

로봇 Endeffector 인식을 위한 모듈라 신경회로망

김영부*, 박동선

전북대학교 정보통신공학과

전화 : (0652) 270-2465 / 팩스 : (0652) 270-2461

A MNN(Modular Neural Network) for Robot Endeffector Recognition

Young-Bu Kim*, Dong-Sun Park

Department of Information and Communication Engineering

Chonbuk National University

E-mail : kyb@multilab.chonbuk.ac.kr

Abstract

This paper describes a modular neural network(MNN) for a vision system which tracks a given object using a sequence of images from a camera unit. The MNN is used to precisely recognize the given robot endeffector and to minimize the processing time. Since the robot endeffector can be viewed in many different shapes in 3-D space, a MNN structure, which contains a set of feedforwarded neural networks, can be more attractive in recognizing the given object. Each single neural network learns the endeffector with a cluster of training patterns. The training patterns for a neural network share the similar characteristics so that they can be easily trained. The trained MNN is less sensitive to noise and it shows the better performance in recognizing the endeffector. The recognition rate of MNN is enhanced by 14% over the single neural network. A vision system with the MNN can precisely recognize the endeffector and place it at the center of a display for a remote operator.

Keywords: Recognition System, Robot Vision System, Neural Network

1. 서론

사용자가 직접 접근하여 작업을 하지 못하는 극한의 작업환경에서 안전하고 신뢰성 있는 작업을 수행하기는 어려우며, 이러한 문제점을 보완하기 위해서 로봇이 활용되고 있다. 이러한 로봇제어를 위해서 컴퓨터 비전 시스템을 이용하여 로봇을 원격으로 조정하고 모니터에

작업환경을 정확히 추적하여 보여주는 연구가 필수적이다.

비전 시스템을 이용한 인식(Recognition)시스템은 획득된 영상의 전처리에 소요되는 시간이 많은 비중을 차지하게 된다. 이와 같은 처리 시간을 최소화하고 정확한 물체의 Recognition을 위해서 신경회로망을 적용하였다. 신경회로망은 학습이 끝난 후에는 일반화된 출력을 보이는 일반화 특성을 갖는다. 그러나 단일신경회로망 SNN(Single-Modular Neural Network)을 이용할 경우 신경회로망의 입력이 많으면 학습하는데 시간이 많이 걸리고 학습 후에도 잡음에 민감하다. 로봇 Endeffector는 보는 시각에 따라 3차원공간에서 모양과 크기가 무수히 많은 모습으로 변한다. 이렇게 많은 패턴을 하나의 신경회로망으로 학습하는 것은 매우 어렵다.

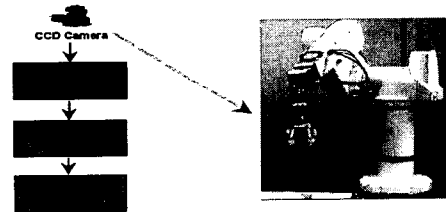


Fig 1. Robot Endeffector Recognition System

로봇 Endeffector 인식시스템(Figure 1)은 영상을 획득하고 획득한 영상에서 전처리를 통해 정보를 파악해 로봇의 Endeffector를 검색하는데, 본 논문에서는 움직임에 따라 다양하게 변화하는 로봇 Endeffector의 형태에 대해 학습이 비교적 쉽고 학습이 끝난 후에

잡음의 영향을 적게 받으며 로봇 Endeffector의 정확한 위치와 크기를 인식하는 MNN(Modular Neural Network)을 구현하였다.

이를 이용하여 로봇 Endeffector의 정확한 위치와 크기를 찾으면 로봇의 인식 및 추적시스템에서 매우 유용하게 이용될 수 있을 것이다.

2. 영상 전처리

로봇 Endeffector의 모습은 로봇의 움직임과 그리퍼(gripper)의 상태에 따라 여러가지 형태를 보이며, Figure 5는 그 예를 보여주고 있다. 이러한 endeffector의 형태를 인식하기 위해 정확한 영상전처리 과정이 요구된다 [1][2].

