

<http://dx.doi.org/10.17703/JCCT.2018.4.3.241>

JCCT 2018-8-31

문자 인식을 위한 영상 복원

Image Restoration for Character Recognition

유석원*

Suk Won Yoo*

요약 영상 기기의 기계적인 문제로 인해 실험 데이터에 발생한 잡음으로 인한 인식 오류를 최소화하기 위해서 영상 복원 과정을 거친다. 영상 복원 방법은 실험 데이터를 구성하는 각각의 픽셀에 대해 Direct Neighbor와 Indirect Neighbor의 개수와 위치를 조사해서 잡음을 해결한다. 결과적으로, 영상 복원 과정을 통해 실험 데이터에 발생한 잡음을 최대한 제거하고, 영역 단위로 학습 데이터와 실험 데이터의 차이를 계산해서 잡음에 의한 인식 오류 가능성을 낮춤으로써 만족할만한 인식 결과를 얻을 수 있다.

주요어 : 지역 경계, 방향성 확장, 영상 복원, 잡음 제거, 영상 처리, 글자 인식, 영역 확장, 누적, 영상 분석, 특징 추출

Abstract Because of the mechanical problems of input camera equipment, image restoration process is performed in order to minimize recognition errors due to the noise problem generated in test data image. The image restoration method resolves the noise problem by examining the numbers and positions of the Direct neighbors and the Indirect neighbors for each pixel constituting the test data. As a result, satisfactory recognition result can be obtained by eliminating the noise problem generated in the test data through the image restoration process as much as possible and also by calculating the differences between the learning data and the test data in the area unit, thereby reducing the possibility of recognition error by the noise problem.

Key words : Regional Boundary, Directional Expansion, Image Restoration, Noise Filtering, Image Processing, Character Recognition, Boundary Expansion, Accumulation, Image Analysis, Feature Extraction

1. 서 론

디지털 영상은 컴퓨터에서 사용하는 영상 데이터 형태이며, 디지털 영상처리는 수학적 분석을 통해 디지털 영상을 더 나은 형태로 변환하는 작업이다. 영상처리와 패턴인식은 컴퓨터 그래픽스 분야에서 중요한 두 축을 형성하고 있다. 디지털 영상에 대해 효율적인 영상처리 과정을 수행함으로써 만족할만한 패턴인식 결과를 얻

을 수 있다.

디지털 영상처리 분야에 속하는 기술들 중에는 영상 개선과 영상 인식 등 다양한 분야들이 있다. 이 중에서 영상 개선은 입력 영상의 밝기를 수정하거나, 영상의 선명도를 수정하고, 기계적인 문제로 인해 만들어진 잡음 문제[1] 등의 수정 작업을 한다. 이 과정들을 통해 보다 나은 수준의 영상 결과물을 얻게 됨으로써 효율적인 영상 인식[2] 결과를 얻을 수 있다. 영상 인식은 입력 영상에서

*정회원, 서경대학교 컴퓨터과학과
접수일: 2018년 5월 15일, 수정완료일: 2018년 6월 17일
게재확정일: 2018년 7월 25일

Received: May 15, 2018 / Revised: June 17, 2018
Accepted: July 25, 2018
*Corresponding Author: swyoo@skuniv.ac.kr
Dept. of Computer Science, SeoKyeong University, Korea

구별되는 특징들을 뽑아낸 뒤, 컴퓨터에 저장된 기초 정보들의 특징들과 비교해서 입력 영상 속의 물체를 인식하는 과정을 말한다.

II. 관련 연구

잡음 문제(Noise)는 영상기기의 기계적 문제나 주변 환경의 영향으로 인해 영상의 임의의 위치에서 픽셀값들이 잘못된 값들을 가지게 되어 영상 화질에 문제를 일으키는 것을 의미한다. 이러한 잡음 문제는 일반적으로 불규칙적으로 발생하며, 이로 인해 인식 과정에서 문제를 야기하게 된다. 이러한 형태의 잡음 문제를 제거하기 위해서 디지털 영상처리 분야에서는 여러 경계값들을 사용하는 필터링[3], 마스크[4], 또는 형태학[5] 등을 사용하여 이 문제를 해결하고 있다.

