

<http://dx.doi.org/10.17703/JCCT.2018.4.4.361>

JCCT 2018-11-47

골격을 이용한 문자 인식을 위한 지역경계 연산

Regional Boundary Operation for Character Recognition Using Skeleton

유석원*

Suk Won Yoo*

요약 학습 데이터를 구성하는 각각의 문자들에 대해 서로 다른 글자체들을 픽셀 단위로 더해서 MASK를 만들고, 해당 MASK에 속하는 픽셀값들을 세 영역으로 나눈다. 실험 데이터를 골격 형태로 수정하고, 지역 경계 연산을 사용하여 수정된 실험 데이터의 배경 중에서 문자의 골격에 인접한 배경 영역을 구분하는 경계를 만든다. 수정된 실험 데이터와 MASK들 간의 불일치 정도를 계산해서 최소값을 가지는 MASK를 찾는다. 이 MASK가 해당 실험 데이터에 대해 최종적으로 인식된 학습 데이터 문자로 선택된다. 문자의 골격과 지역 경계 연산을 사용하는 인식법은 주어진 학습 데이터에 새로운 글자체를 추가해서 학습 데이터를 쉽게 확장할 수 있으며, 구현하기가 간단하면서도 높은 문자 인식률을 얻을 수 있다.

주요어 : 골격, 세선화, 영상 처리, 문자 인식, 경계 확장, 영상 분석, 특징 추출

Abstract For each character constituting learning data, different fonts are added in pixel unit to create MASK, and then pixel values belonging to the MASK are divided into three groups. The experimental data are modified into skeletal forms, and then regional boundary operation is used to create a boundary that distinguishes the background region adjacent to the skeleton of the character from the background of the modified experimental data. Discordance values between the modified experimental data and the MASKs are calculated, and then the MASK with the minimum value is found. This MASK is selected as a finally recognized result for the given experiment data. The recognition algorithm using skeleton of the character and the regional boundary operation can easily extend the learning data set by adding new fonts to the given learning data, and also it is simple to implement, and high character recognition rate can be obtained.

Key words : Skeleton, Thinning, Image Processing, Character Recognition, Boundary Expansion, Image Analysis, Feature Extraction

1. 서론

오늘날 컴퓨터 관련 기술들의 발달로 인해 사무 자동화가 보편화 되었다. 특히, 인간이 사용하는 문자는

정해진 형태의 활자보다는 다양한 형태의 글자체들을 사용하며, 거의 모든 서류들이 책뿐만 아니라 전자책 형태로 출판된다. 영상 형태의 문서[1]에서 다양한 형태의 글자체들을 가지는 문자를 인식[2]하는 과정은 사무

*정회원, 서경대학교 컴퓨터과학과
접수일: 2018년 8월 11일, 수정완료일: 2018년 9월 12일
게재확정일: 2018년 9월 28일

Received: August 11, 2018 / Revised: September 12, 2018
Accepted: September 28, 2018

*Corresponding Author: swyoo@skuniv.ac.kr
Dept. of Computer Science, SeoKyeong University, Korea

자동화 분야에서는 해결하기 어려운 문제이다. 이를 위해 글자체의 뼈대[3]를 추출하여 글자체를 인식하는 연구들이 진행되고 있다. 뼈대를 사용하는 경우에는 글자체들의 다양한 변화로 인한 문제들뿐만 아니라 여러 각도로 회전한 경우도 해결할 수 있는 장점이 있다.

