

# OCR기반의 자동 문자인식\*

임윤지<sup>○</sup>, 에다 카츠토시, 황석형, 김응희, 김민경

선문대학교 글로벌소프트웨어학과

{Yunjir128, katsu97625, shwang, minkyoungkim, ehkim}@sunmoon.ac.kr

## Automatic text recognition system based on OCR

Yunji Lim<sup>○</sup>, Kastsutoshi Eda, Suk-Hyung Hwang, Eung-Hee Kim, Minkyoung Kim

Department of Global Software, Sunmoon University

### 요 약

서류 보관에 대한 중요성이 높아짐에 따라 문서 전자화 시스템을 도입하고 있는 곳이 증가하고 있다. 이러한 가운데 수작업 위주의 입력 저장 방식을 혁신하기 위한 기술로 OCR 이 부상하였다. 이에 따라, OCR 기술을 활용하여 영수증 자동인식을 위한 연구가 진행되어 왔으나, 인식한 텍스트의 정확성이 높지 않다. 따라서, 본 논문에서는 문자 인식률을 높이기 위한 키스톤 보정 처리기술을 적용하여, 기존 전자 가계부 이용자들이 직접 수기로 일일이 입력하는 불편을 개선하고자 한다.

### 1. 서 론

서류 보관에 대한 중요성이 높아짐에 따라 문서 전자화 시스템을 도입하고 있는 곳이 증가하고 있다. 이러한 가운데 수작업 위주의 입력 저장 방식을 혁신하기 위해 OCR(광학문자인식)이 부상하였다. 이 기술은, 스캐너를 통해 입력된 문서 이미지에서 문자에 해당하는 부분의 내용을 인식하는 기술을 말한다. OCR은[2] 동일한 하나 이상의 문자의 복수의 이미지를 캡처 및 디코딩하여 복수의 OCR 디코딩 결과를 획득하는 단계 및 각각의 OCR 디코딩 결과가 복수의 위치를 포함시키는 과정들의 연구를 통해, 인식결과의 신뢰성이 향상되어 왔다. 이에 따라, OCR 기술을 활용하여 영수증 자동 인식을 위한 연구[3]가 진행되어 왔으나, 인식한 텍스트의 정확성이 높지 않다는 문제가 있다.

본 논문에서는, 입력의 편의성을 개선하고 문자 인식률을 높이하고자 OCR에 기반한 키스톤 보정 처리 기술을 적용하는 방법을 제안한다. 또한, 정확한 키스톤 보정 처리를 위해 이진화 과정을 세분화하는 과정을 제시한다. 제안한 방법의 성능평가를 위해, 영수증의 카메라 이미지로부터 문자 인식의 정확도를 측정하였다. 실험결과, 기존의 OCR엔진의 문자 인식률을 89%에서 99%로 향상시켰다. 또한, 키스톤 보정 처리 기술을 적용함으로써, 영수증을 잘못 찍더라도 OCR엔진이 영수증을 잘 인식하는 결과를 얻을 수 있었다.

### 2. 관련 연구

ABBYY의 TextGrabber[4], Goggle의 Cloud Vision API[5] 등은 OCR엔진을 연구, 개발한 대표적인 사례들이다. 먼저, TextGrabber는 ABBYY에서 제공하는 최신 문서 OCR, ICR, OMR, OBR, 문서 이미징 및 PDF 변환 기술을 서버, 데스크탑, 모바일, MFP 또는 기타 장치의 다양한 응용 프로그램에 통합하는 강력한 OCR SDK를 가지고 있는 OCR엔진이다. Google Vision API는 학습된 기계학습 모델을 사용해 이미지의 내용을 파악하고 이미지안의 개별 객체들을 감지하고 추출해주는 OCR엔진이다.

