

# 자연영상에서 교통 표지판의 기울기 보정 및 텍스트 추출

## Skew Compensation and Text Extraction of The Traffic Sign in Natural Scenes

최 규 담\*  
(Gyu-Dam, Choi)

김 성 동\*\*  
(Sung-Dong, Kim)

최 기 호\*\*\*  
(Ki-Ho, Choi)

### 요 약

본 논문은 자연영상에서 얻은 교통표지판의 기울기를 보정하고 텍스트를 추출하는 방법을 제안한다. 본 연구는 명도 이미지를 대상으로 모든 과정이 4단계로 이루어진다. 첫째, 자연 영상에서 에지 검출을 위한 전처리 및 Canny 에지 추출을 수행하며, 둘째, 영상의 기울기를 추출하기 위해 허프 변환에 대한 전처리와 후처리를 한 후, 셋째로 잡음영상과 선을 제거하고 텍스트가 가지고 있는 특징을 이용하여 후보영역 검출을 한다. 마지막으로 검출된 텍스트 후보영역 안에서 지역적 이진화를 수행한 후, 불필요한 비텍스트 연결 요소를 추려내기 위해 텍스트와 비텍스트 간의 연결요소에 나타나는 특징 차이를 이용하여 텍스트 추출을 수행한다. 100장의 샘플영상을 대상으로 실험한 결과 82.54% 텍스트 추출률과 79.69% 추출 정확도를 가짐으로써 기존의 런 길이 평활화 방법이나 푸리에 변환을 이용한 방법보다 더 정확한 텍스트 추출 향상을 보였다. 또한 기울어진 각도 추출에서도 94.3%의 추출률로 기존의 Hough 변환만을 이용한 방법보다 약 26%의 향상을 보였다. 본 연구는 시각 장애인 보행 보조 시스템이나 무인 자동차 운행에 있어 위치 정보를 제공하는데 활용할 수 있을 것이다.

### Abstract

This paper shows how to compensate the skew from the traffic sign included in the natural image and extract the text. The research deals with the process related to the gray image. All the process comprises four steps. In the first part we perform the preprocessing and Canny edge extraction for the edge in the natural image. In the second part we perform preprocessing and postprocessing for Hough Transform in order to extract the skewed angle. In the third part we remove the noise images and the complex lines, and then extract the candidate region using the features of the text. In the last part after performing the local binarization in the extracted candidate region, we demonstrate the text extraction by using the differences of the features which appeared between the text and the non-text in order to select the unnecessary non-text. After carrying out an experiment with the natural image of 100 pieces that includes the traffic sign. The research indicates a 82.54 percent extraction of the text and a 79.69 percent accuracy of the extraction, and this improved more accurate text extraction in comparison with the existing works such as the method using RLS(Run Length Smoothing) or Fourier Transform. Also this research shows a 94.5 percent extraction in respect of the extraction on the skewed angle. That improved a 26 percent, compared with the way used only

\* 회원 : 광운대학교 컴퓨터공학과

\*\* 회원 : 계원조형예술대학 정보통신과 교수

\*\*\* 회원 : 광운대학교 컴퓨터공학과 교수

† 논문접수일 : 2004년 7월 14일

† 이 논문은 2003년도 광운대학교 교내 학술연구비 지원에 의해 연구되었음.

Hough Transform. The research is applied to giving the information of the location regarding the walking aid system for the blind or the operation of a driverless vehicle

**Key Words** : 교통 표지판, 기울기 보정, 텍스트 추출

## 1. 서 론

최근 컴퓨터의 발전과 인간의 시각기관과 유사하게 컴퓨터와 로봇에 부착된 카메라를 통해 인위적으로 삽입하거나 자연적으로 포함되어진 정보들을 적절히 분석하고 이해하고자 하는 많은 연구가 진행되었다. 이러한 정보들은 여러 비전 응용 시스템에 중요한 역할을 하고 있으며 이렇게 수많은 정보를 담고 있는 지능 시스템으로써 꼭 획득해야 하는 것이 바로 텍스트 정보이다.

특히, 교통 표지판의 텍스트 정보는 실시간으로 추출 및 인식 할 수 있다면 어디를 찾아 가고자 하는 위치정보를 얻음으로써 시각 장애인 보행 보조 시스템, 자율 이동 로봇 운행, 무인 자동차 운행 등 다양한 분야에서 널리 활용될 수 있다. 하지만 이런 텍스트를 포함한 영상은 그림 형태로 구성되어 있기 때문에 당장 직접 활용에는 많은 제약점들이 따르게 된다[1-2].

텍스트를 포함한 자연 영상에서 텍스트 추출을 위한 연구는 이제껏 많은 방법들이 제안되고 있다 [3-11]. 기존의 연구로써 대표적으로 영역화 알고리즘 RLSA(run-length smearing algorithm)과 RXYC(recursive X-Y cuts)가 있다.

