



모바일 이미지 기반의 문자인식 시스템

The Character Recognition System of Mobile Camera Based Image

| | |
|--------------------|---|
| 저자 (Authors) | 박영현, 이형진, 백중환 Park, Young-Hyun, Lee Hyoung-Jin, Baek, Joong-Hwan |
| 출처 (Source) | 한국산학기술학회 논문지 11(5) , 2010.05, 1677-1684 (8 pages) |
| 발행처 (Publisher) | 한국산학기술학회 Korea Academy Industrial Cooperation Society |
| URL | http://www.dbpia.co.kr/Article/NODE07210132 |
| APA Style | 박영현, 이형진, 백중환 (2010). 모바일 이미지 기반의 문자인식 시스템. 한국산학기술학회 논문지, 11(5), 1677-1684. |
| 이용정보 (Accessed) | 전문대학교 210.119.34.*** 2018/04/18 16:12 (KST) |

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

모바일 이미지 기반의 문자인식 시스템

박영현¹, 이형진^{1*}, 백중환¹

¹한국항공대학교 정보통신공학과

The Character Recognition System of Mobile Camera Based Image

Young-Hyun Park¹, Hyung-jin Lee^{1*} and Joong-Hwan Baek¹

¹Information and Telecommunication, Korea Aerospace University

요약 최근 모바일 폰의 발달과 스마트 폰의 보급으로 인해서 많은 콘텐츠들이 개발되어지고 있다. 특히, 모바일 휴대장치에 소형 카메라가 탑재되면서부터 카메라로부터 입력되어지는 영상 기반 콘텐츠 개발은 사람들의 흥미뿐만 아니라 활용 면에서도 중요한 부분을 차지하고 있다. 그중 문자인식 시스템은 시각 장애인 보행 보조 시스템, 로봇 자동 주행 시스템, 비디오 자동 검색 및 색인 시스템, 텍스트 자동 번역 시스템 등과 같은 활용영역에서 매우 광범위하게 쓰일 수 있다. 따라서 본 논문에서는 스마트 폰 카메라로 입력되는 자연 영상에 포함되어 있는 텍스트를 추출 및 인식하고 음성으로 출력해주는 시스템을 제안하였다. 텍스트 영역을 추출하기 위해 Adaboost 알고리즘을 이용하고 추출된 개별 텍스트 후보영역의 문자 인식에는 오류 역전파 신경망을 이용하였다.

Abstract Recently, due to the development of mobile phone and supply of smart phone, many contents have been developed. Especially, since the small-sized cameras are equipped in mobile devices, people are interested in the image based contents development, and it also becomes important part in their practical use. Among them, the character recognition system can be widely used in the applications such as blind people guidance systems, automatic robot navigation systems, automatic video retrieval and indexing systems, automatic text translation systems. Therefore, this paper proposes a system that is able to extract text area from the natural images captured by smart phone camera. The individual characters are recognized and result is output in voice. Text areas are extracted using Adaboost algorithm and individual characters are recognized using error back propagated neural network.

Key Words : Character Recognition, Adaboost, Neural Network

1. 서론

카메라로 취득한 이미지나 동영상에 인위적으로 삽입되거나 자연적으로 포함된 텍스트들은 이미지의 내용을 함축적이고 구체적으로 표현하는 중요한 정보들이다. 이러한 정보들을 실시간에 추출하여 인식한다면 시각장애인을 위한 보행안내시스템, 지능 로봇의 자동주행 등과 같은 분야에서 다양하게 활용될 수 있다. 텍스트 추출 연구는 그래픽(Graphic) 텍스트 추출과 장면(Scene)텍스트 추출 연구로 나누어 진행되고 있다. 그래픽 텍스트 추출

은 대부분 이미지의 질이 높은 문서 이미지나 비디오 프레임에 대상으로 영역을 추출하였지만, 최근에는 복잡한 배경을 갖는 자연 이미지에서 텍스트의 크기, 색, 폰트 종류, 방향 등이 다양하게 나타나는 장면텍스트를 추출하는 연구가 응용을 고려한 제한적인 환경에서 활발하게 진행되고 있다.

