

## 실세계 영상에서 경계선과 영상 분할을 이용한 기울기 검출 및 보정

백재경\*, 서영건\*\*

### 요약

본 논문에서는 문자열과 배경이 혼합된 장면에서 영상을 분할하여 기울기를 추출하고 보정하는 방법을 제안한다. 제안된 방법은 모폴로지를 이용하여 전처리를 하고 캐니 연산자를 이용하여 경계선을 검출한다. 그리고 경계선이 검출된 영상을 분할하여 경계선이 포함되어 있지 않는 영역은 배제하고 경계선이 포함되어 있는 영역만을 이용하여 여러 방향의 기울기에 따른 투영 히스토그램을 생성한다. 이를 이용하여 각 영역의 최대 경계선 집중도를 갖는 기울기를 구하고 장면의 기울기를 보정한다. 문자열과 배경이 혼합된 장면의 기울기 검출에서 제안된 방법은 경계선이 없는 무의미한 부분을 배제하기 때문에 기존의 방법보다 0.7% 더 좋은 결과를 얻을 수 있었다.

키워드: 장면 영상, 기울기 검출, 기울기 보정, 경계선 검출

## Extracting the Slope and Compensating the Image Using Edges and Image Segmentation in Real World Image

Jaegyung Paek\*, Yeong Geon Seo\*\*

### Abstract

In this paper, we propose a method that segments the image, extracts its slope and compensates it in the image that text and background are mixed. The proposed method uses morphology based preprocessing and extracts the edges using canny operator. And after segmenting the image which the edges are extracted, it excludes the areas which the edges are included, only uses the area which the edges are included and creates the projection histograms according to their various direction slopes. Using them, it takes a slope having the greatest edge concentrativeness of each area and compensates the slope of the scene. On extracting the slope of the mixed scene of the text and background, the method can get better results as 0.7% than the existing methods as it excludes the useless areas that the edges do not exist.

Keywords : scene image, slope extraction, slope compensation, edge extraction

### 1. 서론

현대 사회에서 정보는 매우 중요한 요소이고, 우리 주위의 장면들은 정보를 담고 있다. 특히 문자열이 포함되어있는 장면은 문자열을 추출, 인식하는 것만으로도 많은 정보를 얻을 수 있다. 장면에 포함된 문자열은 때때로 기울기를 가지고 있으며 이러한 기울어진 장면에서 문자열의 추출 및 인식의 정확도를 높이기 위해 기울기의 추출 및 보정은 중요하다. 기울기를 추출 및 보정 방법으로는 경계선과 최소자승법을 이용한 기울기 보정 방법[1], 객체의 위치와 최소자승법을 이용한 기울기 보정 방법[2], 투영 히스토그

※ Corresponding Author : Yeong Geon Seo

Received : September 30, 2016

Revised : October 25, 2016

Accepted : October 30, 2016

\* Gyeongsang Univ., Graduate School of Education, Dept. of Computer Education

\*\* Gyeongsang Univ., Dept. of Computer Science and Graduate School of CCBM

Tel: +82-55-772-1392

email: young@gnu.ac.kr

램을 이용한 방법[3,4]등이 있다. 이 방법들은 문서 영상 또는 문자열이 다수 포함된 장면에서는 빠르고 정확도가 높으나 문자열과 배경이 혼합된 장면에서는 상대적으로 낮은 정확도를 보이고 있다. 특히 최소자승법을 이용한 방법은 정상적인 데이터 분포에 속하지 않는 데이터가 존재한다면 잘못된 근사 결과를 낼 수 있는 최소자승법의 한계로 인해 문자열과 배경이 혼합된 장면에서는 적용하기 힘든 방법이다.

