

모바일 환경에서의 명함인식 성능 향상에 관한 연구

신현섭¹ · 김차중^{2*}

A Study on Performance Improvement of Business Card Recognition in Mobile Environments

Hyunsub Shin¹ · Chajong Kim^{2*}

¹Department of Computer Engineering, Graduate School of Information & Communications, Hanbat National University, Daejeon 305-719, Korea

²Department of Computer Engineering, Hanbat National University, Daejeon 305-719, Korea

요 약

본 논문은 모바일 환경에서의 명함 인식 성능 향상을 위한 방안으로 서로 다른 알고리즘과 학습 데이터를 갖는 이종(異種)의 명함 인식 엔진을 병렬처리 하여 데이터를 결합하는 하이브리드 OCR 에이전트를 제안하였고, 모바일 카메라의 특성상 촬영자의 환경에 따라 변하는 조명, 촬영방향, 명함의 배경에 적응하는 모바일 카메라에서의 명함 이미지 전처리 기법을 제안하였다. 본 논문에서 제안한 방법으로 하이브리드 OCR 에이전트를 구성할 경우 단일 엔진을 구성하였을 때 보다 국문명함의 명함 인식률이 평균 90.69%에서 95.5%로 향상되었고, 이미지 전처리 기법을 적용함으로써 이미지 용량이 50% 수준으로 줄어들었으며 이미지 전처리 기법을 적용하기 전보다 인식률이 83%에서 92.48% 수준으로 약 9.4%의 향상 효과를 얻을 수 있었다.

ABSTRACT

In this paper, as a way of performance improvement of business card recognition in the mobile environment, we suggested a hybrid OCR agent which combines data using a parallel processing sequence between various algorithms and different kinds of business card recognition engines which have learning data. We also suggested an Image Processing Method on mobile cameras which adapts to the changes of the lighting, exposing axis and the backgrounds of the cards which occur depending on the photographic conditions. In case a hybrid OCR agent is composed by the method suggested above, the average recognition rate of Korean business cards has improved from 90.69% to 95.5% compared to the cases where a single engine is used. By using the Image Processing Method, the image capacity has decreased to the average of 50%, and the recognition has improved from 83% to 92.48% showing 9.4% improvement.

키워드 : 명함인식, 모바일, 영상처리, 이진화

Key word : Business Card, Mobile, Image Processing, Binarization

접수일자 : 2013. 12. 16 심사완료일자 : 2014. 01. 07 게재확정일자 : 2014. 01. 20

* **Corresponding Author** Chajong Kim(E-mail:cjkim@hanbat.ac.kr, Tel:+82-42-821-1206)

Department of Computer Engineering, Hanbat National University, Daejeon 305-719, Korea

Open Access <http://dx.doi.org/10.6109/jkiice.2014.18.2.318>

print ISSN: 2234-4772 online ISSN: 2288-4165

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서 론

스마트폰의 출하량이 PC와 노트북을 넘어서면서 인류에 유용한 도구로서 대중화되고 있다. 인맥을 관리해야 하는 현대인들에게 편리하게 입력할 수 있도록 하는 명함인식기술은 스마트폰 시장에서 다양한 앱으로 유통되고 있으며 이에 따라 명함인식기술의 수요는 점차 증가하고 있다[1]. 그러나 모바일 환경에서 효과적인 명함인식성능을 얻기 위해서는 하드웨어 자원의 한계성으로 인해 경량 인식엔진을 탑재해야 했던 문제를 극복해야 하고[2], 문자인식기술의 특성상 글자를 선명하게 촬영하기 위해 근접촬영을 하게 될 때 조명에 의해 발생하는 그림자의 원인으로 최상의 영상을 획득하지 못하는 문제점을 해결해야 한다. 본 논문에서는 이러한 문제점을 해결하기 위해 모바일 환경에서의 명함인식성능을 향상시킬 수 있는 방안을 제안한다.

현재 모바일 환경에서 요구되는 저전력, 경량이라는 특성 때문에 모바일 환경의 컴퓨팅 파워 및 저장 능력은 PC에 비해 많이 떨어진다. 또한 높은 파워를 가진 단말의 상대적으로 높은 가격으로 인해 모든 사용자가 필요 이상의 컴퓨팅 파워를 가진다고 보장할 수도 없다. 모바일 단말에서도 고사양 PC 수준의 연산처리가 필요한 서비스 제공이 가능해지면 기존 모바일 사용 환경보다 진보된 성능을 기대할 수 있다[3].

표 1. 모바일 환경에서의 제약사항
Table. 1 Limitations of the Mobile environments

모바일 단말	무선 인터넷 환경
CPU 성능의 한계	대역폭(Bandwidth)의 제약
메모리 용량의 한계	데이터 전송속도 성능의 한계
애플리케이션 용량의 부담	데이터 통신에서 보안
대용량 데이터베이스	

이처럼 모바일 단말은 제한된 자원으로 인해 일반적으로 모바일 단말이 감당할 수 있는 것보다 더 많은 자원을 요구하게 된다. 특히 명함인식과 같은 복잡한 연산 처리를 요하는 기술은 모바일 단말이 배터리로 동작되어야 하는 특성상 저전력 설계로 인하여 부동소수점 연산이 전통적 PC에 비해 느린 ARM 계열의 CPU를 사용하므로 대용량 학습 데이터를 가진 인식엔진을 이식(porting)하였을 때 효과적인 성능을 기대하기 어렵다.