그래픽 텍스트 추출에 관한 연구로서 Jain 등은[1] 이진 이미지, 웹 이미지, 색 이미지 및 비디오 프레임의 네 가지 종류의 이미지에서 텍스트를 추출하는 방법을 제안하였다. 이 방법은 네 가지 종류에 대해 서로 다른 특징

본 연구는 경기도지역협력연구센터 (GRRC) 프로그램에 의해 한국항공대학교 차세대방송미디어기술 연구센터의 지원으로 수행되었음.

*교신저자 : 이형진(libae04@naver.com)

접수일 10년 04월 23일

수정일 10년 05월 11일

게재확정일 10년 05월 13일

과 임계값을 적용하기 때문에 이진 이미지, 웹 이미지, 비디오 프레임에서는 높은 추출률을 보이는 반면, 색이미지에 대해서는 낮은 추출률을 보인다. 또한 대부분의 실험 이미지가 비디오 프레임에 국한되어 있다.

Zhong 등은[2] 텍스트의 색은 일정하며, 명도이미지에서 텍스트 영역은 공간적 분산(Spatial Variance) 값이 크다는 특징을 이용하고, 두 방법을 순차적으로 결합한 방법을 제안하였다. 그러나 길이가 짧거나 색이 일정하지 않은 문자열, 수직 방향 또는 필기된 문자열 등에서 오류가 발생하는 단점이 있다.

H. K. Kim은[3] 비디오 프레임으로부터 자동으로 텍스트 영역을 추출하기 위해서 문자들이 수평 방향으로 놓여 있고, 균일한 색과 일정한 크기를 갖는다는 가정으로 색 연속성 특징을 이용한 방법을 제안하였다. 색의 대비가 크지 않은 텍스트와 크기가 작은 텍스트에 대해서는 정확히 찾지 못하며, 16개의 경험적 임계 값을 정해야 하는 단점이 있다.

P. K. Kim은[4] 색 이미지를 대상으로 텍스트의 색연속성 특징을 이용한 추출 방법을 제안하였다. 제안한 방법은 24비트(Bit) 색을 모두 이용하는 전역적 양자화의 단점과 텍스트와 배경 영역이 잘 분리되지 않고 시간이 많이 소요되는 단점을 보완하기 위해서 R, G, B 각각의 상위 2비트를 이용하여 총 6비트를 이용하는 지역적 양자화 방법을 적용하였다. 그러나 지역적 양자화의 처리 시간이 오래 걸리고 기울어진 텍스트 영역을 추출하기 어려운 단점이 있다.

장면텍스트 추출 연구로서 Ohya 등은[5] 실외 자연 이미지의 텍스트 영역을 추출하기 위해서 대부분의 텍스트는 명도 값이 일정하다는 특징을 이용하였다. 이 논문에서의 텍스트 영역은 기울어짐이 없고 배경과 뚜렷하게 구분된다는 가정을 전제하였다.

Gu 등은[6] 미분 탑-햇츠(Differential Top-hats) 연산자를 이용한 추출 방법을 제안하였으며, 문자획을 크기에 따라 8단계로 나누고 흰 문자와 검은 문자로 구분하여 연산자를 적용하였다. 제안한 알고리즘은 조명의 변화에 영향을 많이 받으며 처리 시간이 오래 걸리는 단점이 있다.

