

자연영상에서 문자의 형태 분석을 이용한 문자영역 추출에 관한 연구

양재호¹, 한현호¹, 김기봉², 이상훈^{3*}

¹광운대학교 플라즈마 바이오 디스플레이학과 전임연구원

²대전보건대학 컴퓨터정보학과 교수

³광운대학교 인제니움학부대학 교수

A Study on Extraction of text region using shape analysis of text in natural scene image

Jae-Ho Yang¹, Hyun-Ho Han¹, Ki-Bong Kim², Sang-Hun Lee^{3*}

¹Associate Research Engineer, Dept of Plasma Bio Display, KwangWoon University

²Professor, Department of computer information, Daejeon health institute of technology

³Professor, Ingenium college of liberal arts, Kwangwoon university

요 약 본 논문에서는 일상에서 획득할 수 있는 자연 영상에서 문자를 검출하기 위해 영상 개선 및 문자의 형태를 분석하여 문자를 검출하는 방법을 제안한다. 제안하는 방법은 자연 영상에서 문자로 인식될 영역의 검출률을 향상시키기 위해 객체부분의 경계를 언샤프 마스크를 사용하여 강조하였다. 향상된 객체의 경계 부분을 이용하여 영상의 문자 후보영역을 MSER(Maximally Stable External Regions)을 이용하여 검출하였다. 검출된 문자 후보영역에서 실제 문자로 판단될 영역을 검출하기 위해 각 영역들의 형태를 분석하여 글자의 특성을 갖는 영역외의 비 문자영역을 제거하여 실제 문자영역 검출률을 높였다. 본 논문의 정량적 평가를 위해 문자 영역의 검출률과 정확도를 이용하여 기존의 방법들과 비교하였다. 실험 결과 기존의 문자 검출 방법보다 제안하는 방법이 비교적 높은 문자영역의 검출률 및 정확도를 보였다.

주제어 : 문자 검출, 객체의 경계 향상, MSER 알고리즘, 문자패턴, 언샤프 마스크

Abstract In this paper, we propose a method of character detection by analyzing image enhancement and character type to detect characters in natural images that can be acquired in everyday life. The proposed method emphasizes the boundaries of the object part using the unsharp mask in order to improve the detection rate of the area to be recognized as a character in a natural image. By using the boundary of the enhanced object, the character candidate region of the image is detected using Maximal Stable External Regions (MSER). In order to detect the region to be judged as a real character in the detected character candidate region, the shape of each region is analyzed and the non-character region other than the region having the character characteristic is removed to increase the detection rate of the actual character region. In order to compare the objective test of this paper, we compare the detection rate and the accuracy of the character region with the existing methods. Experimental results show that the proposed method improves the detection rate and accuracy of the character region over the existing character detection method.

Key Words : Text Extraction, Improved Object Edge, MSER Algorithm, Text Pattern, Unsharp Mask

*The study is sponsored by the 2018 research fund of Kwangwoon University.

*Corresponding Author : Sang-Hun Lee (leesh58@kw.ac.kr)

Received November 13, 2017

Revised November 17, 2017

Accepted January 20, 2018

Published January 28, 2018

1. 서론

오늘날 멀티미디어의 발전으로 다양한 문서나 영상에서 문자를 검출하는 알고리즘의 필요성이 요구되고 있다 [1]. 영상에서 문자는 문서 및 영상 분석을 위한 중요한 정보를 포함하고, 이를 위한 문자 검출과 문자 인식은 영상 내의 내용 분석과 같이 다양한 분야에서 사용된다.

영상으로부터 문자 검출 및 인식은 크게 인공영상[2]에서의 문자 검출과 자연영상[3]에서의 문자검출로 나눌 수 있다. 인공영상에서의 문자는 멀티미디어 속에 포함된 임의로 문자로, 영화나 뉴스, 영화포스터 등에서 쉽게 볼 수 있다. 이러한 문자들은 보통 영상 내에 일정한 위치에 있으며, 외부요인에 의한 변수가 작다. 그러나 일상 생활에서 촬영한 교통 표지판이나 간판이 포함된 영상에서 볼 수 있는 자연영상은 문자의 글꼴, 색상, 자간, 조명 등의 여러 가지 특성과 환경적인 조건에 따라 민감하기 때문에 정확한 문자 검출을 위한 연구가 다수 진행중이다.