III. 본 론

과학 기술의 발달에 있어서 정보처리 기술은 매우 중요한 분야이다. 특히, 영상 처리 및 인식 기술은 사무 자동화, 물체 인식 및 추적[6], 의학[7], 무인 시스템 등 다양한 분야에서 사용되고 있다. 병원에서 환자의 건강 상태를 알기 위해 사용하는 의학 기기들이 대부분 영상 기기들이며, 보안을 위해서도 여러 곳에 감시 카메라가 설치되고 있다. 화재나 재난 구조 과정에서도 인명을 구하기 위해서 카메라가 달린 드론을 사용한다.

영상 기기의 기계적인 문제로 인해 입력 영상에 잡음 문제가 발생하는 경우, 입력 영상의 중요 정보가 잡음으로 인해 훼손 되어 인식 과정에서 오류를 범하게 된다. 이러한 인식 오류는 후에 심각한 결과를 초래하게 되고, 주변의 상황에 따라 영상 기기마다 여러 형태의 잡음들이 발생하게 되어 서로 상이한 인식 결과를 얻기도 한다. 이런 문제들을 해결하기 위해서, 잡음으로 인한 인식 오류를 최소화하는 다양한 영상 복원 방법들이 현재 연구되고 있다.

1. 기본 개념

실험 데이터에 점 형태의 잡음 문제 (Salt and Pepper Noise)가 발생했다고 가정하자. 이런 형태의 잡음은 영상 기기에서 발생할 수 있는 일반적인 문제이지만 이런 잡음으로 인해 인식률이 낮아지게 된다. 이런

잡음 문제는 영상 기기에 따라 임의의 위치에서 임의의 형태로 발생하므로 예측이 불가능하며, 영상 복원 과정을 통하더라도 제거되지 않는 잡음이 남을 수 있다. 그리고 학습 데이터를 구성하는 문자들과 글자체들에 따라 실험 데이터의 인식 결과가 다를 수 있다. 그래서 잡음 문제로 인한 인식 오류를 최소화하고 문자 인식률을 높이기 위해서, 학습 데이터와 실험 데이터를 비교할 때는 픽셀 단위가 아닌 영역 단위로 비교하는 방법을 사용한다.

2. 문자 인식을 위한 영상 복원

다음의 알고리즘을 통해 문자 인식을 위한 영상 복원 방법에 대해 보다 자세히 설명하고자 한다.

Step 1) 100개의 학습 데이터는 0부터 9까지 10개의 서로 다른 문자들로 구성되며, 각 문자는 10개의 서로 다른 글자체를 가지고 있다. 각각의 학습 데이터는 픽셀값 0과 1로 구성된 15x15 크기의 흑백 영상이며, 0은 배경을 나타내고 1은 글자 영역을 나타낸다.

Step 2) 학습 데이터는 좌우에 빈 공간을 가지며 세로 방향으로 긴 형태의 글자 영역들을 가진다. 그래서 100개의 학습 데이터 각각을 3x3 크기의 25개 tile들로 나누고, 그중에서 글자 영역을 나타내는 가운데 위치한 15개 tile 영역에 속하는 픽셀값 1의 개수들을 조사한다. 그리고 나서, 각각의 학습 데이터를 15개 tile 영역에 속하는 픽셀값 1의 개수들과 자기 자신의 문자 종류로 구성된 16개의 값들로 표현한다.

Step 3) 실험 데이터도 0부터 9까지 10개의 서로 다른 문자들로 구성되며, 학습 데이터가 사용하지 않은 새로운 글자체를 가지고 있다. 그런데 영상 기기의 기계적인 문제로 인해 실험 데이터에 점 형태의 잡음 문제가 발생했다고 가정한다.

Step 4) 실험 데이터에 발생한 잡음 문제로 인한 인식 오류를 최소화하기 위해서, 먼저 영상 복원 과정을 거친다. 영상 복원 과정은 실험 데이터를 구성하는 각각의 픽셀에 대해 다음과 같이 진행한다.

Step 4.1) 실험 데이터를 구성하는 각각의 픽셀에 대해, Step 4.1.1) 픽셀값이 글자 영역을 나타내는 1인 경우에는 Direct Neighbor의 개수를 조사한다.