이러한 로봇 Endeffector의 특징점을 정확히 획득하기 위해 흑백 CCD 카메라를 이용하여 256 레벨의 그레이 값으로 영상을 획득한다. 영상의 크기는 640×480의 영상을 획득하여 480×480의 영상을 호스트로 전송하며 다시 이 영상을 80×80 영상으로 줄여 영상 처리의 양을 줄인다. 획득된 영상에서 원하는 정보를 찾아내기 위해서는 영상 전처리 단계를 거치게 되는데, 먼저 에지를 검출한 후 thresholding 과정을 거친다. 에지검출 알고리즘은 Laplacian of Gaussian을 사용한다. Laplacian of Gaussian은 smoothing filter로 Gaussian을 사용하고 Laplacian으로 에지를 구한다. 이렇게 함으로써 시각에 따라 다양하게 변화하는 로봇 Endeffector 패턴을 정확히 묘사할 수 있으며, 잡음에 대해서도 덜 민감하게 된다 [3]. Figure 2는 로봇 Endeffector의 원형영상과 Laplacian of Gaussian Edge를 보여주고 있다.



Fig 2. Original Image & Laplacian of Gaussian Edge Image

3. MNN을 이용한 물체인식

로봇 Endeffector의 형태는 보는 시각에 따라 3차원 공간에서 다양하게 변화하므로(Figure 3) 모든 패턴을 학습시킬 수는 없다. 신경회로망의 일반화특성을 이용하여 모든 형태를 대표할 수 있는 Endeffector들을 학습시킨다면 그런 다양한 Endeffector들을 인식할 수 있다. 그런데 대표할 수 있는 Endeffector들이 하나의 그룹만으로는 모두를 대표할 수가 없다. 그래서 신경회로망이

비교적 배우기 쉬운 것들을 Clustering 하여 각각의 신경회로망 모듈을 학습시킨다. 로봇 Endeffector의 정확한 위치를 검출하기 위해 신경회로망의 구조를 위치검출 신경회로망과 크기검출 신경회로망으로 각각 모듈화하였고, 위치검출 신경회로망을 여러 개의 신경회로망 모듈로 다시 모듈화하여 역전파 알고리즘으로 설계하였다(Figure 4)[4][5][6].

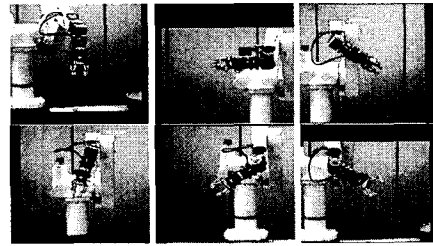


Fig 3. Various Robot Images

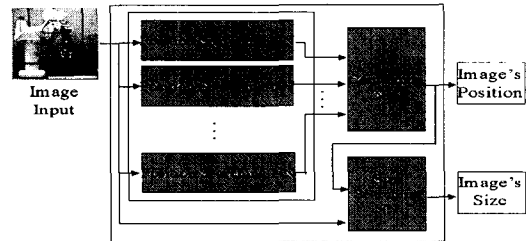


Fig 4. MNN(Modular Neural Network)

여러 개의 신경회로망 모듈로 이루어진 위치검출 신경회로망은 다양한 로봇 Endeffector들 중 비교적 배우기 쉬운 것들을 묶어 각각의 신경회로망 모듈로 학습시킨다. 즉, 모든 패턴들을 하나의 신경회로망이 배우는 것이 아니라 그 패턴들을 여러 개의 배우기 쉬운 그룹들로 Clustering 하여 나눈다음 각 그룹들을 각각의 위치검출 신경망이 배우도록 하는 것이다. 이렇게 각 신경망 모듈들의 학습이 끝나면, 어떤 영상이 입력으로 들어왔을 때 하나의 신경망 모듈만이 Positive 출력을 내야 하는데 여러 개의 신경망이 Positive 출력을 낼 경우 어떤 신경망의 출력을 선택해야 할지를 결정해야 한다. 이 역할을 하는 것이 바로 Position Combining NN인데, 아무래도 자신이 배웠던 패턴이나 그와 유사한 패턴들이 들어오면 출력이 크기 때문에 각각의 Position Searching NN에서 얻은 출력들 중 가장 큰 값을 선택하도록 하고 어느정도의 값 이하이면 출력이 0이 되도록 학습시킨다.

검색된 위치에 있는 Endeffector를 화면중앙에 같은 크기로 위치 시키려면 Figure 5와 같이 Endeffector

들의 크기가 변화하더라도 어느 일정한 크기로 변화시킨 뒤 화면에 보여줘야 한다. 이를 위해서 위치검출 신경회로망 모듈이 Endeffector의 위치를 찾아내면, 그 찾아낸 위치에 있는 Endeffector의 크기가 무엇인지를 알아야 한다. 크기검출 신경회로망 모듈은 이 위치에서 Endeffector의 크기를 알아낸다. 크기검출 신경회로망은 Endeffector의 크기정보를 선형적으로 출력해 주는데 이와 같은 정보를 이용하면 Zoom 기능을 사용할 수 있다 [7][8].