2. 자연 영상에서의 문자 검출

일반적으로 영상에서의 문자는 Caption text와 Scene text로 분류 할 수 있다. 이 중 Caption text는 멀티미디어에서 임의로 포함된 문자로서 영화의 자막, 영화포스터, 뉴스 등으로 일정한 위치에 존재하거나 외부요인으로 인한 변수가 작다. 그러나 일상생활에서 획득할 수 있는 카메라나 핸드폰으로 촬영한 교통 표지판이나 간판 등의 영상인 Scene Text는 Text의 방향, 크기, 스타일, 색상 위치 등 다양한 외부요인이 존재하여 검출이 어렵다. Fig. 1은 문자가 포함된 영상이다.



Fig. 1. Texts in the image.

(a) Scene Text, (b) Caption Text.

최근에 이를 해결하기 위해 자연영상으로부터의 문자 검출을 위한 많은 방법이 제안되었다[3-8]. 자연영상에서 문자 검출을 위한 방법은 대표적으로 텍스트 기반의

SWT 알고리즘과 MSER 알고리즘이 있다.

2.1 SWT알고리즘

B Epshtein, E Ofek, Y Wexler가 제안한 SWT 알고리즘[9]은 SW(Stroke Width)를 이용하여 비 문자영역과 문자 영역을 구분하는 방법이다. 이 방법은 영상에서 각 문자는 한 획(Stroke)으로 구성되고, 한 획 문자의 너비는 일정하다는 규칙을 이용하여 한 획의 너비를 통해 문자를 검출한다. SWT 알고리즘에서 한 획의 너비를 구하는 방식은 Fig. 2와 같다.

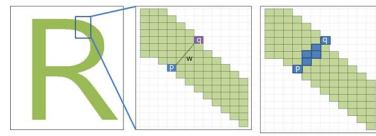


Fig. 2. SWT Method.

먼저 영상에 Canny 연산자를 적용해 생성한 Edge Map을 기반으로 에지 픽셀 p 에서 경사도 방향(Gradient x,y)으로 탐색을 진행하여 만나는 픽셀 q 를 찾는다. 찾은 p 와 q 의 사이의 길이 값 w (width)를 한 획의 폭으로 결정하고, 선을 그어 내부를 채워준다. 길이 값 w 를 구하는 방법은 다음과 같다.

$$w = p + n \times d_p \quad (1)$$

식(1)의 d_p 는 픽셀 p 의 경사도 방향이며, n 은 픽셀 q 의 위치를 찾을 때 까지 순차적으로 증가하는 상수 값이다. w 은 픽셀 q 의 위치를 찾을 때까지 진행된다. 이러한 SWT 알고리즘은 은 폭의 문자들을 쉽게 검출할 수 있으며 처리 시간이 빠른 장점이 있다. 하지만 영상의 문자에서 추출한 윤곽선이 폐곡선이 아니거나 다른 문자와 연결된 형태인 경우 문자의 폭이 일정하지 않아 오검출을 할 수 있다. 또한 문자의 크기 변화, 조명의 밝기 변화에 따라 문자가 손실될 수 있어 검출하지 못하는 단점이 있다.

2.2 MSER 알고리즘

MSER(Maximally Stable Extremal Regions) 알고리즘[10-12]은 영상에서 명도 정보를 이용하여 영상 내의 문자들의 질감이나 색상이 비슷한 값을 가지는 규칙을 이용해 유사한 값을 지닌 영역들을 검출하는 방법으로 문자 영역 검출 과정에 많이 사용되는 알고리즘이다. 이

알고리즘은 유사한 값을 갖는 영역을 구하기 위해 영상을 그레이스케일로 변환한 뒤 임계값을 0부터 255 또는 역순으로 그레이스케일 값을 점진적으로 변화시키면서 영상 내 연결된 각 픽셀들의 변화가 최소인 경우 해당 영역을 검출하는 방법이다.

$$f(t) = \frac{abs(I(t) - I_0)}{\max(\frac{1}{t} \int_0^t abs(I(t) - I_0) dt)} \quad (2)$$

식(2)의 $I(t)$ 는 임의의 임계값 t 에서의 영상의 밝기 값이며, I_0 는 극값(Extremum)에서의 밝기 수치이다. 이를 사용하여 검출된 영역은 내부 값은 1이고 외부 값은 0이 된다. 1의 값을 갖는 픽셀 개수의 총 합은 영역의 넓이를 나타낸다. MSER 알고리즘의 구현 방식은 Fig. 3과 같다.

