

# 스마트폰 기반 OCR 시스템의 구현

허진규 · 김현곤 · 장민석 · 장석환 · 권순각 · 박유현  
동의대학교 컴퓨터소프트웨어공학과

## Implementation of OCR System Based on Smart Phone

Jinkyu Heo · Hyungon Kim · Minseok Jang · Seokhwan Jang · Soonkak Kwon · Yoohyun Park  
Department of Computer Software Eng, Donggeui University

### 요 약

스마트폰은 폭발적인 성장세와 더불어 어플리케이션의 시장규모가 크게 증대될 것으로 예상된다. 본 논문에서는 직접 사진을 캡처 시킨 문자들을 텍스트로 변환시켜주고 데이터베이스를 이용해서 그 문자들의 의미를 표현하는 어플리케이션을 구현한다. 어플리케이션의 기능으로는 텍스트로 변환하고자 하는 사진을 찍거나 혹은 앨범에 있는 사진을 불러내어 텍스트로 변환을 시켜서 상대방에게 메시지나 블루투스 등을 이용하여 전송을 가능하게 해줄 수 있으며, 글꼴을 바꾼다거나 크기 등 편집 기능까지 추가적으로 가능하도록 할 수 있다. 이 어플리케이션으로 인한 효과로 인터넷이나 거리를 돌아다니다가 원하는 문구가 있으면 바로 원하는 상대방에게 쉽고 간편하고 빠르게 전달할 수 있는 편리함 등이 있다.

### 키워드

OCR(Optical Character Recognition), 미디언 필터(median filter),  
히도로프 거리 알고리즘, 바이큐빅 보간 알고리즘

## I. 서 론

요즈음은 스마트 시대이다. 스마트 폰 하나면 대부분 안 되는 일이 없을 정도로, 반대로 스마트 폰이 없으면 우리 일상에 많이 지장을 줄 만큼 중요한 요소로 자리 매김을 하고 있다.

스마트폰으로 인터넷 서핑 중에 정말 중요한 문구를 친구들에게 같이 공유를 하고 싶은데 복사가 금지되어 있는 글들을 어떻게 하면 텍스트로 변환하여 같이 의견을 나눌 수 있을까 하는 궁금증이 생길 수 있다.

본 논문은 그림에 있는 문자들을 텍스트로 변환시켜주는 Application을 개발하는 것이다.

하지만, 무턱대고 어플리케이션 개발을 한다면 2가지의 문제점이 발생할 수 있다.

문제점 1. 마음에 드는 문구를 저장하기 위해 번거롭게 사진을 찍어야 하는가?

문제점 2. 사진을 찍는데도 한계가 있다.(용량문제)

본 논문에서는 이와 같은 방식의 문제점을 해결하고 자동으로 인식하며, 사용하기 편리한 OCR Application을 구현한다.

## II. OCR 인식 및 측정 방법

2-1 OCR을 통한 인식, 문자 측정

|   |  |
|---|--|
| 객체 인식 문제점<br>필터기능을 이용하여<br>잡음과 왜곡 등을<br>제거한다. | 문자 측정 문제점<br>단위 블록의<br>집합으로 분할 후,<br>영역을 분리한다. |
|---|--|

## 2-2 객체 인식

### (1) 사진 객체를 이진화

문서 영상의 이진화는 입력 영상내의 문자 영역은 강조시키고 배경이나 음영은 비교적 포함되지 않도록 하는 것과 문서 영상의 해석이 빠른 속도로 이루어질 수 있게 하는데 목적을 두고 있다.

이진화는 그레이(gray) 영상을 0 또는 1의 이진 영상으로 변환하는 과정으로 인식을 하기 위하여 컬러(RGB)영상을 그레이영상으로 변환하고, 이를 다시 0~255 단계로 구성된 그레이 영상을 이진(binary)영상으로 변환하여야 한다.

이진 영상으로 변환하는 방법은 문서 영상 전체에 대하여 전역적 임계값을 이용한 전역적 이진화 방법과 영상의 각 부분마다 임계값을 이용한 지역적 이진화 방법이 있다. 기본적으로 영상 전체의 평균을 임계값(threshold)으로 사용하는 전역적 이진화(global binarization) 방법에서 화소들의 밝기 분포 히스토그램에서 전경과 배경에는 큰 피크가 존재하는데, 이 두 피크 사이의 중간 값을 임계값으로 하여 이진화를 수행한다.