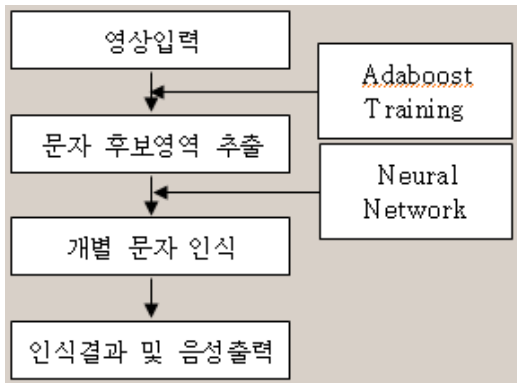
위에서 살펴본 것처럼 그래픽 텍스트를 추출하기 위한 대부분의 방법들은 색 이미지에서의 색 연속성 특징을 이용한 색 병합 방법의 시도와 명도이미지에서는 명도 값이 균일하다는 특징을 이용하기 위한 이진화 방법을 시도하였다. 장면 텍스트 추출 연구에서도 유사한 방법을 시도하였으나, 그래픽 텍스트 추출에 비해 색, 방향, 크기, 폰트, 기울어짐 등에 제약을 두었다. 본 논문에서는 장면 텍스트 추출에서의 제약을 어느 정도 해제하며 기존 연구들의 문제점들을 해결하기 위해서 AdaBoost 학습 알고리즘을 이용하여 특징점들을 선택하고 분류기(classifier)를 학습시켜 문자 후보영역을 추출하고, 개별 후보영역의 문자인식에 오류역전과 신경망을 이용하여 획득한 이미지의 문자영역 추출 및 인식을 하여 음성으로 출력해주는 시스템 그림 1 을 제안하였다.

본 논문의 구성은 다음과 같다. 2장에서는 Adaboost 알고리즘을 이용한 자연영상에서의 문자 후보영역 추출 방법을 제시한다. 3장에서는 추출된 개별 문자 후보영역 인식을 위한 오류역전과 신경망에 대하여 기술한다. 4장에서는 실험과 실험에 대한 결과를 분석한다. 5장에서는 실험에 대한 고찰 과정을 거치고 마지막으로 결론을 내린다.

2. 문자 후보영역 추출

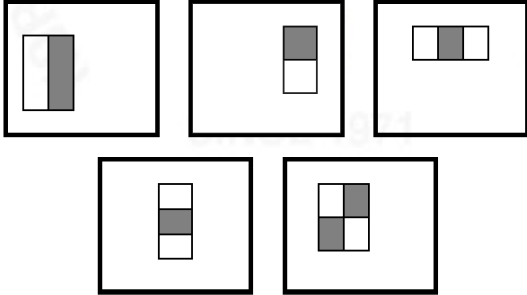
2.1 사각 특징점(Rectangle Feature)

카메라로부터 들어오는 입력 이미지에서 텍스트 후보영역을 정확하게 검출하기 위해서 제안하는 문자인식 시스템에서는 간단한 형태를 가지는 특징점 들의 값을 바탕으로 텍스트 유무를 판단한다. 이미지의 픽셀 값을 직접 이용하는 픽셀 기반의 시스템 보다 특징점 기반의 시스템이 알고리즘의 수행 속도면에서 훨씬 빠르다는 장점이 있다. 제안하는 문자인식 시스템에서는 다섯 가지 종류의 특징점들을 이용한다. 두개의 사각형으로 구성된 특징점의 값은 두개의 사각형 영역내의 밝기 값들의 합의 차이이다. 각각의 영역은 동일한 크기와 모양을 가진다. 세개의 사각형으로 구성된 특징점의 값은 바깥쪽 두 개의 사각형 영역내의 밝기 값의 합에서 가운데 사각형 영역의 밝기 값의 합을 뺀 값이다. 끝으로 네개의 사각형으



【그림 1】 전체 시스템 흐름도

로 구성된 특징점의 값은 대각선 방향의 두 쌍의 사각형 영역의 밝기 값의 합의 차이이다. 간단히 말해서 그림 2에서 보여지는 각각의 특징점의 값은 흰색 영역의 밝기 값의 합에서 회색 영역의 밝기 값의 합을 뺀 값을 의미한다. 제안하는 문자인식 시스템에서는 그림 2에 보이는 다섯 가지의 특징점을 이용하여 텍스트 후보영역을 검출하고자 한다.