II. 관련 연구

영상 처리[4]와 문자 인식 기술은 영상 형태의 문서에 있는 글자들에서 특징을 추출하여 인식하는 기술이다. 인식 과정에서 일반적으로 사용하는 방법들로는 일정 영역 안에 들어 있는 문자를 나타내는 픽셀들의 개수를 조사하는 법[5], 수평이나 수직 방향으로 픽셀들을 합해서 발생 빈도수를 비교 조사하는 법, 학습 데이터로 사용하는 문자들의 차이를 구분 지을 수 있는 특징을 사용하는 법, 문자를 구성하는 요소들의 특성을 벡터로 표현해서 연관관계를 설정하고 비교하는 법[6], 비슷한 형태의 문자들의 유사성을 비교하는 법 등이 있다. 이 방법들은 인식 과정이 빠르다는 장점이 있지만, 인식을 위한 준비 과정이 복잡하고, 학습 데이터로 사용하는 글자체의 종류가 다양할수록 인식률이 낮아진다는 문제가 있다. 이런 문제들을 해결하기 위해서, 문자의 구조를 나타내는 뼈대[7]를 사용하는 방법들이 연구되고 있다. 문자의 뼈대는 해당 문자의 기본적인 형태를 제공하므로 다양한 글자체들의 변화에 의한 문제점들을 해결할 수 있으며, 영상 기기의 기계적인 문제로 인해 입력 영상에 문제가 발생하는 경우에도 부분적인 정보만을 가지고도 다른 방법들에 비해 더 높은 인식 결과를 얻을 수도 있다.

III. 본 론

산업화의 발달로 인해 예술 개념이 강조되다 보니 다양한 글자체들이 제공되고 있다. 그런데 기존의 문자 인식 방법들은 대부분 특정 글자체에 국한되어 있어 다양한 형태의 문자들을 인식하기가 어렵다. 요즘에는 다양한 형태의 글자체들을 인식하기 위해서 기계적 학습이라는 새로운 방식을 사용하기도 한다 [1]. 기계적 학습의 장점으로는 다양한 형태의 글자체들을 인식할 수 있다는 것이다. 그러나 단점으로는 많은 양의 학습 데이터가 필요하고, 학습 시간이 오래 걸리고, 새로운 학습 데이터를

추가하기가 어렵다. 이러한 문제점들을 해결하기 위해서, 본 논문에서는 지역 경계 연산을 사용하여 문자의 골격과 배경 영역을 구분하는 경계를 만들어서 다양한 크기와 형태의 글자체를 가진 문자들을 인식하는 기법을 소개한다.

1. 기본 개념

학습 데이터가 특정 글자체를 가진 경우에는 주어진 글자체의 특성만 고려하면 되므로 실험 데이터를 학습 데이터와 비교해서 인식하는 것은 쉬운 과정이다. 그러나 학습 데이터가 여러 개의 다양한 글자체들을 가진 경우에는 학습 데이터의 다양한 글자체들 간의 변화 정도를 고려해야 하므로 새로운 글자체를 가진 실험 데이터를 학습 데이터와 비교해서 실험 데이터를 인식하는 것은 어려운 과정을 요구한다. 주된 문제점은 학습 데이터의 기본적인 특성을 나타내는 구조는 유지하면서 동시에 다양한 글자체들에 의한 변화까지 고려해서 실험 데이터를 인식해야 한다는 것이다.

본 논문에서는 다양한 글자체들을 가진 학습 데이터의 기본적인 특성을 나타내는 구조를 위해 MASK 개념을 사용한다. 그리고 학습 데이터의 다양한 글자체들의 변화로 인한 문제점을 해결하기 위해서, 실험 데이터의 배경과 문자의 골격 사이에 새로운 지역경계를 추가하는 지역경계 연산을 사용한다.

2. 문자 인식을 위한 지역경계 연산

다음의 알고리즘을 통해 골격을 이용한 문자 인식을 위한 지역경계 연산에 대해 보다 자세히 설명하고자 한다.

Step 1) 학습 데이터는 0부터 9까지 10개의 문자들로 구성되며, 각 문자는 10개의 서로 다른 글자체들을 가지고 있다. 학습 데이터는 전체 크기가 170x170 크기이며, 각각의 학습 데이터 문자는 15x15 크기의 흑백 영상이다. 배경은 0의 픽셀값을 가지며, 글자 영역은 1의 픽셀값을 가진다.

