

감각 정보를 이용한 뱀 로봇의 행동구현

강 준 영, 김 성 주, *조 현 찬, 전 흥 태
중앙대학교, *한국기술교육대학교
전화 : 02-820-5297 / 핸드폰 : 017-228-7154

Snake Robot Motion Scheme Using Image and Voice

Joon-Young Kang, Seong Ju Kim, *Hyun-Chan Cho, Hong-Tae Jeon
ChungAng University, *Korea university of Technology and Education
E-mail : bodysoul@orgio.net

Abstract

Human's brain action can divide by recognition and intelligence. recognition is sensing voice, image and smell and Intelligence is logical judgment, inference, decision. To this concept, Define function of cerebral cortex, and apply the result.

Current expert system is lack, that reasoning by cerebral cortex and thalamus, hippocampal and so on.

In this paper, With human's brain action, wish to embody human's action artificially. Embody brain mechanism using Modular Neural Network, Applied this result to snake robot.

I. 서론

인간의 의식이나 정신활동의 형태적 기반에 대해 여러모로 연구되어 왔다. 하지만 아직까지 정확한 구명은 하지 못한 상태이다. 다만, 1870년 이후로 독일로부터 시작한 임상실험 등을 통하여 대뇌기능이 거시적 또는 미시적으로 임상실험을 통해 어느 정도 밝혀졌고, 공학적으로는 1940년 초부터 학습이론 및 신경회로망 이론을 사용하기 시작하였다. 이런 공학적인 부분에서는 특히 전문가시스템(expert system) 분야에서 활성화되었고, 이는 컴퓨터가 현재 인간이 하고 있는 여러 가지 전문적인 작업들을 대신할 수 있도록 하는 것이었다.

하지만 이와 같은 전문가시스템의 경우는 아직까지는 대뇌피질 및 시상, 해마 등의 상호작용에 의한 추론이라는 면에서 많이 미흡하다. 이에 인간의 두뇌작용을 바탕으로한 행동기능을 인공적으로 구현하고자 한다.

이의 구현을 위하여 뇌 행동 메카니즘은 모듈러 신경회로망의 학습 기능을 이용하고, 추론한 결과를 뱀 로봇에 적용해 보고자 한다.

II. 신경망이론

1940년대 초에 뉴런(neuron)을 모델로 공학적 응용을 한 이후 학습 이론 및 신경 회로망 이론이 활성화하기 시작하였다. 신경망에 대한 중요한 성질을 살펴보면 다음과 같다.

- (1) 비선형 시스템 제어 능력
- (2) 병렬 분산 처리 방식
- (3) 학습과 적응성
- (4) 다변수 시스템

2.1 전방향 네트워크

가장 널리 쓰이는 다층신경망 구조는 그림 1에 보이는 바와 같다. 이 전 방향 네트워크의 뉴런의 출력은 식(1)과 같이 표시할 수 있다.

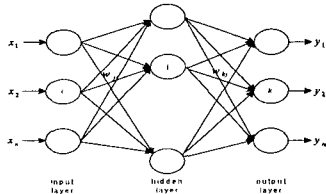


그림 1. 다층 신경망의 구조

$$v_k(n) = \sum_{j=0} w_{kj}(n) y_j(n) \quad (1)$$

$$y_k(n) = f(v_k(n)) \quad (2)$$

여기서 v_j 는 j 번째 뉴런에 출력에 연결강도가 곱해서 모두 더한 값이다. 이 값이 활성화함수의 입력으로 들어간다. y_k 는 k 번째 뉴런의 출력이다

2.2 오차 역전파 알고리즘

역전파 학습 방법은 가장 널리 이용되는 감독학습 (supervised learning)으로 동일한 입력에 대해서 현재 연결강도에 의한 신경망 출력이 비선형 함수 출력을 근사화하도록 한다. 이것은 여러 가 신경망에 대해 역전파되어 연결강도에 따른 여러 기울기가 계산되고 신경망의 연결강도가 조정된다

연결강도들의 변화량은 Gradient descent 방법에 의해 최종적으로 다음 (3), (4)식에 의해 조정된다.

