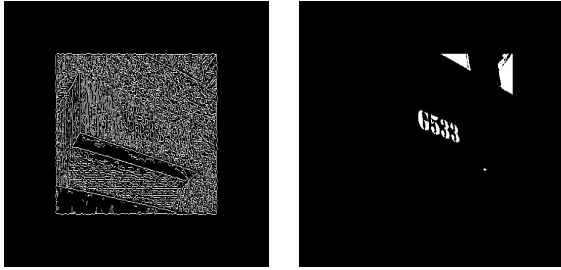


그림 13. 제안 방법과 기존의 MSER의 결과
 Fig. 13. Results of the proposed method and the existing MSER

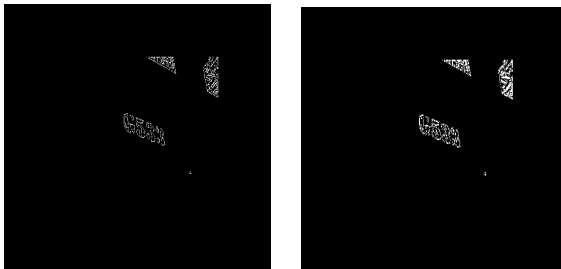


그림 14. 캐니 영상과 MSER 영상의 교차 영상(좌)와 확장된 경계를 갖는 영상(우)
 Fig. 14. Crossed image of canny image and MSER image(left), and image with enlarged boundaries(right)

그림 13의 제안 방법에 의한 결과와 기존 MSER의 결과를 교차시켜 그림 14의 캐니 영상과 MSER영상의 교차 영상을 얻고 이 영상의 경계를 늘려 경계를 늘린 영상을 얻는다. 그림 13의 MSER을 적용한 영상과 그림 14의 경계를 늘린 영상을 반전시켜 다시 교차시키면 그림 15의 교차영상을 얻을 수 있고 이 영상에 기하학적인 처리를 하여 영상을 얻는다. 그림 14의 경계를 늘린 영상을 반전시키게 되면 각 픽셀들이 갖는 값들이 2의 보수형태로 바뀌고 이를 다시 기존 MSER을 적용한 결과에 각 화소들과 AND 연산하는 것이다. 그 다음에 그림 13의 경계 강화 영상을 반전시켜 MSER 영상과 교차시킨 영상을 기반으로 연결요소들을 만들어 이 연결요소들의 기하학적 특성과 SW 정보를 이용하여 비텍스트 영역이라고 판단되는 연결요소들을 제거한다.

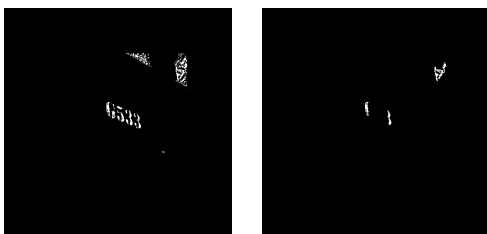


그림 15. MSER 영상과 반전 영상의 교차 영상(좌)와 기하학적 처리를 추가한 영상(우)
 Fig. 15. Crossed image of reversed image and MSER image(left), and image adding geometric processing

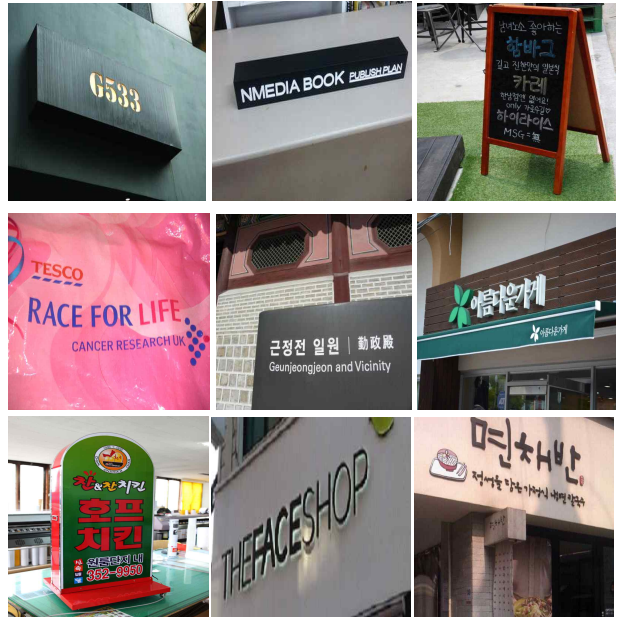


그림 16. 실험 영상
 Fig. 16. Experimental images

IV. 실험 및 평가

4-1 실험방법

제안 방법의 유효성을 보이기 위해 그림 16의 영상을 이용하여 실험하였으며 추출된 문자 영역에 대해서 정확도를 나타내기 위하여 다음과 같은 식을 사용했다. 식 (4)에서 p는 정확도를 의미하고, 분자는 제안 방법을 이용하여 추출된 문자 영역이며, 분모는 실제 문자 영역의 영역을 의미한다.