본 연구에서는 Google Cloud Vision API의 OCR엔진을 활용한다. 이 엔진을 처음 사용한 결과, 인식률이 89%로 나타날 정도로, 한글에 대한 인식률이 높고, 활용을 위한 인터 페이스가 잘 정의되어 있기 때문이다. 본 논문에서는 해당 인식률을 높이기 위해, OCR엔진의 인식단계 이전에 키스톤 보정 처리를 함으로써 영수증의 문자 인식률을 향상시키고자 한다.

기존의 영수증 가계부 어플리케이션에는 꼼꼼 영수증, 비주얼 가계부, 레픽 등이 있다. 꼼꼼 영수증은 영수증을 찍으면 영수증 이미지 저장을 해주고 직접 영수증의 내용을 입력해야 하는 응용 프로그램이다. 비주얼 가계부는 영수증의 이미지만을 저장해준다. 본 연구와

\* "본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학지원사업의 연구결과로 수행되었음"(2018-0-01865)

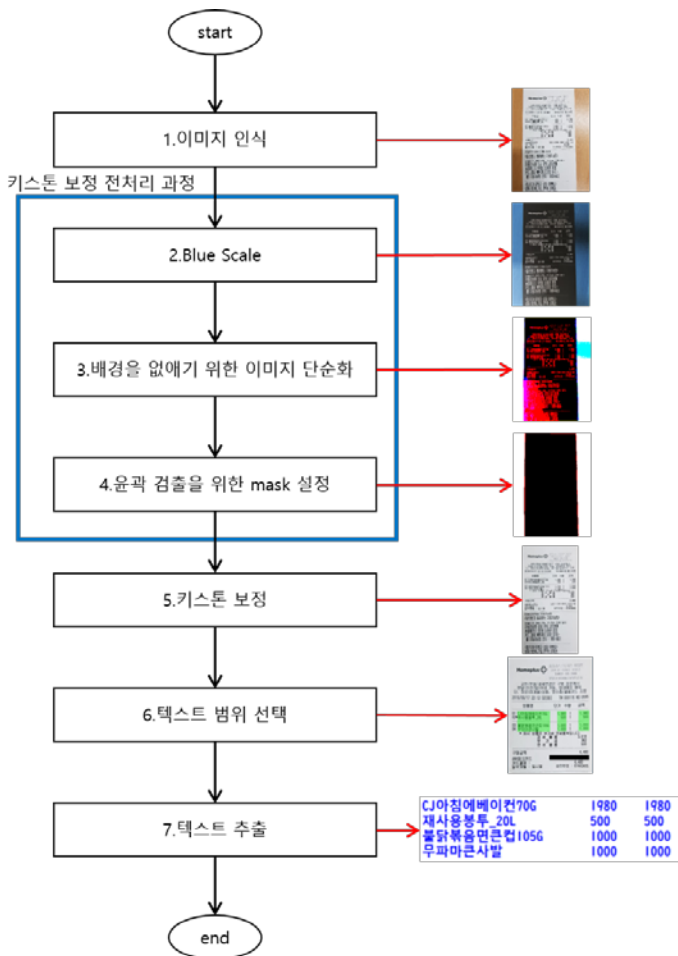


그림1. 시스템 처리 과정 및 결과

가장 유사한 어플리케이션인 레픽은 영수증을 찍으면 영수증의 내용을 인식하여 추천, 가계부, 그래프 등의 서비스를 제공하지만 영수증을 인식하는데 시간이 오래 걸린다. 그 이유는, 타이피리스트들이 사용자가 찍은 영수증을 보고 직접 영수증의 내용을 타자로 치는 수동적 방식으로 어플리케이션을 운영하고 있기 때문이다. 따라서, 본 연구가 기존의 유사 어플리케이션과 차별화되는 점은 영수증을 촬영하여 촬영된 이미지를 OCR엔진을 사용하여 자동으로 인식할 수 있도록 하여 입력의 편의성을 개선했다는 점이다.