문서 영상의 글자에 잡음이 섞이면 전혀 다른 글자로 인식되는 경우가 발생한다. 이러한 잡음이 글자의 한 부분인지 판단하여 글자부분이 아니라면 잡음을 제거해야 한다. 잡음제거를 위해서 기본적으로 평활화, 선형화, 저주파 통과 필터(low pass filter), 미디언 필터(median filter)등이 사용이 되는데, 잡음제거 방법 중 원영상의 강한 에지와 상세 한 부분을 보존하기 위해서는 미디언 필터가 가장 효과적이기 때문에 미디언 필터를 이용한다.

### 2-3 문자 측정

문자 측정 분석은 일반적인 문서나 특정한 형식의 문서에서 지식을 사용하여 문서 영상으로부터 문서의 기하학적인 구조를 추출하는 것이다. 즉, 문자를 단위 블록의 집합으로 분할한 후, 각

블록의 특성을 조사하여 문자 영역과 비문자 영역으로 분리하는 작업이다. 이러한 문자 구조 분석을 위하여 사용되는 방법은 일반적으로 투영 프로파일(projection profile)방법을 적용하거나 연결요소 분석(connected component analysis)방법 등을 적용한다. 연결요소 분석 방법은 임의의 흑(black) 화소 또는 백(white) 화소에 대하여 8방향 또는 4방향의 인접화소가 같을 경우 이 화소를 모두 연결하여 화소의 집합을 얻는다. 그림 영역은 연결 요소의 부피가 큰 반면에 문자 영역은 작고 일정한 크기의 연결 요소를 갖는다. 이 방법은 유사한 특성을 갖는 연결 요소를 만들기 위해 각 요소를 병합하는데, 연결 요소의 종, 횡비 등의 특징을 분석하여 음영을 제거하거나 필요한 영역을 추출할 수 있으며, 활자체나 문자 크기의 변화에 무관하게 임의의 방향에 위치한 문자 행을 분리할 수 있다. 그러나 계산량이 많이 필요하고 경험적 지식이 많이 사용되는 단점이 있다.

투영 프로파일은 투영축상의 흑색화소의 수를 계산함으로써 얻어지는 물결 모양의 윤곽으로 문서의 전역적인 특성을 나타내어 기울어진 문서의 교정이나 요소 추출, 그리고 문자 추출에 중요한 역할을 한다. 투영 윤곽의 값이 부분적으로 최소가 되는 지점을 문서 영역 분할의 기준으로 결정하는데, 대부분의 문서가 수직 또는 수평 방향의 블록으로 구성되므로 반복적으로 적용할 경우 매우 효율적으로 각 영역을 분리할 수 있다.

마지막으로 문서 구조 분석을 통해 추출된 텍스트 영역은 인식 대상인 문자 단위로 분할하게 되는데, 분할된 문자의 가로, 세로의 크기는 일정하지 않으며, 분할된 문자를 일정한 형태로 만들기 위해 정규화 과정을 수행한다.

특정 화면의 에지 영상에 대한 허도로프 거리 알고리즘은 다음과 같이 이용하여 매칭 시켜 영역을 정확하게 검출한다.

$$d(a,B)= \min_{b \in B} ||a-b||$$

$$\begin{aligned} H(A,B) &= \max(h(A,B),h(B,A)) \\ h(A,B) &= \min_{a \in A} d(a,B) \\ &= \max_{a \in A} \min_{b \in B} d(a,b) \\ &= \max_{a \in A} \min_{b \in B} ||a,b|| \end{aligned}$$

$$h(B,A) = \max_{b \in B} \min_{a \in A} ||a-b||$$

이 검출한 결과 영상에 대한 확대된 영상을 보간하기 위해서 바이큐빅 보간 알고리즘과 기울기 고정하기 위한 로테이션 적용하였고 이진화, 잡음 제거, 문자 분할, 정규화 순서로 원하는 문자를 정확하게 인식하였으며, 인식한 결과는 텍스트 화면으로 변환되어진다.  
변환된 텍스트는 상대방에게 공유뿐만 아니라 수정까지 이루어진다.