RLSA는 1974년 Jahnson[3]에 의해 제안되었고, 그래픽 영상으로부터 텍스트 영역을 구별하기 위해 1982년 Wong[4]에 의해 확장되었다. 이 방법은 텍스트 영역을 찾는 연결 성분 기반으로 근접 픽셀을 연결하여 이진 영상을 형성하고 이들을 수직으로 결합하여 그래픽과 텍스트 영역을 구분한다. 그리고 RXYC는 1984년 Nagy[5]에 의해 제안된 알고리즘으로 수직과 수평으로 영상을 투사하여 깊은 골에 해당하는 부분을 잘라주는 방법으로 대개 문서에서 텍스트와 그래픽을 분리하는데 적용된다.

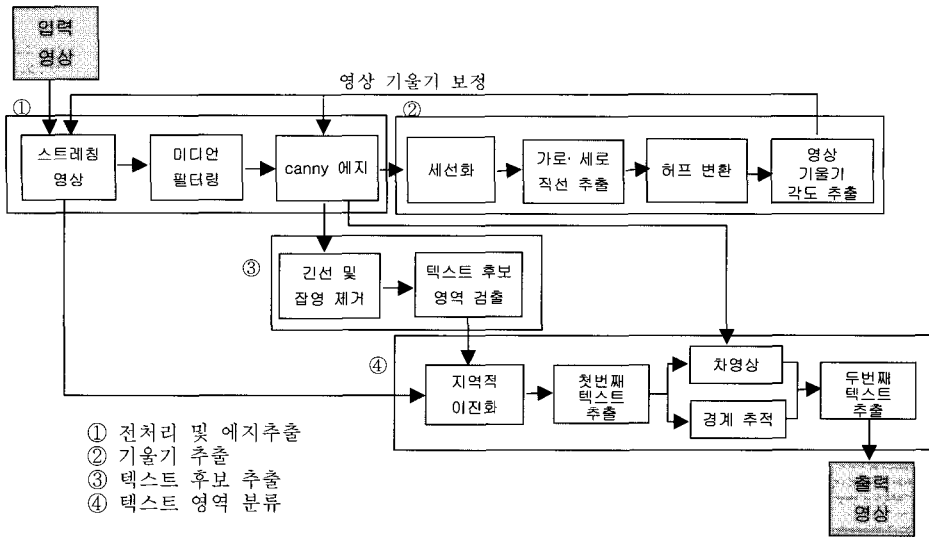
최근에는 복잡한 배경과 여러 종류의 잡영들이 포함되어 있는 자연 영상에서 텍스트 추출에 관한

연구가 활발히 진행되고 있다. Hoya[6]에 의해 제안된 방법은 지역적 이치화를 수행 후 8x8 sub block을 통해 인접 영역의 연결요소의 특징을 이용하여 텍스트 영역을 추출하였고, Bong-Kee Sin[7]는 고속 푸리에 변환(Fast Fourier Transform:FFT) 및 자기 상관계수법(auto correlation)을 이용하여 텍스트 영역을 검출하였다. 그 밖에도 포폴로지를 이용한 방법[8], 투영 윤곽 분석 방법[9] 등 여러 연구들이 제안 되었다. 하지만 이러한 기존 연구들은 단순한 영상 또는 문서에서 텍스트를 추출 하는데 유용하고, 복잡한 자연영상에서 텍스트를 추출하는데 있어 많은 어려움이 있다.

또한 기울어진 영상에 대한 텍스트는 문자인식 과정에서 인식률이 저하시키고 텍스트 영역을 추출하기 위해서도 기울기를 보정해 줄 필요가 있다. 기울기 추출을 위한 기존 연구로써 대표적으로 허프 변환(Hough transform)[10], 투영 윤곽(projection profile) 분석에 의한 방법[11] 등을 사용하고 있으나 이는 단순한 직선과 규칙적인 패턴을 가지고 있는 문서에서 기울기 추출이 가능하고 복잡한 배경을 가진 자연 영상에서는 기울기 각도를 추출하기가 어렵다.

본 논문에서는 위에서 살펴 본 기존의 연구들의 단점들을 보완하여 자연영상에서 자연스럽게 얻어진 교통 표지판의 텍스트를 추출하는 방법을 제안한다. 이를 위해 우선 텍스트 후보영역을 검출하고 검출된 후보 영역 안의 텍스트와 비텍스트들의 특징을 비교하여 텍스트만을 추출하는 방법을 제안하고 기울어진 영상에 대해서는 허프 변환을 위한 전처리와 후처리 알고리즘을 제안하여 기울기 각도를 구한다.

본 논문의 구성은 1장 서론에 이어 2장에서는 자연영상에서 교통 표지판의 텍스트를 추출하는 방법과 기울어진 영상에 대해 기울기를 추출하고 보정하는 알고리즘을 제안하며, 3장에서는 실험결과 및 고찰을 하고, 4장에서 결론을 맺는다.