[그림 2] 사각 특징점(Rectangle features)

2.2 적분 영상(Integral Image)

앞 절에서 제시한 특징점 값은 적분 영상(integral image)이라 불리는 이미지의 표현 방식을 이용하여 쉽고 빠르게 계산할 수 있다. 이미지의 위치 (x, y) 에서의 적분 영상의 값은 식 (1)과 같이 좌표 (x, y) 의 왼쪽과 위쪽에 위치한 픽셀들의 밝기 값의 합을 나타낸다.

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad (1)$$

여기서 $ii(x, y)$ 는 적분 영상을 나타낸다. 그리고 $i(x, y)$ 는 원래의 입력 이미지를 나타낸다. 그리고 식 (1)의 적분 영상은 식 (2), (3)의 순환식을 이용하여 쉽게 구할 수 있다.

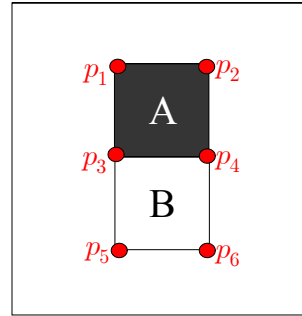
$$s(x, y) = s(x, y-1) + i(x, y) \quad (2)$$

$$ii(x, y) = ii(x-1, y) + s(x, y) \quad (3)$$

여기서 $s(x, y)$ 는 열 방향의 픽셀 값들의 합이고 $s(x, -1) = 0$, $ii(-1, y) = 0$ 이다.

적분 영상을 이용하면 어떠한 사각형 영역내의 픽셀 값들의 합도 적분 영상 상의 네 점을 참조함으로써 빠르게 구할 수 있다. 그림 3에서 나타난 B영역 내의 픽셀 값들의 합을 구하기 위해서는 사각형 B의 네 꼭지점에서의 적분 영상 값을 참조하여 다음과 같이 구할 수 있다[7].

$$B = ii(p_1) + ii(p_4) - ii(p_2) - ii(p_3) \quad (4)$$



[그림 3] 두개의 사각 특징점 계산 예

따라서 인접한 두 개의 사각형 각각의 합 사이의 차이 값은 적분 영상 상의 6개의 점을 참조함으로써 다음 식과 같이 간단하게 구할 수 있다.

$$\text{Two-rectangle feature value} \quad (5)$$

$$\begin{aligned} &= \sum_{(x,y) \in B} i(x,y) - \sum_{(x,y) \in A} i(x,y) \\ &= (ii(p_6) + ii(p_3) - ii(p_4) - ii(p_5)) \\ &\quad - (ii(p_4) + ii(p_1) - ii(p_2) - ii(p_3)) \\ &= ii(p_6) + 2ii(p_3) - ii(p_5) - 2ii(p_4) \\ &\quad - ii(p_1) + ii(p_2) \end{aligned}$$

결과적으로 말하면 앞 절에서 제시한 두개의 사각형으로 구성된 특징점 값을 구하기 위해서는 적분 영상 상의 여섯개의 점을 참조함으로써 구할 수 있고 3개의 사각형으로 구성된 특징점 값은 여덟 개의 점을 참조하면 구할 수 있다. 또한 4개의 사각형으로 구성된 특징점 값은 아홉 개의 점을 참조함으로써 쉽고 빠르게 구할 수 있다.

2.3 AdaBoost 알고리즘

제안하는 문자 인식 시스템에서는 AdaBoost 학습 알고리즘을 이용하여 특징점들을 선택하고 분류기(classifier)를 학습시킨다. AdaBoost 학습 알고리즘은 간단한 학습 알고리즘의 판단 성능을 향상시키기 위해서 사용된다[8]. 텍스트 검출의 기준이 되는 24×24 크기의 창(window) 내에는 앞에서 언급한 사각형의 특징점들이 대략 180,000개 존재할 수 있다. 비록 각각의 특징점의 값이 적분 영상을 이용하여 매우 효과적으로 계산될 수 있다고 할지라도 180,000개의 특징점들을 계산한다는

것은 상당한 계산량을 필요로 한다. 하지만 실험을 통해서 적은 수의 특징점들을 결합하여 이용함으로써 효과적으로 을 검출할 수 있음을 확인할 수 있었다. 여기서 해결해야 되는 문제가 복잡하고 다양한 배경에서 텍스트를 강인하게 검출할 수 있는 특징점들을 찾는 것이다[9].

