



영상처리 기법을 이용한 문자인식 시스템 개발

Development of character recognition system based on the image processing techniques

저자 (Authors)	이원용 Won-Yong LEE
출처 (Source)	융복합지식학회논문지 5(2) , 2017.7, 99-103 (5 pages) The Society of Convergence Knowledge Transactions 5(2) , 2017.7, 99-103 (5 pages)
발행처 (Publisher)	융복합지식학회 The Society of Convergence Knowledge
URL	http://www.dbpia.co.kr/Article/NODE07220322
APA Style	이원용 (2017). 영상처리 기법을 이용한 문자인식 시스템 개발. 융복합지식학회논문지, 5(2), 99-103.
이용정보 (Accessed)	선문대학교 210.119.34.*** 2018/04/18 16:12 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

영상처리 기법을 이용한 문자인식 시스템 개발

이 원 용*

Development of character recognition system based on the image processing techniques

Won-Yong LEE
wylee@hj.ac.kr

요 약

문자 인식기술은 사용자가 임의의 영역에 그린 이미지로부터 가장 유사한 문자를 자동으로 식별해주는 기술을 의미한다. 최근 들어 스마트 기기들의 활용이 커지므로, 그 활용 영역이 점점 더 커지고 있다. 본 논문에서는 기존의 영상처리분야 에서 사용되던 알고리즘을 기반으로 한 문자인식 시스템 개발을 소개한다. 우선 사용자가 그린 영상을 이진화 하고, 골격화, 세 선화 및 히스토그램 분석을 순차적으로 수행하였다. 그리고 K 최근접 알고리즘을 적용하여 기 정의된 문자 중 가장 유사한 K개의 문자를 추천하여 최적의 문자를 추천한다. 제안하는 시스템은 실시간으로 가장 유사한 문자 결과를 도출해 주며 정답 률 전 평균 60%정도 되는 것을 확인하였다.

ABSTRACT

The character recognition technology refers to the techniques that automatically identifies the most similar character from an image drawn in an arbitrary region. Recently, the utilization of smart devices has become more and more widespread. In this paper, we introduces the development of a character recognition system based on the algorithm used in existing image processing field. First, we apply binarization based on the user drawn image. Then, we performed Image skeleton, thinning and histogram analysis sequentially. Third, K-NN(K-Nearest Neighbor) clustering algorithm is applied to recommend the best character by the most similar K characters among predefined characters. The proposed system identified the most similar character results from the DB with real-time performance, which averaged around 60% of the answers.

Keywords : Character recognition, Binarization, Thinning, Skeleton, Image Processing

I. 서 론

21세기 들어 다양한 스마트 기기가 개발 및 대중화 되었으며, 이와 더불어 터치패드 및 스마트 펜을 기반으로 한 정보 입력방법 또한 널리 이용되고 있다. 이와 같은 변화는 문자인

식 기술의 발달을 통해 가능해 졌다. 문자 인식이란 사용자가 펜 또는 손으로 그린 문자 영상으로부터 의도한 문자가 무엇 인지를 분석하여 추천해 주는 기법을 의미한다. 문자 인식은 1990년 후반 본격적으로 연구되기 시작하였으며, 현재는 인공 지능 및 머신러닝 기법 등을 적용하여 매우 높은 인식률을 보

* 제1저자, 해전대학교 인터넷보안과

논문접수일:2017.06.27. 논문심사일:2017.07.02. 게재확정일:2017.07.11.

이며 음성인식 기술과 더불어 다양하게 활용되고 있다.

본 논문에서는 다양한 문자인식 처리 기법 중 영상처리 기법에서 사용되는 다양한 이미지 처리 알고리즘들을 기반으로 한 문자인식 시스템을 제안하며, 제안된 알고리즘에 기반한 문자인식 시스템의 성능을 구현하여 확인하였다.