투영 히스토그램을 이용하여 기울기를 추출 및 보정 할 때에는 경계선 추출 또한 중요하다. 경계선 추출에 사용되는 방법으로는 소벨 연산자, 라플라시안 연산자, 캐니 연산자, FIS(Fuzzy Inference System) 등이 있다[5,6,7]. 이 방법들을 이용하여 경계선을 추출할 때 정확도를 높이기 위한 전처리 방법으로 모폴로지를 이용한 전처리 방법[8]과 히스토그램의 다중 임계영역을 이용한 전처리 방법[9]이 있다. 본 논문에서는 문자열과 배경이 혼합된 장면에서 기울기를 검출하기 위해 모폴로지를 이용하여 전처리하고 전처리 후 결과물에 캐니 연산자를 적용하여 경계선을 검출한다. 검출된 경계선을 분할하고 분할된 영상에서 경계선이 포함된 유의미한 영역만을 사용하여 투영 히스토그램을 생성한다. 투영 히스토그램을 이용하여 최대 경계선 집중도를 구한 후 이를 사용하여 기울기를 검출하는 방법을 제안한다. 제안된 방법을 이용하여 문자열과 배경이 혼합된 장면에서 기울기를 검출한 결과 높은 정확도로 보정이 된다는 것을 볼 수 있다.

## 2. 관련연구

### 2.1 기울기 검출

Jae-hyon 외 1인은 문서 영상에서 고속 기울기 검출을 위해 3단 coarse-to-fine 탐색 알고리즘을 사용한 기울기 검출을 제안하였다[4]. 이 방법은 문자 영역에서 캐니 연산자를 사용하여 경계선 영상을 구한다. 이후 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.

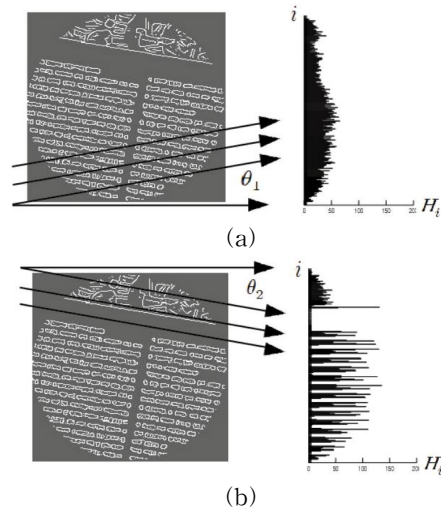
단계1)  $-45^\circ \sim +45^\circ$ 에서  $3^\circ$  간격으로 최대 경계선 집중도를 갖는  $\theta_{max1}$ 를 검출.

단계2)  $\theta_{max1}$ 을 중심으로  $-2^\circ \sim +2^\circ$ 에서  $1^\circ$  간격으로 최대 경계선 집중도를 갖는  $\theta_{max2}$ 를 검출.

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

(그림 1)은 경계선 투영의 예로 기울기에 따른 경계선 투영과 투영 히스토그램을 보여주고 있다. (그림 1)의 (a)는 문서의 기울기와 투영 방향이 달라 모든 행이 투영 구간에 비례한 경계선 집중도를 갖는 반면에 (그림 1)의 (b)는 문서의 기울기와 투영 방향이 일치하여 행에 따라 경계선이 집중된 결과를 보여주고 있다.

(그림 1) 경계선 투영의 예  
(a) 투영각  $\theta_1$  (b) 투영각  $\theta_2$

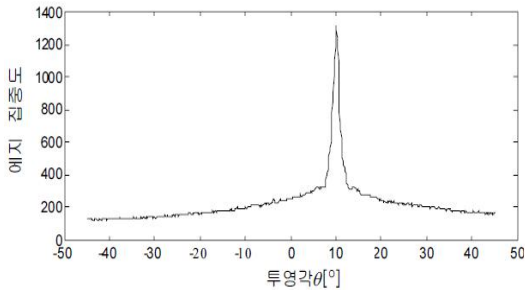


(Figure 1) Examples of edge projection  
(a) projection degree  $\theta_1$  (b) projection degree  $\theta_2$

(그림 2)는 문서의 기울기에 따른 경계선의 집중도를 그래프로 나타낸 것이다. coarse-to-fine 알고리즘을 사용한 기울기 검출은 기울기가 일정한 문서 영상에서 아주 빠르고 강력하다. 하

지만 문자와 배경이 혼합된 영상에서는 상대적으로 최적의 기울기가 추출되지 않을 수 있다.

