

## 신경망을 이용한 비전 시스템의 2차원 물체의 인식에 관한 연구

\*홍진철\*, 김연태\*, 정경재\*, 이해영\*, 이석규\*, 이달해\*

\* 영남대학교 전기공학과

### A Study on 2-Dimensional Objects Recognition of Vision System using Neural Network

\*J.C. Hong\*, Y.T. Kim\*, G.C. Jeong\*, H.Y. Lee\*, S.G. Lee\*, D.H. Lee\*

\*Dept. of Electrical Eng., Yeungnam Univ.,

#### Abstract

This paper proposes a method to recognize object with 2-dimension image. In most cases, it takes too many processes, complicate algorithm and time to recognize object with expert system because of inherent configuration of the object.

This paper includes some processing steps such as pre-processing method, recognition method with neural network and learning algorithm of multi-layer perceptron using error backpropagation.

#### 1. 서 론

화상은 한번에 많은 양의 정보를 얻을 수 있고 무엇보다도 유리한 것은 이러한 정보가 시작적이라는 것이다. 그러나 불행히도 화상은 3차원 정보가 2차원의 형태로 변형되어 인식되기 때문에, 물체를 인식하는 거리, 위치, 조명정도, 기타의 왜곡이 심하여 정보를 분석하는데는 복잡한 알고리즘과 반복 처리가 필요하다.<sup>[1][2]</sup>

효과적으로 물체의 정보를 인식하는데, 다중센서 융합 방식, 스텝레오 비전<sup>[3]</sup>등의 연구가 진행되고 있으며, 최근에는 하나의 비전을 이용한 물체의 연구가 진행되고 있기도 하다. 화상처리를 이용하여 물체를 인식하는 경우, 인식 수준에 따라서 다중의 화상이 필요할 수 있다. 2차원 수준의 물체인식에는 적당한 왜곡의 보정이 있다면 단일 화상으로도 물체를 인식할 수 있으나, 3차원 수준의 물체 인식에는, 왜곡을 보정하거나 한정된 분해능을 가진 카메라의 정밀도를 높이기 위해서, 또는 3차원 정보는 2.5차원의 정보를 정확한 3차원의 정보로 확대하기 위해서는 서른 다른 위치에서의 다중화상을 필요로 한다.<sup>[4][5]</sup>

일반 산업현장에서는 다음의 그림 1와 같이 로보트의 몸체 중 일부에 카메라를 부착하여 작업대상물을 추적하며 작업을 한다. 그러므로 흔히 대상물의 형태나 작업환경은 이미 알려

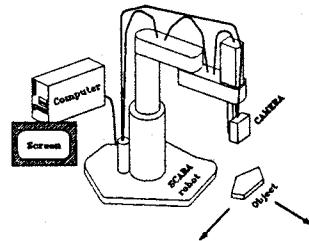


그림 1. 스카라 로봇에 카메라를 장착한 경우

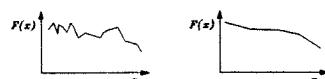
져 있다고 볼 수 있다. 따라서 본 논문에서는 화상의 처리속도나 인식의 수준을 높이기 위해서, 대상물에 대한 정보를 알고 있을 경우, 화상을 통해 2차원 수준의 물체를 인식하는 실시간 처리가 가능한 물체의 인식을 연구하였다.

#### 2. 전처리

본 논문에서 사용한 전처리 단계에는 주로 두가지 작업이 이루어지는데 첫째는 불필요한 잡음을 제거하기 위해 각각의 필터링을 수행하며, 둘째는 세그멘테이션과 외곽선 추출 등이다.

##### 2.1 저역 필터링(Low-pass filtering)

입력영상에 따라서는 카메라의 낮은 분해능, 불량, 또는 입력시의 내부적 문제로 인하여 원하지 않는 잡음이 나타나게 된다. 이러한 잡음에서 고주파 성분을 제거할 경우에는 저역 필터를 사용하면 효과적이다. 이러한 저역 필터링의 효과는 그림 2과 같이 영상이 평활화되어서 고주파의 잡음이 어느 정도 제거되지만, 화상이 경계선에서 흐미해지는 현상(smothing)이 발생하여 영상의 침여도(尖銳度)가 떨어진다는 것을 알 수 있다.



(a) 실제 영상      (b) 저역 필터된 영상  
그림 2. 저역 필터링의 효과

$G(m-1, n-1)$	$G(m, n-1)$	$G(m+1, n-1)$
$G(m-1, n)$	$G(m, n)$	$G(m+1, n)$
$G(m-1, n+1)$	$G(m, n+1)$	$G(m+1, n+1)$

$$F(m, n) = \sum_{i=m-1}^{m+1} \sum_{j=n-1}^{n+1} w_i G_{ij} \quad (1)$$

$w_i$  : 가중치 함수

## 2.2 고역 필터링(High-pass filtering)

컬러 영상의 경우 조명이나 촛점에 의해 영상이 희미하게 나타날 수 있는데, 이 경우 고역 필터를 사용하여 영상의 휘도(contrast)를 증가시켜 선명하게 할 수 있다. 그러나 고주파의 잡음이 증폭될 수는 있다.