$$p = \frac{\text{Extracted Region}}{\text{Real Text Region}} \quad (4)$$

4-2 실험 결과 및 평가

제안된 방법의 성능을 평가하기 위한 척도는 기존의 방법과 제안된 방법이 영상에서 텍스트 영역을 얼마나 정확하게 찾아내는가를 정확도로 표현한다. 그림 17은 그림 16의 실험 영상을 이용하여 기울기 보정 후 실험한 문자 영역을 표시한 결과이며, 이를 바탕으로 기존의 EEMSER과 MSER 방법과 비교하여 표 1에 보이고 있다. 제안 방법이 기존의 방법에 비하여 성능이 향상되었으며, 이에 대한 평균을 도표로 나타낸 것이 그림 18이다. 그림 18을 통해 제안 방법이 기존의 EEMSER보다 10%, MSER 보다는 15% 성능이 향상된 것을 알 수 있다.



그림 17. 실험 영상의 기울기 보정 및 텍스트 영역
 Fig. 17. Slope compensation and text area of experimental images

실험 영상 5는 실제 글자가 각인된 영역만 생각하면 오히려 보정한 것이 더 기울어져 보인다. 이는 실험 영상 내에 존재하는 직선 영역들이 많아서 이를 기준으로 보정이 되었기 때문에 표 1의 결과에서도 보이듯이 더 안 좋은 결과를 보이기도 하였다. 표 1에서 실험 영상 3과 8은 안 좋은 결과가 나왔는데, 이는 기울어진 것을 보정하는 측면에서는 좋은 결과를 보였으나 문자 영역은 다소 크게 추출되는 경우를 보이고 있다.

표 1. 각 실험 영상에 대한 제안 방법과 기존 방법의 비교
 Table 1. Comparison for each experimental image

image	precision		
	Proposed method	EEMSER[7]	MSER
1	0.12	0.08	0.07
2	0.20	0.13	0.15
3	0.45	0.50	0.48
4	0.75	0.38	0.50
5	0.30	0.42	0.38
6	0.87	0.61	0.45
7	0.71	0.70	0.70
8	0.20	0.45	0.38
9	0.98	0.89	0.85

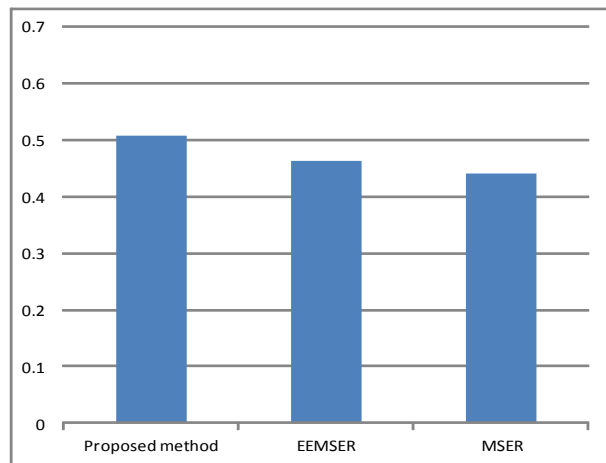


그림 18. 각 실험 모델에서 텍스트 영역을 추출하는 평균 정확도
 Fig. 18. Average precision rare on extracting text area for each model

V. 결 론

실세계 장면 영상에서 문자를 추출할 때, 문자열이 기울어져 있으면 문자가 포함된 영역의 범위가 넓어지는 문제점을 개선하기 위한 방법으로 영상의 기울기를 추출 및 보정하는 방법을 본 연구에서 제안했다. 이 방법은 영상의 경계선을 구하고 3X3으로 영상을 분할한 후 기울기의 정확도를 위해 경계선이 포함되지 않은 영역은 연산에서 배제하고 각 영역의 경계선 집중도를 구해 영상 전체의 최적의 기울기를 구하였다. 기울기 보정은 영상에 포함되어 있는 문자 영역을 수평으로 정렬하기 위함이다. 실험을 통해 제안한 문자 영역 추출 방법이 기존의 MSER보다는 15%, EEMSER보다 10% 더 정확하게 문자 영역을 추출할 수 있었다.