제안하는 시스템에서의 문자인식의 대상은 영문자 대문자 A~Z, 소문자 a~z 및 숫자 0~9까지, 총 62개의 문자로 한정하였다. 그리고 다양한 영상처리의 알고리즘 중 이진화(Binarization), 세선화(Thinning), 골격화(Skeleton) 및 히스토그램 분포(Histogram projection) 분석 기법을 적용하여 문자의 특징을 추출하는데 활용하였다. 본 시스템은 기 수집한 수만장의 샘플 이미지들을 기반으로 각 문자별 특징점을 정리한 문자별 데이터베이스를 구축하였다. 그리고 이를 기반으로 새로운 문자 이미지가 들어왔을 때 특징 점을 기반으로 현재 데이터베이스의 이미지들과 가장 차이가 적은 문자 값을 추천해 주도록 시스템을 개발하였다.

제안한 시스템은 기존에 많이 알려진 영상처리 알고리즘을 기반으로 비전문가들도 쉽게 문자인식의 기본 개념을 확인하고 구현해 볼 수 있는 장점을 가진다.

II. 관련 연구

문자 인식에 대한 개념을 확립하기 위해서는 [1]의 논문을 참고하였다. 이 논문에서는 전반적인 문자인식에 대한 개념을 설명하며 전통적인 문자인식의 알고리즘들을 제시하였다. [2]는 최초의 연구부터 현재까지의 문자인식 기법에 대한 발전 과정을 소개하고 앞으로 나아가야 할 방향에 대해서 소개를 한 논문이다. [3]은 흑백 영상을 이진화 하는데 있어서 최적의 값을 찾는 과정에 대해서 설명하고 있다. 단순히 0(검은색)~255(흰색)으로 변환된 영상을 그 중간인 128(검은색과 흰색의 중간 색)을 기준으로 나누는 것이 아닌, 이미지의 픽셀 값의 분포를 기반으로 한 임계값 설정을 하여 최적의 이진화 영상을 얻을 수 있음을 제시한 논문이다. [4]는 영상으로부터 얻을 수 있는 주요 특징(Feature)점에 대해 소개하고 있으며, 자세한 예시를 기반으로 특정 기법들에 대한 소개를 보여준다. [5]는 우리가 추출한 특징 중에 과연 어떤 것을 사용하는 것이 효율적인지에 대해 통계적 방법으로 접근한 연구이다. 최근의 연구들은 보다 높은 정확도 및 다양한 문자로 그 연구 영역을 넓혀 가고 있다. [6]은 한글 문자에 대한 문자 인식 시스템을 위한 알고리즘을 제시하였으며 [7]은 보다 높은 인식률을 위해 신경망 기반의 문자 인식 기법을 소개하고 있다.

III. 영상처리기법을 이용한 문자인식 시스템

<그림 1>은 본 시스템의 알고리즘 플로우를 보여주고 있다. 각 단계에 대해서는 아래에서 자세히 설명한다.

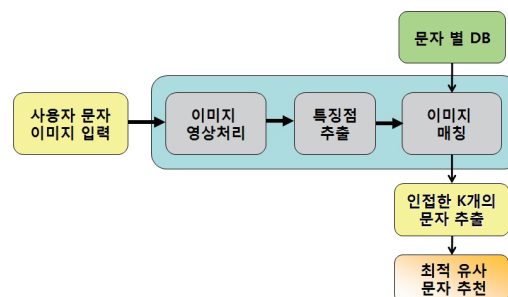


그림 1. 본 문자인식 시스템의 기능 구성도
Fig 1. Function flow of proposed system

3.1. 특정 데이터 추출을 위한 영상 처리

입력된 이미지를 가지고 정보를 추출하기 위해 본 논문에서는 흑백 영상변환, 이진화, 세선화, 골격화 알고리즘을 적용하였다.

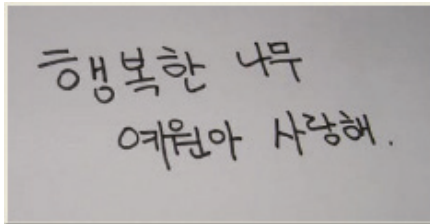
입력 영상을 흑백 영상으로 변환하는데 있어서는 RGB 색상 공간상에서 각 화소의 R(Red), G(Green), B(Blue) 값의 평균값을 구하여 사용하였다. 이후 제시되는 기법들은 모두 흑백 영상을 기반으로 알고리즘을 적용한다.

