

# 시각입력을 갖고 음성으로 제어되는 로봇트 팔의 제어

洪 錫 教

( 亞 洲 大 工 大 助 教 授 )

■ 차 례 ■

- |                                |               |
|--------------------------------|---------------|
| 1. 서론                          | 3.3 과제조정계층    |
| 2. 계층적 지능제어                    | 3.4 조정된 준최적제어 |
| 3. 시각입력을 갖고 음성으로 제어되는<br>로봇트 팔 | 3.5 시각 시스템    |
| 3.1 시스템 하드웨어                   | 3.6 실험결과      |
| 3.2 과제조직과 음성입력                 | 4. 결론         |
|                                | 참고문헌          |

## 1 서론

신체장애자를 위한 재활공학 (Rehabilitation Engineering)의 시도는 부러진 다리 대신 지팡이로 혹은 없어진 손 대신 집게로 대치하면서부터 시작되었다. 그러나 이러한 시도는 몇세기동안 큰 발전을 가져오지는 못하였고 2차 대전후 이런 장애자들을 위한 의수, 의족등 많은 보조기구들이 생산되기 시작하였다<sup>1)</sup>

최근 로봇트와 manipulator의 발달과 이들의 응용이 전기로 구동되는 wheel chair 및 전자의수 등의 출현까지 가능하게 해 주었다.

또 디지털 계산기의 발달로 이들 보조기구에 상당히 많은 지능적인 기능(Intelligent Function)을 부여할 수 있게 되었고 이제 거의 실용화단계에 이르고 있다.<sup>3),7),8)</sup>

이러한 연구들이 세계 각 곳에서 활발히 진행되고 있고 그중 대표적인 예로 팔이 없는 사람을 위한 TV 카메라가 부착되고 음성으로 제어되는 로봇트 팔에 관한 연구<sup>6),8)</sup>를 소개하고자 한다. 이를 위해 독립적인 로봇트의 제어를 위해 제시된 계층적 지능 제어시스템을 소개하고 이 방법을 로보

트 팔의 제어에 적용하고자 한다.

## 2 계층적 지능 제어(Hierachically Intelligent Control)

인식시스템 (Cognitive System)은 전통적으로 인간행동에서 나타나는 것과 비슷한 기능을 전자계산기로 실현시키기 위한 인공지능(Artificial Intelligence)의 한 분야로 개발되어 왔다.<sup>1)</sup> 음성인식 및 해석, 영상 해석, learning 그리고 의사결정 (Decision Making) 등의 이러한 기능은 단순한 논리적인 연산으로부터 발달된 이론이 만들어진 예들이다. 주어진 명령 (Command)을 해석하기 위한 의사결정, Learning 등의 지적기능을 갖는 시스템을 제어하기 위하여 요구되는 현대의 기술적인 문제들을 풀기위해서 지능제어 (Intelligent Control)가 개발되어 왔다. 그것은 여러가지 수학적인 프로그래밍 제어기법과 함께 인식시스템의 연구결과들을 이용한다. Manipulator와 Robotics의 분야에도 이러한 방법들이 도입되어 생산공장의 조립라인, 작업환경이 나쁜 곳을 위한 원격조정장치, 화성등에서의 무인 자동차로의 탐사, 불구자를 위한 병원용 보조기구, 기타 자동 로봇트의 제어등의 목적으로

사용되고 있다.<sup>2)</sup> 대부분의 경우 제어 프로세스는 조작자로부터 원격 조정되며 그 기능은 반 독립적이다. 제어 시스템은 행하여야 할 일을 이해하고 주위환경을 파악하고 그 일을 가장 잘 수행하기 위하여 인식 시스템들이 사용되어야 한다. 또 여러 가지 패턴인식 방법이 인식시스템의 각 부분인 각종 센서로부터 나타나는 음성신호, 영상신호 혹은 각종 정보를 분류하고 해석하는데 사용된다. 의사결정 (Decision Making) 및 동작제어 (Motion Control) 는 Dynamic Programming 혹은 Optimal Control 등에 기초를 둔 Kinematic Method, 궤적추적 (Trajectory Tracking) 혹은 동역학적인 방법을 전자계산기로 실현시켜 행해진다. 계층적 지능제어 (Hierachically Intelligent Control) 는 인식 및 제어시스템의 이론적인 접근방법의 하나로 G.N. Saridis에 의해 제시되었다.<sup>11)</sup> 제어지능 (Control Intelligence) 은 "지능의 증가와 함께 정밀도는 감소한다 (Decreasing Precision with Increasing Intelligence)" 는 원리에 의해 계층적으로 분포되어 있다는 것이다. 그 계층적 분포는 다음과 같이 세개의 계층으로 되어 있다.