<그림 1> 교통 표지판에서의 텍스트 추출 과정

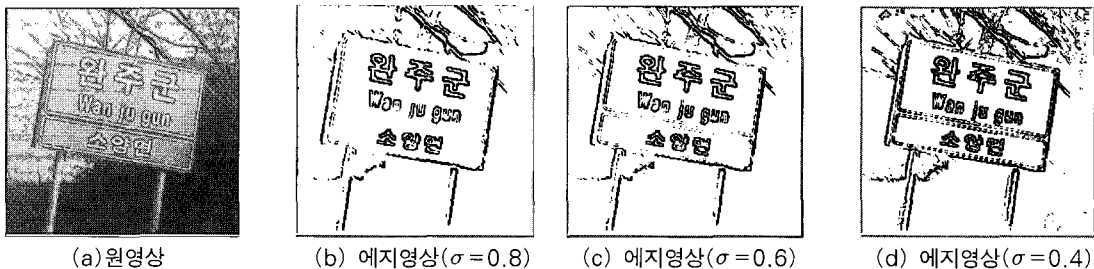
## II. 자연 영상에서 교통 표지판의 텍스트 추출 설계

본 논문에서 제안하는 텍스트 추출과정은 네 단계를 거친다. 첫 번째는 전처리와 에지 검출을 수행하고, 두 번째는 영상의 기울기 추출 및 보정단계를 세 번째는 텍스트 후보영역화(segmentation) 단계를 거치고 마지막으로 텍스트 영역 분류(Block classification) 단계로 비텍스트와 텍스트를 분류해낸다. <그림 1>은 이 과정을 요약하여 나타낸 그림이다.

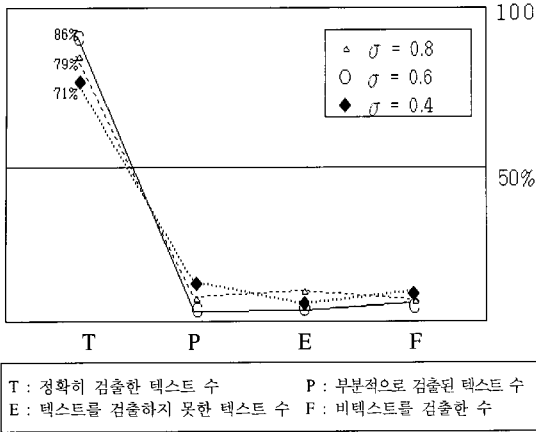
### 1. 전처리 및 에지 추출

자연영상에서 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 부각시키기 위해 어떤 특정부분에 집중되어 있는 화소값을 명암대비가 골고루 분포 되도록 화소 밝기를 스트레칭(Stretching)시키고 잡영의 일부를 제거하기 위해 3×3 마스크 미디언(Median)필터를 적용한다. 다음 단계로 에지 수행을 하는데 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 본 연구에서 다섯 장의 샘플로 테스트 한 결과 prewitt 마스크, 가우시안의 표준편차  $\sigma=0.6$ , 이



<그림 2>  $\sigma$ 에 따른 Canny 마스크의 프리윗 에지 영상

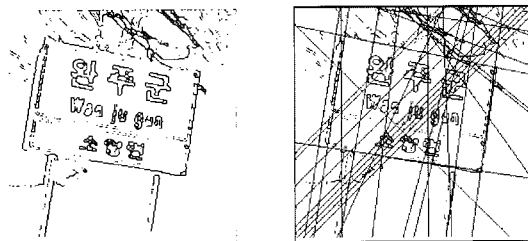


<그림 3>  $\sigma$ 에 따른 텍스트 검출 그래프

중 경계값, high값=70, low값=40을 이용한 Canny 에지 검지기를 사용하였을 때, <그림 2>의 (c)와 같이 텍스트를 검출을 하기 위한 가장 좋은 영상을 얻는다. <그림 2>, <그림 3>은  $\sigma$ 값 변화에 따른 에지 변화를 나타낸 그림과 그래프를 나타낸다.

## 2. 텍스트 기울기 추출 및 보정

기울어진 영상에 대해 기울기를 추출 하는데 가장 많이 사용하는 것이 허프변환(Hough Transform)이다. 그러나 기존의 허프 변환만을 단지 수행하면 <그림 4>의 (b)와 같이 너무나 많은 기울기를 가지고 있는 직선이 추출되어 정확한 기울기 각도를 추출하는데 어렵다. 따라서 본 논문에서는 허프변환을 수행하기 위한 전처리로서 가로·세로  $-45^\circ \sim +45^\circ$ 의 직선을 추출하는 5x5 윈도우와 허프변환 후처리로 기울기 각도 추출 방법을 제안한다.