본 연구에서는 다양한 특징점 추출을 위해 최우선적으로 이진화 기법을 적용하였다. 이진화는 영상의 값을 검정색(0) 또는 흰색(255)의 두 가지 값만 가지도록 변형하는 기법이다. 이를 위해 우리는 Otsu의 방법[3]을 이용한다. Otsu의 방법을 사용 적용하기 위해 영상의 각 화소 정보를 히스토그램으로 분포시키고, 해당 분포에서 가장 명확하게 영역이 구분되는 지점을 선택해서 임계값으로 설정해 주는 방법이다.

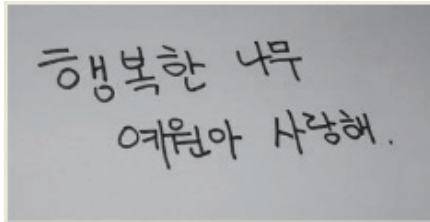
다음 과정으로 세선화 기법을 적용한다. 세선화란 두꺼운 라인을 하나의 화소로 이어진 가는 라인을 생성하는 기법이다. 본 논문에서는 장쎈(Zhang-Suen)이 제시한 세선화 기법[8]을 사용하였다. 마지막으로 골격화 기법을 적용한다. 골격화는 세선화와 다르게 두꺼운 라인의 양 끝의 윤곽선만을 얻는 알고리즘이다. 영상의 골격 구조를 추출하기 위해 본 과제에서는 체인코드를 사용하여 윤곽선을 얻어내었다.

<그림 2>는 위의 각 기법들을 입력 영상에 적용하여 얻은 흑백 영상, 이진화 결과 영상, 세선화 알고리즘 적용 결과 영

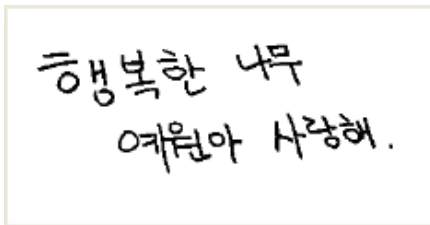
상, 골격화 알고리즘 적용 결과 영상을 보여주고 있다.



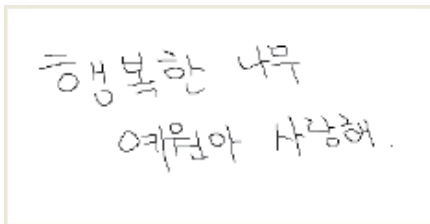
입력 영상



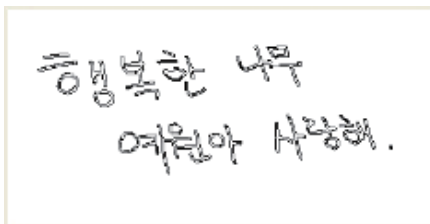
그레이 영상변환



이진화 영상



세선화 영상



골격화 영상

그림 2. 본 시스템에서 적용한 다양한 영상처리 기법을 적용한 결과 영상

Fig 2. Various image processing results using our system

3.2. 영상의 특징점 추출

위의 영상처리 기법을 적용한 결과들을 바탕으로 영상의 특징점(Feature)을 추출한다. 이때 입력된 이미지의 크기와 위치가 저마다 다르기 때문에 이를 일정 크기와 일정 위치로 정규화 시켜주는 작업을 먼저 수행해 주어야 한다.

입력 받은 이미지를 가지고 글자가 있는 영역을 찾아낸다. 찾아낸 글자의 영역이 클 경우는 축소시켜주고 작을 경우는 확대시켜 준다. 본 시스템에서는 이미지의 크기를 60*60 화소로 고정시켰다. 또한 이때 문자를 인식하는데 불필요한 노이즈를 제거해 준다. 노이즈 제거는 체인코드를 사용하여 체인코드가 일정 길이만큼 연결되지 않을 경우 노이즈로 간주하고 제거하였다.