Step 2) 실험 데이터 또한 0부터 9까지 10개의 문자들로 구성되며, 현재 사용되고 있는 자동차 번호판의 글자체 이면서 동시에 학습 데이터에서는 사용하지 않은 새로운 글자체를 가지고 있다. 실험 데이터는 전체 크기가 1000x100 크기의 흑백 영상이며, 각각의 실험 데이터 문자는 90x90 크기이다. 학습 데이터와 마찬가지로 배경은

0의 픽셀값을 가지며, 글자 영역은 1의 픽셀값을 가진다. Step 3) 10개의 실험 데이터 문자들 각각에 대해, 글자 영역이 더 이상 축소되지 않을 때까지 다음 두 Step 3-1과 3-2를 반복해서 문자의 골격을 형성하는 픽셀들을 찾는다.

Step 3-1) 실험 데이터의 글자 영역을 내부 영역과 주변 영역으로 구분한다. 내부 영역은 Direct Neighbor들이 모두 글자를 나타내는 1의 픽셀값을 가지는 픽셀들로 구성되며, 주변 영역은 내부 영역이 아닌 글자 영역을 의미한다.

Step 3-2) 주변 영역에 속하는 각각의 픽셀에 대해, 해당 픽셀을 제거했을 때 실험 데이터의 글자 영역이 서로 공통부분이 없는 두 개 이상의 영역으로 나누어지는지 조사한다.

Step 3-2-1) 글자 영역이 두 개 이상의 영역으로 나누어지는 경우, 해당 픽셀은 실험 데이터의 골격을 형성하는 중요 픽셀이므로 제거하지 않고 그대로 둔다.

Step 3-2-2) 해당 픽셀을 제거해도 글자 영역이 두 개 이상의 영역으로 나누어지지 않는 경우, 해당 픽셀은 실험 데이터의 골격을 구성하는데 아무런 의미가 없는 불필요한 픽셀이다. 해당 픽셀을 제거하기 위해서 픽셀의 값을 0으로 변경한다. 그 결과로, 글자 영역은 축소된다.

Step 4) 90x90 크기인 실험 데이터 문자들 각각은 이제 골격 형태로 수정되었으며, 골격은 1 또는 2의 두께를 가진다. 15x15 크기의 학습 데이터와 비교해서 인식하기 위해서, 90x90 크기인 실험 데이터 문자들을 15x15 크기가 되도록 누적형태로 축소시킨 뒤에, 1보다 크거나 같은 값을 가지는 픽셀들은 다시 1로 변경한다.

Step 5) 학습 데이터를 구성하는 각각의 문자들에 대해, 10개의 서로 다른 글자체들을 픽셀 단위로 더해서 MASK를 만든다. 이렇게 만들어진 10개의 MASK는 해당 문자의 서로 다른 글자체들을 대표해서 사용한다.

Step 6) 학습 데이터를 구성하는 각각의 MASK에 대해 최고값을 찾는다. 그리고 나서 해당 MASK의 값들을 0부터 최고값의 1/3, 최고값의 1/3부터 최고값의 2/3, 그리고 최고값의 2/3부터 최고값까지 세 영역으로 나눈 뒤에, 0부터 최고값의 1/3에 속하는 값은 0으로 변경하고, 최고값의 1/3부터 최고값의 2/3에 속하는 값은 1로 변경하고, 최고값의 2/3부터 최고값에 속하는 값은 2로 변경한다.

참고로, 각각의 학습 데이터의 여러 글자체들을 픽셀 단위로 더해 얻어진 결과인 MASK를 세 개의 영역으로

다시 나누는 이유는 첫 번째 영역인 0부터 최고값의 1/3 영역은 바탕에 해당하는 것으로 간주하고, 두 번째 영역인 최고값의 1/3부터 최고값의 2/3 영역은 다양한 글자체를 가진 학습 데이터의 변화에 해당하는 것으로 간주하고, 세 번째 영역인 최고값의 2/3부터 최고값 영역은 글자 영역에 해당하는 것으로 간주하기 때문이다. 이 과정을 통해서 MASK는 다양한 글자체들을 가진 학습 데이터의 기본적인 구조적 특성을 나타내며, 학습 데이터들의 미세한 변화들의 차이에 의한 인식과정의 오류를 낮출 수 있다.