향후 연구 과제로는 본 논문에서 MSER의 특징으로 인해 영상의 크기에 따라 결과가 상이하게 나올 수 있다. 이를 보완하기 위해 영상의 크기에 따른 적응적 임계값을 적용할 수 있는 방법에 대한 연구가 필요하고 본 논문의 연구 및 실험을 할 때 정지된 실세계의 영상으로만 하여 현실에서 사용 시 효용성이 낮을 것으로 보인다. 이를 보완하고 현실에서 효용성을 높이기 위해 모바일, 또는 카메라를 이용한 실시간 영상에서 제안한 방법의 적용에 관한 연구가 필요하다.

참고문헌

[1] L. Yao, J. Wenjing, S. Chunhua, and A. Hengel, "Character3ness: an indicator of text in the wild", *IEEE Trans. Image Process.* Vol. 23, No. 4, pp. 1666-1677, 2014.

[2] L. Kang, Y. Li, and D. Doermann, "Orientation robust text line detection in natural images". *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.* pp. 4034-4041, 2014.

[3] M. C. Sung, B. Jun, H. Cho, and D. Kim, "Scene text detection with robust character candidate extraction method", *Proc. 13th ICDAR*. pp. 426-430, 2015.

[4] B. Epshtein, E. Ofek, and Y. Wexler, "Detecting text in natural scenes with stroke width transform", *CVPR*, pp. 2963 -2970 2010.

[5] M. W. Kim, and I. S. Oh, "Stroke Width Transform and Feature Extraction for Maximally Stable External Regions(MSER)", *Journal of KIISE : Computing Practices and Letters*, Vol. 20, No. 1, pp. 21-25 2014.

[6] Y. Park, S. Park, and Y. G. Seo. "An Extracting Text Area Using Adaptive Edge Enhanced MSER in Real World Image", *Journal of Digital Contents Society*, Vol. 17, No. 4, pp. 219-226, 2016.

[7] Chen H., Tsai S., Schroth G., Che, D., Grzeszczuk R. and Girod B., "Robust text detection in natural images with edge-enhanced maximally stable extremal regions", *In Image Processing (ICIP), 18th IEEE International Conference on*, pp. 2609-2612, 2011.

[8] P. Malathi. "Skew Detection based on Bounding Edge Approximation", *IOSR Journal of Computer Engineering*, Vol. 16, pp. 136-139, 2014.

[9] B. Drako, "The evaluation of the initial skew rate for printed text", *Journal of Electrical Engineering*, Vol. 62, No. 3, pp. 134-140, 2011.

[10] Ju J. H., and Oh J. S., "Skew Correction of Document Images using Edge", *Journal of the Korea Institute of Information and Communication Engineering*, Vol. 16, No. 7, pp. 1487-1494, 2012.

[11] Y. Ishitani, "Document Skew Detection Based on Local Region Complexity", *ICDAR*, pp. 49-52, 1993.

[12] U. N. Gyan, and K. Lakhwani, "Edge Detection Using Fuzzy Approach Involving Automatic Threshold Generation", *IJSTR*, Vol. 2, No. 7, pp. 128-131, 2013.



백재경(Jaegyung Back)
 2010년 : 호남대학교 게임에니메이션
 학과 학사
 2017년 : 경상대학교 교육대학원
 컴퓨터교육과 석사



장재혁(Jaehyuk Jang)
 2006년 : 경상대학교 컴퓨터과학과
 공학박사
 2012년~현재 : 경상대학교
 산학협력중점 교수

관심분야 : 유비쿼터스 컴퓨팅, 소프트웨어 공학, 영상처리, 네트워크보안



서영건(Yeong Geon Seo)
 1987년 : 경상대학교 전산과 학사
 1997년 : 숭실대학교 전산과 박사
 1989년~1992년 : 삼보컴퓨터
 1997년~현재 : 경상대학교
 컴퓨터과학과 교수

2014년~현재 : 경상대학교대학원 문화융복합학과 교수
 관심분야: Med. Image, IT융복합, Computer Network