$$w_{kj}(t+1) = w_{kj}(t) + \Delta w_{kj} \quad (3)$$

$$w_{ji}(t+1) = w_{ji}(t) + \Delta w_{ji} \quad (4)$$

$$\Delta w_{kj} = -\eta \cdot \frac{\partial E}{\partial w_{kj}} = \eta \cdot \delta_k \cdot y_j \quad (5)$$

$$\delta_k = (d_k - y_k) \cdot f'(v_k) \quad (6)$$

$$\Delta w_{ji} = -\eta \cdot \frac{\partial E}{\partial w_{ji}} = \eta \cdot \delta_j \cdot y_i \quad (7)$$

$$\delta_j = f'(v_j) \cdot \sum_k (\delta_k \cdot w_{kj}) \quad (8)$$

여기서, d_k 는 출력단 k 번째 뉴런의 목표 값이며, y_k 는 출력단 k 번째 뉴런의 실제 출력 값, η 는 학습률을 나타내고 δ_k 는 역방향으로부터 전달되어 오는 오차 그리고 E 는 cost-function을 나타낸다.

$$E = \frac{1}{2} \cdot \sum_k (d_k - y_k)^2 \quad (9)$$

여기서, 위의 식을 출력 단에서 입력 단까지 계속적

으로 반복하면 각 뉴런의 출력 오차가 감소하도록 연결강도가 변하게 되며, 마지막에는 오차가 없게 되어 학습이 완료된다.

2.3 Hierarchical Neural Network(HNN)

단순신경망의 경우 다음과 같은 단점을 갖는다.

(1) 시스템이 복잡할수록 신경망 구조가 커진다. 따라서, 신경망을 학습하는데 많은 시간과 계산량을 필요로 한다.

(2) 신경망 구조에서, 전체적인 지식은 연결 강도에 의해서 내부적으로 표현된다. 새로운 지식을 기존의 신경망에 첨가 또는 수정해야 할 경우, 전체 신경망이 다시 학습되어야 한다.

신경망에서 학습은 은닉층의 노드 수에 의해서 결정되는데 노드수가 충분하지 않을 때에는 원하는 출력을 얻을 수가 없고, 입력이 증가한다면 신경망에서 원하는 출력을 얻기 위해서는 은닉층의 노드 수를 충분히 늘려야 한다. 노드수의 증가는 신경망의 내부 구조를 복잡하게 할 뿐만 아니라 내부적으로 계산되는 양을 지수적으로 증가하게 한다. 즉, 신경망에서 노드 수의 증가는 학습시간에 있어서 비효율적이다. 신경망의 입출력 데이터 역시 바뀔 경우에 전체 신경망을 다시 학습해야만 한다.

HNN은 고려해야 할 조건이 많은 복잡한 시스템을 여러 개의 작은 부분으로 세분화해서 종합할 수 있다. 각 부분은 출력에 대한 입력 데이터의 관계에 의해 결정된다. 이것은 전체 시스템을 단순화시키고 정확도를 향상시킨다. 각각의 신경망은 각 계층에서 특정 입출력 데이터에 의해 학습이 이루어진다. 만약 학습해야 할 데이터가 바뀐다면 특정 부분의 신경망만 학습하면 되므로 이러한 경우에 학습 시간이 감소된다. HNN에 의해서 전체 신경망 구조는 보다 작아지고 더 적은 데이터에 의해서 학습이 가능하다.

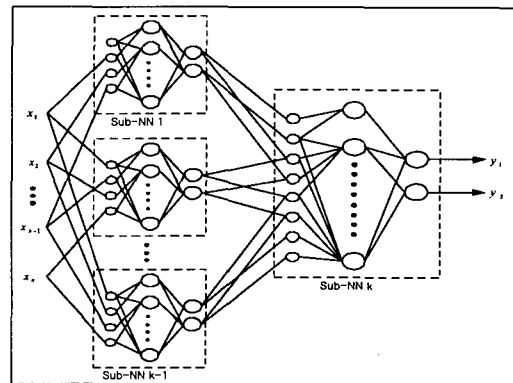


그림 2. Hierarchical Neural Network architecture

III. 감각정보를 이용한 복합적 추론

대뇌기능은 거시적으로는 내장기관의 기능을 통제하는 항상성의 유지, 적응행동을 일으키게 하는 정보처리, 기억·감각·운동 등의 발현, 창조성의 발현, 의식 등의 메커니즘이 밝혀졌고, 미시적으로는 전자현미경의 발달로 신경세포의 기능이 밝혀지기 시작하면서, 신경세포 사이의 기능적 연결부분인 시냅스의 촉진과 억제제를 일으키는 상세한 메커니즘이 밝혀졌다.

이런 인간의 두뇌작용을 크게 지각과 지능으로 나눌 수 있는데, 지각은 음성이나 도형 그리고 냄새의 지각 등이고, 지능은 논리상의 판단, 연역, 귀납 등을 꼽을 수 있다.