Step 4.1.1.1) Direct Neighbor의 개수가 1보다 크거나 같으면, 잡음 문제가 발생하지 않은 것으로 간주하고 수정하지 않는다.

Step 4.1.1.2) Direct Neighbor 개수가 0이면, Indirect Neighbor의 개수를 조사한다.

Step 4.1.1.2.1) Indirect Neighbor의 개수가 0이거나 1이면, 잡음 문제가 발생한 것으로 간주하고 해당 픽셀값을 0으로 수정한다.

Step 4.1.1.2.2) Indirect Neighbor의 개수가 2보다 크거나 같으면, 잡음 문제가 발생하지 않은 것으로 간주하고 수정하지 않는다.

Step 4.1.2) 픽셀값이 배경을 나타내는 0인 경우에는 Direct Neighbor의 개수를 조사한다.

Step 4.1.2.1) Direct Neighbor의 개수가 3보다 작거나 같으면, 잡음 문제가 발생하지 않은 것으로 간주하고 수정하지 않는다.

Step 4.1.2.2) Direct Neighbor의 개수가 4이면, 잡음 문제가 발생한 것으로 간주하고 1로 수정한다.

Step 5) 학습 데이터와 마찬가지로, 실험 데이터도 좌우에 빈 공간을 가지며 세로 방향으로 긴 형태의 글자 영역을 가진다. 그래서 실험 데이터 각각을 3x3 크기의 25개 tile들로 나누고, 그중에서 글자 영역을 나타내는 가운데 위치한 15개 tile 영역에 속하는 픽셀값 1의 개수들을 조사한다. 그리고 나서, 각각의 실험 데이터를 15개 tile 영역에 속하는 픽셀값 1의 개수들로 표현한다.

Step 6) 10개의 실험 데이터 각각에 대해 최적으로 인식된 학습 데이터 문자를 찾는 방법은 다음과 같이 진행한다.

Step 6.1) 10개의 실험 데이터 각각에 대해,
 Step 6.1.1) 학습 데이터를 구성하는 10개의 문자 각각에 대해,

Step 6.1.1.1) 각 문자가 가지는 10개의 글자체 각각에 대해, 해당 글자체와 해당 실험 데이터를 각각 구성하는 15개 tile 영역들에 속하는 픽셀값 1의 개수들의 차이를 계산한다.

Step 6.1.1.2) 10개의 차이값(Difference)을 오름차순으로 정렬한 뒤에, 그중에서 제일 작은 차이값 5개의 평균을 구한다. 이 평균을 학습 데이터의 해당 문자와 해당 실험 데이터의 불일치값(Discordance)으로 사용한다.

Step 6.1.2) 10개의 불일치값들 중에서 최소값을 가지는 문자가 해당 실험 데이터에 대해 최적으로 인식된 학습 데이터의 문자로 선택된다.

3. 실험 결과

그림 1은 영상처리 시스템에 저장되어 있는 100개의 학습 데이터들(Learning Data Set)[8]을 사용하여 10개의 실험 데이터들(Test Data Set)[8]을 인식하는 과정에서 잡음 문제가 발생한 실험 데이터들에 대해 문자 인식을 위한 영상 복원 방법을 실행해서 얻어진 결과를 보여준다. 실험 데이터들은 학습 데이터들이 사용하지 않은 새로운 글자체를 가진다. 학습 데이터와 실험 데이터 각각은 15x15 크기의 흑백 영상이며, 바탕은 흰색으로 픽셀값 0을 가지고 글자 영역은 검은색으로 픽셀값 1을 가진다.