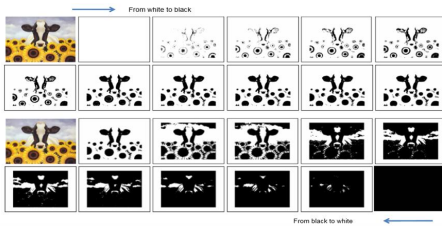


Fig. 3. MSER Method.

MSER 알고리즘은 영역 간의 인접성이나 보존 및 임계 값 범위 내에서 비슷한 영역만을 선택하는 안정성, 다중 영역 검출 등의 장점이 있다. 하지만 영상에서 영역들의 값이 동일한 인접 영역들이 많을 경우 불필요하게 많은 영역들이 검출되고, 이러한 경우 비 문자영역을 제거하는 후처리 단계를 이용하여 영상에서 문자 영역만을 검출을 하는 과정이 필요하다.

3. 제안하는 방법

본 논문에서는 자연영상에서의 문자 부분 손실을 최소화하기 위해 경계부분에서 대비를 강조 처리하여 영상을 개선하였다. 개선된 영상에서 흑백 값의 변화에 따라 연결영역의 변화가 적은 곳을 문자 후보 영역으로 검출한다. 그리고 문자 영역의 패턴을 분석하여 단계별로 비 문자 영역을 제거함으로써 영상 내 문자를 검출하게 된

다. 또한 비 문자 영역 패턴 필터링을 통해 다국어 문자의 문자검출이 가능하다. 제안하는 문자 검출 방법은 Fig. 4와 같다.

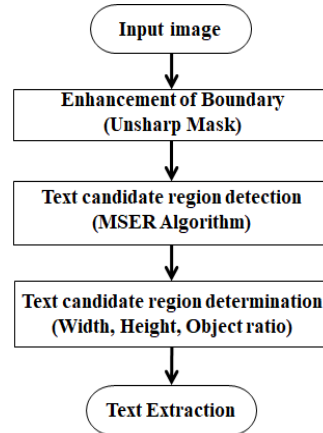


Fig. 4. Flow chart

3.1 객체 경계 향상

자연영상은 외부 촬영한 영상들로 외부 요인으로 인해 문자검출에 많은 어려움이 있다. 그 중 카메라 등에서 디지털 영상을 획득 할 때 흔들림이나 빛에 의한 문자 손실이 발생 할 수 있다. 이러한 외부요인은 영상에서 문자 검출 알고리즘의 성능을 저하시키는 주요 원인이므로 개선이 필요하다. 이러한 문제를 해결하기 위해 본 논문에서는 객체의 경계부분의 대비를 전처리 과정을 통해 향상시켜 손실된 문자들의 경계를 강화한다.

언샤프 마스크[13]는 객체의 경계부분을 향상시키는 방법이다.

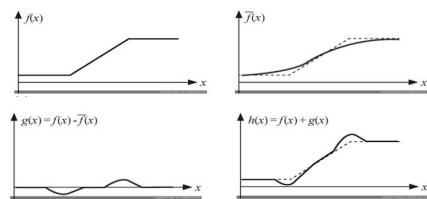


Fig. 5. Graph of Unsharp mask. (a) Boundary of Input image, (b) Boundary of Blur Image, (c) Different in Boundary part, (d) Result of Unsharp mask

Fig. 5는 입력 영상에서 윤곽선 부분의 그레이스케일 값 변화를 1차원 함수 $f(x)$ 로 표현한 것이다. x 값이 증가함에 따라 그레이스케일 값의 변화가 큰 것을 볼 수

다. 이 함수에 흐려짐 효과를 적용한 결과가 (b)의 $\bar{f}(x)$ 이다. $\bar{f}(x)$ 는 x 축을 따라 값이 완만하게 증가하게 된다. 입력 함수 $f(x)$ 에서 $\bar{f}(x)$ 를 뺀 결과는 (c)의 $g(x)$ 함수로 나타난다. $g(x)$ 는 입력 함수의 윤곽선 부분에서 0이 아닌 값을 갖는 함수 형태이다. 마지막으로 $g(x)$ 를 원본 함수 $f(x)$ 와 더하면 그림 (d)의 $h(x)$ 와 같이 강조된 경계를 갖는 함수가 생성된다.



Fig. 6. Boundary contrast enhancement results using unsharp mask. (a) Original Image (b) Blur image (c) Detected Boundary (d) Contrast enhancement of object boundaries.