본 논문에서는 소리정보와 시간정보라는 복합적인 지각의 판단을 이용하여, 인간의 두뇌를 모방한 행동 추론을 위한 구조를 제안한다.

입력정보를 시각과 청각을 바탕으로 하여 이 두 개의 정보를 각각의 모듈을 통해 1차 판단을 하고, 판단한 특징을 바탕으로 복합적으로 판단하는 모듈 개념의 구조이다.

시각 정보로는 형태와 색상이고 청각정보는 학습된 소리정보와 크기 등을 입력받는다. 이 입력을 바탕으로 모듈라 네트워크를 통해 복합적으로 판단하고, 그 판단결과에 따라서 뱀 로봇이 행동을 한다. 전체 구성도는 그림 3과 같다.

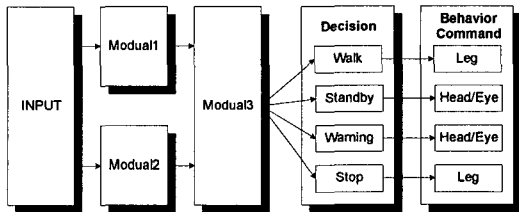


그림3 감각정보를 이용한 행동결정/하달 블록도

3.1 감각정보 행동판단 모듈

첫 번째 모듈은 시각정보를 입력으로 한다.

IBM-PC의 USB 카메라를 이용하여 영상을 획득하고, 획득한 영상으로부터 영상의 색상과 모양정보로부터 특징을 추출한다.

영상 정보는 Pixel의 집합이며 이때 각각의 Pixel에는 빛의 3가지 요소인 Red, Green 그리고 Blue의 데이터를 Byte단위로 가지고 있다. 이 RGB 신호로부터 먼저 세 가지 색상정보를 얻을 수 있고, 이들의 조합으로 모든 색이 표현된다. 이를 이용해 일정한 Boundary를 정하여 일정색상으로 규정하고 그를 찾을 수 있으며, 또는 이 RGB 신호로부터 명도정보도 얻을 수 있

기 때문에 간단한 도형의 구분도 가능하다.

두 번째 모듈의 음성 정보는 마이크로부터 들어오는 소리를 음성인식칩(센서리칩)을 이용하여 미리 인식되었던 음성정보와 비교할 수 있다. 또한 소리의 크기 정도를 입력받는다.

3.2 최종 행동결정 모듈

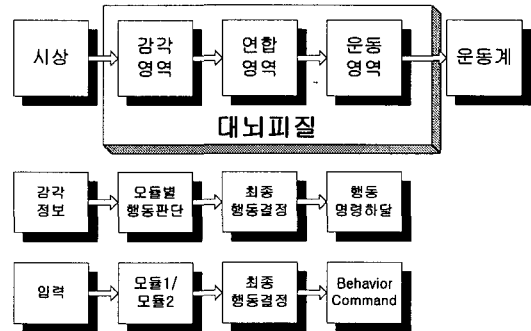


그림 4 대뇌피질 기능 및 영역정의

시상을 통해 들어온 정보를 바탕으로 대뇌피질에서는 종합적으로 판단하고 명령을 내리게 된다. 들어온 정보들은 감각영역, 연합영역, 운동영역을 통해 운동계에 전해지게 되는데 그림4에서처럼 들어온 정보를 일차로 판단하고, 이를 바탕으로 최종 행동이 결정된다. 이 결과를 하달하는 구조로 되어있다.

즉, 모듈1은 신경망을 이용하여 시각정보를 학습하고, 모듈2는 청각정보를 학습하고, 각 모듈의 출력을 바탕으로 종합적인 행동을 추론하여 행동기관에 행동 명령을 하달하게 된다.

3.3 행동명령 하달 모듈

대뇌피질의 연합영역에서 최종 행동이 결정되게 되면 운동피질에 그 명령이 전달되고, 학습된 형태로 행동이 일어나게 된다. 이 모듈에서는 행동이 결정되면, 사전에 기억되었던 내용대로 행동하도록 되어 있다.

구현을 위해 사용된 뱀로봇은 8개의 모듈로 구성되어 있다. 뱀 로봇은 메인 컨트롤러와 8개의 서브 컨트롤러 부분으로 구성되며, 각 Sub Controller는 두 개의 Actuator를 내장하고 있다.