이러한 문제점을 완화하기 위해서는 대부분의 연산 수행 능력을 외부의 자원에서 얻어야 한다[4-5].

따라서 명함인식엔진을 서버환경에서 동작되도록 구성할 경우, 서버의 자원을 활용할 수 있게 됨에 따라 서로 다른 장단점을 가진 명함인식엔진의 병렬화를 위한 멀티 프로세싱(Multiprocessing)이 가능해지므로 서로 다른 알고리즘과 학습 데이터를 가진 복수의 명함인식엔진을 동시에 수행하도록 하고, 인식된 결과에서 각 항목별로 계산된 가중치를 기준으로 더욱 높은 인식결과를 선택하여 최종적으로는 각각의 명함인식엔진에서 추출된 결과보다 더욱 향상된 인식률을 얻을 수 있게 된다[6].

영상처리 관점에서 모바일 단말은 고정된 밝기와 위치에서 영상을 획득할 수 있는 스캐너와 달리 촬영자의 촬영 방향과 장소, 조명에 의해 발생할 수 있는 그림자와 복잡한 배경이 최상의 영상을 획득하는 방해요소가 된다. 이를 위해서는 그림자를 회피하여 촬영할 수 있도록 하면서 촬영 시 어쩔 수 없이 발생된 그림자가 있을 경우 이를 제거한 이진화 영상을 생성해야 한다. 또한 명함이 놓여진 배경에 상관없이 명함 영역을 추출하는 이미지 전처리 기법이 필요하다[7].

본 논문에서는 다양한 이미지 처리 알고리즘을 사용하여 모바일 단말에서 실시간으로 명함의 영역이 검출되는 것을 촬영자에게 시각화하여 보여줄 수 있도록 하고, 촬영된 영상에서 그림자가 제거된 명함 영상을 생성하는 이미지 전처리 기법을 제안한다.

II. 관련연구

2.1. 모바일 CPU의 한계

전통적 PC에 비해 모바일 기기의 성능차이가 발생하는 이유는 사용 목적이 다르고 CPU의 설계부터 차이가 있기 때문이다. ARM과 x86 아키텍처를 상세히 분석한 연구에 따르면 인텔의 x86은 CISC(Complex Instruction Set Computer)로 설계되어 명령어셋이 복잡하고 가변 길이지만, ARM은 RISC(Reduced Instruction Set Computer) 설계로 명령어셋이 단순하고 고정길이를 갖고 있다. 이렇게 ARM 아키텍처는 간소화에 초점을 맞췄기 때문에 전력 소비량에서는 우수하지만, 성능 향상에 제한을 받게 된다[8].

현재 대부분 음성인식과 문자인식 분야에서는 신경망을 이용한 역전파 학습 알고리즘을 사용하고 있다. 특히 신경망은 비선형적인 특성 때문에 정수 기반의 연산으로는 구현하기가 어렵다. 특히 시그모이드 함수나 지수 함수와 같은 비선형 함수들을 연산하기 위해서는 부동소수점 연산이 필요한데 이를 하드웨어적으로 지원하는 부분이 FPU(Floating-Point Unit)이다. x86 아키텍처와 달리 ARM 기반 CPU는 전반적인 퍼포먼스와도 차이가 있지만, FPU 성능이 상대적으로 떨어지는 단점이 있다. 또한 본 논문에서 제안하는 병렬 인식을 위해서는 멀티 프로세싱 해야 한다. 이 때 문맥 교환(Context Switch)이 빈번히 발생한다. 문맥 교환이란 하나의 프로세스가 CPU를 사용 중인 상태에서 다른 프로세스가 CPU를 사용하기 위해, 이전의 프로세스의 상태(문맥)를 보관하고 새로운 프로세스의 상태를 적재하는 작업을 말한다. RISC는 CPU 명령어의 개수를 줄여 하드웨어 구조를 좀 더 간단하게 만드는 방식이기 때문에 RISC의 경우 레지스터의 용량이 CISC보다 상대적으로 크기 때문에 좀 더 큰 오버헤드가 발생한다[9].

2.2. 모바일 환경에서의 영상처리

모바일 기기의 카메라를 활용한 촬영 방법은 매우 편리함을 주는 반면 기존의 스캐너처럼 일정한 밝기와 고정된 상태에서 영상을 얻는 것과는 차이가 크다.

모바일 기기의 카메라를 통해 명함을 인식하는 방식에서 최적의 인식률을 얻기 위해서는 몇 가지 제약조건이 따른다.

첫째, 이미지의 왜곡이 없어야 한다. 문자를 선명하게 촬영하기 위해서는 접사 모드에서 촬영하기 때문에 카메라의 렌즈 특성상 근접촬영 시 이미지의 왜곡이 생길 가능성이 높다.

둘째, 흔들림이 없어야 한다. 사람의 손으로 카메라를 촬영하는 특성상 렌즈와 피사체가 가까이 다가갈수록 피사체가 흔들려 찍힐 수 있고, 피사체 주변의 광량이 부족해져 카메라의 셔터속도가 느려지기 때문에 사람의 미세한 흔들림이 선명한 사진 촬영을 방해하는 요소가 되기 쉽다.