Fig. 6은 언샤프 마스크를 이용하여 객체의 경계(c)를 구하고 대비가 향상된 결과(d)를 볼 수 있다.

3.2 MSER 기반의 문자 후보 검출

MSER은 영역 검출 알고리즘으로, 주변에 비해 밝기 값이 밝거나 어두운 영역에서 변화가 유사한 값들을 가진 영역을 분류하는 방법이다. 일반적으로 영상에서 문자의 밝기는 일정하고 배경과의 차이가 분명한 특징이 있다. 이러한 특징이 MSER을 이용한 문자 검출 연구에서 연결요소를 찾기 위한 과정에 자주 사용되고 있다. MSER 알고리즘은 extremal 영역을 구하는 과정과 밝기의 변화값이 비슷한 영역을 구하는 과정이 필요하다. 먼저 extremal 영역을 구하는 방법은 언샤프 마스크 결과 $h(x)$ 영상의 각 픽셀을 그레이스케일로 변환한다. 변환한 영상에서의 각 픽셀 값은 $h'(x)$ 로 정의한다. 또한 각 영역을 저장할 level set $S(x)$ 은 식(3)을 통하여 구한다.

$$S(x) = y \in A : h(y) \leq h'(x) \quad (3)$$

식(3)의 A 는 $h(x)$ 의 화소의 집합, $S(x)$ 에서 조합 검색(Union find)을 통해 연결 구성요소 C 를 계산하며, 영역(Q)을 형성한다. 이때 연결 구성요소 C 는 $Q_i \subset Q_{i+1}$ 를 만족하며, 이를 만족하는 C 를 maximal region R 로 정의한다. 이렇게 주어진 영상 $h(x)$ 에 대해

생성된 모든 영역 R 의 집합을 $R(h)$ 로 정의한다. 다음은 생성된 모든 $R(h)$ 사이에서 가장 변화가 안정적인 영역을 구하는 방법이다.

$$\begin{aligned} R_{+\Delta} &= \operatorname{argmin}\{ |Q| : Q \in R(h), Q \supset R, h(Q) \geq h(R) + \Delta \} \\ R_{-\Delta} &= \operatorname{argmax}\{ |Q| : Q \in R(h), Q \subset R, h(Q) \leq h(R) - \Delta \} \\ &\text{where } \Delta > 0 \end{aligned} \quad (4)$$

다음 식(4)를 통해 $R_{-\Delta}$, $R_{+\Delta}$ 를 구하고 영역 변화율을 계산한다.

$$\sigma = \frac{|R_{+\Delta}| - |R_{-\Delta}|}{|R|} \quad (5)$$

식(5)의 σ 는 가장 안정화된 영역을 찾는 기준인 영역 변화율이며, σ 의 값이 최소화되는 영역 R 이 값의 변화가 가장 작은 영역이 된다. Fig. 7는 언샤프 마스크를 적용하여 경계부분을 강조한 뒤 MSER 알고리즘을 적용한 결과이다.

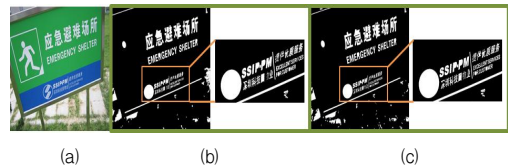


Fig. 7. Flowchart of text detection in image (a) Original Image (b) Result of MSER (c) Result of Contrast Enhancement MSER.

Fig. 7의 (b)는 간판의 문자 중 일부 경계가 손실돼 검출되지 않은 문제점이 있다. 이러한 문제를 해결하기 위해 제안하는 방법은 자연 영상에서 언샤프 마스크를 사용하여 객체의 경계를 향상시키고 기존의 MSER 결과에서 손실된 문자영역을 검출(c) 하는 것을 볼 수 있다.

3.3 문자후보 패턴 분석을 통한 비 문자 영역 제거

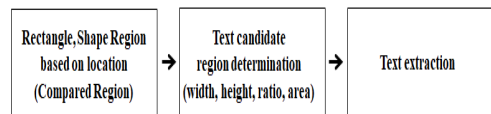


Fig. 8. Process of text area detection using text pattern analysis.

3.2절에서 MSER의 결과는 유사한 값을 지닌 영역들을 모두 검출하기 때문에 검출된 영역이 문자 영역이라고 판단하기 어렵다.