## 2.3 외곽선 추출(Edge detection)

에지는 화소의 변화량이 급격한 곳이다. 그러므로 빛의 차이를 나타내는 경사면이 길수록 에지를 나타내는 선은 굵어진다. 이를 방지하기 위해서는 필터링등의 방법을 사용하여 원화상을 개선 하여야 한다.

$(i-1, j-1)$	$(i, j-1)$	$(i+1, j-1)$
$(i-1, j)$	$(i, j)$	$(i+1, j)$
$(i-1, j+1)$	$(i, j+1)$	$(i+1, j+1)$

$$G(i, j) = \frac{\partial F(i, j)}{\partial i} + \frac{\partial F(i, j)}{\partial j} \quad (2)$$

$\partial$  : 미분 연산자

$G(i, j)$  : 픽셀(i,j) 주변의 8픽셀의 그레디언트  
 $F(i, j)$  : 픽셀(i,j)의 화소값

식 (2)과 같이 하나의 화소에 대하여 행방향과 열방향의 그레디언트 값을 조합하여 중심화소에 대한 기울기를 구할 수 있는데 8방향에 대한 그레디언트도 구할 수 있으며 다음의 그림 3과 같이 나타낼 수 있다.

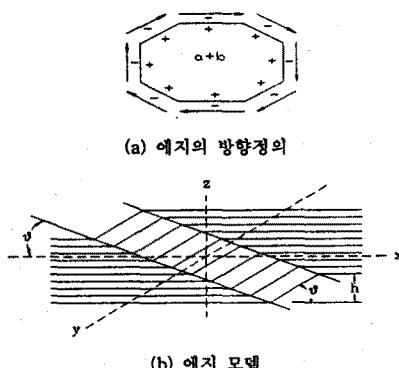


그림 3. 2차원 영상에서의 8방향 에지 모델

그림 3(a)에서 8각형의 각 변은 에지의 방향을 나타내는데, 즉 8방향에 에지가 있음을 알 수 있다. 8각형 도형의 좌측 하단부분의 에지 방향에 대한 2차원 에지 모델을 (b)에서 보이고 있다. 2차원 모델에서는 1차원과는 달리 에지의 방향 정보  $e$ 가 필요하다.

## 3. 물체 인식 처리

### 3.1 물체인식을 위한 신경망

물체의 인식단계는 2장에서 설명한 기본적인 화상처리를 거친 물체의 화상을 부호화하는 과정이다. 화상은 주위환경에 아주 민감하기 때문에 전처리 단계에서의 처리가 잘 이루어져더라도 그 인식수준이 낮을 수 있다.

그리고 인식 단계에 있어서 전문가 시스템을 고려할 수 있는데 이는 뚜렷한 인식 툴이 있어야 하고 개발 시간이 길 뿐만 아니라 비교판단 문제에서의 많은 시간을 소비하지 않으면 안되기 때문에 본 논문에서는 인식시간을 개선하기 위해 신경회로망을 적용하기로 했다.

신경회로망을 이용할 경우, 화상에서 가장 큰 문제가 되는 잡음의 영향을 줄이고, 병렬처리가 쉬워, 일단 퍼셉트론이 충분히 훈련되고 난 후라면 처리속도가 빠르므로 현재 영상 인식 시스템이나 음성 인식시스템에 많이 사용되고 있다.

신경회로망을 이용하여 물체의 인식 시스템을 구현하는데는 두가지 방법이 있다. 첫째로는 물체의 특징점을 추출하지 않고 다중 퍼셉트론(multi-layer perceptron)을 훈련시키는 방법이다. 이것은 화상의 데이터 양이 적고, 실시간 처리가 필요없는 곳에 적당해서 주로 문자의 인식이나 간단한 형상인식에 많이 사용되고 있다. 둘째는 물체의 특징점을 찾아서 이를 이용하여 퍼셉트론을 훈련하는 방법인데, 이것은 화상의 데이터 양이 많고 복잡한 형상의 처리에 적당하다. 본 논문에서 하고자 하는 실험은 대개 큰 사이즈의 화상이 사용되므로 물체의 특징점 추출이 불가피 하다. 즉 그림 6과 같은 순서로 처리되는 알고리즘이 제시될 수 있을 것이다.

