

# OpenCV 를 이용한 검침기 문자 인식 시스템

남승완\*, 이계환\*, 황광일\*  
 \*인천대학교 임베디드시스템 공학과  
 e-mail : seungwany3@gmail.com

## Character Recognition System in Meter Reading with using OpenCV

Seung-wan Nam\*, Gye-Hwan Lee\*, Kwang-il Hwang\*  
 \*Dept. of Embedded Systems Engineering, Incheon National University

### 요 약

본 논문에서는 OpenCV 에서 제공하는 라이브러리 중 K-Nearest Neighbor 알고리즘을 이용하여 검침기 안에 문자들을 인식하는 방법을 제안하였다. 텍스트 이미지에서 인식률은 정확하였으나, 실제 검침기 사진에서 취약한 인식률을 보였다. 그러나 기계 학습을 통한 영상처리가 가능하다는 점과 정확성 있는 학습 데이터들만 확보가 된다면 매우 전망이 높은 분야일 것으로 판단된다.

### 1. 서론

현재는 전기, 수도, 가스 등의 사용량을 알기 위하여 계량기의 숫자를 검사하는 검침원들이 각 가정의 검침 값을 측정하고 있다. 이 과정 속에선 많은 인력과 비용이 소모된다. 이 많은 인력과 비용을 줄이기 위하여 검침 시스템에 대한 연구는 활발히 진행되어오고 있으며, 검침기 안에 문자들을 정확하게 인식하는 방법들이 다양하게 시도되고 있다. 이러한 방법들 중엔 형태분석 또는 템플릿 매칭으로 인식하는 Tesseract-OCR[1], 이미지 픽셀의 분포를 분석하여 인식하는 PCA[2] 방법 등이 있다. 본 논문에서는 라즈베리파이에서 OpenCV(영상 처리 라이브러리)의 K-Nearest Neighbor 알고리즘[3]을 이용하여 검침기에 문자를 인식하는 방법을 제안한다.

### 2. 연구 내용

#### 2.1 검침기 문자 영상 처리

검침기 문자 영상 처리에 사용되는 OpenCV(Open Computer Vision)[4]는 인텔에서 개발한 실시간 이미지 프로세싱에 중점을 둔 라이브러리이다. OpenCV 에는 기초 영상 처리부터 고급 수준의 영상 처리까지 상당한 양의 알고리즘들이 함수로 구현이 되어있다. 본 논문에서는 OpenCV 에서 제공하는 라이브러리들을 이용하여 4 단계 과정[5]을 거쳐 이미지의 잡음제거 및 필터링, 윤곽선 검출을 구현하였다.

##### 2.1.1 그레이 영상 처리 (GrayScale)

미리 필요한 부분을 캡처 한 이미지는 컬러 정보를 포함하고 있다. 그러나 본 논문에서는 불필요한

컬러 정보를 제거하여 계산량을 줄이기 위해 원본 이미지(RGB)를 그레이 영상으로 변환하여 처리한다.

##### 2.1.2 노이즈 제거 (Blurring)

노이즈(잡음) 픽셀들만 정확히 발견하여 제거하는 방법이 있다면 좋겠지만 이런 방법을 구현하기엔 거의 불가능하다. 따라서 모든 픽셀에 대하여 주변 픽셀들과 비슷한 색을 갖도록 하는 것이 블러링 방법이다. 블러링 하는 방법은 여러 종류가 있지만 본 논문에서는 여러 가지 블러링 방법을 적용한 결과 가우시안 블러링 방법이 가장 적합하여 적용을 하였다.

##### 2.1.3 이진화 방법 (Binarization)

이진화 방법은 설정한 임계 값을 기준으로 픽셀 값이 임계 값보다 크면 결과 영상의 같은 위치의 픽셀 값을 흰색(1)로 하고, 그 반대이면 검은색(0)으로 한다. 본 논문에서는 여러 이진화 방법 중에서 주변 픽셀 값에 의해 결정되는 임계 값을 이용한 영상 이진화 방법을 적용하였다.

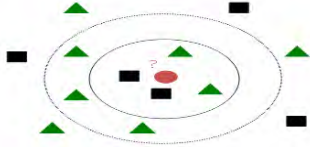
##### 2.1.4 윤곽선 검출

영상을 처리하는데 있어서 중요한 목적 중 하나는 객체들을 식별하고 추출하는데 있다. 따라서 다양한 윤곽선 검출하는 방법들을 통해 객체를 식별 및 추출하는데, 본 논문에서는 OpenCV 에서 제공하는 라이브러리 중 findContours 라는 알고리즘을 적용하였다.

#### 2.2 K-Nearest Neighbor(K-NN) 알고리즘을 통한 문자 인식

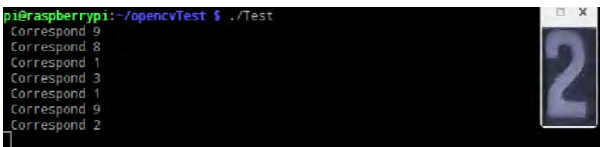
K-NN 알고리즘이란 분류 알고리즘이다. 예를 들어 (그림 1)와 같이 빨간색 원이랑 가장 근접한 이웃을 찾는다면 당연히 검은색 사각형이 된다. 그러므로

빨간색 원은 검은색 사각형으로 분류 기준에 적합하게 된다. 그런데  $k=7$  즉, 7개의 이웃과의 거리를 비교한다면 5개의 초록색 삼각형과 2개의 검은색 사각형이 가까이에 있기 때문에 빨간색 원은 초록색 삼각형 클래스에 속하게 된다. 이 방법을 K-최근접 이웃 알고리즘(K-Nearest Neighbor)이라 부른다.