이러한 목표를 위해서 AdaBoost 학습 알고리즘은 텍스트 이미지와 텍스트가 아닌 이미지를 가장 잘 판단할 수 있는 앞 절에서 보여진 형태의 사각형 특징점을 찾도록 설계된다. 각각의 특징점에 대해서 판단의 에러를 최소화 할 수 있는 최적의 임계값을 결정해야 한다.

약분류기(weak classifier) $h_j(x)$ 는 특징점 f_j , 임계값 θ_j 그리고 부등호의 방향을 표시하는 parity p_j 로 구성된다.

$$h_j(x) = \begin{cases} 1 & \text{if } p_j f_j(x) < p_j \theta_j \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

여기서 x 는 24×24 크기의 sub-window 이다.

Adaboost 학습 알고리즘의 과정을 간단하게 요약해보면 다음과 같다.

1. 학습에 필요한 텍스트 이미지들과 텍스트가 아닌 이미지 $(x_1, y_1), \dots, (x_n, y_n)$ 를 준비한다. 여기서 텍스트가 아닌 이미지와 텍스트 이미지 각각에 대하여 $y_i = 0, 1$ 이다.

2. $y_i = 0, 1$ 각각에 대해서 가중치

$$w_{1,i} = \frac{1}{2m}, \frac{1}{2l} \text{를 초기화한다. 여기서 } m \text{과 } l$$

은 텍스트가 아닌 이미지와 텍스트 이미지 각각의 개수이다.

3. $t = 1, \dots, T$

$$\text{가중치를 정규화한다. } w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^n w_j}$$

각각의 특징점 j 에 대해서 분류기(classifier) h_j 를 학습시킨다. 각각의 에러는

$$\varepsilon_j = \sum_i w_i |h_j(x_i) - y_i| \text{으로 계산된다.}$$

최저의 에러값 ε_t 를 가지는 분류기를 선택한다.

가중치를 갱신한다. $w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$ 만약 이미지가 x_i 가 정확하게 분류되었다면 $e_i = 0$ 그렇지 않으면 $e_i = 1$ 이다. 그리고 $\beta_t = \frac{\varepsilon_t}{1 - \varepsilon_t}$ 이다.

4. 최종적으로 결정되어지는 강분류기(strong classifier)는 다음과 같다.

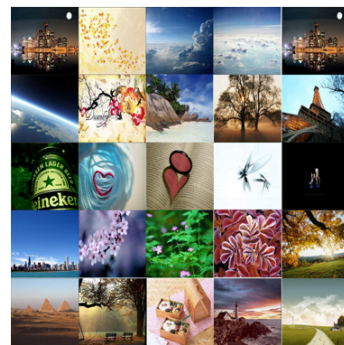
$$h(x) = \begin{cases} 1 & \sum_{t=1}^T \alpha_t h_t(x) \geq 0.5 \times \sum_{t=1}^T \alpha_t \\ 0 & \text{otherwise} \end{cases}$$

여기서 $\alpha_t = \log \frac{1}{\beta_t}$ 이다.

제안하는 시스템에서는 그림 4에서 보여지는 텍스트 이미지들과 그림 5의 텍스트가 아닌 이미지를 이용하여 텍스트 후보영역 검출에 필요한 특징점들을 학습을 통해서 얻을 수 있었다. 학습에 사용된 텍스트 이미지의 개수는 200개이고 텍스트가 아닌 이미지는 600개를 사용하였다. 총 800장의 이미지를 이용하여 텍스트 후보영역 검출에 필요한 특징점을 얻을 수 있었다.