셋째, 피사체에 그림자가 없어야 한다. 보통 실내 또는 대낮에는 광원이 사람 위에 있기 때문에 위에서 아래로 명함을 촬영 할 경우에는 기기나 신체의 일부에

의해 피사체에 그림자가 생길 가능성이 높다. 이 때 그림자로 인해 어둡게 촬영된 영역에는 노출의 차이가 있어 글자의 구분이 명확하지 않거나 이진화 과정에서 문자영역이 사라지는 현상이 발생한다. 따라서 인식이 불가능한 상태가 되거나 인식률이 현저히 떨어지는 문제점이 있다[10].

본 논문에서 사용된 모바일 카메라에서의 촬영된 영상의 이진화의 효율이 높은 Sauvola 이진화 알고리즘과 OpenCV의 Adaptive Threshold 알고리즘은 주변 픽셀들의 분포를 분석하여 자동으로 임계값을 결정하는 적응형 이진화 알고리즘으로 조명에 의해 밝기가 급격하게 변화하는 영상에 유리하다[11].

III. 하이브리드 OCR 에이전트

하이브리드 OCR 에이전트는 기존의 단일 엔진을 구성했던 OCR 서버[12]에 비해 두 형태의 인식엔진에 작업을 할당하기 때문에 동시 처리 능력이 반으로 줄어든다. 하지만 제안하는 하이브리드 방식을 통해 서로 다른 알고리즘을 가진 인식엔진에서 출력된 결과를 분석하여 상호간 단점을 보완하기 때문에 진보된 성능을 제공할 수 있다.

3.1. 하이브리드 OCR 에이전트의 시스템 구성

하이브리드 OCR 에이전트를 구성하기 위해서는 클라이언트와의 통신을 위해 그림 1과 같은 경로를 통하여 사용자로부터 요청된 작업을 전달 받기 위해 데이터베이스와 연결된다. 서버와 클라이언트간의 OCR 인터페이스는 REST(Representational State Transfer)기반의 Open API가 담당한다.

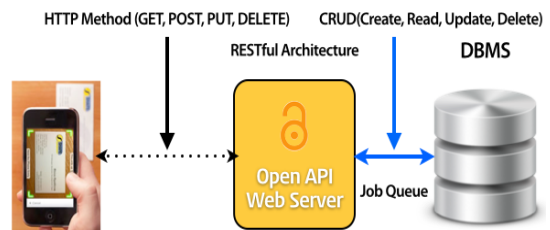


그림 1. RESTful OpenAPI를 활용한 작업 할당
Fig. 1 Task allocation using the RESTful OpenAPI

Open API 사용목적은 사용자가 촬영한 영상을 하이브리드 OCR 에이전트에 전송하고, 클라이언트의 요청에 따라 에이전트에서 처리되는 일련의 프로세스를 전달받아 클라이언트에 응답하는 역할을 수행하는 데 있다. 구현된 API의 URI는 Resource 클래스를 구현한 것으로 그림 1에서의 HTTP Method 호출에 의해 제공된 명함 영상 및 데이터를 데이터베이스에 CRUD하고 그 결과를 HTTP Response에 담아서 되돌려 주는 역할을 한다. 이와 같은 작업은 대표적으로 다음 4가지 API를 통하여 동작된다.

1) http://URI/OCRService/api/tokens

Token API는 API의 사용이 허가된 모바일 단말이 신규로 작업을 요청할 때 인증의 절차와 함께 작업 요청에 필요한 토큰(token) id를 부여받는다. 이 때 토큰 id는 암호화 되며 이미지와 XML 데이터의 암호화에 사용된다. API 서버는 API Key를 비롯한 인증에 필요한 정보를 데이터베이스를 통해 체크하고 기록한다.

2) http://URI/OCRService/api/images

토큰 ID와 인식할 영상 데이터를 전송하면 데이터베이스의 작업 큐(Queue)에 작업을 추가하고, 오류 발생 시 에리코드를 반환한다. 이 때 그림 2에서의 로드 밸런스 서버에 작업이 추가되었음을 알리고 로드 밸런스 서버는 즉시 작업을 작업 분배 알고리즘에 따라 작업을 분산 처리한다.

3) http://URI/OCRService/api/status/{tid}

status API는 어플리케이션에서 작업이 진행되는 상태를 총 6가지로 분리하여 진행상황을 사용자가 인지할 수 있도록 값을 반환하고, 서버의 과부하로 작업이 많이 밀려있을 때는 Busy상태를 알리고 사용자가 기다리지 않고 다음 작업을 진행할 수 있도록 한다. 작업 상태를 알기 위해서는 토큰 ID를 URI에 포함하면 현재 작업 상태를 반환한다. 작업상태가 Complete 상태로 전환되면 인식이 완료된 것을 뜻한다.

4) http://URI/OCRService/api/result/{tid}

전자명함의 표준방식인 vCard를 RDF 형식[1]으로 구성하여 XML 내용을 반환 받을 수 있다. 이 때 서버에서는 이미 토큰 ID를 기반으로 XML을 암호화 하여 XML을 저장되어 있으므로, 클라이언트에서는 토큰 ID를 복호화 하여 복호화된 토큰 ID를 기반으로 XML을 복호화 한다.