(그림 2) 경계선 집중도



(Figure 2) Edge concentrativeness

Ishtani는 흑백 문서 영상에서 투영 히스토그램을 사용한 기울기 검출 방법을 제안하였다[3]. 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.

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

최대 전이점 집중도를 구하기 위하여 각각의 기울기 마다 전이점 집중도를 구해야 한다. 전이점 집중도는 전이점 투영 히스토그램의 분산(V)이고 n은 투영한 행의 합이다. 이 외에 Drako가 제안한 문자의 위치와 최소자승법을 이용한 방법이 있다[2]. 이 방법은 문자열에 포함된 각 문자에 사격형의 영역을 생성한다. 생성한 각 영역의 무게중심을 추출하고 추출된 무게중심 집합과 최소자승법을 이용하여 기울기를 구한다.

2.2 모폴로지를 이용한 전처리 방법

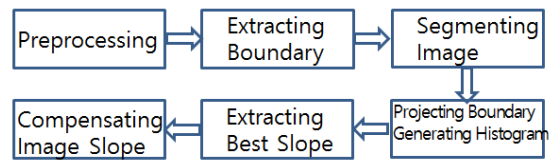
Chang-Young 외 1인은 경계선 검출 결과를 보완하기 위한 영상 개선 방법으로 그레이 기반의 모폴로지 연산을 수행하여 영상을 개선하는 방법을 제안하였다[8]. 모폴로지 연산에서 대표적인 방법으로는 침식, 팽창, 열기, 닫기가 있다. 여기서 열기는 침식 후 팽창 연산으로 주로 잡영을 제거할 때 사용하며 닫기는 팽창 후 침식

연산으로 주로 패인 부분을 채워줄 때 사용한다. 모폴로지 연산 중 영상 뿔셈을 열기 및 닫기와 결합하여 얻을 수 있는 top-hat 변환 및 bottom-hat 변환 중 bottom-hat 변환과 열기의 차를 이용하여 모폴로지 연산을 수행하였다.

3. 경계선과 영상 분할을 이용한 기울기 검출

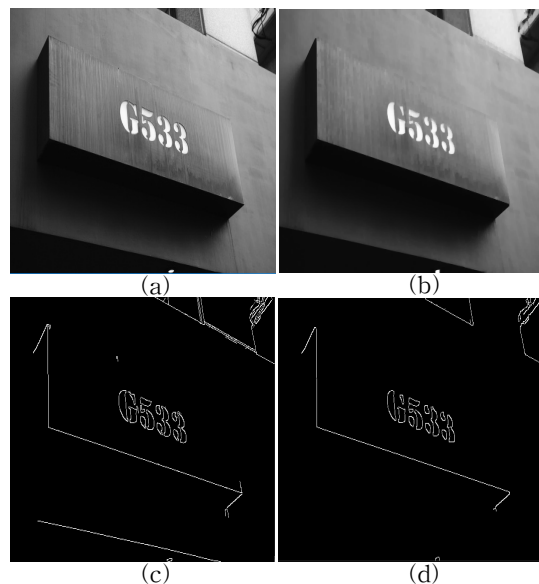
본 장에서는 장면 영상에서 문자열 인식을 위한 기울기 검출 및 보정을 제안한다.