[그림 4] 학습에 사용된 텍스트 이미지



[그림 5] 학습에 사용된 비텍스트 이미지

2.4 AdaBoost 알고리즘을 이용한 텍스트 후보영역 검출

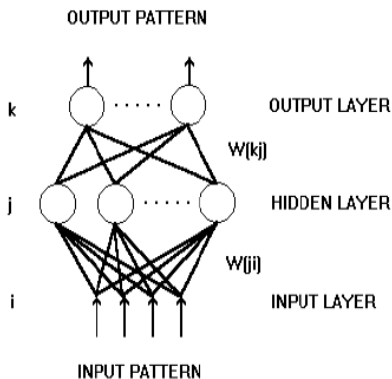
Adaboost 학습 알고리즘을 통해서 얻어진 특징점을 이용하여 다양한 입력 이미지에 대하여 텍스트 후보영역을 검출하는 과정은 다음과 같다.

입력 이미지가 들어오면 이 이미지에서 텍스트 후보영역을 구하게 된다. 텍스트 후보영역에서 24×24 크기의 sub-image를 추출해내고 다음으로 앞서 설명한 학습을 통해서 얻어진 약분류기를 직접이미지에 대입해서 텍스트임을 판단한다. 제안한 시스템에서는 특징점의 값을 사용하여 계산한 후 그 값이 텍스트임을 판단하는 임계값을 넘게 되면 입력 이미지를 텍스트라고 판단하게 된다. 다양한 텍스트의 크기에 대응하기 위해서 이미지 피라미드 과정을 거쳐 앞의 과정을 되풀이하면서 텍스트 후보영역을 검출한다.

3. 개별 문자 인식

3.1 역전파 신경회로망

역전파 신경회로망은 그림 6과 같이 입력층, 은닉층, 출력층으로 구성된 계층적 구조를 지닌 다층 신경 회로망으로써 학습알고리즘을 적용하여 제안한 신경회로망이다. 이 모델은 퍼셉트론에 은닉층을 추가하여 복잡한 논리나 지식을 학습할 수 있는 신경회로망으로서 동작과정은 전방향 처리과정과 학습과정으로 나눌 수 있는데 그 과정은 다음과 같다[10-12].



[그림 6] 역전파 신경회로망

3.2 전방향 처리과정

신경회로망의 가중치가 선정된 경우, 임의의 입력에 대한 출력의 전방향 처리 과정은 목표 출력을 산출하기

위하여 각 단의 동작값을 순차적으로 변경하면서 수행한다. 역전파 신경회로망의 계층수를 L , K 층의 P 번째 노드를 p^k , 입력을 U_p^k , 출력을 a_p^k 라 한다면 $K-1$ 층의 출력은 식 (7)과 같다.

$$U_p^k = \sum_{p=1}^L W_p^k a_p^{k-1} \quad (7)$$

각 단의 노드의 활성화 상태를 결정하는 함수로 주로 시그모이드(sigmoid) 함수가 사용되는데 식 (8)와 같이 기술할 수 있다.

$$a_p^k = \frac{1}{1 + \exp(U_p^k + \theta_j)} \quad (8)$$

신경회로망의 실제출력 a_i 와 목표출력 T_i 와의 차이를 구하기 위하여 평균자승오차(LMS) 알고리즘을 적용하면 식 (9)과 같다.

$$E = \frac{1}{2} \sum_{t=1}^L (T_i - a_i^L)^2 \quad (9)$$

오차 값 E 는 가중치 갱신 시 기준이 되는 값이다.

3.3 학습과정

임의의 입력과 훈련패턴에 의한 학습과정은 신경회로망의 내부연결 가중치를 순차적으로 변경하면서 수행하며, 그 연결강도는 학습 알고리즘에 기본을 둔다. 가중치 변경문제는 입력과 훈련패턴의 집합을 기준으로 하여 오차 평가함수 E 의 크기에 따라 가중치를 변경하며, 식 (10)와 같이 기술할 수 있다.

$$\Delta W_p^k = \eta \delta_p^k a_p^{k-1} + a W_p^k \quad (10)$$

식 (3.4)에서 ΔW_p^k 는 $k-1$ 층과 k 층 노드간의 가중치 W_p^k 의 변화량을 나타내며, η , a 는 각각 학습계수, 판정계수로서 1회의 학습에 따른 변화량을 계산하는데 이용된다. 각 층에서 계산된 오차는 역방향으로 가중치를 순차적으로 갱신한다.