뱀 로봇의 Main Controller에는 8K Bytes Flash Memory를 가지는 Atmel사의 89c52를 사용하였으며, 각 모듈의 서브 컨트롤러는 89c2051를 사용하며, 89c52로부터 받은 정보를 토대로 각각의 Actuator를 제어하게 된다. 89c52와 89c2051은 9600bps의 속도로 Serial 통신을 한다. 89c52는 89c2051로 Actuator ID, 속도 그리고 각도 명령을 내리게 된다. 또한 PC와

89c2051도 Serial 통신을 하는데, PC에서는 화상 정보를 처리하여 동작 명령을 내보내 주게 된다. 또한 단일 펄스로 Actuator를 제어하면 펄스에 상응하는 각도로 움직인 후 Torque가 소멸되기 때문에 89c2051의 Timer를 이용하여 PWM파형을 만들어 지속적으로 펄스를 보냄으로 토크를 유지할 수 있게 하였다.

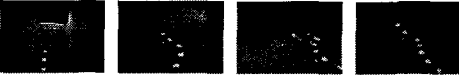
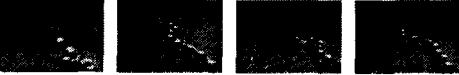
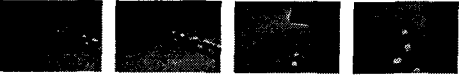


IV. 실험평가

뱀 로봇의 구현과 동작 움직임을 구현한 후, 뱀 로봇의 머리에 위치에 있는 USB 카메라의 시각 정보, 음성인식칩을 이용한 음성정보를 PC에서 입력 받았다.

뱀 로봇 머리에 USB 카메라를 장착하여 영상을 획득하였다. 색상정보는 RGB 각 색상신호를 이용하였고, 320*240의 해상도를 이용하여 실험하였다. 이 색상정보를 바탕으로 색상과, 간단한 모양을 구분할 수 있도록 학습하였다.

음성정보는 음성인식칩을 이용해 기차소리, 자동차소리, 보행신호소리만 이용하였다.

이 테스트를 위해 몇 가지 상황을 시나리오식으로 정하고, 실험해 보았다. 시나리오 구성은 아래와 같다.

●(파랑) + 소리입력없음	: Go
●(파랑) + 자동차소리	: Watch out
	
▲(파랑) + 보행신호	: Go
▲(빨강) + 보행신호	: Go
	
▲(녹색) + 기차소리	: Stop
▲(빨강) + 자동차소리	: Watch out
	
■(빨강) + 보행신호	: Go
■(파랑) + 자동차소리	: Stop
	
●(파랑) + 자동차소리	: Watch out
●(빨강) + 보행신호	: Stop
	

V. 결론

본 논문에서는 소리정보와 시간정보라는 복합적인 지각의 판단을 이용하여, 인간의 두뇌를 모방한 행동 추론을 위한 구조를 제안하였다. 이를 바탕으로 입력 정보로 시각과 청각을 이용하였고, 이 두 개의 정보를 각각의 모듈을 통해 1차 판단을 하고, 판단한 특징을 바탕으로 복합적으로 판단하는 모듈 개념의 구조를 이용하여 뱀로봇에 적용하여 보았다.

지금까지의 연구의 경우는 구조의 제안과 이의 구현이라는 부분에서 초보적인 단계이다. 아직은 대뇌피질의 기능의 일부를 제안하고, 적용해본 것에 불과하다. 그리고 각 모듈별로도 아주 간단한 분류만 하였고, 그에 따른 추론만 하였다. 향후 과제로는 해마에 의한 영상부분과, 협조적 분산 지능의 강화 학습에 관한 연구와 더불어, 감각 정보의 처리 메커니즘의 구조에 대한 지속적인 연구가 필요하다.

본 연구는 '과학기술부 뇌신경 정보학 연구사업'에 의해 지원 받았습니다.

참고문헌(또는 Reference)

- [1] Simon Haykin, *Neural Networks - A Comprehensive Foundation*, Macmillian College Publishing Company Inc., 1994.
- [2] R. K. Easley, "A learning architecture for control based on Back-Propagation neural network", *Proc. of the IEEE Conf. on Neural Networks*, vol. 2 .pp 587-594. 1988
- [3] S. R. Chi, R. Shouresshi, and M. Tenorio, "Neural Networks for system identification ", *IEEE Contr. syst. Mag.*, vol. 10, pp. 31-34, 1990.
- [4] Jordan, M. I., and R. A. Jacobs, 1994. "Hierachical mixtures of experts and the EM alogrithm," *Neural Computation*, vol. 6, pp. 181-214.
- [5] Hodge, L., Auda, G. and Kamel, M., "Learning decision fusion in cooperative modular neural networks", *Neural Networks, 1999. IJCNN '99*. International Joint Conference on Vol 4, pp. 2777 -2781