(그림 3) 제안 방법의 구성도



(Figure 3) Flowchart of the proposed method

(그림 4) (a) 회색조 (b) 모폴로지 연산자 (c) 캐니 연산자 (d) 모폴로지 후 캐니 연산자



(Figure 4) (a) Grayscale (b) Morphology operator (c) Canny operator (d) Canny operator after morphology operator

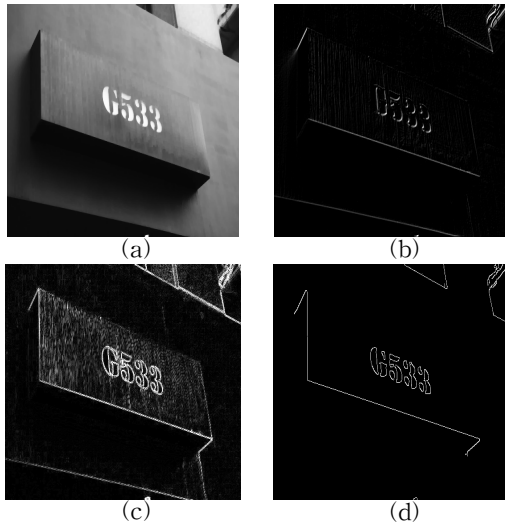
3.1 모폴로지 연산을 이용한 전처리

본 연구에서는 영상에서 경계선을 효과적으로 추출하기 위한 전처리 방법으로 모폴로지 연산을 사용한다. (그림 4)는 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.

3.2 경계선 추출

영상에서 경계선을 추출하기 위해 캐니 연산자를 사용하며, 캐니 연산자는 소벨 연산자와 라플라시안 연산자에 비해 잡음에 민감하지 않으며 강한 경계선을 추출할 수 있다. (그림 5)는 영상을 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.

(그림 5) (a) 모폴로지 연산자 (b) 소벨 연산자 (c) 라플라시안 연산자 (d) 캐니 연산자

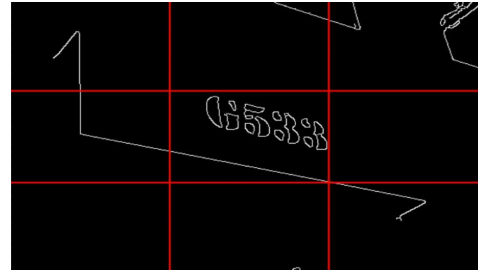


(Figure 5) (a) Morphology operator (b) Sobel operator (c) Laplacian operator (d) Canny operator

3.3 영상 분할

본 연구에서는 문자와 배경이 섞여있는 영상의 특성을 고려하여 추출한 경계선 영상을 (그림 6)과 같이 분할한다.

(그림 6) 3X3 영상 분할



(Figure 6) 3X3 image segmentation

3.4 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 다음과 같은 3단계 검출 알고리즘을 사용한다.

(단계 1)

입력: 경계선 투영 히스토그램의 분산(V)

출력: 최대 경계선 집중도를 갖는

기울기(max1)

For  $i = -45$  to  $45$

If  $maxV < V(i)$  then

$maxV <- V(i)$

$max1 <- i$

End

End

(단계 2)

입력: 분할된 영역의 최대 경계선 집중도를

갖는 기울기의 집합(max\_array)

출력: 최빈값을 갖는 기울기(max2)

$max2 <- getMode(max\_array)$

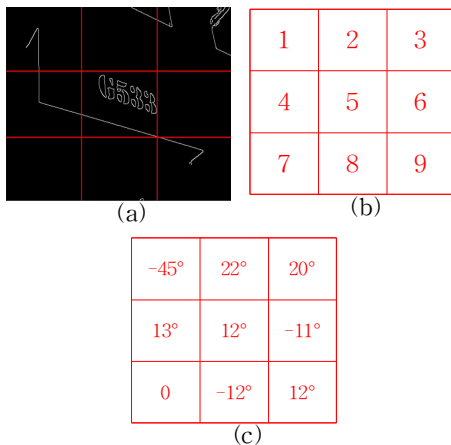
(단계 3)

```

입력: 중복된 최빈값을 갖는 기울기의
      집합(mode_array)
출력: 중앙값과 가장 근접한 기울기
(max3)
//mean 분할된 영역의 최대 기울기
      집중도를 갖는 기울기 집합의 중앙
      값
//temp 기울기와 중앙값의 차이를 저장하
      는
      임시저장장소
For i = 1 to mode_array.length
  If mean < mode_array[i] then
    temp <- mode_array[i] - mean
  else temp <- mean -
mode_array[i]
  End
  If max3 > temp then max3 <- temp
End
End
    
```