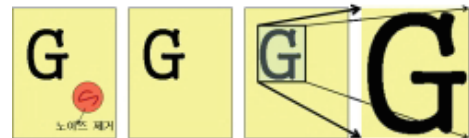


그림 3. 정확한 특징점 추출을 위한 노이즈 제거 및 문자 크기 정규화 과정

Fig 3. Denoising and size normalization for feature extraction

정규화 된 영상을 기반으로 특징 점을 추출한다. 본 시스템에서는 영역 구분을 통해 각 영역들의 정보를 분석하는 조닝(zoning) 기법과 전체 영상의 히스토그램 투영(histogram projection), 그리고 가로와 세로의 비율을 사용하였다.

이진화/세선화/골격화 조닝: 이진화 영상을 가로, 세로로 4등분 하여 16개의 공간을 구분한다. 전체 이미지의 검은 화소의 수를 대상으로 각 영역의 검은 화소수를 비율로 나타낸다. 즉 전체 16개의 영역의 비율을 다 더하면 1이 된다. 이진화 조닝을 적용하므로 16개의 특징점이 추출된다. 세선화 조닝도 위의 바이너리 조닝과 마찬가지로 이다. 대신 바이너리 이미지가 다르더라도 세선화를 시키면 비슷한 모양으로 변할 수 있기 때문에 세선화 조닝도 한다. 위에서는 4x4의 영역을 나눴으므로 세선화 조닝에서는 3x3으로 영역을 나눈다. 이것에서도 전체 검은 화소수를 가지고 각 영역 별로 가지고 있는 검은 화소수를 찾아서 영역의 비를 구한다. 여기에서는 총 9개의 특징점이 추출된다. 골격화 조닝 역시 영역별 비율을 사용하는 특징점이다. 윤곽선에서 추출한 4방향의 체인코드 값을 가지고 각 영역에서 어느 방향으로 몇 개의 방향 값을 가지고 있는지를 판별하는 피쳐이다. 각 영역별로 4개의 방향 값들을 가지고 영역이 총 16개 이므로 여기에서 특징점은 4x4x4로 총 64개의 특징점을 얻을 수 있다

가로와 세로의 비율: 각 문자는 가로와 세로의 길이가 저마다

다 다르다. 또한 이는 문자의 고유한 특성으로 볼 수 있다. 따라서 가로길이와 세로길이의 비율 또한 하나의 피쳐로 사용되었다.



그림 4. 이진화/세선화/골격화의 조닝을 위한 영역 구분

Fig 4. Region segment of binarization, thinning and skeleton result for zoning techniques

히스토그램 투영: 정규화된 영상을 기반으로 가로 및 세로로 가지고 있는 화소의 수를 따져본다. 예를 들어 1과 같은 문자이미지였다면 가로로 히스토그램 투영을 한다면 왼편과 오른편에서는 하나의 화소도 나오지 않다가 중간부분에서만 급격하게 많은 화소들이 분포 하는 것을 알 수 있고 세로로는 일정한 화소수를 가지고 길게 분포하는 것을 볼 수 있다. 따라서 본 과제에서는 가로 세로 60*60 화소(Pixel)의 각 화소 열 및 화소행을 하나의 피쳐로 사용하였고 이를 통해 총 120개의 특징점을 추출 할 수 있었다.

결론적으로 문자 인식을 위해 본 시스템에서 사용된 특징점은 바이너리조닝(16개) + 세선화 조닝(9개) + 가로세로 길이비(1개) + 컨투어 조닝(64개)+ 히스토그램 프로젝션(120개)로 총 210개를 사용하였다.

3.3. 문자 DB 데이터 구축

위의 특징점 정보를 각 문자별 다양한 샘플 영상에 적용하여 특징점 DB를 데이터로 구축한다. 본 시스템에서는 각 문자별로 적게는 500장에서 많게는 5000장의 정답 영상을 대상으로 앞에서 정의한 210개의 특징점을 각각의 영상으로부터 추출하였다. 각각의 이미지는 210개의 정보를 가지는 텍스트 파일로 저장하였으며, 이를 특정 폴더에 저장하고 DB로 사용하였다. 본 시스템에서 사용한 전체 이미지는 약 150,000장을 활용하였다.