### III. 구현 내용

#### (1) 개발환경

|                        |  |
|------------------------|--|
| <u>Product Version</u> | Xcode Version 3.2.5 ( 64bit )<br>Xcode IDE 1760.0,<br>Xcode Core 1763.0,<br>ToolSupport 1758.0 |
| <u>System</u>          | Mac OS X Version 10.6.8  |

#### (2) 구현 화면

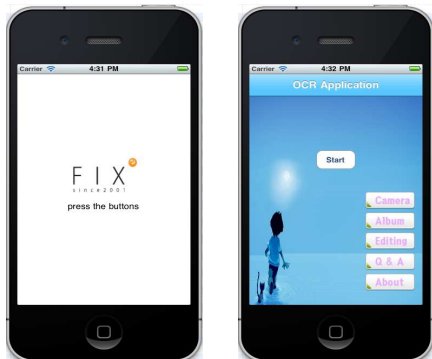


그림 3-1. 어플리케이션 실행 후 화면

왼쪽의 화면이 최초 실행하였을 때의 초기화면. press the buttons를 누르게 되면 오른쪽에 보이는 메인화면으로 전환하게 되며 이 OCR Application의 기능이 이루어진다.

딱딱한 OCR Application을 방지하기 위해서 bgm기능 까지 사용자가 이용할 수 있다.

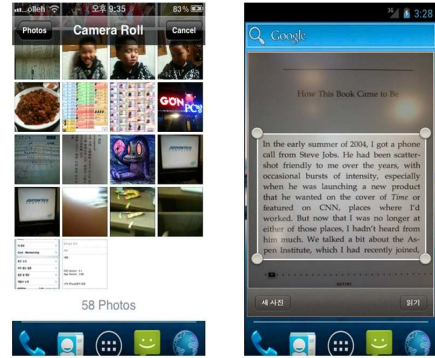


그림 3-2. 추출할 문자를 터치하여 지정

추출할 문자를 Album에서 지정할 수도 있고, 사진을 직접 찍어서 추출할 텍스트 부분을 터치하여 지정할 수도 있다.



그림 3-3. 추출 후 공유기능

텍스트로 변환이 이루어진 다음에는 위 그림처럼 메시지나 블루투스, e-mail을 통하여 상대방에게 공유기능도 가능하다.

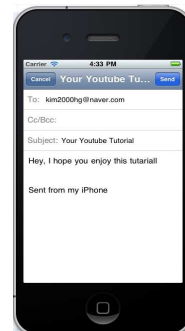


그림 3-4. 개발자에게 문제점 보고

사용 중 불편한 점이 있으면 언제든지 개발자에게 편리한 UI를 통하여 건의사항 등을 mail로 보낼 수가 있다.



그림 3-5. 사용방법을 사용자에게 정의

처음 사용하는 개발자를 위해서 사용법까지 친절하게 설명이 되어 있어서 사용하는 데에는 문제가 없을 것이다.

### 참고문헌

- [1] Salomon, "Curves And Surfaces of Computer Graphics", 2005.04
- [2] Mortenson, "Geometric Modeling", 2006.05.01
- [3] Richard E. Neapolitan, Foundations of Algorithms USING C++ PSEUDOCODE, 2004.03.
- [4] <http://www.roborealm.com/>
- [5] <http://vision.cs.arizona.edu/>



그림 3-6. 개발자 소개 및 연락처 제공(건의)

개발자 소개 및 프로젝트 버전, 연락처 등을 제공한다.

## IV. 결 론

바이큐빅 보간법을 적용한 OCR알고리즘으로 구현을 하게 되면 속도 및 인식률 측면에서는 타 알고리즘에 비해 향상되는 것을 알 수 있다.

허도로프 거리 알고리즘과 필터기능을 이용하여 잡음, 문자 복원, 흐림 영상을 복원하기 위해 바이큐빅 보간법을 구현하였다.

인식되어졌을 때, 한자에 대해서는 획을 많이 필요로 하고 데이터베이스의 양이 엄청나기 때문에 구현하기에 어려움이 있는 반면 특수문자 및 숫자, 영어에 대해서는 높은 인식률 측면에서 향상이 되었다. 향후 제안된 알고리즘을 이용하여 구현된 시스템의 성능을 테스트하기 위하여 금융, 명함, 차량번호 등 실생활 분야에 적용 할 수 있다.