Step 7) 축소된 뼈대 형태의 10개의 문자들로 구성된 실험 데이터 각각에 대해,

Step 7-1) 학습 데이터를 대표하는 10개의 MASK 각각에 대해,

Step 7-1-1) 0과 1로 구성된 실험 데이터의 픽셀값 중에서 픽셀값 1을 2로 수정한다.

Step 7-1-2) 수정된 실험 데이터의 2에 해당하는 픽셀들 각각에 대해, 해당 픽셀의 Direct Neighbor들 중에서 어느 것 하나라도 배경인 경우에는 배경에 해당하는 픽셀의 값을 0에서 1로 변경한다. 지역경계 연산을 사용하는 이 과정은 실험 데이터의 문자 영역 주변에 지역경계를 만들어서 배경 중에서 문자 영역에 인접한 배경 영역을 구분한다. 지역경계 연산은 MASK의 문자 영역 안에서 실험 데이터의 문자 영역이 움직일 수 있는 여유를 제공함으로써 인식률을 높이는 역할을 한다.

Step 7-1-3) 수정된 실험 데이터와 MASK에 대한 불일치 정도를 계산하기 위해, 이들을 픽셀단위로 비교하는 과정에서 1) 수정된 실험 데이터의 픽셀값이 0이면서 MASK의 해당 픽셀의 픽셀값이 2이거나, 2) 수정된 실험 데이터의 픽셀값이 2이면서 MASK의 해당 픽셀의 픽셀값이 0인 경우에는 해당 픽셀에서 수정된 실험 데이터와 MASK 간에 불일치가 발생한 것으로 간주해서 불일치값을 1 증가시킨다.

Step 7-2) 각각의 실험 데이터는 10개의 MASK들과 비교한 10개의 불일치 값들을 가지고 있다. 10개의 불일치 값들 중에서 최소값을 가지는 MASK를 찾는다. 이 MASK가 해당 실험 데이터에 대해 최종적으로 인식된 학습 데이터 문자로 선택된다.

3. 실행 결과

다음 그림들은 골격을 이용한 문자 인식을 위한 지역

경계 연산을 실행해서 얻어진 결과를 보여준다. 그림 1은 현재 사용되고 있는 자동차 번호판의 글자체이면서 학습 데이터가 사용하지 않은 새로운 글자체를 가진 10개의 실험 데이터를 보여준다. 10개의 글자체를 가진 100개의 학습 데이터[8]와 실험 데이터 모두 바탕은 흰색으로 픽셀값 0을 가지고 글자 영역은 검은색으로 픽셀값 1을 가진다. 그림 2는 10개의 실험 데이터의 문자들의 골격을 보여주고, 그림 3은 15x15 크기의 학습 데이터와 비교 과정을 통해 인식하기 위해서, 90x90 크기인 실험 데이터 문자들의 골격들을 15x15 크기가 되도록 1/6 크기로 축소시킨 결과를 보여준다.

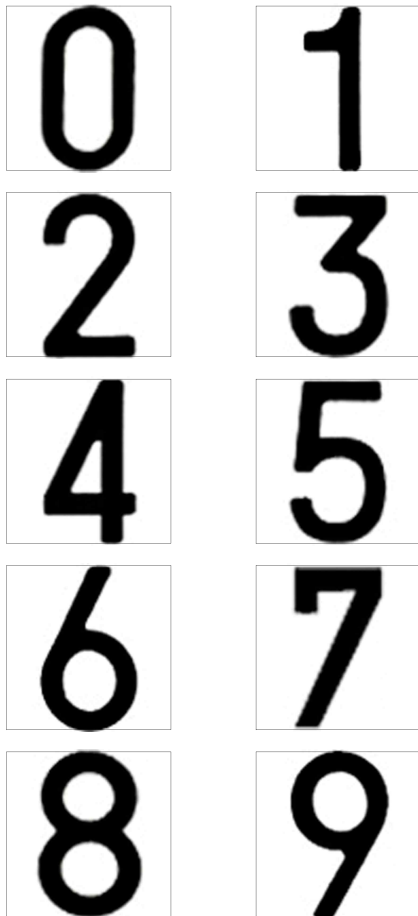


그림 1. 10개의 자동차 번호판 숫자들
Figure 1. 10 Number of Plate Numbers

10개의 학습 데이터 문자들 각각에 대해, 10개의 글자체들을 픽셀 단위로 모두 더해 MASK를 구한 뒤에 MASK를 구성하는 픽셀값들 중에서 최고값을 찾는다. 그리고 나서 MASK를 구성하는 픽셀값들을 0부터 최고값의 1/3, 최고값의 1/3부터 최고값의 2/3, 그리고 최고값