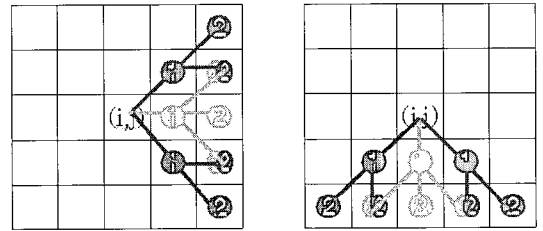


(a) 세선화    (b) 허프변환

<그림 4> 허프변환을 이용한 기울기 추출

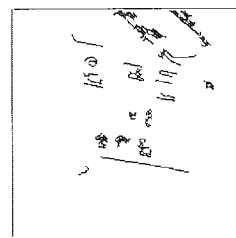
우선 허프변환의 전처리로서 위에서 얻은 에지 추출 영상에 세선화를 수행하여 <그림 4>의 (a)와 같은 영상을 얻는다. 다음으로 세선화 영상에 각 가로·세로  $-45^\circ \sim +45^\circ$ 의 직선을 추출하는 <그림 5>의 5x5 윈도우를 이용하여 (i,j)에 있는 후보소에 대해 색상이 같은 ①과 ②가 연속되어 있는 연결요소만 남기고 나머지 화소는 제거함으로써 <그림 6>의 (a),(b)와 같은 가로·세로 직선을 얻을 수 있다. 이렇게 얻은 각 직선에 허프변환을 수행하여 <그림 6>의 (c),(d)와 같이 각 가로·세로 직선에 대한 기울기들을 추출한다.

추출된 각 기울기의 빈도수를 가지고 다음과 같은 과정을 거쳐 기울어진 영상의 각도를 구하게 된다.

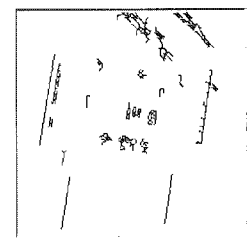


(a) 가로 추출 윈도우    (b) 세로 추출 윈도우

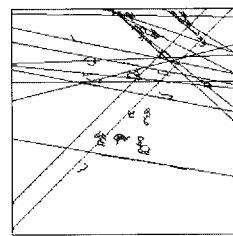
<그림 5> 가로·세로 직선 추출 윈도우



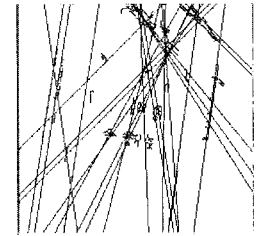
(a) 가로 직선 추출



(b) 세로 직선 추출

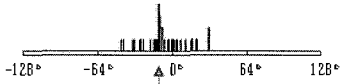


(c) 가로 각도 추출(허프변환)



(d) 세로 각도 추출(허프변환)

<그림 6> 기울기 추출 과정



〈그림 7〉 가로·세로 기울기 누적 히스토그램



〈그림 8〉 기울기 보정 예

- (1) 세로 각도의 빈도수와 가로 각도의 빈도수를 가장 많은 것부터 적은 것으로 차례로 서로 비교해 가면서 차가 90도를 가지면 그때 가로 각도는 영상의 기울기가 된다. (단 빈도수가 1인 것은 연산에서 제외시킴)
- (2) (1)에서 기울기 각도를 찾지 못하면 <그림 7>과 같이 가로·세로 기울기의 전체 빈도수 히스토그램을 구해 첨두값을 영상의 기울기를 정한다.

위의 과정을 통해 추출된 영상의 각도는 양선형 보간(bilinear interpolation) 방법을 사용하여 기울기 보정을 한다. 양선형 보간은 현재위치(실수값) 근처의 이웃하는 네 점(실수값)의 밝기값을 이용하여 현재 위치와 각 지점과의 거리에 따라 반비례하여 가중치를 달리하는 기법이다. 앞에서 수행한 스트레칭 영상과 Canny(Prewitt) 영상에 식(1)을 이용하여 기울어진 영상을 바로 보정함으로써 <그림 8>과 같은 결과를 얻을 수 있다.

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x+c_x \\ y+c_y \end{pmatrix} + \begin{pmatrix} c_x \\ c_y \end{pmatrix} \quad (1)$$

- $(x, y)$  : 회전 전의 원 영상 좌표값
- $(x', y')$  : 회전 후의 새로운 영상 좌표값
- $(c_x, c_y)$  : 원 영상 회전 중심의 좌표값
- $\theta$  : 회전각