따라서 본 논문에서는 문자 후보 영역내의 문자 영역만을 검출하기 위해 문자영역의 패턴을 분석하여 비 문자 영역제거 과정을 진행하였다. 먼저 제안하는 방법에서는 MSER 결과 후 각각의 영역의 패턴 분석을 위해 Fig. 9와 같이 사각형 모양 및 외곽선 모양으로 라벨링 [14]을 진행하였다.



Fig. 9. Results of MSER region labeling
(a) Original Image (b) Result of Contour Labeling
(c) Result of rectangle Labeling

MSER결과 영역의 특징을 살펴보면 영상에서의 문자들은 서로 유사한 너비와 높이 혹은 면적을 가지고 있다. 제안하는 방법은 이러한 특성을 이용해 각각의 영역의 너비와 높이, 면적 등을 구하여 각 영역의 종횡비와 객체의 크기 비율을 구한다.

3.3.1 종횡비 분석 기반의 비 문자 영역 판별

Fig. 9의 영상을 보면 영상에서 문자들은 너비와 높이의 비율이 일정한 특징을 가지고 있다. 보통 종횡비를 이용해 비 문자 영역을 판단하는 경우, 문자의 일부는 종횡비 값이 높아 비 문자 영역으로 판단된다.

$$\epsilon_1 = \frac{R_w}{R_h} \quad (6)$$

식(6)은 기본 기하학 특징으로 사용하는 종횡비에 대한 수식으로 R_w 는 사각형 라벨링에 대한 너비이고 R_h 는 높이이다. Fig. 10은 문자 후보영역의 종횡비를 적용하여 비 문자 영역을 판단한 영상이다.



Fig. 10. Incorrect removal non-text results using aspect ratio. (a) Original Image (b) Comparison area (c) Result of Labeling.

Fig. 10의 결과에서 대문자 I는 식(8)에서 종횡비 값이 높아 비 문자영역으로 인식되는 결과를 볼 수 있다. 이러한 문제점을 해결하기 위해 종횡비방식을 추가하여 식(7)로 나타내었다.

$$\epsilon_2 = \frac{R_h}{R_w} \quad (7)$$

식(6)과 식(7)의 ϵ_1 와 ϵ_2 를 문자후보영역의 패턴 분석을 이용해 문자영역들이 가지는 종횡비의 비율을 $\omega_1=2$, $\omega_2=4$ 로 적용한다. 비 문자 후보영역을 제거하기 위한 수식은 다음과 같다.

$$T_{area} = \begin{cases} \epsilon_1 < \omega_1 \text{ and } \epsilon_2 < \omega_2 \\ NT \end{cases} \quad (8)$$

식(8)을 적용하여 종횡비 값이 높은 비 문자 영역 제거 결과를 Fig. 11에서 보여준다.



Fig. 11. Result of non-text area removal using aspect ratio. (a) Original Image (b) Result of Contour Labeling (c) Non-text area (d) Result of removing the non-text area.

3.3.2 객체의 크기 분석 기반의 비 문자영역 판별

다음 단계로 각 영역 윤곽의 넓이와 윤곽의 외접 사각형의 넓이의 비로 객체의 크기 비율을 계산하여 비 문자 후보 영역을 판단하였다. 각 영역들 구하기 위한 방법은 식(9)와 같다.

$$\epsilon_3 = \frac{\text{Object Area}}{\text{Bounding Rectangle Area}} \quad (9)$$

종횡비를 적용하여 얻은 영상에 각 영역들의 ϵ_3 을 분석함으로써 객체의 크기 비율을 구해 비 문자 영역을 제거한 식은 다음과 같다.

$$T_{area} = \begin{cases} \omega_3 < \epsilon_3 < \omega_4 \\ NT \end{cases} \quad (10)$$

식(10)에서 객체크기의 비율을 각각 $\omega_3=0.3$, $\omega_4=0.9$ 를 적용하여, 비 문자 영역이 제거된 결과는 다음과 같다.



Fig. 12. Result of non-text area removal using Extent. (a) Result of aspect ratio, (b) Result of Aspect ratio, (c) Non-text area, (d) Result of removing the non-text area.

Fig. 12는 객체의 크기를 이용해 비 문자 영역을 제거한 영상으로 제안하는 방법의 중횡비를 사용한 결과(a)에 윤곽선 라벨링을 진행하여 문자 후보영역 객체의 크기를 구한 뒤, 식(10)을 적용하여 최종적으로 자연 영상에서 문자영역을 검출한 결과를 볼 수 있다.