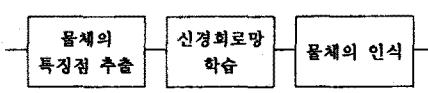


그림 4. 제안된 물체 인식 불러도

### 3.2 오류 역전파의 사용

신경망을 이용한 때던 인식은 알고리즘적이지는 못하지만, 훈련이 가능한 블랙박스 전략으로 볼 수 있다. 인공 신경망이 폐린인식 시스템 설계자의 관심을 끄는 이유는 사전 지식과 내부 시스템의 동작에 대한 상세한 지식이 거의 필요하지 않기 때문이다.

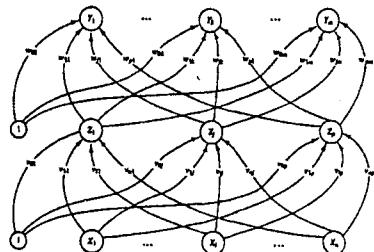


그림 5. 오류 역전파 알고리즘으로 훈련되는 다층 퍼셉트론의 구조.

1980년대 초에 온뇌충을 가지는 다층 퍼셉트론을 훈련시키는 오류 역전파 알고리즘이 발표된 이후로, 여러 분야에서 많은 각광을 받고 있다. 특히, 화상 인식 분야에서, 많은 물체를 인식하여야 하고, 짐음이 많으며, 인식 대상의 유사성이 비교적 큰 경우에는 이 오류 역전파 알고리즘으로 훈련된 다층 퍼셉트론이 비교적 양호한 인식 결과를 나타낸다.

본 논문에서 사용한 신경 회로망도 일반화된 멜타 규칙을 이용하여 훈련되는 하나의 온뇌충을 가진 다층 퍼셉트론이다. 모든 웨이트들과 바이어스는 작은 임의의 값으로 초기화되며, 국부적 최소점에 빠지거나, 유사한 패턴이 반복적으로 제시될 경우의 학습시간을 단축하기 위하여 모멘텀 항을 추가하였다.

본 논문에서는 이러한 신경망을 이용하여 2차원 물체의 인식에 관한 연구를 하였는데, 화상의 실시간 처리를 위해 Database는 2차원의 정보를 갖는 수개의 물체를 보유하고 있으며, 다른 뷰-포인트의 물체를 신경망 기법으로 Database와 비교하여 물체를 인식하는 알고리즘을 제시한다. 물체의 인식에는 95% 이상의 인식률과 실시간 처리가 가능한 적당한 시간이 제시되어야 할 것이다.

## 4. 실험장치 및 비전 시스템의 구조

### 4.1 실험장치의 구성

실험장치는 크게 3개 부분으로 나뉘어져 있는데 CCD 카메라, 프레임 그래비, 컴퓨터 등이다.

실험장치	제품명	제조회사
CCD 카메라	Superior 3Lx Videca-9	
프레임 그래비	VISIONplus-AT	IMAGING Tech.
컴퓨터	IBM Pentium PC	IBM

카메라는 흑백영상(gray level)을 촬영할 수 있는 CCD 카메라이며, 프레임 그래비는  $512 \times 512$ 의 해상도와 4096(12bit)의 월러를 지원한다.

### 4.2 비전 시스템의 구조

프레임 그래비는 주로 3부분으로 나뉘어져 있는데 디지타

이저, 프레임 메모리, 디스플레이 제너레이터가 그것이다. 디지타이저는 카메라로부터 영상을 입력받아 NTSC 신호방식을 RGB로 변환하는 역할을 한다. 프레임 메모리는 영상을 저장하는 곳이며, 외부의 명령에 의해 디지타이저를 통한 신

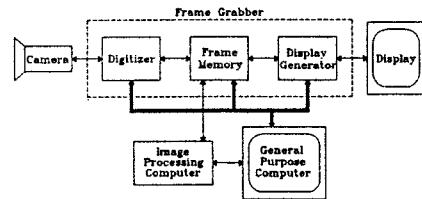


그림 6. 실험장치 구성도

호를 저장하게 된다. 디스플레이 제너레이터는 프레임 메모리의 영상자료를 NTSC신호로 바꾸어 일반 TV 수신용 디스플레이에 전송한다.

본 논문에서 사용한 프레임그래비의 메모리는 12비트의 깊이와 8비트의 너비를 가지고 있다. 한 개의 화소를 저장하기 위해 8~12비트를 사용할 수 있으며, 한번에 8개의 화소를 동시에 읽거나 쓸 수 있다. 그림 7은 프레임 그래비의 메모리 구조와 액세스 방법을 보여주고 있다. 이웃한 8개의 화소값을 동시에 읽거나 쓸 경우에는 X모드를 사용할 수 있으며, 1개의 화소의 월러값(혹은 그레이 레벨값)을 얻을 경우에는 Z모드를 사용할 수 있다.