학습이 끝난 후 임의의 영상이 입력으로 들어오면 먼저 위치검출 신경망 모듈들이 이 영상 안에서 로봇 Endeffector의 위치를 검색한다. 검색이 끝난 후에 검색된 위치 중에서 가장 신빙성 있는 위치를 찾아내고, 이 찾아낸 위치에서 크기검출 신경망은 Endeffector의 크기를 검출하게 된다.



Fig 5. Translated Robot Images and endeffectors

4. 실험결과 및 토의

Down sampling 된 80× 80의 영상에 대해 24× 24의 윈도우를 이동시키면서 그 안에 로봇 Endeffector가 있는지를 위치검출 신경회로망이 검색한다. 위치검출신경회로망은 로봇의 Endeffector를 입력 Feature로 사용하여 언제나 화면중앙에 Endeffector가 위치할 수 있도록 하였다.

본 논문에서는 위치검출 신경회로망을 세개의 모듈로 나누어, Normal Endeffector Searchiong NN은 Normal Endeffector를, Left-Rotated Endeffector Searchiong NN은 Left-Rotated Endeffector를, Right-Rotated Endeffector Searchiong NN은 Right-Rotated Endeffector를 학습하도록 했다. 또 로봇 Endeffector의 크기가 다양하므로 Normal Size의 80%,100%,120%로 변화된 영상을 입력으로 사용하였다(Figure 5). 각각의 신경망 모듈은 48,000개의 패턴들에 대해 학습을 시켰다. Position Combining NN는 이들 세개의 신경망에서의 출력중 가장 큰값을 선택하도록 하고 그 값이 0.5 이하인 경우에는 출력값이 0이 되도록 했다.

크기검출 신경망 모듈은 위치검출 신경망 모듈들이

Positive 출력을 내는 위치에서의 Endeffector들을 입력으로 사용하였는데, Normal Image인 경우에는 출력이 0, Normal Image의 80% Image인 경우에는 -1, Normal Image의 120% Image인 경우에는 +1을 출력하도록 학습시켰다.

학습이 끝난 후, 어떤 영상이 입력으로 들어왔을 때 Figure 6과 같은 패턴이 윈도우의 중심에 위치하고 있으면 1에 가까운 값을 출력으로 내보내게 된다. Figure 7의 위 3개의 그림은 크기와 모양이 다른 영상이 입력으로 들어왔을 때 위치검출 신경회로망의 결과를 보여준다

Robot Input Image를 MNN의 각 신경망으로 학습시킨 결과 기존의 단일신경회로망으로 학습시켰을 때보다 학습시간이 단축되었고(Figure 8), Table 1과같이, Robot Input Image를 Size가 75-125% 각각 150개에 대해 테스트했을 때, 기존의 단일신경회로망으로 테스트한 결과는 80%Image는 80%, 100%Image는 81%, 120%Image는 79%의 정확도를 보이는데 비해 본 문에서 제안한 MNN으로 테스트한 결과는 80%Image는 95%, 100%Image는 95%, 120%Image는 93%의 정확도를 보였다(Figure 9)[9].

일단 로봇 Endeffector를 검출하면 크기검출신경회로망은 그 위치에서 검색된 Endeffector의 크기가 얼마인지를 알아낸다. 크기검출 신경회로망은 Figure 7의 아래 3개의 그림과 같이, 입력 영상의 크기가 Normal Size의 80%이면 -1을 100%이면 0을, 120%이면 +1을 출력하고, Input Image가 Normal Image의 80%-120%사이의 크기이면 그 크기를 -1~+1 사이의 선형적인 값으로 출력한다. Figure 10은 로봇 Endeffector의 위치를 찾은 다음 그 위치에서 Size를 검색했을 때 정확도를 나타낸다.