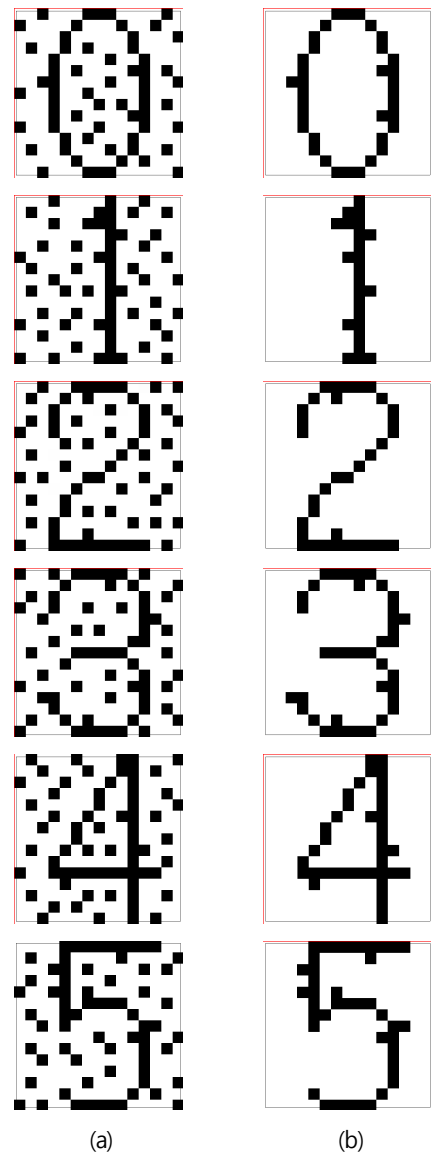


그림 1. (a) 잡음 문제가 발생한 실험 데이터들
 (b) 영상 복원 방법을 실행해서 얻어진 결과물
 Figure 1. (a) Test Data with error
 (b) Resulting images after Image Restoration

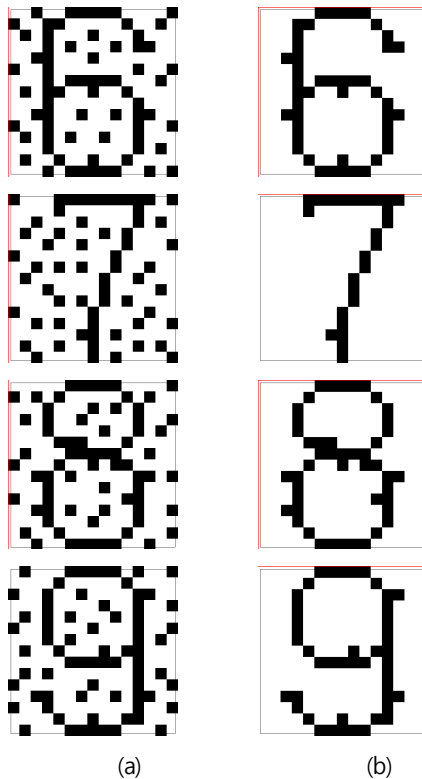


그림 1. (a) 잡음 문제가 발생한 실험 데이터들
(b) 영상 복원 방법을 실행해서 얻어진 결과물
Figure 1. (a) Test Data with error
(b) Resulting images after Image Restoration

표 1은 그림 1에 있는 잡음이 발생한 10개의 실험 데이터들에 대해 영상 복원 방법을 사용한 후, 0부터 9까지 10개의 학습 데이터의 문자들과 계산한 불일치값들, 최소의 불일치값, 그리고 최종적으로 인식된 학습 데이터 문자를 보여준다. 결과적으로, 문자 인식을 위한 영상 복원 방법은 잡음이 발생한 10개의 실험 데이터들에 대해, 영상 복원 방법을 실행해서 얻어진 결과에 대해서 100% 문자 인식 결과를 만든다.

표 2는 그림 1에 있는 잡음이 발생한 실험 데이터들에 대해 문자 인식을 위한 영상 복원 방법을 사용하지 않은 인식 결과를 보여준다. 표 2에서 보듯이, 실험 데이터 3은 학습 데이터 문자 3과는 24.0 크기의 불일치값을 가지며, 학습 데이터 문자 8과는 20.6 크기의 불일치값을 가진다. 그래서 실험 데이터 3은 학습 데이터 문자가 8일 때 최소의 불일치값 20.6을 가지므로 학습 데이터 문자 8로 인식되어 오류가 발생한다. 마찬가지로, 실험 데이터 5는 학습 데이터 문자 0으로 인식되고, 실험 데이터 6은 학습 데이터 문자 8로 인식된다. 결과적으로, 잡음 문제가 발

생한 실험 데이터들에 대해 문자 인식을 위한 영상 복원 방법을 사용하지 않는 경우, 10개의 실험 데이터에 대해 3개의 오류가 발생하여 70%의 문자 인식 결과를 만든다.