- 1) 조직계층 (Organization Level)
- 2) 조정계층 (Coordination Level)
- 3) 제어계층 (Hardware Control Level)

조직계층은 이 시스템의 최고 지휘 계층이다. 여기서는 입력과 시스템으로부터의 Feedback를 받고 분석하여 행하여야 할 과제를 정의하고 그 과제를 할 순서에 따라 부과제로 분해한다. 적당한 부과제 목록 (Subtask Library) 과 Learning Scheme 이 조직자 (Organizer) 에게 지적기능을 더해주기 위해 준비될 수 있다.

조정계층은 조직자로부터의 명령과 실행한 부과제를 위한 프로세스로부터의 Feedback 신호를 받아 아래 계층으로의 작업을 조정한다. 조정자 (Coordinator) 는 평가함수와 종단조건 (End Condition) 혹은 가능한 Penalty Function 등을 지정하기도 한다.

최하위 계층은 움직이는 동작을 수행하기 위한 Process를 제어한다. 조작자로부터 받은 평가함수 혹은 종단조건을 해결하기 위해 최적 혹은 준최적 제어이론이 적용되기도 한다. 이런 방법을 시각과 음성입력을 갖는 일반적인 로봇의 제어에 적용하려는 연구가 진행되고 있다.<sup>8)</sup> 이와같은 계층적 지능제어시스템의 블록선도는 그림 1과 같다.

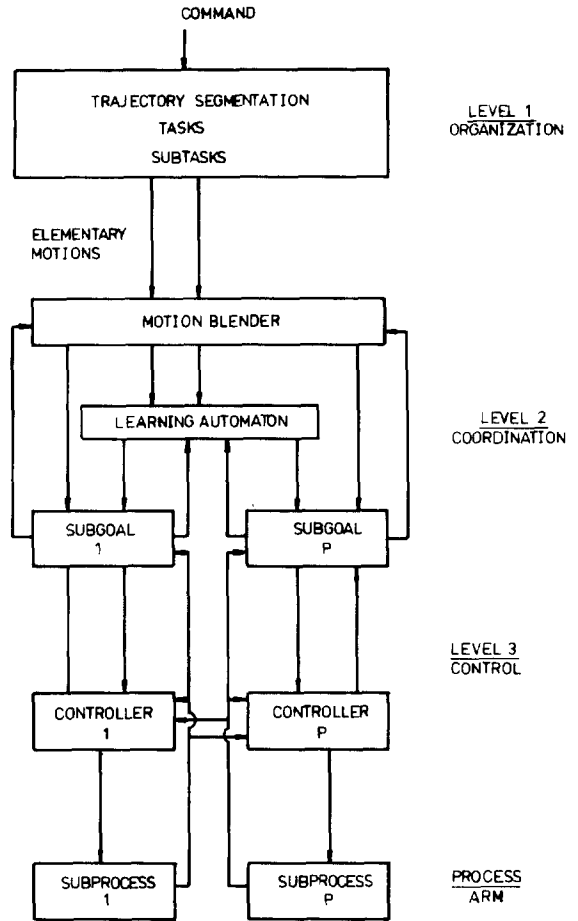


그림 1. 계층적 지능제어

### 3 시각입력을 갖고 음성으로 제어되는 로봇팔

계층적인 지능을 갖는 로봇 시스템은 7개의 자유도를 갖는 로봇 팔 (3개는 팔의 위치 선정, 다른 세개는 손의 방향선정, 그리고 하나는 손가락의 잡는 동작을 나타낸다), TV 카메라, 팔의 센서, 그리고 전자계산기 및 이들 로봇과의 interface 등으로 구성되어 있다. 계층적 지능시스템으로 표시한 블록선도는 그림 2와 같고 기계적인 부분 및 다른 Hardware 는 그림 3에 표시되어 있다.

사람이 원하는 일을 하기 위해서 계산기로 제어되는 이 시스템은 다음의 특성을 갖는다.

- 1) 사람과 로봇과의 통신-사람의 명령을 알아듣고 그에 맞게 행동해야 한다.
- 2) 조정된 동작제어-사람의 도움없이 각종 제

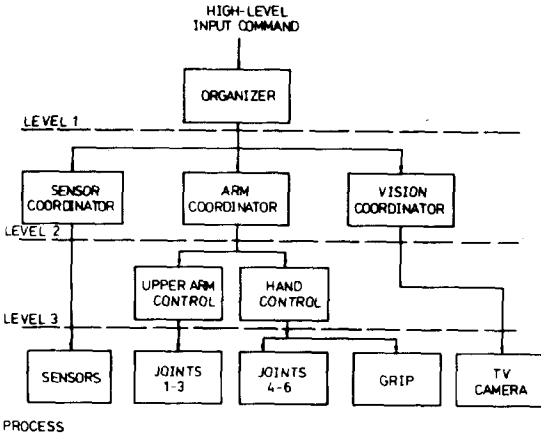


그림 2. 시각 입력력을 갖고 음성으로 제어되는 로봇트 팔의 계층적 지능제어

어가 행해져야 한다.

3) 주위 환경에의 적응-TV 카메라 및 다른 외부 센서로부터의 Feedback 신호를 읽고 거기에 알맞는 일을 수행해야 한다.