의 2/3부터 최고값 등 세 개의 영역으로 나누고, 각 영역에 속하는 픽셀값들을 0, 1, 2로 수정한다.

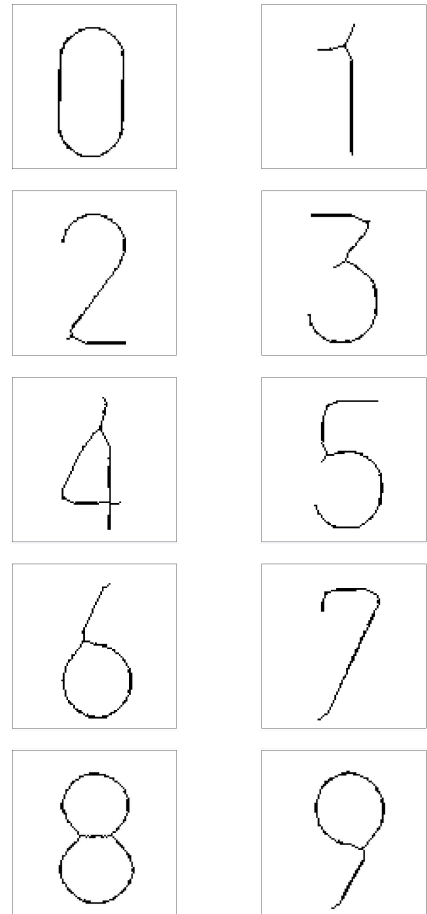


그림 2. 10개의 자동차 번호판 숫자들의 골격들
Figure 2. 10 Number of Skeletons of Plate Numbers

표 1은 최고값의 1/3과 최고값의 2/3을 사용하여 MASK들을 세 개의 영역으로 분리하는 경우, 10개의 실험 데이터 각각에 대해 지역 경계 연산을 사용한 결과와 0부터 9까지 10개의 MASK들과 계산한 불일치 값들, 최소의 불일치 값, 그리고 결과적으로 인식된 MASK를 보여준다. 예를 들어, 실험 데이터 3은 MASK 3과는 7 크기의 불일치값을 가지며, MASK 9와는 29 크기의 불일치값을 가진다. 결과적으로, 실험 데이터 3은 MASK가 3일 때 최소의 불일치 값 7을 가지므로 실험 데이터 3은 MASK 3으로 인식된다. 결과적으로, 최고값의 1/3과 최고값의 2/3을 사용하여 세 개의 영역으로 분리하는 경우, 본 논문에서 제안한 지역경계 연산은 학습 데이터에서는 사용하지 않은 새로운 문자체를 가진 10개의 실험 데이터에 대해 100%의 문자 인식 결과를 만든다.