표 3은 잡음 문제가 전혀 발생하지 않은 10개의 실험 데이터 각각에 대해, 문자 인식을 위한 영상 복원 방법을 사용하여 최종적으로 인식된 학습 데이터의 문자를 보여준다. 결과적으로, 문자 인식을 위한 영상 복원 방법은 100개의 학습 데이터와 잡음이 발생하지 않은 10개의 실험 데이터에 대해서도 100%의 성공적인 문자 인식 결과를 만든다.

4. 알고리즘의 장단점

기존의 문자 인식 기법들과 마찬가지로 본 논문에서 제안한 방법에도 장단점들이 있다. 장점으로는 1) 잡음 문제가 발생한 경우에도 만족할만한 인식 결과를 얻을 수 있다. 2) 다른 방법들에 비해 영상 복원 과정이 이해하기 쉽고 간단하다. 3) 새로운 문자나 글자체를 학습 데이터에 추가해서 쉽게 확장할 수 있다. 단점으로는 1) 잡음이 임의의 위치에서 발생하므로 잡음을 완벽하게 제거하지 못할 수도 있어서 경우에 따라서는 다른 인식 결과를 얻을 수 있다. 2) 학습 데이터를 구성하는 문자들과 글자체들의 종류에 따라 인식 결과가 달라질 수 있다.

IV. 결론 및 추후 연구방향

본 논문에서 제안한 문자 인식을 위한 영상 복원법은 실험 데이터에 잡음 문제가 발생한 경우에도 영상 복원 과정을 통해 잡음을 최대한 제거함으로써 만족할만한 인식 결과를 얻을 수 있다. 그리고 픽셀 단위가 아닌 3x3 크기의 tile 단위로 학습 데이터와 실험 데이터의 차이를 계산함으로써 잡음에 의한 인식 오류 가능성을 낮춘다. 추후 연구 방향으로는 입력 영상 속의 글자 영역을 몇 개의 점(Vertex)들의 집합으로 표현하는 방법에 대해 연구 중이다.

References

- [1] J. Parker, Algorithms for Image Processing and Computer Vision, Wiley, pp.137-175, 2010.
- [2] C. Chen, L. Pau, P Wang, Handbook of Pattern Recognition and Computer Vision, World Scientific, pp.61-84, 1993.

- [3] T. Adali, S. Haykin, Adaptive Signal Processing, John Wiley & Sons, pp.40-58, 2010.
- [4] E. Hall, Computer Image Processing and Recognition, Academic Press, pp.76-153, 1979.
- [5] F. Shih, Image Processing and Mathematical Morphology, CRC Press, pp.25-35, 2009.
- [6] Y. Yang, S. Lee, "An Object Tracking Method for Studio Cameras by OpenCV-based Python Program", Journal of Convergence on Culture Technology(JCCT), Vol.4, No.1, pp.291-297, 2018.
- [7] G. Dougherty, Digital Image Processing for Medical Application, Cambridge, pp.155-189, 2009
- [8] S. Yoo, "Character Recognition Algorithm using Accumulation Mask", International Journal of Advanced Culture Technology(IJACT), Vol.6, No.2, pp.123-128, 2018.

표 1. 잡음이 발생한 실험 데이터들에 대해 영상 복원 방법을 적용한 결과들과 학습 데이터들의 불일치 정도
 Table 1. Discordance between Learning Data and Test Data with error (after Image Restoration Process)