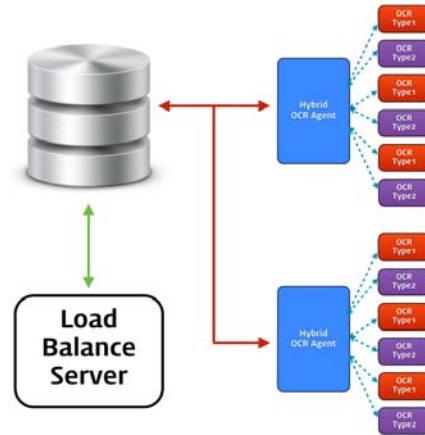


그림 2. 로드밸런스 서버의 구조
Fig. 2 Structure of Load Balance Server

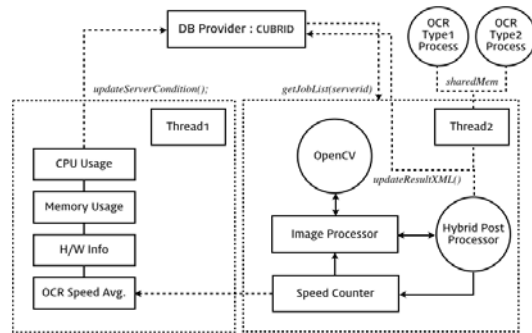


그림 3. 하이브리드 OCR 에이전트의 구조
Fig. 3 Structure of the Hybrid OCR Agent

그림 3은 하이브리드 OCR 에이전트의 동작방식을 도식화 하였다. 두 형태의 OCR 엔진을 갖는 프로세스와 공유된 메모리 영역을 통해 작업 진행 상태와 같은 일련의 과정을 주고받는다. 에이전트 내부에는 총 2가지의 스레드로 구성되어 있으며 2번 스레드의 Speed Counter는 요청된 작업 시간을 계산하여 누적한다.

$$CPU\% = 100 \times \frac{User\ Time_{Process} + Kernel\ Time_{Process}}{User\ Time_{CPU} + Kernel\ Time_{CPU}} \quad (1)$$

1번 스레드는 에이전트의 상위 프로세스에서 작업을 분배하기 위해 로드밸런싱의 보조 자료의 제공을 담당한다. 누적된 인식 수행시간의 평균치와 식(1)에서 계산된 CPU 사용량과 메모리 사용량, 하드웨어 정보를 데

이터베이스에 업데이트되며 로드밸런스 서버는 서버의 우수한 성능에 요청된 작업을 우선 할당한다.

3.2. CPU 성능에 따른 OCR 프로세스 구성

하이브리드 OCR 에이전트는 CPU 성능에 코어의 수량만큼 OCR 프로세스를 생성하여 멀티 프로세스 방식으로 동작한다. 따라서 생성된 OCR 프로세스 수만큼 동시처리가 가능하다. 만약 서버의 프로세서가 옥타 코어라면 총 8개의 OCR 프로세스를 생성하게 된다. 이 때 1개의 작업당 2개의 프로세스를 사용하게 됨으로써 동시 처리가 가능한 작업 수는 4가 된다. 또한 하이브리드 OCR 에이전트는 두 형태의 인식 프로세스를 관리하기 때문에 그림 4와 같이 작업을 순차적으로 균등하게 배분하기 위한 작업이 필요하다.

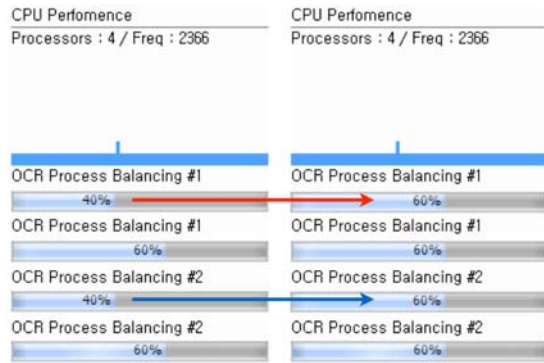


그림 4. 작업 할당을 위한 OCR 프로세스 밸런싱
Fig. 4 Task allocation for OCR Process Balancing

3.3. 하이브리드 후처리 기법

요청된 작업을 수행하는 OCR 프로세스의 결과에 따라 하이브리드 후처리기(Hybrid Post Processor)에서는 필드별 데이터 병합을 위해 표 2에서 정의한 기준으로 점수를 누적하여 가중치를 설정한다.

기본적으로 인식이 되지 않아 공백으로 처리된 필드의 경우 상대 필드에 대한 계산을 수행하지 않고 상대 필드에 인식된 값이 존재할 경우 해당 값을 참조한다. 이러한 생략과정은 양쪽 모두 인식된 값이 동일한 경우에도 적용된다.

표 2의 기준은 기본적인 규칙만 나열한 것이다. 필드별 점수는 해당 필드의 기준이 되는 중요도에 따라 결정된다. 예를 들어 이름의 경우 “홍길동”과 “홍길동”

으로 인식되었다고 가정했을 때 “홍”이라는 성은 존재하지 않기 때문에 인식이 정상적으로 수행되었는지 판단하는 기준에서 중요도가 높다고 볼 수 있다. 또한 경우에 따라서는 점수를 증가해야 상황과 점수를 감소해야 상황이 있다. 예를 들어 회사명의 경우 “(주)”라는 단어로 시작했다면 회사명을 올바르게 인식했을 가능성이 높기 때문에 점수를 증가한다. 반면 회사명에 불규칙적으로 숫자나 특수기호등이 들어간 경우 오인식된 단어가 누적되었을 확률이 높기 때문에 점수를 감소해야 한다.