4. 실험 및 결과 고찰

본 실험은 Windows 10 환경에서 Visual Studio 2015 프로그램으로 OpenCV를 사용하여 개발하였고, 자연 영상에서의 문자 검출 연구에 쓰이는 공개데이터 셋인 MSRA-TD500[15]을 이용하였다.

실험은 MSRA-TD500의 테스트 영상(200개)와 학습 영상(300개)를 사용하여 영상에서 문자검출을 하였으며, 각 알고리즘의 결과를 비교하였다. 비교 대상으로 기존의 Connected Component 기반 알고리즘과 SWT(Stroke Width Transform) 기반 알고리즘을 비교하여 실험을 진행하였다.

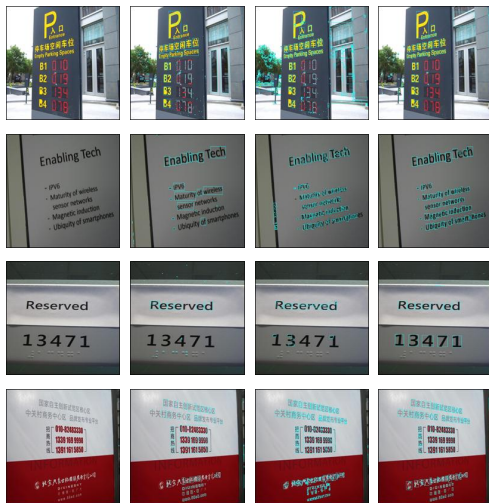


Fig. 13. Test set Images

Fig. 13은 테스트 영상에서의 문자 검출 결과로, 첫 번째 행의 영상에서 연결기반 방법과 SWT 알고리즘에서는 문자영역뿐만 아니라 비 문자영역 또한 문자영역으로 검출되는 문제가 있다, 하지만 제안하는 방법에서는 비 문자영역으로 판단하여 문자영역만 검출하는 결과를 보여준다. 두 번째 행 영상과 세 번째 행영상의 결과에서 연결기반방법은 이진화 과정에서의 임계 값에 따라 문자영역의 일부분이 검출되지 않은 결과를 보였고, SWT 방법에서도 Canny 연산자의 결과에 따라 문자 영역의 일부가 검출되지 않는 결과를 보였다. 그러나 제안하는 방법에서는 모든 문자영역을 검출하는 결과를 볼 수 있다. 네 번째 행 영상의 간판에서는 SWT방법에서 문자의 폭이 일정하지 않을 경우 일부 문자만을 검출하였고, 제안하는 방법에서는 모든 문자영역을 검출하는 결과를 볼 수 있다.

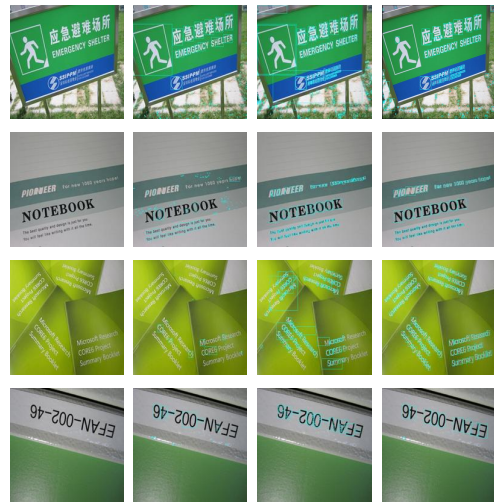


Fig. 14. Training set Images.

Fig. 14는 학습영상에서의 문자 검출 결과로 첫 번째 행의 결과에서는 연결기반 방법과 SWT 결과 모두 비 문자 영역인 로고 영상 또한 문자영역으로 판단하는 결과를 보였다. 하지만 제안하는 방법에서는 문자영역만을 검출하는 결과를 볼 수 있다. 두 번째 행와 세 번째 행의 결과에서 SWT알고리즘은 글자의 폭이 일정하지 않거나 밝기 값에 의한 임계값의 차이로 인해 문자 영역의 일부가 검출되지 않았고, 연결 기반방법에서도 문자의 그레이스케일 값이 일정하지 않아 문자의 일부분이 검출되지 못하는 결과를 보였다. 네 번째 행의 결과에서 제안하는 방법은 문자의 일부분이 중횡비의 비율이 높아 일부

영역이 검출되지 못하는 결과를 보였고, SWT방법에서는 문자의 폭이 일정하지 않은 잘못된 영역 또한 검출되는 결과를 보였다.