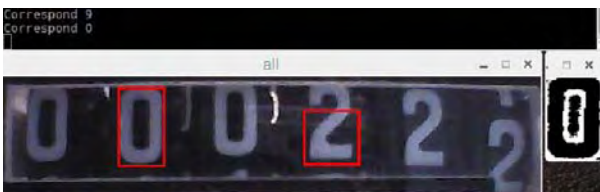
(그림 1) K-Nearest Neighbor 알고리즘

이 원리를 통하여 K-NN 알고리즘이 문자를 인식하는데 사용된다. 본 논문에서는 0부터 9까지 숫자 이미지들(학습 데이터)을 가지고 시작을 하였다. 전처리 과정을 거친 각 이미지들은 K-NN 알고리즘에서 각 이미지들은 소수 점 값으로 분류 된다. 각 이미지들이 알맞은 숫자로 분류가 되었는지 확인하기 위하여 무작위로 숫자 이미지를 읽어와서 인식하는 테스트를 진행하였다. 또한, 최종적으로 여러 숫자들이 섞여 있는 이미지를 가지고 테스트를 진행하였다.



(그림 2) 검침기 개별 문자 인식 과정

(그림 2)에서는 각 학습된 검침기 사진들을 개별적으로 테스트하는 과정인데, 테스트 결과는 잘 인식하여 성공적이었다. 개별적인 숫자 이미지 테스트 후엔 최종적으로 여러 숫자들이 섞여있는 전체 검침기 사진을 가지고 (그림 3), (그림 4)와 같이 진행하였다.



(그림 3) 전체 검침기 문자 인식 과정 1

(그림 3)에선 숫자 0은 정확하게 Rect를 잡아서 인식하지만 숫자 2는 (그림 4)에서와 비슷하게 Rect를 부정확하게 잡는다.



(그림 4) 전체 검침기 문자 인식 과정 2

(그림 4)에선 숫자 7은 인식이 되었으나, 숫자 9는 부정확하게 Rect를 잡아 6으로 인식하는 모습이다. 오류의 원인을 정확하게 검증하려면 조금 더 많은 사진들을 가지고 진행을 해야 할 수 있겠지만, 전처리 과정에서 숫자 부분 추출, 부정확한 검침기 사진 등으로 추정 되어진다.

### 3. 성능 평가

본 장에서는 K-NN 알고리즘을 이용하여 실험한 검침기 문자 이미지 인식률을 살펴 보았다. 그러나 실제 검침기 사진으로 실험을 하였을 경우엔 개별적인 숫자 이미지는 인식을 하였으나, 전체적인 검침기 문자 인식은 극히 일부만 성공하였다. 이러한 오류들을 개선하기 위해 향후 연구 방향을 학습할 이미지 파일의 개수를 늘려서 진행하는 방향도 고려하여 진행할 예정이다. 다양한 학습 데이터들을 기반으로 한 KNN 알고리즘은 좋은 성능을 보여줄 것으로 기대된다.

### 4. 결과 및 고찰

영상처리 분야에서 기계학습은 오랜 기간 동안 함께 연구되어 왔다. 기계학습과 영상처리를 같이 수행하면 다양한 작업들이 가능하다. 그 결과, OpenCV에서 제공하는 KNN 기계 학습 알고리즘을 사용하여 기존에 학습한 데이터를 기반으로 새로운 입력 영상에 대해 예측하여 인식을 시도하였다. 비록 실험한 결과 검침기에서 인식률은 높지 않았지만, 기계학습을 통한 문자인식이 구현되었다는 점에서 향후 지속적인 연구를 통해 정확성 있는 학습 데이터를 늘려가면서 인식률을 높인다면 매우 전망이 밝은 분야라고 판단된다.

사사 : 본 논문은 2016년도 중소기업융복합기술개발사업 (과제명: LoRaWAN 표준기반 Application Protocol Framework 및 Meter OCR 최적화에 관한 연구)에 의해 지원되었음.

### 참고문헌

- [1] Anthony Kay, "Optical Character Recognition by Open Source OCR Tool Tesseract: A Case Study," International Journal of Computer Applications, Vol. 55, No. 10, pp. 50-56, 2012.
- [2] 유재만, 한정훈, 김우생, "PCA 를 이용한 온라인 문자인식 기법," Journal of Korea Multimedia Society Vol.9, No.4, April 2006.
- [3] Maurício Marengoni, Denise Stringhini, "High Level Computer Vision using OpenCV," 24th SIBGRAPI Conference on Graphics, Patterns, and Images Tutorials, 2011.
- [4] S.Hwang, Learning OpenCV(Korean translated), Hanbit Media, Seoul, 2010.
- [5] Manisha V.Shinde, Pradip W.Kulkarni, "Reading of Energy Meter based on Image Processing Technology," 1st International Conference at SITS, Narhe, Pune on April 5-6, 2014.
- [6] OpenCV, <http://opencv.org>