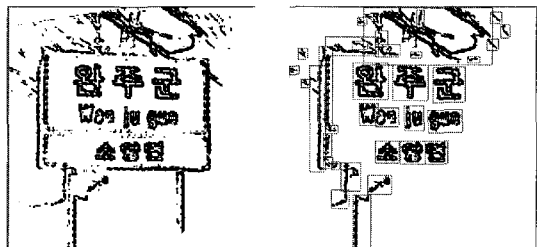
### 3. 텍스트 후보 영역화

기울기가 보정된 에지영상의 텍스트 주변에는 텍스트 후보영역을 추출하는데 불필요한 잡영과

필요없는 긴선 들이 많이 발생하게 된다. 작은 잡영의 경우 비록 이것이 텍스트라고 하더라도 인식이 어렵기 때문에 제거해 줄 필요가 있다. 이런 잡영 및 복잡한 선들을 제거하기 위해 라벨링(Labeling)의 알고리즘을 사용하는데 라벨링을 수행할 때 연결요소의 수가 매우 작거나 연결요소를 둘러싸고 있는 블록의 넓이보다 5분에 1보다 작으면 제거한다. 텍스트 후보 영역(블록)은 에지 영상을 조사하는 과정에서 연결요소가 처음 발견된 위치  $Xstart = Xc, Ystart = Yc, Xend = Xc + \Delta X, Yend = Yc + \Delta Y$  ( $\Delta X$  : 블록의 길이,  $\Delta Y$  : 블록의 높이)로 표현된다. 계속적으로 연결요소를 좌->우, 위->아래로 조사하면서 같은 연결성의 화소의 블록을 찾아 낼 수 있다. 각 연결요소의 블록을 찾는 식은 다음과 같다.

$$\begin{aligned} Xstart &= \min(Xmin, Xc) \\ Ystart &= \min(Ymin, Yc) \\ \Delta X &= \max(Ymax, Xc) \\ \Delta Y &= \max(Ymax, Yc) \end{aligned} \quad (2)$$

비교 대상 블록 :  $(Xmin, Ymin, Ymax, Ymax)$



〈그림 9〉 텍스트 후보 영역화

<그림 9>는 잡영 제거와 함께 라벨링을 수행한

후 텍스트 후보 영역을 찾은 결과를 보여준다.

#### 4. 텍스트 분류

본 논문에서 텍스트와 비텍스트를 분류하기 위해 우선 스트레칭 된 영상에 각 후보영역의 위치 정보를 가지고 있는 블록을 매칭시켜 각 영역별로 지역적 이진화 작업을 수행한다. 이렇게 이진화 된 영역은 두 번의 텍스트 추출 작업을 수행 한다. 첫 번째는 텍스트 영역은 후보 영역의 블록선 상에 흑화소가 닿지 않는다는 특징을 이용하고 두 번째는 스트레칭 된 영상에 Canny(Sobel) 에지를 수행한 영상, 첫번째 텍스트 추출을 거친 텍스트 영상과 텍스트 경계추적을 하여 얻은 영상을 연산 조합하여 텍스트와 비텍스트를 분류한다.

##### 1) 연결요소의 블록 특징을 이용한 텍스트 추출

이진화 된 후보 텍스트 영역은 <그림 10>과 같이 텍스트는 블록 선 안에 존재하지만 비텍스트인 경우는 선에 존재하는 경우가 있다. 따라서 각 블록 안에 라벨링을 위해 연결요소들을 조사하여 선에 붙어있는 연결요소는 식(3)를 이용하여 삭제함으로써 <그림 10>의 오른쪽 영상과 같은 결과를 얻을 수 있다.

$$BC_k(i, j) = \begin{cases} 255 & \text{if } BC_k^{count} > 2 \\ B(i, j) & \text{else} \end{cases} \quad (3)$$

$BC_k$  : 텍스트 후보 블록 안의 k번째 연결요소

$B$  : 텍스트 후보 블록

$BC_k^{count}$  : 텍스트 후보 블록 안의 k번째 연결요소가 블록선에 닿은 화소 수

##### 2) 영상의 차를 이용한 텍스트 추출

첫 번째 텍스트 추출을 수행 후 나머지 비텍스트를 없애기 위해 영상의 차를 이용한 두 번째 텍스트 추출 알고리즘을 수행한다. 텍스트 추출을 위해 <그림 2>의 Canny(Prewitt) 대신 Canny(Sobel) 연산을 수행한 에지추출 영상과 위에서 블록 특징을 이용하여 텍스트를 추출 한 영상 및 각 연결요소의 경계 추적을 수행한 영상을 이용한다. 두 번째 텍스트 추출 방법은 다음과 같은 과정을 거쳐 수행한다.

step-1 : 식(4)를 이용하여 Canny(Sobel) 에지 영상과 첫 번째 텍스트 추출 영상의 차를 연산하여 <그림 12>의 (a)와 같은 영상을 얻는다.