표 2. 필드에 대한 후처리 규칙

Table. 2 The rules of post-processing for Fields

필드	조건	점수
이름	우리나라 성씨 시작 여부에 따라 증감	±2
	특수기호 포함 여부에 따른 증감	±1
	글자 수가 3자 또는 4자인지 여부	±1
회사명	㈜를 인식했는지 여부	+2
	괄호를 제외하고 특수기호가 들어갔는지 여부	-1
	2개 이상의 문자(영문, 숫자, 특수기호)가 연속성 없이 삽입 되어있는지 여부 (예. 한국정1p동T학회)	-2
부서명	후보 단어로 끝나는지 여부 (예. XX연구소, XX청, XX식당, XX대학교, XX은행 등)	+3
	후보 단어가 포함되었는지 여부 (예. 영업부, 사법부, 총무부, 부설연구소, 개발팀, 디자인 팀 등)	+3
직위	특수기호가 포함되었는지 여부	-1
	후보 단어가 포함되었는지 여부 (예. 부장, 과장, 대리, 대표이사 등)	+3
휴대폰	010, 011 등 휴대폰 번호로 시작하는지 여부	±2
	숫자 수가 11자~13자에 해당하는지 여부	±1
이메일	@를 정상적으로 인식했는지 여부	±2
	@뒤에 도메인이 형식에 맞는지 여부	±1
홈 페이지	숫자로 시작되거나 특수기호가 들어가는 등의 이메일 형식에 부합하지 않을 경우	-1
	http:// 또는 www로 시작했는지 여부	+2
전화	.com, .co.kr 등 도메인 형식으로 끝나는지 여부	±1
	지역번호로 시작되었는지 여부	±2
주소	영문, 특수기호등이 들어가거나 길이가 길어 전화번호 형식에 맞지 않는 경우	-1
	국내 지명으로 시작되었는지 여부	±1
	우편번호가 올바르게 추출되는지 여부	±2

표 3. 실험 이미지에 대한 후처리와 높은 점수를 기준으로 병합한 결과

Table. 3 Accumulated score for the sample image post-processing and High scores based on the result of merging

필드	OCR Type1	점수	OCR Type2	점수	Hybrid	점수
이름	류제경	3	류제경	3	류제경	3
회사명		0	이니텍	1	이니텍(주)	1
부서명	보안사업본부/서비스팀	-1	보안사업본부 서비스팀	1	보안사업본부 서비스팀	1
직위	차장	3		0	차장	3
휴대폰	0010-9911-7xx7002-6445-7090	-3	010-9911-7xx7	2	010-9911-7xx7	2
이메일	jexxxx@gmail.com	3	jexxxx@gmalloom	1	jexxxx@gmail.com	3
전화	002-6445-7000	-2	02-6445-7000	2	02-6445-7000	2
홈페이지	www.initech.com	3	www.initech.com	3	www.initech.com	3
우편번호	152-050	1	152-050	1	152-050	1
주소	서울특별시 구로구 구로동 222-14 에이스하이엔드타워 2차 12층	3	서울특별시 구로구 구로동 222-14 에이스하이엔드타워2차12층	3	서울특별시 구로구 구로동 222-14 에이스하이엔드타워 2차 12층	3

표 3은 그림 5를 인식하여 얻은 결과를 점수 누적 기준에 따라 병합한 결과이다.

인 후처리를 위해서 더욱 많은 실험을 통해 정의된 후처리 규칙과 후보단어를 필요로 한다.



그림 5. 후처리 실험 이미지
Fig. 5 Sample image for post-processing

휴대폰 번호와 전화번호의 경우 Type1에 인식한 결과를 살펴보면 전화, 모바일, 팩스의 구분을 위한 기호를 숫자 0으로 처리하여 핸드폰 번호를 0010으로 인식하였고, 전화번호를 002로 인식하였다. 이것은 전화번호 규칙의 지역번호로 시작되었는지 여부의 규칙에 어긋남에 따라 -2의 점수가 누적되고 인식된 문자의 길이가 일반적인 전화번호 길이에 벗어나 -1의 점수가 추가 누적되었다. 따라서 명함의 병합 프로세스에서 Type2의 인식 데이터를 기준으로 처리하게 된다. 하지만 부서명과 같이 특수기호가 들어간 경우 -1의 점수를 누적시키게 되어 육안으로 보았을 때는 Type1의 인식결과가 올바르지만 Type2의 인식결과가 채택된 것을 볼 수 있다. 따라서 이와 같은 점수 누적 방식에 의한 후처리 기술은 효과적

IV. 모바일 카메라의 명함영상처리

4.1. 일반 명함 촬영 방식

모바일 카메라는 촬영 환경에 대한 변화가 크기 때문에 일정한 밝기와 형태로 촬영하기 어렵다. 그림 6은 명함을 수직상태에서 촬영할 때 실내조명에 의한 그림자가 발생하여 촬영 방향과 조명의 위치에 따라 이진화에 방해요소가 되는 것을 알 수 있다. 따라서 모바일 단말을 기울여 촬영하는 방법을 통해 조명의 방해요소를 일부 개선할 수 있게 된다.



그림 6. 일반 명함 촬영 방식
Fig. 6 Method of the general shooting

4.2. 그림자 회피 촬영의 개선

본 논문에서는 기존의 연구[10]에서 그림자를 회피하기 위한 목적으로 명함을 촬영할 때 모바일 단말의 기울기에 따라 촬영 프레임을 사다리꼴 형태로 시각화 하여 사용자가 명함을 프레임에 맞춰 촬영하도록 유도하는 방식에서 향상된 기법으로 실시간으로 입력되는 프리뷰 영상에서 Canny Edge 알고리즘을 이용해 외곽영역을 검출하여 사용자가 어떤 방향으로 촬영해도 명함의 영역이 검출되었음을 시각화 되도록 개선하였다.