4. 실험 및 결과

본 논문에서는 문자 후보영역의 검출과 개별 문자인식을 위해 Motorola MS-700에 장착된 모바일 카메라로 촬영된 영상을 입력영상으로 사용하였으며, Microsoft Visual C++ 6.0을 사용하여 MFC로 구현하였다.

4.1 문자 후보영역 추출

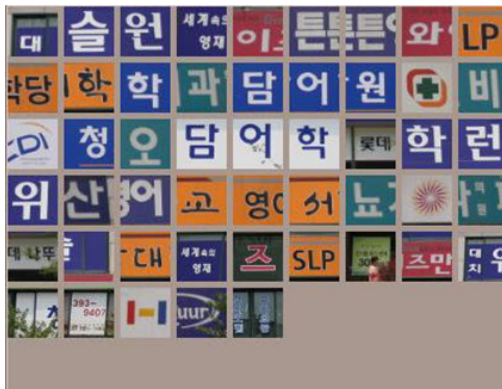
문자 후보영역을 추출하는 과정에서 표 1과 같이 Adaboost 알고리즘을 사용한 학습 데이터의 양을 증가시킴으로서 더욱 정확한 특징점을 판별하여 추출 결과의 정확도를 증가시킬 수 있었다.

[표 1] 학습량에 따른 문자후보영역 검출

| | 정검출 | 부검출 |
|------|-----|-----|
| 200장 | 67% | 24% |
| 400장 | 73% | 26% |
| 800장 | 86% | 19% |



[그림 7] 입력영상과 문자 후보영역 표시



[그림 8] 개별 텍스트 후보영역

그림 7은 입력 영상으로부터 문자들의 후보 영역들을 표시한 것이고 그림 8은 각각의 문자 후보영역 텍스트들을 분류하여 표시한 것이다. 그림 8에서 보는 것처럼 후보 영역으로 분류되어 지는 텍스트들이 비교적 정확하게 텍스트 부분만을 추출해 내고 있는 것을 볼 수 있다.

4.2 개별 문자 인식

오류역전과 신경망을 이용한 개별문자 인식에서는 그림 9에서 보는 것과 같이 사전에 학습한 문자의 경우 100%의 인식률을 보였다.



[그림 9] 문자 인식을 위한 학습 및 이진화

그림 10은 실제 프로그램이 구동 되어지는 모습을 캡처한 것이다. 모바일 구동 장치로부터 입력 영상을 얻은 후 영상으로부터 유효한 텍스트들을 추출한 후 각각 추출된 텍스트들을 인식한 후 음성으로 인식된 문자를 출력해주는 시스템 모형이다.



[그림 10] 인식결과 및 음성출력 화면

5. 결론

본 논문에서는 모바일 카메라로 획득한 영상에 포함되어 있는 텍스트를 추출하고 인식하며 인식결과를 음성으로 출력 해주는 시스템을 제안하였다. 문자 후보영역을

빠르고 정확하게 추출하기 위해 Adaboost 알고리즘을 사용하였고, 추출된 개별 텍스트 후보영역의 문자 인식에 오류역전과 신경망을 이용하여 시스템의 효율성과 정확성을 높이고자 하였으며 음성 출력을 통해 사용자의 편리성 또한 추구하였다.

모바일 카메라로 촬영된 간판 영상의 텍스트들에 대한 실험결과 높은 텍스트 영역 검출률과 사전에 학습한 문자에 대한 인식률을 보였다. 하지만 오류역전과 신경망은 특징값이 많아질 경우 학습시간이 길어 질 수 있다는 단점을 가지고 있기 때문에 향후 이를 보완할 보다 효율적인 학습 분류 알고리즘 및 특징추출에 대한 연구를 진행할 예정이다.

