

신경망을 통한 숫자 검출 및 인식

조현구* · 김남호* · 김찬수**

*호남대학교 인터넷소프트웨어학과, **호남대학교 게임애니메이션학과

A number detection and recognition through a neural network

Hyun-gu Cho* · Nam-ho Kim* · Chan-soo Kim**

*Dept. of Internet Software, Honam University

**Dept. of Game Animation, Honam University

E-mail : dream0950@nate.com, nhkim@honam.ac.kr, cskim@honam.ac.kr

요 약

문자 인식이란 시각 정보를 통하여 문자를 인식하고 의미를 이해하는 것으로 인간의 능력을 컴퓨터로 실현하는 패턴인식의 한 분야이다. 본 논문에서는 문자 인식 중 가장 많이 사용되고 있는 숫자 검출과 인식을 소개하고자 한다. 또한 숫자 인식을 위해서 인간의 두뇌를 모델로 하여 만들어진 신경망에 대한 기본적인 원리와 신경망의 학습을 위한 역 전파(Back propagation) 알고리즘에 대하여 알아보하고자 한다.

ABSTRACT

Character recognition is one field of pattern recognition which comes true the ability of the human being with the computer. In this paper, we performed a comparative study on mostly used method of number detection and recognition. Also number recognition from hazard brain the human being with the model. We research about fundamental principle and back propagation algorithm for studying of neural networks.

키워드

문자인식, 숫자인식, 신경망, back propagation

I. 서 론

문자인식(Character Recognition)이란 시각 정보를 통하여 문자를 인식하고 의미를 이해하는 인간의 능력을 컴퓨터로 실현하려는 패턴인식의 한 분야로서, 광학 문자 인식, 우편물 자동 분류, 문서 인식 등의 분야에서 부분적으로 실용화가 이루어지게 되었으며, 그 중 숫자 인식은 현재 가장 많이 사용되고 있는데 특히 자동차 번호판 인식 등에 활용되어 사용되고 있다.

또한 숫자 인식은 많은 발전을 이루어서 근래에 들어서는 99% 가까운 인식률을 보이고 있지만 아직도 많은 외부 환경에 대한 제약이 따르고 있으며 개발자가 정해 놓은 환경에서만 높은 인식률을 보인다는 한계점이 있다.

이번 논문을 통해서 더욱 쉽고 정확하게 숫자

를 검출하고 인식할 수 있는 방법을 알아보고 또한 숫자를 인식함에 있어서 인간에 두뇌를 모델로 하고 있는 신경망에 대한 기본적인 원리와 신경망의 학습을 위한 역 전파 알고리즘에 대해서 엘리베이터 버튼 숫자를 통해서 알아본다. 연구결과는 로봇의 미션 수행을 위한 시각정보 처리에 적용하고자 한다.

II. 숫자 검출

2.1 숫자의 위치 추적

숫자를 검출하기 위해서는 해당 숫자에 정확한 위치를 먼저 알아내는 것이 필요로 한다. 그 이유는 영상에서 숫자가 위치하는 것은 항상 일정한 것이 아니기 때문이다. 또한 엘리베이터 안에서

숫자를 찾는 것은 여러 숫자가 함께 존재하고 또한 위치에 따라 숫자에 크기가 다르기 때문에 전처리 과정에서 숫자의 위치와 숫자의 크기 정합이 필요로 한다.

2.2 엘리베이터 안에서의 숫자 검출

제안한 알고리즘은 영상에서 엘리베이터의 특성을 이용하여 숫자 영역을 추출하는 구성도로는 그림 1과 같다.

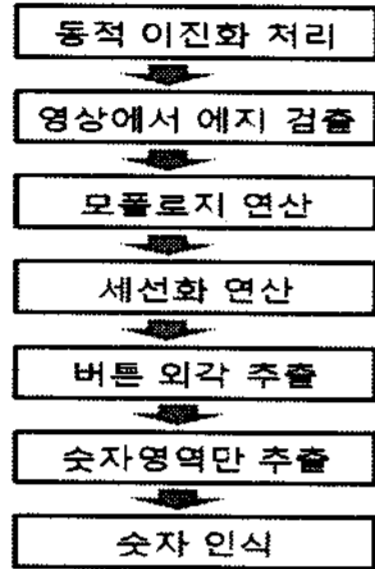


그림 1. 구성도

2.2.1 동적 이진화 처리

먼저 빛에 양에 따라서 칼라 영상에서 흑백 영상으로 바꾼 후 이진화를 통하여 흰색과 검은색 두 가지 색으로만 영상에 나타나도록 하여야 한다.