$$DT(i, j) = CS(i, j) - T1(i, j) \quad (4)$$

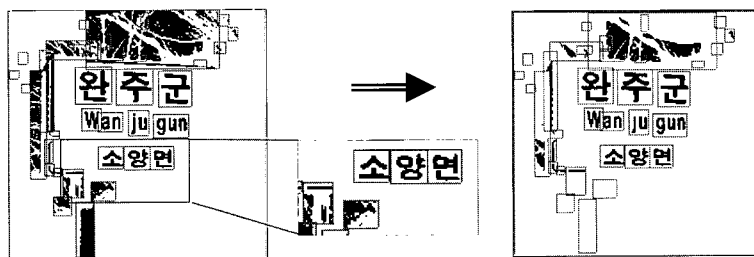
if ( $T1(i, j) < 0$ )  $DT(i, j) = 0$

$CS$  : Canny-소벨 에지 영상

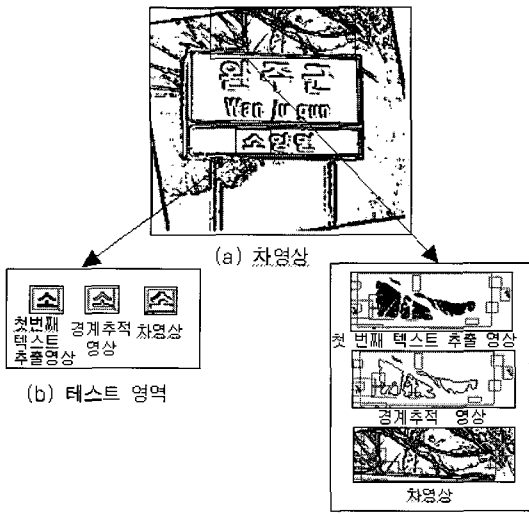
$T1$  : 첫 번째 텍스트 추출 후 영상

(가정 : 1 - 흑화소, 0 - 백화소)

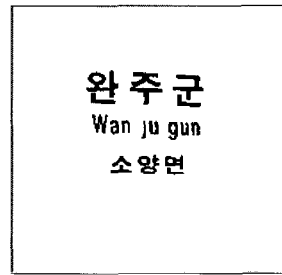
step-2 : <그림 11>의 (b)와 <그림 11>의 (c)에서 보면 비텍스트 영역은 차영상에서 에지가 많이 분포되어 있는 것을 볼 수 있다. 텍스트는 그래디언트 변화가 적어 에지가 적게 발생하는 반면 수풀이나 나뭇가지와 같은 비텍스트의 경우 복잡한 그래디언트 때문에 에지가 많이 발생하게 된다. 이러한 특징을 이용하여 식(5)에서 임계값(T)에 <그림 12>와 같이 임계값의 변화에 따른 실험 결과로 2를 대입시켜 식을 만족하는 블록을 삭제함으로



<그림 10> 블록의 특징을 이용한 텍스트 추출



<그림 11> 차영상에 따른 텍스트와 비 텍스트 비교



<그림 13> 최종 텍스트 추출 결과

### III. 실험 결과 및 고찰

#### 1. 실험 환경

제안한 알고리즘을 이용하기 위해 하드웨어 : Pentium IV 2.4GHz PC, 운영체제 : Windows XP, 컴파일러 : Microsoft Visual C++.NET, 영상 입력 장치 : 캐논 A80 카메라를 사용하였다.

#### 2. 실험 결과

본 논문에서는 자연영상에서 얻은 교통 표지판의 텍스트를 추출하기 위해 카메라로 명도로 입력받아 여러 필터링과 연산을 통해 기울어진 영상은 보정하고, 텍스트와 상관없는 비텍스트는 삭제함으로써 순수 텍스트만을 추출 하였다.

실험은 거리에 상관없이 교통 표지판을 포함한 자연 영상 100장을 가지고 실험을 하였다. 단, 실험자의 눈으로 인식 가능한 크기의 대상으로 단순한 영상과 복잡한 영상들을 크기에 상관없이 실험에 사용하였다.

텍스트 추출에 대한 실험 결과는 <표 1>과 같으며, Total은 전체 텍스트의 총개수, True는 정확히 찾은 텍스트 개수, Part는 텍스트의 일부분만을

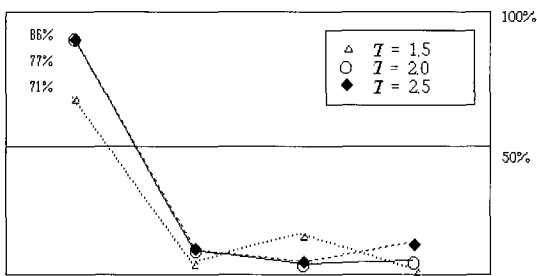
<표 1> 텍스트 추출 결과

텍스트 : 문자 1개로 설정

	Total	True	Part	Error	False
텍스트 개수	1203	993	95	115	158
결과	Recall : 82.54 Precision : 79.69				

$$T2B(i, j) = \begin{cases} 255 & \text{if } DTB_{cnt} > DBB_{cnt} \times T \\ T1B(i, j) & \text{else} \end{cases} \quad (5)$$

- T2B : 영상의 차를 이용한 텍스트 추출 후 텍스트 후보 블록
- T1B : 첫번째 텍스트 추출 후 텍스트 후보 블록
- DTB<sub>cnt</sub> : 차영상 텍스트 후보 블록의 흑화소 수
- DBB<sub>cnt</sub> : 경계추적영상 텍스트 후보 블록의 흑화소 수