그림 7. 기존 그림자 회피 촬영 방식
Fig. 7 Method of shooting for the avoid shadows

그림 7의 기존 방식을 개선한 그림 8의 촬영 방식은 카메라의 프리뷰 상황에서 실시간으로 엣지를 검출하고 사용자에게 명함의 영역이 인지될 수 있도록 시각화하는 개선된 방법이다.

본 논문에서는 안드로이드 환경에서 OpenCV를 활용하여 구현되었으며 실시간으로 검출된 명함 영역은 촬영 직전까지의 좌표를 기억하여 촬영 이후에 명함 영역을 추출하면 투영변환을 통해 명함을 평판 스캐너에서 출력된 영상과 같은 형태로 변환 할 수 있다.



그림 8. 개선된 명함 촬영 방식
Fig. 8 Improved method shooting business card

명함 영역의 검출을 위해서는 프리뷰 영상에서 입력되는 YCbCr 색공간 영상을 RGB 채널로 변환하여 그림 9과 같이 각 채널을 분리해야 한다.

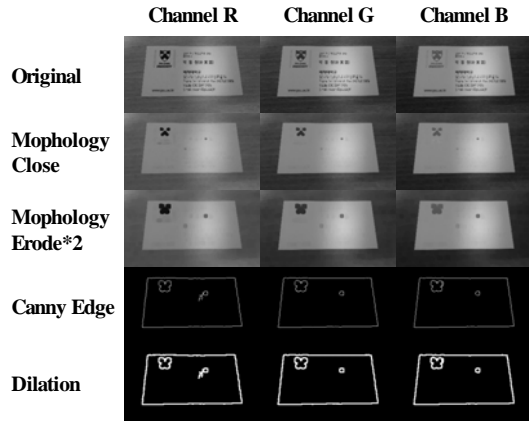


그림 9. 프리뷰 영상에서 외곽선 검출
Fig. 9 Edge Detection from Preview Image



그림 10. 명함의 투영변환 과정
Fig. 10 Projective transform of business card

컬러영상을 바로 그레이스케일로 변환하여 처리하지 않는 이유는 Canny Edge 알고리즘을 적용할 때 명함과 배경의 색상의 차이가 있더라도 컬러 영상과 달리 8비트 영상에서는 배경과 명함의 영역의 구분이 모호하여 엣지 검출이 제대로 되지 않을 수 있기 때문이다. 본 논문에서는 명함과 배경을 단순화 하여 엣지 검출 시 명함의 영역을 효과적으로 구분하기 위하여 분리된 채널별 영상을 모폴로지(Mophology) 기법인 침식과 팽창 연산을 반복하여 배경과 명함 영상을 단순화 하였다.



그림 11. 명함 영상의 이진화 기법 (a) 원본 영상 (b) Sauvola 이진화 (c) 적응형 이진화
 Fig. 11 Binarization for Business Card Images (a) Original Image (b) Sauvola Thresholding (c) Adaptive Thresholding

일반적으로 영상의 사각영역을 추출하기에 앞서 Smooth 또는 Blur 필터를 사용하여 영상의 잡음을 제거하는데 모폴로지 기법을 적용해본 결과 잡음을 제거하거나 영상에서 객체의 모형을 표현하는데 더욱 효과적인 결과를 얻을 수 있었다. 영상을 단순화하는 단계를 마치면 Canny Edge 알고리즘을 통해 외곽선을 검출한다. 특히 검출된 선 사이에 빈공간이 발생할 수 있기 때문에 그림 9의 마지막 단계와 같이 팽창 연산을 통해 빈공간을 채워야 한다. 검출된 외곽선은 투영변환을 통해 그림 10과 같이 보정할 수 있다.

4.3. 그림자 제거를 위한 이진화

본 논문에서는 앞서 변환된 영상을 통해 Sauvola[11] 이진화 알고리즘과 OpenCV의 AdaptiveThreshold[13] 알고리즘을 구현하여 모바일 단말에서 촬영된 영상에 두 알고리즘을 적용하여 인식률을 비교해 보았다.

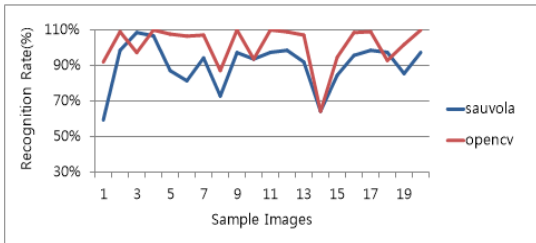


그림 12. 두 이진화 알고리즘의 인식률 비교
 Fig. 12 Recognition rate of two Binarization algorithms

Sauvola 알고리즘의 경우, 그림 11처럼 선명한 문자 표현으로 작은 글씨의 뭉개짐을 완화할 수 있는 장점이 있으나 두꺼운 글자나 심벌의 내부 영역을 침투하는 현상이 발생할 수 있어 이름이나 회사명의 인식률이 AdaptiveThreshold에 비해 저조한 단점이 있었다.