3.4. 문자 DB로부터 최적 문자 매칭 및 추천

사용자가 그린 영상을 입력 받아서 입력 영상의 특징점을 분석하고, 기존에 구축되어 있는 문자 DB의 각 이미지와 하나하나씩 비교해 나간다. 그리고 210개의 특징점이 가장 유사한 영상을 후보군으로 관리한다. 본 시스템에서는 K-최 근접 이웃(K-NN) 알고리즘을 사용한다. 일반적으로 K는 5를 사용하였다. 즉 입력된 문자 영상과 특징점 정보가 가장 유사한 5

개의 이미지를 찾고, 그 5개중 가장 많이 제시된 문자를 현재 들어온 입력 데이터의 문자라고 출력하는 것이다. 예를 들어 추천된 후보가 (L, L, I, 2, C) 라면, 최종 결정은 L이 되는 방식이다.

IV. 본 시스템의 수행 결과

본 시스템의 구현은 Visual Studio 2013을 사용하였으며, MFC기반[9, 10]의 OpenCV 라이브러리를 활용하여 영상처리하였다. 시스템의 구동 환경은 펜티엄 2.8 메모리는 4G를 사용한다. 우선 다양한 문자 이미지 150,000장에 대해 210개의 특징 점을 계산하는 DB 구축을 수행하는데 있어서 약 12시간 정도가 소요되었다. 프로그램이 시작되고 파일로 구축한 DB 데이터파일을 메모리로 업로드 하는데는 약 1분정도의 시간이 소요되었으며 사용자가 입력한 문자로부터 가장 유사한 문자를 추천해 주는 데는 1초 이하로 소요된다.

<표 1>은 본 시스템을 기반으로 일부 문자들에 대한 정답률 및 정답이 아닌 오답을 제시하였을 때 많이 추천되었던 오답 문자를 보여준다.

표 1. 문자별 정답률 및 많이 추천된 오답
Table 1. Matching results of each character and highly recommended wrong character

문자	정답률(%)	많이 추천된 오답 문자
A	65.2	h,p,P,R
E	59.7	e,F,B,G
K	51.4	h,k,X
L	85.82	6,4,h
O	3.77	0(숫자)
Q	43.15	9,0,o,q
T	89.49	7,r

앞서 제시된바와 같이 속도측면에서는 매우 빠르게 문자를 추천해 주는 것을 확인해 볼 수 있었다. 하지만 문자 인식을 부분에서는 각 문자마다 편차가 심하며 60% 전후의 인식률을 보이는 것을 확인할 수 있었다. 구체적으로 보면, 컴퓨터로 출력한 문자를 입력으로 사용할 때는 매우 높은 인식률을 보였으나 사람이 손으로 쓴 문자에 대해서는 인식률이 낮음을 확인할 수 있었다. 이 중에서도 의도적으로 혼동시키지 않고 적은 이미지에 대해서는 상대적으로 높은 인식률을 가지는 것을 확인하였다. 그리고 5,7,6,S 등과 같이 상대적으로 다른 문자들과 차이가 있는 글자들에 대해서는 인식률이 높으나 애매한 문자에 대해서는 인식률이 낮음을 볼 수 있었다.

K-최 근접이웃 알고리즘에서 K를 1로 사용했을 때보다 5로 사용했을 때 약 20% 정도 높은 정답률을 얻을 수 있었다.

V. 결론 및 향후 연구

본 논문에서는 영상처리에서 사용하는 다양한 알고리즘을 기반으로 문자 인식이 가능한 시스템을 개발하고 결과를 확인하였다. 제안한 시스템은 현재 개발 및 연구된 다른 기법들에 비해 상대적으로 낮은 문자 인식률을 보였다. 하지만 본 시스템은 매우 기본적인 영상처리 기법들을 기반으로 시스템을 구성하였기 때문에 문자 인식 기법을 배우고자 하는 입문자들이 쉽게 구현해 보며 원리를 이해하는데 좋은 기초자료로 사용될 수 있을 것으로 사료된다.