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

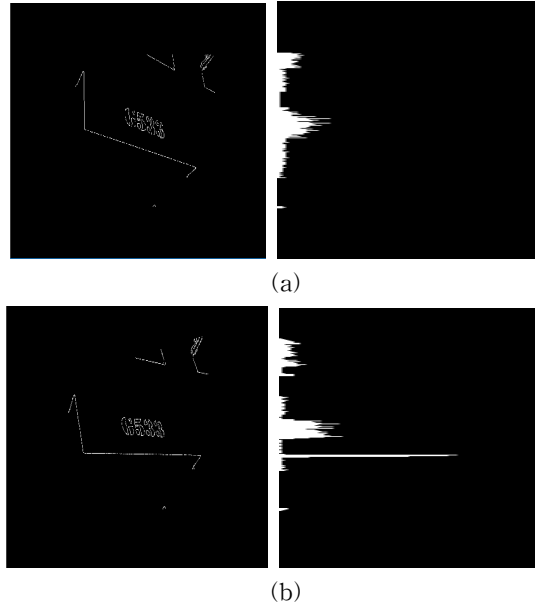
(그림 7) (a) 3X3 분할된 영상 (b) 영역 순서 (c) 각 영역의 최적 기울기



(Figure 7) (a) 3X3 segmented images (b) Order of each area (c) Best slope of each area

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

(그림 8) 기울기에 따른 경계선 투영 히스토그램 (a) 기울기 보정 전 (b) 기울기 보정 후



(Figure 8) Edge projection histogram according to the slopes (a) Before slope compensation (b) After slope compensation

최대 경계선 집중도를 검출하기 위해서는 투영 기울기에 따른 경계선 집중도가 필요하다. 경계선 집중도는 경계선 투영 히스토그램의 분산 (V)으로 n은 경계선 투영 히스토그램의 값이 0이 아닌 행의 개수이고, A는 경계선이 없는 행을 제외한 경계선 투영 히스토그램 값의 집단이다. 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. 여기서 getMode()란 분할된 영상의 최대 경계선 집중도를 갖는 기울기를 각각 비교하여 기울기마다 빈도수를 측정 후 최대 빈도수를 가지는 기울기를 추출한다. 이때 (그림 7)의 분할된 영상에서 정확도를 높이기 위해 경계선이 없는 영역7을 제외한다. 단계 3은 분할된 영역의 최대 경계선

집중도를 갖는 기울기의 집합에서 중복된 최빈 값이 있을 경우 중앙값에 가장 가까운  $\max 3^\circ$ 을 검출하여 정확성을 높인다. 여기서 검출된 최적 기울기를 이용하여 영상 전체에 적용, 기울기를 보정한다. (그림 9)는 기울기 보정 전 영상의 경계선 누적 히스토그램과 기울기 보정 후 영상의 경계선 누적 히스토그램의 비교이다. 3단계 검출 알고리즘의 단계 1은 Jae-hyon 의 1명[4]이 제안한 것으로 투영 기울기 변화는 장면 영상의 경계선 분포 특성을 고려하였다

### 4. 실험 및 평가

#### 4.1 실험 방법

본 연구는 (그림 10)의 영상을 이용하여 실험 하였으며 보정된 기울기의 정확도를 추출하기 위하여 다음과 같은 식을 사용했다.

$$d = \tan^{-1}\left(\frac{\Delta y}{\Delta x}\right)\left(\frac{180}{\pi}\right), p = 100 - d \quad \text{식(3)}$$