V. 성능분석

본 논문에서 제안한 하이브리드 OCR 에이전트의 성능을 살펴보기 위해 총 100장의 명함을 모바일 단말에서 촬영한 명함 이미지를 하이브리드 OCR 에이전트에서 인식하였으며, 두 형태의 인식엔진에서 추출된 인식 결과와 하이브리드 후처리기를 통해 병합된 결과로 Type1, Type2, Hybrid로 분류하였다. 인식률의 평가는 각 필드별로 인식한 글자 수를 대상으로 백분율에 따라 명함단위의 인식률을 계산하였다.

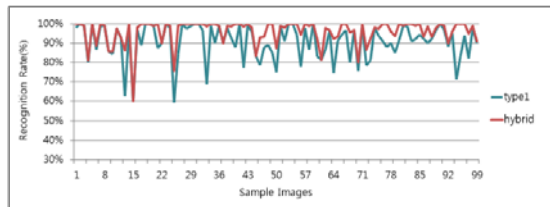


그림 13. Type1 OCR과 Hybrid의 결과 비교
 Fig. 13 Comparing the results of Hybrid and Type1

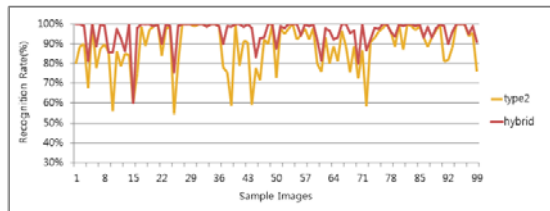


그림 14. Type2 OCR과 Hybrid의 결과 비교
 Fig. 14 Comparing the results of Hybrid and Type2

그림 13, 14는 각 알고리즘의 인식률과 하이브리드 후처리기를 통한 결과 차트이다. Type1 대비 평균 5%, Type2 대비 평균 9% 향상된 것을 알 수 있었다.

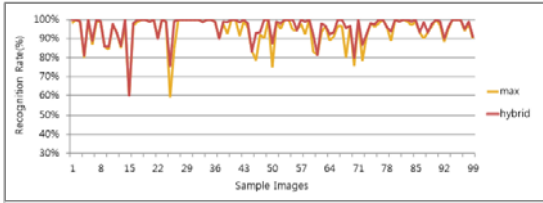


그림 15. 최상의 인식률과 하이브리드의 비교
Fig. 15 Comparison of the best case and hybrid

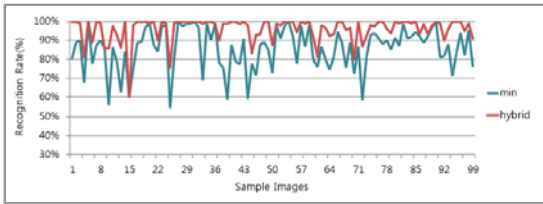


그림 16. 최악의 인식률과 하이브리드의 비교
Fig. 16 Comparison of the worst case and hybrid

표 4. 하이브리드 인식 성능의 결과
Table. 4 Result of hybrid recognition performance

구분	type1	type2	min	max	hybrid
평균 인식률	90.69%	86.50%	83.68%	93.51%	95.52%
hybrid 대비	4.83%	9.02%	11.84%	2.01%	0%

그림 15, 16의 차트는 Type1과 Type2에서 각 명함단위로 최상의 인식률과 최악의 인식률의 케이스를 골라 하이브리드 방식이 두 결과 대비 어느 정도의 향상효과를 얻을 수 있는지 분석해보았다. 그 결과 표 4와 같이 최악의 인식률 대비 12%의 성능향상이 있었고 최상의 인식률 대비 2%의 성능향상이 있었다. 따라서 인식률

표 5. 스마트폰에서의 인식 성능 비교
Table. 5 Compared recognition performance on smart-phones

구분	3G(KT)		3G(SKT)		LTE(SKT)		LTE(LG)		KT(Wibro 4G)	
	갤럭시S2		갤럭시S		갤럭시S3		옵티머스G Pro		아이폰4	
	인식	전송	인식	전송	인식	전송	인식	전송	인식	전송
명함1(65.1KB)	0.703	1.134	0.568	1.401	0.506	1.021	0.566	1.242	0.542	1.742
명함2(72.2KB)	0.945	1.572	0.683	1.905	0.551	1.752	0.675	2.205	0.597	2.105
명함3(58.3KB)	0.724	1.124	0.623	1.624	0.343	1.381	0.355	1.714	0.502	1.652
명함4(53.1KB)	0.721	1.037	0.572	1.519	0.421	0.984	0.461	1.029	0.502	1.015
명함5(59.8KB)	0.687	1.348	0.712	1.02	0.593	1.325	0.612	0.75	0.621	0.758
평균(Sec)	0.76	1.24	0.63	1.49	0.48	1.29	0.53	1.39	0.55	1.45
전체(Sec)	2		2.12		1.77		1.92		2	

을 향상 시키는데 하이브리드 방식이 단일 엔진을 구성하였을 때 보다 높은 인식결과를 얻을 수 있다는 사실을 알 수 있다.

최종적으로 모바일 단말에서 앞서 평가된 인식 결과를 얻기 위해 3G, LTE, Wibro로 분류된 모바일 네트워크 환경에서의 실험 결과 표 5에서와 같이 본 논문에서 제안한 영상처리에 의한 이진화 전처리 이미지를 전송하고 서버에서 명함을 인식하는데 걸린 시간이 약 2초로 측정되었다. 이미지 전송과 명함 인식외의 일련의 통신과정을 포함하면 단말기의 환경과 서버와의 통신량에 따라 시간이 늘어날 수 있다.