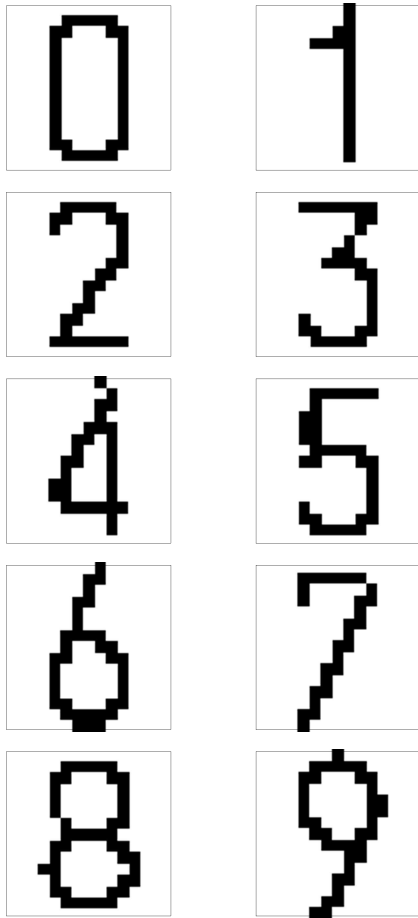


그림 3. 10개의 자동차 번호판 숫자들의 골격들의 축소형
 Figure 3. Shrunken Versions of Skeletons of 10 Plate Numbers

그래서 학습하지 않은 새로운 글자체의 문자 인식이 가능하며, 만족할만한 문자 인식 결과를 얻는다. 추후 연구 방향으로는 신경망 이론을 사용하여 문자를 인식하는 알고리즘 개발에 대해 연구 중이다.

References

- [1] E. Hall, Computer Image Processing and Recognition, Academic Press, pp.76-153, 1979.
- [2] G. Dougherty, Pattern Recognition and Classification, Springer, pp. 123-134, 2012.
- [3] D. Eberly, "Skeletonization of 2D Binary Images", www.geometrictools.com, pp.1-11, 2008.
- [4] K. Jang, S. Lee, "Analysis of Digital Video with a Focus on the Documentary Colored Hearing", Journal of Convergence on Culture Technology(JCCT), Vol.2, No.2, pp.23-27, 2016.
- [5] W. Pratt, Introduction to Digital Image Processing, CRC Press, pp. 139-154, 2013.
- [6] J. Parker, Algorithms for Image Processing and Computer Vision, Wiley, pp.137-175, 2010.
- [7] C. Suen, P.Wang, Thinning Methodologies for Pattern Recognition, World Scientific Press, pp.113-136, 1994.
- [8] S. Yoo, "Character Recognition Algorithm using Accumulation Mask", International Journal of Advanced Culture Technology(IJACT), Vol.6, No.2, pp.123-128, 2018.

4. 알고리즘의 장단점

기존의 문자 인식 기법들과 마찬가지로 본 논문에서 제안한 방법에도 장단점들이 있다. 장점으로는 1) 지역 경계 연산을 사용하여 문자를 인식하는 과정이 이해하기 쉬우며 알고리즘을 구현하기가 간단하다. 2) 기존의 기법들에 비해 높은 문자 인식률을 얻을 수 있다. 3) 주어진 학습 데이터에 새로운 글자체를 추가해서 학습 데이터를 쉽게 확장할 수 있다. 단점으로는 주어진 학습 데이터의 글자체 종류와 MASK를 세 개의 영역으로 분리하는 값들에 따라 문자 인식 결과가 다를 수 있다.

IV. 결론 및 추후 연구방향

본 논문에서 제안한 지역 경계 연산은 학습 데이터의 배경과 글자 영역 사이에 학습 데이터의 다양한 글자체들의 변화를 받아들이는 새로운 경계를 생성한다.

표 1. 자동차 번호판들과 학습 데이터의 MASK들과의 연산 결과

Table 1. Calculation Results between MASKs of Learning Data and Plate Numbers

	M#0	M#1	M#2	M#3	M#4	M#5	M#6	M#7	M#8	M#9	Result M#	Min
PN#0	0	45	23	15	23	15	7	32	9	8	0	0
PN#1	48	7	40	41	37	42	48	26	55	51	1	7
PN#2	25	24	3	19	37	28	34	13	32	22	2	3
PN#3	21	29	25	7	29	11	18	23	20	29	3	7
PN#4	45	33	41	34	6	45	42	37	40	45	4	6
PN#5	20	38	32	16	32	3	5	34	13	21	5	3
PN#6	31	32	40	27	27	21	14	43	24	37	6	14
PN#7	39	25	15	25	48	33	42	9	37	37	7	9
PN#8	12	46	27	12	27	16	6	34	4	9	8	4
PN#9	29	31	25	31	41	34	37	19	32	15	9	15