

제한된 자원을 가진 휴대용말단기에서 문자인식을 실현하기 위한 한가지 방법

장금성, 김장수

위대한 령도자 김정일동지께서는 다음과 같이 교시하시였다.

《프로그램을 개발하는데서 기본은 우리 식의 프로그램을 개발하는것입니다. 우리는 우리 식의 프로그램을 개발하는 방향으로 나가야 합니다.》(《김정일선집》 증보판 제21권 42페이지)

오늘날 세계적으로 이동통신망이 광범히 구축되어 손전화를 비롯한 휴대용말단기들은 인간활동에서 없어서는 안될 중요한 부분으로 되고있다.[1, 2]

우리는 이러한 휴대용말단기들의 촬영기로 입력된 문서의 문자인식방법에 대한 연구를 진행하였다.

일반적으로 휴대용말단기들은 일반PC에 비하여 일련의 제한성을 가지고있다.

우선 CPU동작주파수가 낮다.

또한 기억기용량이 제한되어있고 기억기와 CPU사이의 자료전송이 느리므로 기억기의 자료접근에 소요되는 시간이 길며 자료접근회수가 많다.

그리고 류점수연산을 보장하지 못하거나 그 성능이 매우 낮다.

이러한 인자들은 휴대용말단기를 리용하여 문자인식을 진행하는데서 인식시간을 증가시키므로 사용자들의 요구를 만족시킬수 없다.

우리는 인식성능을 크게 떨구지 않으면서도 실시간성을 보장할수 있도록 인식속도를 개선하기 위한 한가지 방법을 제안한다.

론문에서 제안한 방법은 크게 세 단계로 이루어진다.

① 전처리

이 과정은 원천화상에 대한 잡음제거, 두값화, 토막화를 진행하는 과정인데 화상처리의 정확성과 속도를 높이기 위하여 OpenCV서고를 리용하였다.

단계 1 계조화상으로 넘기기

촬영기로부터 입력된 색화상은 다음의 식을 통하여 계조화상으로 변환된다.

$$Y = 0.299R + 0.587G + 0.114B$$

단계 2 화상에서 문자만을 따내기 위하여 3×3 크기의 저역통과여과기를 통하여 평활화를 진행하며 다음 Sobel마스크를 적용하여 경계선을 검출한다.

단계 3 Otsu방법을 리용하여 화상의 두값화를 진행한다.

$$Y = \begin{cases} 1, & x > k \\ 0, & x \leq k \end{cases}$$

여기서 x 는 주어진 점에서의 화상의 화소값이고 k 는 두값화알고리즘의 역값이다.

단계 4 다음으로 형태론적추출을 진행한다. 즉 $m = \text{dilation}(\text{image}) - \text{erosion}(\text{image})$.

이때 구조적원소의 크기를 화상의 크기(촬영기로부터 화상까지의 거리)에 따라 설정해야 하며 여기서는 12×8 로 정한다.

단계 5 새로 얻어진 화상 m 에 대하여 등고선처리를 진행하여 문자들의 예상구역들을 찾으며 그에 해당하는 4각형영역들을 추출한다.

단계 6 4각형들에서 필요없는 영역들은 여러가지 조건을 적용하는 방법으로 없애고 주목하는 영역만을 따낸다.

우리는 이러한 조건을 다음과 같이 주었다.

우선 4각형들은 화상의 경계에 위치하지 말아야 하며 가로대 세로의 비가 0.5를 넘어야 한다.

그리고 4각형의 크기가 글자를 포함할만큼 충분히 커야 한다.

다음으로 4각형의 히스토그램이 기본적으로 2개의 극점을 가져야 한다.

② 문자인식진행

우리는 문자인식엔진으로서 정확도가 높은 문자인식엔진 Tesseract를 리용하였다.

론문에서 제안한 문자인식체계의 구성도는 그림 1과 같다.

그림 1에서 단어인식과정은 그림 2와 같다.