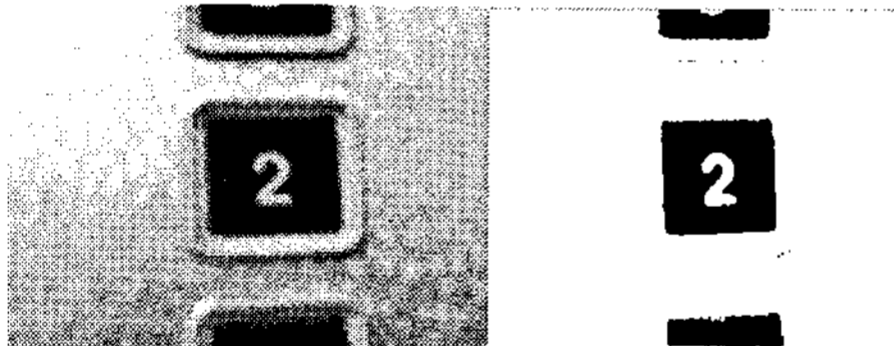


그림 2. 입력 영상 및 이진화 처리된 영상

2.2.2 에지 검출

이진화 되어 있는 영상에서 먼저 각 영상에 에지를 추출하여야 하는데 에지(Edge)추출에서는 Sobel 마스크를 적용하였다.

Mask1	Mask2
-1 0 1	-1 2 1
-2 0 2	0 0 0
-1 0 1	-1 2 1



그림 3. Sobel Filter 연산후(Edge 추출)

Sobel Filter는 윤곽선을 검출하는 필터로써 미분 연산자에 의한 밝기 값의 변화를 이용해서 찾아내는 것인데, 편미분 연산을 이용하여 미분 값을 구한다. Sobel Filter는 비선형 연산자로서 사용하는 마스크 윈도우 영역에서 양 끝단에 속한 화소들 사이의 합의 차이를 구한 후, 이를 수평과 수직 방향에 대하여 평균 크기를 구함으로써 경계 부위를 강조하는 역할을 한다.

2.2.3 모폴로지 연산

에지를 추출한 영상을 보면 조금씩 선분들이 끊어져 있는 것을 볼 수 있다. 이러한 선분을 이어 주기 위해서 모폴로지 연산 중 채움(closing) 연산을 하게 된다. 채움 연산의 경우 영역 내의 갭들을 확장하는 효과가 있다.

채움 연산은 침식 연산(erosion)을 수행한 후 그 결과 영상에다가 팽창 연산(dilation)을 하면 된다.

이진 영상 침식

$$A \otimes B = \{ x: B+x < A \}$$

이진 영상 팽창

$$A \oplus B = \{ A \ominus (-B) \}$$

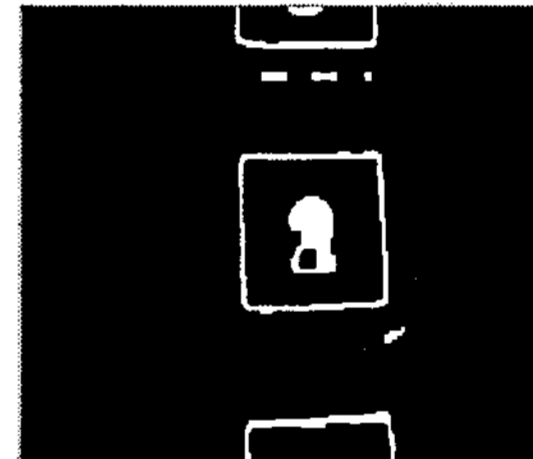


그림 4. 모폴로지 연산

2.2.4 세션화 연산

모폴로지 연산에서 두꺼워진 선들을 세션화 과정을 통하여 하나의 선분으로 만들어야 한다.

세션화란 대상 물체의 표면을 조금씩 벗겨내어 최종적으로 두께를 1로 만들도록 하는 과정으로 여기서는 Zhang Suen 세션화 알고리즘을 사용하였다.

<Zhang Suen 알고리즘 순서>

- ▶ 1루프>처리대상인 검은 픽셀 I(i,j)에 대해서 다음과 같은 조건을 만족하면 지운다.
- ① 픽셀 I(i,j)에 대해서 그 주위의 픽셀들의 connectivity가 1이고,
- ② 픽셀 I(i,j)에 대해서 그 주위의 픽셀들 중 검은 픽셀이 적어도 2개에서 6개사이에 있어야 하며,
- ③ 적어도 픽셀 I(i,j+1), I(i-1,j), I(i,j-1)중의 하나는 배경 픽셀 즉 255이어야 하며,
- ④ 적어도 픽셀 I(i-1,j), I(i+1,j), I(i,j-1)중의 하나는 배경 픽셀 즉 255이어야 한다.
- ⑤ 조건에 맞으면 픽셀을 제거한다.