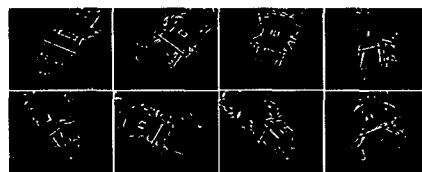


Fig 6. Robot Endeffector Input Featurs

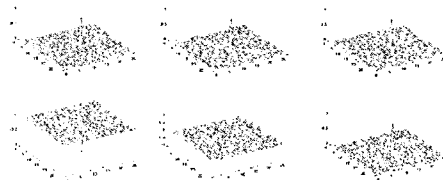


Fig 7. The Results of Position(up) and Size(base)
(80%, 100%, 120%)

크기(%)	75	80	85	90	95	100	105	110	115	120	125
MNN	137	142	142	141	142	142	140	139	137	139	128
SNN	114	120	120	119	119	122	120	119	118	119	113

Table 1. 각 150 개의 영상에 대한 인식결과

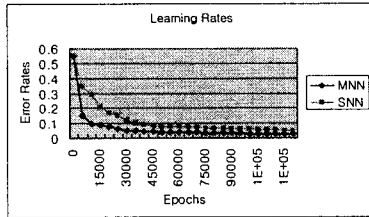


Fig 8. The Comparison MNN with SNN

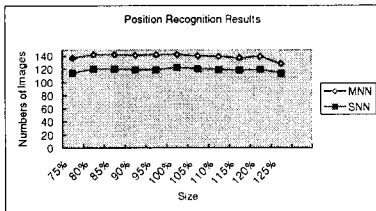


Fig 9. The Test Results of MNN and SNN's position searching

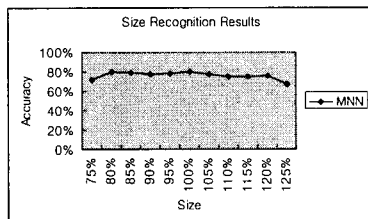


Fig 10. The Recognition Results of Size information

5. 결론

시각에 따라 무수히 많은 모습으로 변하는 로봇 Endeffector를 단일신경회로망 SNN (Single-Modular NN)을 이용하여 학습하고 인식시키는 경우 학습시간이 길고 잡음에 민감하여 정확한 인식이 매우 어렵다.

본 논문에서 제안한 MNN은 SNN보다 학습시간이 많이 단축되었고, 학습이 끝난 후 테스트패턴에 대한 인식의 정확도가 약 94%로 SNN의 80%보다 높은 것을 볼 수 있다. 또 크기에 대한 인식률도 약 80%로 높고 MNN은 SNN에 비해 복잡도가 높는데 비해 그 정확도 역시 상당히 높다.

앞으로 로봇인식 및 추적시스템에서 보다 복잡한 로봇 Endeffector에 대하여 보다 정확한 인식을 한다면 로봇비전시스템에 매우 유용하게 이용될 것이다.

REFERENCES

- [1]. H. J. Lee, L. F. Luang and Z. Cheng. "Multi-frame ship detection and tracking in an infrared image sequence", *Pattern recognition*, Vol. 23, pp. 785-798. 1990
- [2]. N. J. Cutaia and J. A. O'Sullivan. May, "Identification of Maneuvering Aircraft Using Class Dependent Kinematic Models, Electronic Signals and Systems Research Laboratory Monograph ESSRL-95-13", St Louis, MO. 1995.
- [3]. Ramesh Jain, Rangachar Kasturi, Brian G. Schunck. *Machine Vision*, pp. 149-172. 1995.
- [4]. Ming Zhang and John Fulcher. May. "Face Recognition Using Artificial Neural Network Group-Based Adaptive Tolerance (GAT) Trees", *IEEE* Vol 7, No 3. 1996.
- [5]. Dinesh Nair and J. K. Aggarwal. August 25-30. "Hierarchical, Modular Architectures for Object Recognition by Parts", *Proc. 13th Intl. Conference on Pattern Recognition*, Vienna, Austria, pp. A:601-606. 1996.
- [6]. Stephen I. Gallant. *Neural Network Learning*. 1994.
- [7]. Y. Yamada and M. Ishikawa. July. "High Speed Target Tracking Using Massively Parallel Processing Vision", *Proc. IEEE Int. Conf. on Int'l. Robots and Systems*, pp. 267-272. 1993.
- [8]. Granlund, G. H., H. Knutsson, C-j. Westelius and J. Wiklund. Apr. "Issue in robot vision", *Image and Vision Computing* Vol. 12, No. 3. 1994.
- [9]. SeongHyun Baek, DongSun Park, JaeWan Cho, YongBum Lee. "A Robot Endeffector Tracking System Based on Feedforward Neural Network". *FLINS*. pp. 324-331. 1998.