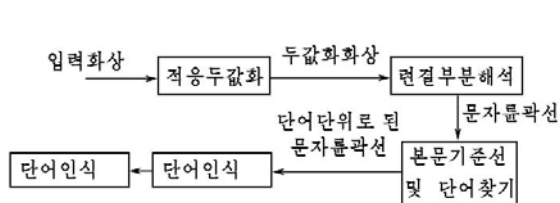


그림 1. 문자인식체계의 구성도

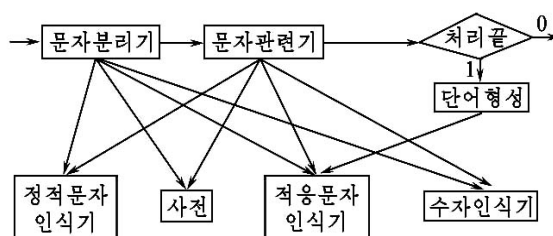


그림 2. 단어인식과정

그림 2에서 계산량이 가장 많은 부분은 정적문자인식기로서 여기서는 특징추출에 의하여 이미 정의된 20~30개의 원형(prototype)들과 입력문자화상으로부터 추출된 50~100개의 특징점들과의 비교를 진행한 다음 유사성정도를 계산하여 최대값을 가지는 문자후보를 선택하게 된다. 이때 매 특징점들은 3차원 자료구조(x , y 자리표, 방향각)로 정의되어있으며 1개 문자의 원형은 4차원 자료구조(x , y 자리표, 방향각, 길이)로 정의되어있다.

우리는 문자인식엔진의 인식률을 크게 떨어지지 않으면서도 휴대용말단장치들의 장치적제한성을 극복할수 있도록 특징점들의 개수를 30~50개로 제한하였다. 이때 유사성척도를 반영하는 거리는 $d_f = d^2 + w\theta^2$ 과 같이 계산된다. 여기서 θ 는 원형과의 각도, w 는 무게값이다.

이로부터 유사성척도 E_f 는 다음과 같이 정한다.

$$E_f = 1/(1 + kd_f^2)$$

결과 기억기리용량은 약 53%로 줄어들었으며 인식률은 종전에 비하여 2% 떨어진 91%로 되었다.

제안한 방법으로 문자인식을 진행하는 과정은 그림 3과 같다.

그림 3의 ㄷ)에서 굵고 짧은 선들은 특징들이며 얇고 긴 선들은 원형들이다.

다음으로 우리는 JNI기술을 리용하여 실행속도를 높이었다.

일반적으로 Android체계가 설치된 말단장치의 응용프로그램들은 Android SDK를 기반으로 한 JAVA언어로 작성된다. 그러나 우리는 C++로 작성된 Tesseract를 JNI(Java Native Interface)기술을 리용하여 처리하였다. 즉 엔진을 콤팩트화하여 외부서고(*.so)로

만든 다음 JAVA언어로 작성된 Activity에서 호출하는 방식을 취하였다. 이때의 호출 및 실행속도는 전부 JAVA언어로 작성하였을 때보다 평균 2배이상으로 높아지게 된다.

③ 사전을 리용한 후처리

Tesseract엔진을 통하여 인식된 문자열은 직접 사전에 입력되어 그에 해당하는 의미를 현시한다. 이때 사전은 사전탐색나무(Trie)를 리용하여 구현하는데 그 복잡도는 $O(N \log_2 N)$ 이다.

이렇게 구현된 사전에서 문자인식의 틀린 결과에 대하여 수정을 할수 있도록 함으로써 체계의 정확도가 높아지게 되었다.

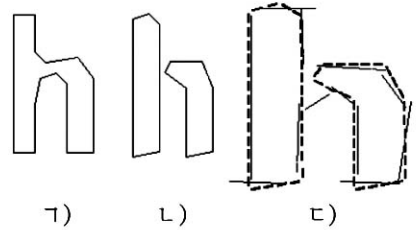


그림 3. 원래 문자(ㄱ)), 입력문자(ㄴ)), 특징점들과 형태와의 비교(ㄷ))

맺 는 말

OpenCV서고와 Tesseract엔진을 리용하여 휴대용말단장치들에서의 문자인식을 실현하기 위한 한가지 새로운 방법을 제안하였다.

참 고 문 헌

- [1] J. Liang et al.; IJDAR, 7 , 84, 2005.
- [2] Ing. Juraj Stefanovic; TranSign Android Sign Translator, Slovak University of Technology in Bratislava, 50~116, 2013.
- [3] A. B. Cambra et al.; IEEE International Conference on Computer Vision Workshops, 11, 64, 2011.

주체103(2014)년 12월 5일 원고접수

A Method for Realizing Character Recognition in Handheld Terminals with Limited Resources

Jang Kum Song, Kim Jang Su

A new method is proposed for optical character recognition in handheld terminals with limited resources by using OpenCV library and Tesseract engine.

Key words: optical character recognition, Tesseract