▶ 2루프>처리대상인 검은 픽셀 $I(i,j)$ 에 대해서 다음과 같은 조건을 만족하면 지운다.

- ① 픽셀 $I(i,j)$ 에 대해서 그 주위의 픽셀들의 connectivity가 1이고,
- ② 픽셀 $I(i,j)$ 에 대해서 그 주위의 픽셀들 중 검은 픽셀이 적어도 2개에서 6개 사이에 있어야 하며,
- ③ 적어도 픽셀 $I(i-1,j)$, $I(i,j+1)$, $I(i+1,j)$ 중의 하나는 배경 픽셀 즉 255이어야 하며,
- ④ 적어도 픽셀 $I(i,j+1)$, $I(i+1,j)$, $I(i,j-1)$ 중의 하나는 배경 픽셀 즉 255이어야 한다.
- ⑤ 조건에 맞으면 픽셀을 제거한다.

위의 두 subiteration을 계속하여 더 이상 지울 픽셀이 남아 있지 않을 때까지 계속 적용 시킨다.

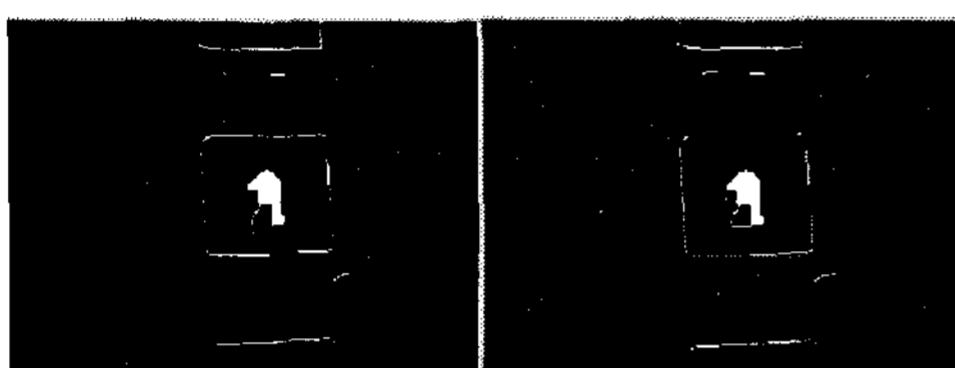


그림 5 세션화 연산 및 사각영역 검출

2.2.5 사각형 검출

숫자 영역을 검출하기 위해서 마지막으로 사각영역을 검출 하여야 하는데 사각영역을 검출하기 위해서 라벨링 작업과 Hough Transform 작업을 병합하여 작업하였다.

먼저 하나에 바탕(검정색)과 다른 픽셀(흰색)을 찾으면 그 픽셀을 따라 상, 우상, 우, 우하, 하 5방향으로 검색해 나아가며 지나간 점은 지나갔다고 표시하고 그 픽셀들에 위치를 새로운 공간에 저장 하면서 픽셀들을 따라 이동한다. 더 이상 움직일 수 있는 공간이 없다면 반대로 상, 좌상, 우, 좌하, 하 이 5방향을 검색하고 이동한다. 이때 더 이상 갈 곳이 없다면 이전에 이동 했던 공간에 좌표를 불러와 그 위치에서 다시 검색 하여 이동하며 이런 순으로 연속적으로 탐색 하여 처음 시작 위치로 다시 돌아오면 도형으로 판단하고 판단된 좌표들을 Hough Transform 작업을 통하여 직선인가를 판단 한 후 직선이 맞다면 사각영역을 추출해 낸다. 영상에서 엘리베이터의 버튼이 가장 큰 사각 영역이므로 엘리베이터 버튼을 제외한 나머지 부분은 바탕으로 판단하면 된다.

그 다음에는 해당 사각영역 이외의 바탕을 전부 바탕으로 만들어 잡음을 제거한다.