T : 정확히 검출한 텍스트 수      P : 부분적으로 검출된 텍스트 수  
 E : 텍스트를 검출하지 못한 텍스트 수      F : 비텍스트를 검출한 수

<그림 12> 임계값 T에 따른 텍스트 검출 그래프

찾은 개수, Error은 텍스트를 찾지 못한 개수이며 False는 텍스트가 아닌 비텍스트를 찾은 개수이다. 그리고 Recall(텍스트 추출률)과 Precision(정확도)은 다음과 같은 식으로 계산하였다.

$$Recall = \frac{True}{Total} \times 100 \quad (5)$$

$$Precision = \frac{True}{True + Part + False} \times 100 \quad (6)$$

또한, 기울기 추출에 대해서는 위와 같은 샘플 이미지로 실험한 결과 <표 2>와 같이 나타났다.

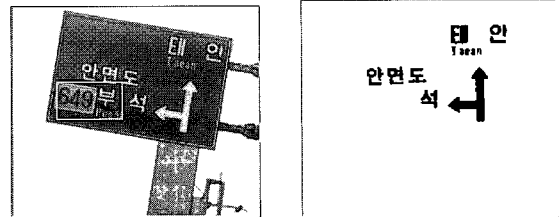
<표 2> 기울기 각도 추출 결과

	허프 변환만 실행시	제안한 방법
정확도	68%	94%

에지를 세선화 한 영상에 대해 바로 허프 변환을 수행했을 때 보다 전처리로 3x3 윈도우 마스크를 사용하여 가로와 세로 직선을 각각 추출하고 허프변환 후처리로써 앞에서 제안한 추출 단계를 사용하였을 때 더 정확한 각도를 추출하는 것을 볼 수 있다.

### 3. 고찰

기존 연구를 살펴보면 텍스트와 비텍스트를 분리하기 위해 밀도 또는 연결요소를 이용한 방법, 런길이 평활화 방법, 모폴로지 방법 등을 이용하였으나 이는 텍스트 후보 영역이 문자열이거나 하나의 문자 영역을 가지고 있어야 한다. 그러나 텍스트 크기 및 텍스트와 텍스트 간격이 일정하지 않은 교통 표지판에서 그 텍스트 영역을 추출 한다는 것은 매우 어렵고, 추출 한다고 해도 텍스트와 비슷한 나뭇가지나 벽돌같은 비텍스트를 정확하게 분리 할 수가 없었다. 따라서 본 논문에서는 2장 4절에서 제시한 블록과 차영상을 이용한 두 가지 텍스트 추출 과정을 통해 <표 1>과 같이 텍스트 추출률과 정확도를 높일 수 있었다. 그러나 <그림



<그림 14> 정확히 추출 하지 못한 텍스트 예

14>와 같이 텍스트와 비텍스트가 너무 많이 붙어 있거나 또는 텍스트와 배경사이의 명암 대비의 차가 현저히 적을 때 텍스트를 추출하지 못하는 경향을 보였다.

기울기 추출에 대해서는 전체 이미지에 대해서 94%의 추출률을 보였으며, 기존 연구에서 단지 허프변환을 사용한 방법보다 기울기 추출률의 정확성을 현저하게 높일 수 있었다.

## IV. 결 론

본 논문은 자연영상에 포함되어 있는 교통 표지판의 텍스트가 사람들의 눈에 잘 띄게 하기 위해 텍스트와 주변 배경의 명도의 대비차가 크다는 특징을 착안하여 에지를 이용한 텍스트 추출을 하였다. 또한 기울어진 영상에 대해서는 기울기 각도를 추출하여 보정하는 방법을 연구를 하였다. 연구의 전체적인 과정은 첫째, 자연 영상에서 에지 검출 처리를 위한 전처리 및 Canny 에지 추출을 수행하며, 둘째, 영상의 기울기를 추출 과정을 셋째로 잡음영상과 선을 제거하고 텍스트 특징을 이용한 후보영역 검출을 하였다. 마지막으로 텍스트 후보영역 안에서 지역적 이진화를 수행하여 불필요한 비텍스트 연결 요소를 추려내기 위해 텍스트와 비텍스트 간의 연결요소에 나타나는 특징 차이를 이용하여 텍스트 추출을 수행하였다.

기울기를 추출하기 위해서 5x5 윈도우를 사용하여 가로, 세로 -45° ~ +45°의 직선을 추출 후, Hough 변환과 본 논문에서 제시한 각도 추출 단계를 수행함으로써 단지 Hough 변환만을 사용한 것에 비해 향상 된 것을 보였다.



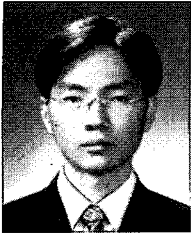
100장의 교통 표지판을 포함한 자연영상을 대상으로 실험한 결과 82.54%의 텍스트 추출률과 79.69%의 정확도를 가졌으며, 기울어진 각도 추출에서도 94.3%로 단지 Hough 변환을 이용한 방법보다 26%가량 높은 추출률을 보였다.