	LD#0	LD#1	LD#2	LD#3	LD#4	LD#5	LD#6	LD#7	LD#8	LD#9	Min. Value	Result LD#
TD#0	12.4	43.2	29.4	28.8	28.4	26.4	24.0	33.0	20.4	22.2	12.4	0
TD#1	47.6	13.4	34.2	40.6	40.0	43.2	46.8	25.8	51.2	45.6	13.4	1
TD#2	25.6	30.4	16.6	25.0	40.8	28.0	26.8	25.8	23.6	28.2	16.6	2
TD#3	22.2	34.8	22.0	17.6	33.4	22.2	23.8	27.0	21.8	19.6	17.6	3
TD#4	31.6	48.4	39.4	38.6	12.0	44.0	41.0	38.4	37.4	38.6	12.0	4
TD#5	22.8	40.2	36.6	30.6	33.6	21.6	27.6	34.0	27.2	23.8	21.6	5
TD#6	20.0	40.6	29.4	27.2	36.0	17.6	15.2	35.0	16.4	21.4	15.2	6
TD#7	35.4	25.2	22.2	28.2	36.8	37.2	41.4	10.6	37.2	34.4	10.6	7
TD#8	17.6	42.8	28.6	24.2	30.8	20.4	18.8	35.0	15.2	17.8	15.2	8
TD#9	18.8	40.4	25.8	20.6	34.4	18.4	21.2	32.2	16.4	16.2	16.2	9

표 2. 잡음이 발생한 실험 데이터들에 대해 영상 복원 방법을 적용하지 않은 결과들과 학습 데이터들의 불일치 정도
 Table 2. Discordance between Learning Data and Test Data with error (without Image Restoration Process)

	LD#0	LD#1	LD#2	LD#3	LD#4	LD#5	LD#6	LD#7	LD#8	LD#9	Min. Value	Result LD#
TD#0	19.6	38.4	30.2	30.6	27.2	27.2	24.4	34.8	21.2	23.0	19.6	0
TD#1	31.6	25.0	29.0	34.8	34.0	31.6	33.8	29.0	34.6	34.6	25.0	1
TD#2	25.6	32.4	20.2	28.6	35.6	25.6	27.2	30.8	23.6	27.8	20.2	2
TD#3	23.0	37.8	24.4	24.0	28.6	22.2	25.4	29.4	20.6	21.2	20.6	8
TD#4	24.6	41.4	31.2	29.6	20.4	31.0	28.0	33.4	24.4	26.4	20.4	4
TD#5	23.8	39.2	32.8	33.2	31.6	24.6	28.6	32.6	26.8	28.0	23.8	0
TD#6	22.0	42.0	28.6	29.2	32.0	23.6	19.6	35.4	18.8	23.0	18.8	8
TD#7	28.0	30.4	24.2	29.0	30.0	29.0	30.6	21.4	27.6	28.4	21.4	7
TD#8	22.6	39.8	31.2	30.6	28.6	25.0	22.0	35.6	20.2	23.6	20.2	8
TD#9	23.8	41.4	28.8	26.8	30.6	23.0	24.0	34.0	19.8	19.6	19.6	9

표 3. 학습 데이터들과 실험 데이터들의 불일치 정도
 Table 3. Discordance between Learning Data and Test Data

	LD#0	LD#1	LD#2	LD#3	LD#4	LD#5	LD#6	LD#7	LD#8	LD#9	Min. Value	Result LD#
TD#0	12.0	41.2	29.0	29.2	29.2	25.2	23.2	32.6	20.0	23.4	12.0	0
TD#1	47.8	12.8	34.8	40.4	39.8	42.2	45.0	26.0	50.6	45.4	12.8	1
TD#2	24.8	32.4	16.2	24.2	39.2	27.2	26.0	26.2	22.0	26.6	16.2	2
TD#3	21.0	34.8	20.0	17.6	32.6	21.0	23.4	25.6	19.8	19.6	17.6	3
TD#4	30.4	45.4	38.8	38.0	12.2	41.8	40.2	37.2	37.6	39.2	12.2	4
TD#5	20.4	35.8	32.2	28.4	32.0	20.0	25.6	32.6	25.6	23.4	20.0	5
TD#6	17.8	40.6	28.4	27.0	33.0	17.4	13.4	35.6	15.8	20.0	13.4	6
TD#7	35.0	26.0	22.6	27.4	34.8	37.6	41.8	12.2	37.6	34.0	12.2	7
TD#8	15.0	40.6	24.4	23.4	30.6	18.6	17.4	31.6	13.8	18.0	13.8	8
TD#9	17.4	40.6	25.6	20.8	33.0	17.4	19.4	31.4	15.8	14.0	14.0	9