(그림 9) 실험에 사용한 영상



(Figure 9) Images used for experimentation

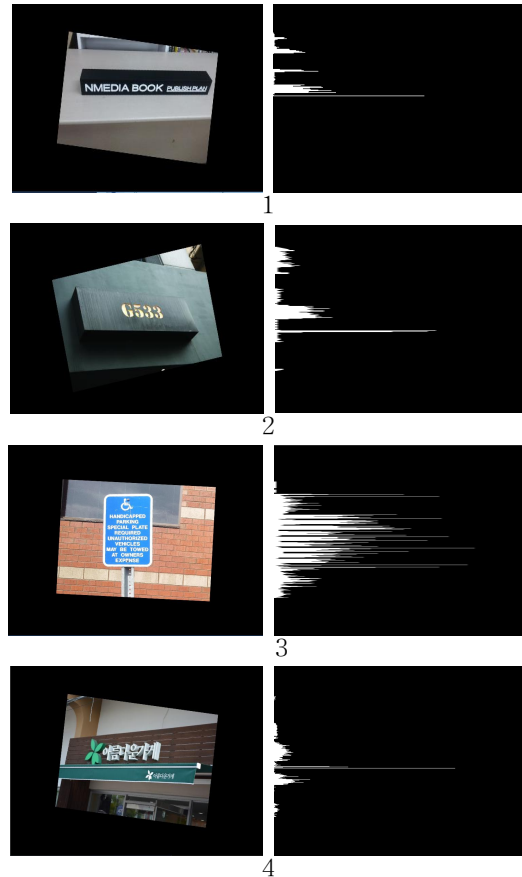
식 (3)에서  $d$ 는 수평선과 문자열 기울기의 각도,  $m$ 은 직선의 기울기,  $p$ 는 정확도를 의미하고 이는 육안으로 식별 가능한 문자열의 기울기와

수평간의 근사치를 의미한다. 본 실험을 진행한 시스템 환경으로는 Windows 10 Enterprise 64 비트 운영체제, 4GB 메모리를 사용하였다.

#### 4.2 성능 평가

제안된 방법의 성능을 평가하기 위한 척도는 기존의 방법과 제안된 방법의 기울기 추출의 정확도를 비교한다. 기존의 Ishitani 알고리즘[3]과 coarse-to-fine 알고리즘[4], 본 연구에서 제안한 알고리즘을 비교하였다. (그림 11)은 본 연구에서 제안한 알고리즘을 적용하여 기울기를 보정한 영상이다. 좌측은 원 영상의 기울기를 보정한 것이며 우측은 그 영상의 수평 히스토그램을 보인 것이다. 히스토그램의 분포를 보면 기울기가 잘 보정되었음을 알 수 있다.

(그림 10) 기울기 보정 결과 영상







(Figure 10) Result images after slope compensation

<표 1>은 제안된 방법과 비교 방법들을 실험 영상 9 개에 적용한 기울기 보정의 성능을 평가한 것이다. 표시된 값은 식 (3)의 p값이며, 식 (3)을 이용하여 문자열의 기울기와 수평간의 차이를 나타낸 것으로써, 기울기 보정의 정확도를 의미한다. 100에 가까운 값일수록 성능이 우수한 것이다. 제안 방법의 평균 정확도는 99.29이며, [4] 방법의 평균은 98.53이며, [3] 방법의 평균은 98.51이다. 결론적으로 제안 방법이 기존