본 시스템의 인식률이 예상보다 낮은 이유는 다음과 같이 분석하였다. 첫 번째, 사람이 쓴 글씨의 경우 대, 소문자가 확실히 구별이 않는다. 또한 숫자 1과 소문자 l, O와 숫자 0과 같은 경우 육안으로도 판별 불가능 한데 이런 것들은 하나의 문자단위로 영상처리 알고리즘을 적용하는 것만으로는 한계를 가진다고 판단하였다. 두 번째, 대부분의 대/소문자의 경우 문자의 크기만 다르지, 각 문자의 영역별 형태 및 패턴이 유사하기 때문에 대문자를 소문자로 추천 또는 그 반대로 추천하는 경우가 많았다.

이것은 본 시스템에서 특징점을 추출하는데 있어서 몇 가지의 한계점이 존재하기 때문인 것으로 분석된다. 조닝의 경우는 데이터를 분석하는데 도움을 주었지만 히스토그램 투영의 경우 210개의 특징점중 이것이 120개를 차지한다. 따라서 이미지의 기울어짐이나 형태 정보가 큰 비중을 두어서 탐색되었다는 분석이 나온다. 그래서 8이라는 글자가 기울어져 있으면 이것 역시 7이라는 문자로 인식해 버리는 한계를 가진다. 따라서 향후 히스토그램 투영 특징점수를 상대적으로 낮추고 이에 따른 문자 인식률의 변화 추이를 확인이 필요할 것으로 보인다.

참고문헌

- [1] S. Mori, C.Y. Suen, and K. Yamamoto, "Historical review of OCR research and development", Proceedings of the IEEE, vol. 80, 1992.
- [2] S. Mori, C. Y. Suen, and K. Yamamoto, "Historical Review of OCR Research and Development," Proceedings of The IEEE 80, pp. 1029-1058, July 1992.
- [3] HJ, Vala, A. Baxi, "A review on Otsuimage segmentation algorithm", International Journal of Advanced Research in Computer Engineering. 2,387-389, 2013.
- [4] O. D. Trier, A. K. Jain and T. Text, "Feature Extraction Methods For Character Recognition-A Survey", Pattern

Recognition, Vol. 29 No. 4, pp.641-662, 1996.

- [5] A. K. Jain, D. Zongker, "Feature Selection: Evaluation, Application, and Small Sample Performance," Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 2, pp. 153-158, 1997.
- [6] 이명훈 외, 간판영상에서 한글 인식 성능향상을 위한 가중치 기반 음소 단위 분할 교정, 한국콘텐츠학회논문지, 제12권, 제2호, pp.105-115, 2012.
- [7] M. Liwicki, A. Graves, S. Fernández, H. Bunke, J. Schmidhuber, A Novel Approach to On-Line Handwriting Recognition Based on Bidirectional Long Short-Term Memory Networks, In Proceedings of the 2007 International Conference on Document Analysis and Recognition (ICDAR2007), pp. 367-371, 2007.
- [8] T. Y. Zhang and C.Y. Suen. "A fast parallel algorithm for thinning digital patterns", Communications of the ACM. Vol.27, No. 6, pp.236-239. Mar. 1984.
- [9] 김성채, "MFC 프로그래밍 Visual Studio 2015(C++)", 기초 알고리즘과 순서도" 복월링
- [10] 김세훈, "MFC 프로그래밍 : 주식 분석 프로그램 만들기", 한빛미디어



이원용 (Won-Yong Lee)

1980년 중앙대학교 전자계산학과 (학사)
1985년 중앙대학교 경영정보학과 (경영학석사)
2002년 순천향대학교 전산학과 (공학박사)

1982년 8월 ~ 1993년 8월 : LG전자 컴퓨터설계실
1993년 9월 ~ 현재 해전대학교 인터넷보안과 부교수
e-mail: wylee@hj.ac.kr

※ 관심분야 : 병렬처리, 멀티미디어, 인터넷보안