### 3. 연구 방법



그림2. 키스톤 보정 처리 전후 영수증



그림3. 텍스트 인식 영역 지정

그림1은 제안 시스템의 처리 과정을 나타낸 것이다. 시스템의 처리과정은 크게 3단계로 구분된다. 첫째, 영수증 이미지 인식, 둘째, 키스톤 보정을 위한 전처리과정 및 키스톤 보정과정, 셋째, 텍스트 인식 및 추출 과정이 그것이다.

첫번째 단계에서는, 카메라로 영수증의 이미지를 인식한다. 두번째 단계에서는, 키스톤 보정 처리를 하기 위해서 OpenCV를 사용하여 (1)영수증 이미지를 인식하고, (2)BlueScale[6]로 바꿔준다. Gray Scale은 인식한 영수증 이미지에서 배경은 흰색으로, 내용은 검정색으로 단순화시키는 과정에서 배경과 내용을 부분적으로 오인식하는 오류를 초래하기 때문이다. (3) 변환한 이미지를 한 픽셀당 8비트 기준으로 0~255 값을 갖는 이미지를 주어진 임계값을 기준으로 임계값 이하는 0, 초과는 255로 변환해준다. 그 다음 (4)단순화한 영수증 이미지를 mask로 설정해 영수증의 윤곽을 검출한다. (5)키스톤 보정 과정에서는, 영수증이 화면속에서 벗어난 상태이면 잘 검출하지 못하기 때문에 첫번째 단계에서 원래의 이미지보다 큰 이미지를 준비해 배경을 확대해서 문제를 해결하였다. 검출한 윤곽을 바탕으로 영수증 내용만을 잘라낸 이미지를 저장한다. 그리고, 잘라낸 영수증에서 필요한 부분만을 인식할 수 있도록 [그림3]과 같이 (6)텍스트 영역을 선택하도록 하였다. 선택한 텍스트의 영역에서 (7)텍스트를 추출해 Google Vision API의 OCR엔진을 활용하여 문자를 식별하고 리스트화 하였다([그림4]). 이 과정에서 기존 OCR엔진의 인식률과 인식한 텍스트의 정확도를 개선하기 위해, 화면의 모서리 부분을 강제적으로 이동시키듯이 투사하여 원래의 형태인 사각형에 가깝도록 화면을 조정하는 기능인 키스톤 보정 처리 기술을 적용하였다[7].

본 연구에서는, [그림2]의 상단 이미지와 같이, 사용자가 영수증을 주의를 기울여 찍지 않더라도, 인식률을 최대한 높일 수 있도록 키스톤 보정 처리 기술을 적용함으로써, 기존의 연구[3]보다 인식률을 6.5% 향상시켰다.

01CJ아침에베이컨70G	1980	1980	싱글레귤러콘	3200
재사용봉투_20L	500	500	>민트초콜릿칩	0
불닭볶음면큰컵105G	1000	1000	싱글레귤러칩	3200
무파마근사발	1000	1000	>민트초콜릿칩	0
			싱글레귤러칩	3200
			>레인보우사베트	0
			싱글레귤러칩	3200
			>31-요거트	0
			싱글레귤러칩	3200
			>그린티	0
			싱글레귤러칩	3200
			>월넛	0
				0
			옥수수식빵	2700
			슈	800

그림4. 이미지를 텍스트로 인식한 결과 화면

금액	
01CJ아침에베이컨70G	
02*재사용봉투_20L	
1980	
500	
19801	
5001	
03불닭볶음면큰컵105G	
04무파마근사발	
*표시상품은	
10001	
10001	
1000	
1000	
부가세면세품목입니다	
과세_물품	
세	
면세물품	
3618	
362	
500	

그림5. 기존의 연구결과와의 비교 (왼쪽: 기존의 OCR엔진 인식결과, 오른쪽: 기존의 연구결과[3])