참고문헌

- [1] Anil K. Jain, Bin Yu, "Automatic Text Location in Images and Video Frames," Pattern Recognition, Vol. 31, No. 12, pp. 2055-2076, 1998.
- [2] Yu Zhong, Kalle Karu, Anil K. Jain, "Locating Text in Complex Images," Pattern Recognition, Vol. 28, No. 10, pp. 1523-1535, 1995.
- [3] H. K. Kim, "Efficient Automatic Text Location Method and Content-based Indexing and Structuring of Video Database," Journal of Visual Communications and Image Representation, Vol. 7, pp. 336-344, 1996.
- [4] Pyeoung-Kee Kim, "Automatic Text Location in Complex Color Images using Local Color Quantization," TENCON 99. Proceedings of the IEEE Region 10 Technical Conference, Vol. 1, pp. 629-632, 1999.
- [5] J. Ohya, A. Shio, S. Akamatsu, "Recognizing Characters in Scene images," IEEE Transactions Pattern Analysis and Machine Intelligence, PAMI-16(2), pp. 67-82, 1995.
- [6] Lixu Gu, Toyahisa Kaneko, "Robust Extraction of Characters from Color Scene Image Using Mathematical Morphology," Proceeding of 7th International Conference on Pattern Recognition, Vol. 2, pp. 1002-1004, 1998.
- [7] Jung Sung Uk, "Efficient Rectangle Feature Extraction for Real-time Facial Expression Recognition based on AdaBoost" : Korea Advanced Institute of Science and Technology, 2005.
- [8] Yoav Freund, Robert E. Schapire, "A Short Introduction to Boosting" : Journal of Japanese Society for Artificial Intelligence, 14(5):771-780,

September, 1999.

- [9] Paul Viola, Michael J. Jones, "Robust Real-Time Face Detection" : International Journal of Computer Vision 57(2), 137-54, 2004.
- [10] Bong-Wha Hong, Jie-Young Lee, "On the Enhancement of the Recognition Performance for Back Propagation Neural Networks", 한국 OA학회 학회지 제4권 4호, 1999.
- [11] B.Widrow and M.A.Lehr, "30 Years of Adaptive Neural Network : Perceptron, Msdaline, and Backpropagation", Proceeding of the IEEE, Vol. 78, pp. 1415-1441, September, 1990.
- [12] Yoh-Han Pao, "Adaptive Pattern Recognition and Neural Network", Addison Wesley Publishing Company Inc., 1989.

박 영 현(Young-Hyun Park)

[정회원]



- 2002년 3월 : 한국항공대학교 항공 공전자 및 정보통신공학부 입학
- 2009년 2월 : 한국항공대학교 정보통신공학과 졸업

<관심분야>

이동통신, 멀티미디어

이 형 진(Hyung-Jin Lee)

[정회원]



- 2003년 2월 : 천안대학교 정보통신학부 졸업 (공학사)
- 2005년 7월 : 천안대학교 정보기술대학원 졸업 (공학석사)
- 2007년 7월 : 한국항공대학교 정보통신학과 (박사수료)
- 2007년 8월 ~ 현재 : 한국항공대학교 정보통신공학과 박사과정 재학중

<관심분야>

영상처리, 패턴인식, 영상압축, 멀티미디어

백 중 환(Joong-Hwan Baek)

[정회원]



- 1981년 2월 : 한국항공대학교 항공통신공학과 졸업 (공학사)
- 1987년 7월 : 오클라호마주립 대학원 전기 및 컴퓨터공학과 (공학석사)
- 1991년 7월 : 오클라호마주립 대학원 전기 및 컴퓨터공학과 (공학박사)
- 1992년 3월 ~ 현재 : 한국항공대학교 항공전자 및 정보통신 공학부 교수

<관심분야>

영상처리, 패턴인식, 영상압축, 멀티미디어