2.2.6 숫자만 검출

숫자만을 검출하기 위해서 사각영역이 검출된 좌표를 가지고 원래 영상에서 (이진화 처리가 된 영상) 해당 영역만을 가지고 와서 그 영역에서 라벨링을 거쳐서 숫자만을 찾아내고 숫자 이외의 부분은 전부 제거해야 하기 때문에 영상에서 상,

하, 좌, 우 4부분에 어느 정도의 영역을 배제하고 나머지 라벨링 된 영상에서 상, 하, 좌, 우 좌표를 찾아서 그 부분만을 잘라 낸다.



그림 6. 버튼자르기, 라벨링연산 및 숫자만 추출

III. 숫자 인식을 위한 신경망

3.1 뉴런의 구조

인간의 두뇌의 신경조직은 많은 신경세포(Neuron)들이 상호 연결된 집합체로, 대략 1011개의 신경세포들로 구성되어 있다. 또한 각 신경세포를 가지고 있는 것으로 알려져 있다. 이러한 생물학적 신경세포의 구조는 그림 7과 같다.

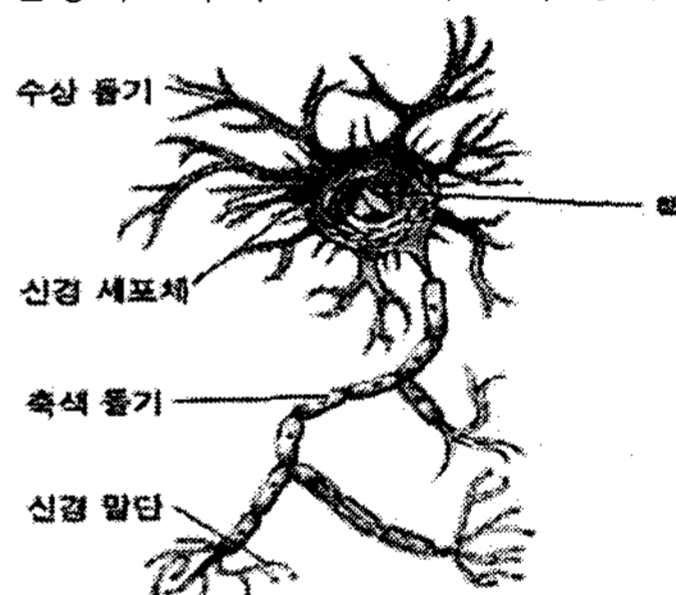


그림 7. 신경세포의 구조

신경세포는 몸체에 해당하는 세포체, 인접 신경세포의 출력이 입력되는 수상 돌기 및 신경 세포에 의해 생성된 신호가 출력되는 축삭돌기(Axon)라 불리는 신경섬유로 구성되어 있다. 또한 신경세포간의 연결은 신경 자극 전도부에서 이루어진다.

3.2 인공 뉴런의 구조

생물학적인 신경세포의 특성을 모형화한 인공 뉴런의 구조는 그림 8과 같다.

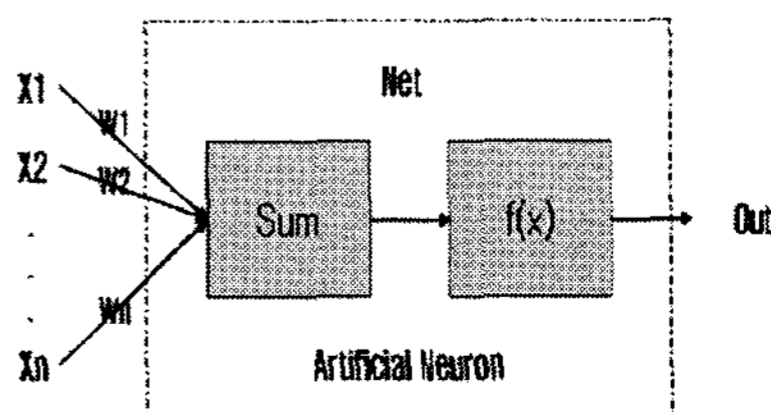


그림 8. 인공 뉴런의 구조

그림에서 X1, X2, X3 ...Xn은 외부 혹은 다른 뉴런으로부터의 입력이며, W1, Q2, ... Wn은 가중치로 로서 신경 자극 전도부의 연결강도에 해당 한다.

3.3 오류 역 전파 학습 알고리즘

3.3.1 원리

역 전파 알고리즘(Back propagation)은 다층이고, feedforward 신경망에서 사용되는 학습 알고리즘이며, 학습의 방법은 지도 학습이다. 즉, 학습을 하기 위해서는 입력 데이터와 원하는 출력 데이터가 있어야 한다.