본 연구는 시각 장애인 보행이나 무인 자동차 운행에 있어 매우 중요한 위치 정보를 제공하는데 활용 할 수 있을 것이며, 앞으로 원근감이 있는 영상에 대하여 원근 정보를 이용한 보정과 텍스트 후보영역에서 텍스트를 보다 더 정확하게 추출하기 위한 연구가 필요하다.

### 참 고 문 헌

- [1] X. Huang, A. Acero, H. W. Hon, "Spoken language processing", Prentice Hall PTR, New Jersey, p.1-5, 2001.
- [2] T.F. Quatieri, "Discrete-time speech signal processing", Prentice Hall, New Jersey, p.1-5, 2002.
- [3] 이호영, "국어음성학", 태학사, 서울, chap. 3-5, 1996.
- [4] S. Young, D. Kershaw, J. Odell, D. Ollason, Valtcher, P. Woodland, "The HTK Book (for HTK Ver.3.2)", Cambridge University Engineering Department, 2002.
- [5] L.R. Rabiner, B.H. Juang, "Fundamentals of speech recognition", Prentice Hall, New Jersey, chap. 6, 1993.
- [6] L.R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," Proceedings of the IEEE , Volume: 77 Issue: 2 , pp. 257-286, Feb. 1989.
- [7] B.H. Juang, S. E. Levinson, and M. M. Sondt, "Maximum likelihood estimation for multivariate mixture observations of Markov chains," IEEE Trans. Information Theory, IT-32(2), pp. 307-309, March 1986.
- [8] S. Young, N.H. Russell, "Token passing: A simple conceptual model for continuous speech recognition systems", Cambridge University Engineering Dept., Technical Report No. 38, 1989.
- [9] H. Ney, D. Mergel, A. Noll, A. Paeseler, "A data-driven organization of the dynamic programming beam search for continuous speech recognition", Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '87., Volume: 12 pp. 833-836, Apr. 1987.
- [10] S. Takahashi. S. Sagayama, "Four-level tied-structure for efficient representation of acoustic modeling", ICASSP-95, International Conference on , Vol.: 1 , pp. 520-523, May 1995.
- [11] K.F. Lee, "Context-dependent phonetic hidden Markov models for speaker-independent continuous speech recognition", Acoustics, Speech, and Signal Processing, IEEE Transactions on, Volume: 38 Issue: 4 pp. 599-609, Apr. 1990.

〈저자소개〉



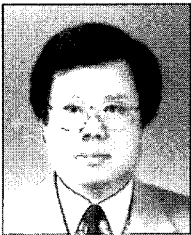
최 규 담 (Gyu-Dam, Choi)

2002년 2월 : 남서울대학교 전산학과 이학사  
2004년 8월 : 광운대학교 대학원 컴퓨터공학과 공학석사  
2004년 9월 ~ 현재 : 광운대학교 공학교육연구소 공학인증 조교  
관심분야 : image processing, network



김 성 동 (Sung-Dong, Kim)

1981년 2월 : 광운대학교 컴퓨터공학과 공학사  
1992년 9월 : 미국 CUNY 대학원 컴퓨터과학과 석사  
2001년 8월 : 광운대학교 대학원 공학박사  
1984년 3월 ~ 1987년 9월 : 한국컴퓨터 연구소 선임연구원  
2002년 2월 ~ 2003년 1월 : 미국 University of Notre Dame. Visiting Research scientist  
1995년 3월 ~ 현재 : 계원조형예술대학 정보통신과 교수  
관심분야 : object tracking, 3D face recognition, 3D simulation game



최 기 호 (Ki-Ho, Choi)

1973년 2월 : 한양대학교 전자공학과 공학사  
1977년 2월 : 한양대학교 대학원 전자공학과 공학석사  
1987년 2월 : 한양대학교 대학원 전자공학과 공학박사  
1977년 3월 ~ 1979년 2월 : 한국과학기술연구소(KIST) 전자공학부 연구원  
1979년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 교수  
1989년 2월 ~ 1990년 2월 : Univ. of Michigan 전기 및 전산과 객원교수  
2001년 2월 ~ 2002년 1월 : San Jose State Univ. 컴퓨터공학과 객원교수  
2001년 1월 ~ 2003년 12월 : ISO/IEC JTC1/SC28 국내 전문위원  
2002년 1월 ~ 현재 : 한국멀티미디어학회 부회장 (수석부회장)  
2002년 6월 ~ 현재 : 한국 ITS 학회 부회장(회장)  
2003년 2월 ~ 현재 : 광운대학교 전자정보대학 학장  
2004년 5월 ~ 현재 : 광운대학교 정보통신대학원 원장  
관심분야 : 영상처리, 내용기반검색