본 논문에서는 제안하는 방법의 성능을 평가하기 위해 기존의 비교 방법인 문자열, 문자, 화소 중 객관적으로 많이 사용하는 문자 단위 평가를 이용하여 측정하였고, 기존의 알고리즘과 검출률(Detection Rate) 및 정확도(Precision)를 비교하였다.

$$Detection\ Rate = \frac{NODT}{TNOT} \times 100 \quad (11)$$

$$Precision = \frac{TNOCT}{NODT} \times 100 \quad (12)$$

식(11)과 식(12)에서 검출률과 정확도에 사용한 TNOT (Total number of texts)는 영상에 존재하는 실제 문자영역의 총 개수이고 NODT(Number of detected texts)는 검출 개체의 수이다. 또한 정확도에 사용한 TNOCT(Total number of correct texts)는 영상 내에서 검출한 정확한 문자개체의 개수이다. Table 1은 MSRA-TD500 영상에서 기존의 방법들과 제안하는 방법의 검출률과 정확도를 비교한 표이다.

Table 1. Performance comparison between proposed method and existing method

Method	Proposed Method	Connected Component Method	SWT Method
TNOT	12,326	12,326	12,326
NODT	10,079	6,580	9,433
TNOCT	9,228	5,869	8,252
Detection Rate(%)	81.8%	53.4%	76.5%
Precision Rate(%)	91.6%	89.2%	87.5%

실험 결과에서 기존의 연결 기반 방법은 53.4%의 가장 부정확한 문자 후보영역 검출 결과를 보였고, 제안하는 방법은 연결 기반을 개선한 SWT 알고리즘 대비 검출률이 향상된 결과를 보였다. 정확도는 비 문자영역을 제거하여 기존 방법 대비 4.1% 향상된 결과를 보였다. 제안하는 방법은 문자 후보영역 검출 뒤 문자영역의 형태를 분석하여 문자영역의 손실을 최소화하였으며, 비 문자영역의 제거를 통해 기존의 알고리즘보다 개선된 결과를 보였다.

5. 결론

본 논문에서는 영상에서 정확한 문자검출 알고리즘을 제안한다. 제안하는 방법은 자연 영상에서의 손실된 객체들을 개선하기위해 언샤프 마스크를 사용하여 객체의 경계부분을 개선하였다. 그리고 문자 후보영역을 검출하기 위해 MSER(Maximally Stable External Regions)알고리즘을 적용하였다. 검출된 문자 후보영역 각 영역들의 높이와 너비, 객체의 크기, 객체의 비율 등의 패턴을 분석하여, 비 문자영역을 제거함으로써 최종적으로 문자영역을 검출 할 수 있도록 제안하였다. 제안하는 문자 검출 기법은 비문자 영역으로 판단될 수 있는 문자영역의 손실을 최소화하고, 비문자 영역을 제거함으로써 기존의 문자 검출 방법보다 향상된 결과를 보였다. 제안하는 문자 검출 기법은 문자 인식 과정의 전처리 과정으로 사용될 수 있어 많은 연구에서 응용이 가능할 것으로 예상된다.

REFERENCES

- [1] Khatib, T., Karajeh, H., Mohammad, H., & Rajab, L. (2015). A hybrid multilevel text extraction algorithm in scene images. *Scientific Research and Essays*, 10(3), 105-113.
- [2] Raza, A., Siddiqi, I., Djeddi, C., & Ennaji, A. (2013). Multilingual artificial text detection using a cascade of transforms. *In Document Analysis and Recognition (ICDAR), 2013 12th International Conference on*, 309-313
- [3] Seeri, S. V., Pujari, J. D., & Hiremath, P. S. (2016). Text Localization and Character Extraction in Natural Scene Images using Contourlet Transform and SVM Classifier. *International Journal of Image, Graphics and Signal Processing*, 8(5), 36.
- [4] Koo, H. I., & Kim, D. H. (2013). Scene text detection via connected component clustering and nontext filtering. *IEEE transactions on image processing*, 22(6), 2296-2305.
- [5] Feng, Y., Song, Y., & Zhang, Y. (2016, December). Scene text detection based on multi-scale SWT and edge filtering. *In Pattern Recognition (ICPR), 2016 23rd International Conference on*. 645-650.
- [6] Sujatha, P., & Sudha, K. K. (2015). Performance analysis of different edge detection techniques for image segmentation. *Indian Journal of Science and Technology*, 8(14).
- [7] Kang, L., Li, Y., & Doermann, D. (2014). Orientation

robust text line detection in natural images. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 4034-4041.