VI. 결론

본 논문에서는 모바일 환경의 하드웨어 사양에서 명함 인식 성능의 향상을 위해 모바일 단말이 갖는 연산 처리 능력의 한계를 극복할 수 있는 서버 기반의 하이브리드 OCR 에이전트와 모바일 카메라에서의 영상처리 기법에 관한 2가지 방안을 제안하였다. 하이브리드 OCR 에이전트는 서로 다른 알고리즘과 학습 데이터를 갖는 2가지 형식의 명함 인식 엔진을 병렬로 처리하도록 구성하여 필드별로 높은 인식결과를 얻을 수 있도록 하는 후처리 기법을 통해 인식률이 최저 인식률 대비 12%의 향상을 보였다. 또한 모바일 카메라의 특성상 발생할 수 있는 조명과 촬영 방향에 상관없이 사용자가 명함영역이 정상적으로 감지되는 것을 시각화하여 명함의 인식영역을 효율적으로 추출할 수 있으며, 조명에 의해 발생한 그림자 영상의 제거에 효율이 높은 이진화 방법을 사용하여 명함 인식 성능 향상을 극대화하였다.

본 논문에서 제안한 하이브리드 OCR 에이전트는 클라우드 컴퓨팅, 빅데이터와 같은 미래의 컴퓨팅 기술의 발전에 대비할 수 있으며, 모바일 단말에 독립적으로 동작되는 명함인식 응용프로그램에 비해 연산 수행 빈도를 줄일 수 있다. 또한 인식엔진의 학습 데이터 증가로 인한 업데이트가 지속적 발생해도 모바일 단말이 갖는 네트워크 사용 빈도를 줄일 수 있게 된다. 그 동안 문자인식기술의 SDK 가격이 고가의 가격으로 유통되어 온 것을 감안했을 때 하이브리드 OCR 에이전트의 서버 환경으로 인하여 명함인식기술의 웹서비스화가 가능해짐에 따라 다양한 응용 분야에 활용되어 모바일 앱 시장의 변화에 기여할 수 있다.

감사의 글

본 연구는 2013년도 한밭대학교 교내학술연구비 지원 사업에 의해 연구되었으며, ㈜한국인식기술로부터 지원된 2종류의 명함인식엔진을 사용하였습니다.

REFERENCES

- [1] Character and speech recognition, and future prospects related technologies Available: <http://www.connectinglab.net/wordpress/?p=9608>
- [2] Mollah, A. F., Majumder, N., Basu, S., & Nasipuri, M. "Design of an Optical Character Recognition System for Camera-based Handheld Devices," in *International Journal of Computer Science Issues*, vol. 8, issue 4, no. 1, pp. 283 - 289, July, 2011.
- [3] Beginning of Mobile Cloud Available: http://www.imaso.co.kr/?doc=bbs/gnuboard.php&bo_table=article&wr_id=38258
- [4] Gao, Y., Jin, L., He, C., & Zhou, G. "Handwriting Character Recognition as a Service: A New Handwriting Recognition System Based on Cloud Computing," in *International Conference on Document Analysis and Recognition*, pp. 885 - 889, Beijing, Sept, 2011.
- [5] Huerta-Canepa, G., & Lee, D., "A virtual cloud computing provider for mobile devices," in *Proceedings of the 1st ACM Workshop on Mobile Cloud Computing & Services Social Networks and Beyond*, New York: NY, pp. 1-5, 2010.
- [6] Hyunsub, S., HIART Corporation, *Cloud OCR business card information management system*, KR1020120111954, 2013.
- [7] Hsueh, M. "Interactive Text Recognition and Translation on a Mobile Device," University of California, Berkeley: CA, Technical Report UCB/EECS-2011-57, 2001.
- [8] Blem, E., Menon, J., & Sankaralingam, K., "A detailed analysis of contemporary arm and x86 architectures," University of Wisconsin-Madison, Technical Report, Feb, 2013.
- [9] Context switch Available: http://en.wikipedia.org/wiki/Context_switch
- [10] Yeiyong, J., Hyunsub, S., Chajong, K., "A Development of Mobile Camera Application for Avoiding Shadows in Business Card from Smart Device," in *Proceedings of the KIIT Summer Conference*, pp. 306-311, 2013.
- [11] Sauvola, J., & Pietikäinen, M., "Adaptive document image binarization. Pattern Recognition", pp. 225-236, 2000.
- [12] Hyunsub, S., Chajong, K., Injun, S., "A Study on the Business card Web Service for Smartphone," in *Proceedings of the KIIT Autumn Conference*, pp. 1-6, 2011.
- [13] Miscellaneous Image Transformations Available: http://docs.opencv.org/modules/imgproc/doc/miscellaneous_transformations.html



신현섭(Hyunsub Shin)

2007년 한밭대학교 정보통신전문대학원 컴퓨터공학과 (공학석사)
 2014년 한밭대학교 정보통신전문대학원 컴퓨터공학과 (공학박사)
 ※관심분야 : 명함인식, SNS, 클라우드 컴퓨팅, 모바일, 영상처리



김차종 (Chajong Kim)

1986년 광운대학교 대학원 컴퓨터공학과 (공학석사)

1991년 광운대학교 대학원 컴퓨터공학과 (공학박사)

※관심분야 : Web Services, SNS, 공학교육, 복잡계이론