3.3.2 학습의 개념

먼저, 입력이 신경망의 가중치(weights)와 곱하고 더하는 과정을 몇 번 반복하면 입력의 결과 값인 출력이 나온다. 이 때 출력은 학습 데이터에서 주어진 원하는 출력과 다르다. 결국, 신경망에서는 원하는 출력 데이터 (목표치)에서 출력 데이터 (나온 값)를 빼서 오차를 찾고 오차에 비례하여 출력 층의 가중치를 갱신하고, 그 다음 은닉 층의 가중치를 갱신한다. 가중치를 갱신하는 방향이 신경망의 처리 방향과는 반대 방향이다. 이런 이유로 역 전파 알고리즘이라 한다.

신경망의 처리는

입력 층 → 은닉 층 → 출력 층
 가중치 갱신은
 출력 층 가중치 → 은닉 층 가중치

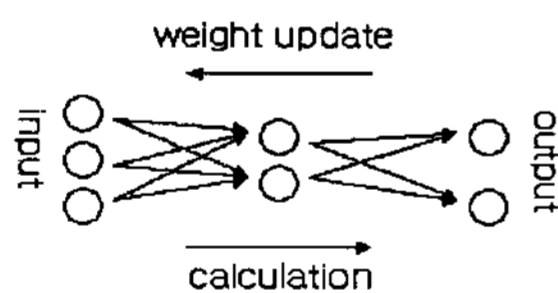


그림 9. BP알고리즘

전체적인 단계는 다음과 같다

- 1 단계 : 초기화, 신경망 상태를 결정하는 가중치를 초기화 (-0.05 ~ 0.05)
- 2 단계 : 은닉 층 계산, 입력 층에 데이터와 가중치의 곱을 가지고 은닉 층을 초기화 한다.
- 3 단계 : 출력 층 계산, 은닉 층의 데이터와 가중치의 곱을 가지고 출력 층을 구한다.
- 4 단계 : 출력 층 오류 계산, 목표치와 실제 출력 차이 값으로 출력 층의 오류를 계산 한다.
- 5 단계 : 은닉 층과 출력 층간의 연결 강도 (가중치) 변화량 계산
- 6 단계 : 은닉층의 오류 계산
- 7 단계 : 입력 층과 은닉 층의 연결 강도 변화량 계산
- 8 단계 : 가중치 값 갱신

IV. 실험 결과 및 평가

4.1 숫자 검출

빛의 양	사각형 검출	숫자 검출
밝은 곳	96/100	92/100
보통	99/100	99/100
어두운 곳	95/100	90/100

4.2 숫자 인식

은닉층의 수	학습 데이터	일반 데이터
3	96/100	92/100
10	98/100	95/100
30	97/100	89/100

사각형 검출이 되면 숫자는 잘 검출 되지만 전처리 과정이 더 필요하고 숫자 인식은 학습 데이터는 잘 인식되는 반면 일반 데이터는 조금 오차가 있는 것으로 나오는데 인식률을 올리기 위해서 더 정확한 은닉층의 수와 학습 데이터를 추가하여야 할 것 같다.

V. 결론

지금까지 숫자 검출과 인식까지의 과정을 살펴 보았다. 숫자 검출에 있어서 전처리 과정에서의 필요성과 여러 방법들을 살펴보았다.

전처리 과정에서 많은 시간이 들지만 정확한 데이터를 찾아내는 데에 있어서 많은 효과를 얻은 것 같다. 그리고 인식률은 잘 나오지만 인식할 데이터가 하나가 늘어날 때 마다 다시 학습을 시켜야 하는 단점과 많은 학습 데이터와 시간이 필요하다라는 단점이 있지만 신경망을 통한 인식에 있어서는 큰 장점을 보이고 있다.

숫자인식은 많은 곳에서 쓰이며 앞으로도 더 발전할 전망이다. 지금은 비록 제한 된 글자만 인식이 가능하지만 필기체 인식도 할 수 있도록 학습 데이터도 추출하고 검출할 수 있는 알고리즘에 대한 추가적인 연구가 필요하다.

참고문헌

- [1]오창석, 뉴로 컴퓨터 개론, 대하출판사, 2000
- [2]http://cafe.naver.com/opencv.cafe, 네이버 openCV 카페
- [3]황선규, 영상처리 프로그래밍, 한빛미디어, 2007
- [4]유현중, 김태우 공역, MATLAB를 이용한 디지털 영상처리, ITC, 2004
- [5]강동중, 하종은 저, Visual C++을 이용한 디지털 영상처리, 사이텍미디어, 2003