[8] Gupta, N., & Banga, V. K. (2012). Localization of text in complex images using haar wavelet transform. *International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN, 2278-3075*.

[9] Epshtein, B., Ofek, E., & Wexler, Y. (2010, June). Detecting text in natural scenes with stroke width transform. *In Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. 2963-2970.

[10] Cao, X., Ren, W., Zuo, W., Guo, X., & Foroosh, H. (2015). Scene text deblurring using text-specific multiscale dictionaries. *IEEE Transactions on Image Processing, 24(4)*, 1302-1314.

[11] Yu, Y. J., (2014). Document Image Binarization Technique using MSER, *Journal of the Korea Institute of Information and Communication Engineering, 18(8)*, 1941-1947.

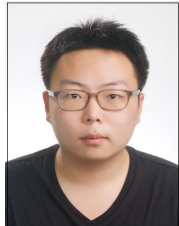
[12] Jeong, J. W & Jun K. K, Noise Robust Document Image Binarization using Text Region Detection and Down Sampling, *Journal of Korea Multimedia Society, 18(7)*, 843-852, 2015.

[13] Lira, J., & Rodríguez, A. (2014). Edge enhancement in multispectral satellite images by means of vector operators. *Geofisica internacional, 53(3)*, 289-308.

[14] Park, J. H., Lee, G. S., Kim, J. S., Ryu, S. H., & Lee, S. H. (2016). Independent Object Tracking from Video using the Contour Information in HSV Color Space. *Indian Journal of Science and Technology, 9(S1)*.

[15] Yao, C., Bai, X., Liu, W., Ma, Y., & Tu, Z. (2012). Detecting texts of arbitrary orientations in natural images. *In 2012 IEEE Conference on Computer Vision and Pattern Recognition*. 1083-1090.

양재호(Yang, Jae Ho) [학생회원]



- 2016년 2월 : 광운대학교 컴퓨터공학과(공학사)
- 2018년 2월 : 광운대학교 대학원 플라즈마바이오디스플레이학과(공학석사)
- 관심분야 : 영상인식, 3D 영상처리, VR, AR

· E-Mail : yjh2005@naver.com

한현호(Han, Hyun Ho) [정회원]



- 2009년 2월 : 광운대학교 컴퓨터공학과(공학사)
- 2011년 2월 : 광운대학교 정보콘텐츠 대학원 유비쿼터스컴퓨팅학과(공학석사)
- 2011년 3월 ~ 현재 : 광운대학교 정보디스플레이학과(박사과정)

· 관심분야 : 무선 네트워크, 영상인식, 3D 영상처리, 기계 학습

· E-Mail : icewine@kw.ac.kr

김기봉(Kim, Ki Bong) [정회원]



- 1993년 2월 : 충남대학교 대학원 전산학과(전산학석사)
- 1998년 8월 : 충남대학교 대학원 전산학과(정보과학박사)
- 1995년 9월 ~ 1997년 2월 : 해전대학교 전임강사

· 1997년 3월 ~ 현재 : 대전보건대학교 컴퓨터정보과 교수

· 관심분야 : 데이터베이스, 헬스케어, 정보시스템

· E-Mail : kbkim@hit.ac.kr

이상훈(Lee, Sang Hun) [중신회원]



- 1983년 2월 : 광운대학교 응용전자공학과(공학사)
- 1987년 2월 : 광운대학교 전자공학과(공학석사)
- 1992년 2월 : 광운대학교 전자공학과(공학박사)

· 1990년 ~ 현재 : 광운대학교 정교수

· 2012년 2월 ~ 2013년 2월 : 광운대학교 정보통신처장

· 2013년 2월 ~ 2015년 2월 : 광운대학교 학생복지처장

· 2015년 2월 ~ 2017년 12월 : 한국융합학회 편집위원장

· 2018년 1월 ~ 현재 : 한국융합학회 회장

· 관심분야 : 영상인식, 3D 영상처리, 기계 학습, 영상 처리와 생체 인식, IOT 영상처리

· E-Mail : leesh58@kw.ac.kr