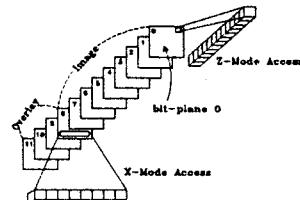


그림 7. 프레임 그래비 메모리 액세스 구조

## 5. 실험 결과

실험에는 1024개의 입력 유닛과 20개의 유닛을 가지는 하나의 온뇌충과 6개의 출력 유닛을 가지는 다층 퍼셉트론 신경망이 사용되어졌고, 오류 역전파 방법을 이용하여 훈련을 하였다.

다음의 그림 8과 같이 6개의 Database를 사용하였으며, 에러를 0.01이 될 때 까지 대략 1500번의 반복 과정을 통하여 훈련을 완성하였다. 훈련된 퍼셉트론의 웨이트와 각 유니트의 바이어스는 디스크에 저장되어 다음에는 훈련없이 바로 사용할 수 있도록 하였다.

입력된 화상에 손실이 없을 경우, 모든 입력에 대하여 물체를 인식할 수 있었으며, 그 확률이 거의 95~96%였다.

다음의 그림 9는 실험에서 사용된 인식 대상 물체중 일부를 나타낸 것이다.

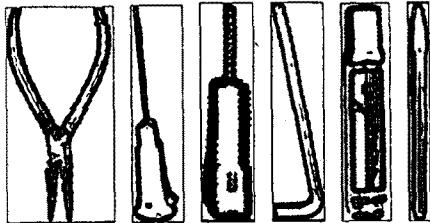


그림 8. 실험 사용된 Database

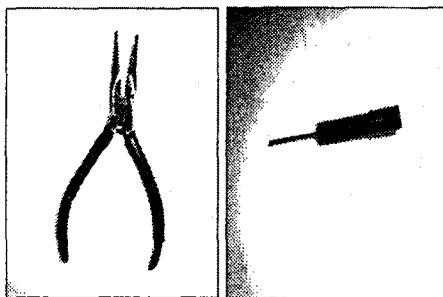


그림 9. 인식 대상물의 예

## 6. 결론

본 논문에서는 일반 산업현장에서 흔히 사용되는 물체의 인식 방법을 실험하여 그 가능성을 제시하였다.

실제 실험에서는 전처리 과정을 제외하면, 아주 빠른 시간에 물체를 인식함을 볼 수 있었으며, 그 인식률도 대단히 높았다.

그러나, 아직 실질적인 인식 시간보다는 전처리 단계에서 인식률의 증대를 위한 다소의 시간이 소비되었으며 이런 시간을 절약할 수 있다면, 매우 훌륭한 계측장치로 사용될 수 있을 것이다. 이를 위해서는 전처리 과정의 개선된 알고리즘이나 하드웨어적인 처리가 필요하리라 생각된다. 또한 2차원 영상에 있어서는 카메라의 위치나 기타의 원인에 의한 좌표의 왜곡은 인식률을 급격히 저하시키기 때문에 이 또한 해결해야 할 과제라 할 수 있을 것이다.

## 5. 참고문헌

- [1] R.A.Benett, Y.C.Shiu and M.B.Leahy.Jr "A Robust Light Invariant Vision System for Aircraft Refueling", IEEE Robotics and Automation Vol.1, p138~p143, 1991
- [2] S.Shina and Y.Nakagwa, "Principal Patches - A Viewport Invariant Surface Description", IEEE Robotics and Automation Vol.1, p226~p231, 1990
- [3] 이종훈·전태현·김도성·이명호, "포토메트릭 스테레오를 이용한 물체표면 방향의 인식", 한국자동제어학회 학술회의 논문집, Vol.1, p816~822, 1990

[4] W.T.Lim and E.K.Teooh, "Segmentation of Range Data Using a Hybrid Approach", SICE, p1281~p1286, 1992

[5] 김도년·조동섭, "3차원 물체의 형상 인식에 관한 연구", 대한전기학회 하계학술 대회 논문집, p542~p545, 1990

[6] 김경민·박중조, "Neural Network를 이용한 고무 타이어의 들출 문자 인식", 한국자동제어학회 학술회의 논문집, p864~869, 1993

[7] D.J.Park, K.M.Nam and R.H.Park, "Edge Detection in Noisy images based on the Cooccurrence matrix", Pattern Recognition Vol.1, p765~p771, 1993

[8] F.Ferreira and S.Ubeda, "Ultra-fast Parallel Contour Tracking with Applications to Thinning", Pattern Recognition Vol.2, p867~p871, 1994

[9] A.Datta and S.K.Parui, "A Robust Parallel Thinning Algorithm for Binary Images", IEEE Robotics and Automation Vol.2, p1181~p1186, 1994

[10] Fank Y.Shih and Wai-Tak Wong, "Fully Parallel Thinning Tolerance to Boundary noise", Pattern Recognition, Vol.2, p1677~p1681, 1994