#### 4. 실험 결과

[그림5]에서 보여주는 바와 같이, 기존 OCR엔진은 영수증을 인식할 때 ‘10001’, ‘5001’, ‘과세\_물품’ 등 영수증의 내용을 제대로 인식하지 못하는 경우가 많다. 또한, 기존 연구결과[3]도 부, 가세, 바, 그매, 신한카드-채크 등 영수증의 텍스트 인식 정확도가 높지 않음을 확인할 수 있다. 하지만, 본 연구에서는 이러한 오인식을 해결하고, 사용자가 보기 쉽게 인식결과([그림4])를 상품과 단가, 금액 순으로 리스트화하였다. 이는, 기존의 OCR기술에 기반하여 키스톤 보정처리를 개선, 적용함으로써 텍스트의 인식률과 정확성을 향상시킨 결과이다. 사용자가 영수증을 잘못 찍을 경우를 포함하여 각기 다른 형식의 영수증 20장을 대상으로 실험한 결과, 키스톤 보정을 통해 사용자가 영수증을 잘못 찍어도 영수증 텍스트를 정확하게 인식할 수 있도록 하는 결과를 얻었다. 또한 기존 연구의 정확도[3] 보다 6.5% 향상된 99%의 정확도를 얻을 수 있었으며 표준편차도 1.47%로 각기 다른 형식의 영수증을 고르게 인식하는 결과를 얻었음을 알 수 있었다.

#### 5. 결론

본 논문에서는, 시중에 출시된 수동 입력식 가계부와 차별화되는 OCR기반의 영수증 인식시스템을 개발하여, 사용자가 영수증을 잘 못 찍어도 영수증을 잘 인식하도록 하였다. 그리고 기존의 연구결과보다 텍스트의 정확성(표준편차: 1.47%)을 6.5% 향상시켰다. 즉, 비정형화된 이미지로부터 문자 인식률을 향상시킬 수 있었고, 본 논문의 연구결과는 보다 정확한 OCR엔진을 필요로 하는 분야에 기여할 수 있을 것이라 기대된다.

향후에는 상품명 DB[8]를 연동하여 영수증 상품의 인식률을 더욱더 높여, 가계부를 수동으로 입력해야하는 불편함을 한층 더 개선시키고자 한다. 이용자가 증가하게 되면, 더 많은 이용로그를 쌓게 되어 데이터가 풍부해짐으로써 더 정확한 텍스트를 제공할 수 있을 것이라 예상한다.

#### 6. 참조문헌

- [1] 이종원, 여일연, 정회경. (2019). 문서 분석 기반 주요 요소 추출 시스템. 한국정보통신학회논문지, 23(4), 401-406
- [2] Neil Corporation, Orlando, FL (US) (72) Inventor: H. Sprague Ackley, Seattle, WA (US) Appl. No.: 15 / 720, 550 Sep. 29, 201 METHODS FOR OPTICAL CHARACTER RECOGNITION (OCR) Applicant: Datamax - O<sup>1</sup>
- [3] 채성은, 정기석, 이정렬, 노영주(2018). OCR기반의 스마트 가계부. 한국인터넷방송통신학회논문지, 18(6), 269-276
- [4] TextGrabber, ABBYY. Retrieved at <http://www.textgrabber.pro/en/> on Oct. 24, 2019
- [5] Google Cloud Vision API: OCR, Google. Retrieved at <https://cloud.google.com/vision/docs/ocr> on Oct. 24, 2019.
- [6] 김종수, 김형배, 남부희. (2010). 이미지 이진화를 사용한 자동차 번호판 문자 검출. 대한전기학회 학술대회 논문집, (), 263-264.
- [7] 어성욱, 홍욱, 이준구. (2017). 모바일 프로젝터에서 화면 비율을 유지하는 디지털 키스톤 보정. 대한전자공학회 학술대회, 653-656.
- [8] 김수현, 이송희, 이상준, 이상호. (2017). 광학 문자인식(OCR)을 통한 가계부 저장 서비스. 한국정보과학회 학술발표논문집, 377-379.

시연동영상 URL

<https://youtu.be/dshSYou7jD8>