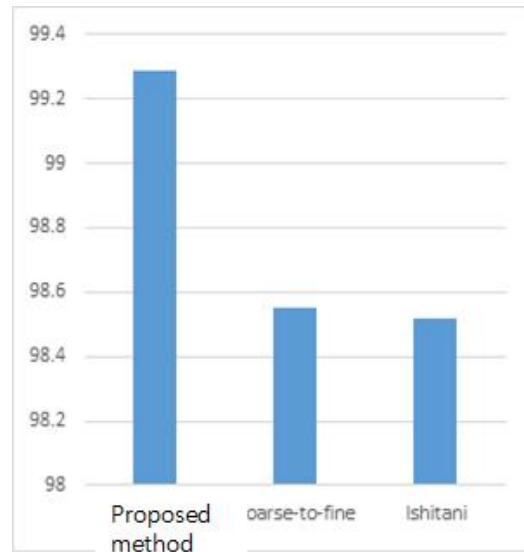
의 방법에 비해 0.7% 우수한 기울기 추출 방법임을 볼 수 있었다.

<표 1> 실험 영상에 대한 실험 결과

Result Image	Precision(p)		
	Proposed Algorithm	coarse-to-fine[4]	Ishitani[3]
1	99.7272	98.9088	98.9088
2	99.6688	99.6348	99.6155
3	99.7619	99.6610	99.7619
4	98.9391	98.9058	98.9391
5	96.9059	96.8722	96.8722
6	99.2258	96.9335	96.7130
7	99.7917	99.5833	99.7917
8	99.7982	98.8356	98.8356
9	99.7699	97.6041	97.2409

<Table 1> Experimentation results for text images

(그림 11) 기울기 보정 평균 정확도



(Figure 11) Average accuracy of slope compensation

## 5. 결론

본 연구는 문자열과 배경이 혼합된 장면에서

문자열을 인식하기 위한 기울기 보정을 제안하였다. 기존의 경계선 투영 히스토그램을 이용한 기울기 보정 방법 coarse-to-fine 알고리즘과 Ishitani 알고리즘은 문서 영상에서 빠르고 정확도가 높은 기울기 검출이 가능하다 하지만 문자열과 배경이 혼합된 장면에서는 상대적으로 낮은 정확도를 보여주고 있다. 본 연구에서 제안한 방법은 영상을 분할하고 무의미한 영역을 배제한 후 유의미한 영역만을 이용하여 0.7% 기울기를 추출, 정확도를 높일 수 있었다.

### References

[1] P, Malathi, "Skew Detection based on Bounding Edge Approximation", IOSR Journal of Computer Engineering, Volume 16, pp136-139, 2014.

[2] Brodić, Drako, "The evaluation of the initial skew rate for printed text", Journal of Electrical Engineering, 62.3, pp134-140, 2011.

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

[4] Ju, Jae-hyon, et al., "Skew Correction of Document Images using Edge.", Journal of the Korea Institute of Information and Communication Engineering, 16, 7, pp.1487-1494, 2012.

[5] Upadhyay, Nishchal Gyan, and Kamlesh Lakhwani, "Edge Detection Using Fuzzy Approach Involving Automatic Threshold Generation.", International Journal Of Scientific & Techonology Research Vol. 2, Iss. 7, pp.128-131, July 2013.

[6] Canny, John, "A computational approach to edge detection." IEEE Trans. Pattern Anal. Mach. Intell., vol. 8, pp.679-698, 1986.

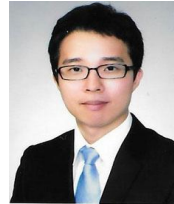
[7] Fang, Mei, et al., "The study on an application of otsu Operator in canny operator." International Symposium on Information Processing (ISIP). 2009.

[8] Lee, Chang-Young, et al., "A Study on Mask-based Edge Detection Algorithm using Morphology.", Jour

nal of the Korea Institute of Information and Communication Engineering, 19.10, pp.2441-2449, 2015.

[9] Woo, Chong-Ho, "Image Preprocessing in Container Identifier Recognition System Using Multiple Threshold Regions.", Journal of Korea Multimedia Society, 16.5, pp.549-557, 2013.

### 백재경



2010년 : 호남대학교 게임애니메이션학과 학사  
 2015년 : 경상대학교 교육대학원 컴퓨터교육과 재학

### 서영건



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