이와 같은 특성을 갖는 일반적인 로봇트를 제어 하는데는 계층적 지능제어가 가장 적당하다. "지능의 증가와 함께 정밀도는 감소한다."는 원리에 의해 최하위 계층 즉 실행단계에서는 가장 정밀하게 일을 수행해야 하며 그러기 위해 정확한 model 이 필요하게 된다. 로봇트의 이런 model 은 상태변수 방법에 의해 취급된다. 그 위의 계층은 조정자로 각각의 부과제가 각자 특성을 갖고 집합되어 아래 계층에서의 의사결정을 감독하고 시스템의 특성을 개선하기 위해 Learning 의 일도 수행한다. 최고위 계층인 조직계층은 주어진 음성입력을 분석하여 그 과제가 무엇인지를 결정하고, 과제를 조직하고, 적당한 제어패턴을 결정하여 준다.

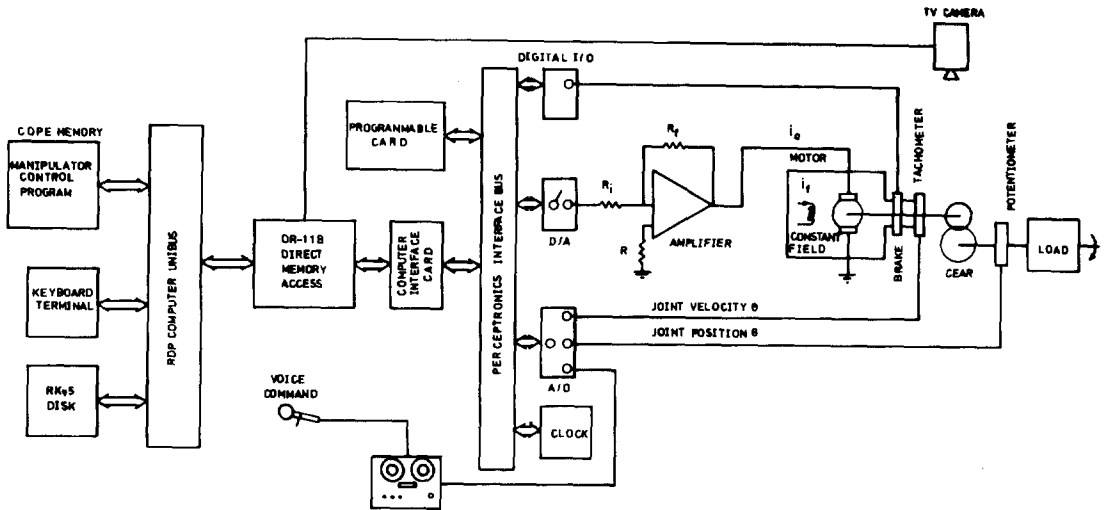


그림 3. 시스템 하드웨어의 블럭선도

### 3.1 시스템 하드웨어

7개의 자유도를 갖는 MIT Scheinman arm 이 본 연구를 위해 사용되었다. 이 로봇트의 손은 팔의 어깨로부터 반경 50cm내에서 일할 수 있고 1.4kg까지 들어 올릴 수 있다. 모든 접속부분은 회전접속이고 직류전동기로 구동된다. 각 접속부분에는 회전계와 포텐시오펀타가 붙어 있어 속도와 위치의 Feedback 신호를 제공한다. 전자계산기는 P DP 11/45로 128K Words 의 Core Memory를

갖는다. I/O 주변장치외에 128 gray level를 갖는 2대의 TV카메라와 Rantek Digitizer 가 부착되어 있다.

### 3.2 과제조직(Task Organization)과 음성입력

조직계층은 다음의 두가지 역할을 한다. 하나는 조작자 즉 사람과의 통신하는 일이고 다른 하나는 서로 다른 주위 환경과 각종 제어방법을 위해 여러 가지 과제를 조직하는 일이다. 로봇트의 외부감시

를 최소화하기 위해 계층적지능제어가 사용되지만 과제를 수행하기 위해서는 최소한 사람의 명령을 받아들이는 일을 해야 한다. 따라서 조직자는 사람으로부터 음성입력을 받아 그 명령을 이해하고 그 일을 수행하기 위한 여러가지 과제를 조직한다. 조직자의 실현은 음성인식방법을 사용하여 이루어진다. 음성을 1초동안 10 KHz로 샘플링하여 30개의 시편(time - segment)으로 나누고 각 시편으로부터 4개의 특성을 뽑아낸다. 4개의 특성이란 극주파수(Pole Frequency), 표준화된 오차(Normalized Error), 영 교차회수(Zero Crossing Rate) 그리고 절대크기(Absolute Magnitude)이다. 이 값들이 서로 다른 말을 구별하는데 사용되고 또 유효하다. 따라서 한 말은 4×30

행렬로 표시되는 "Template"로 나타낼 수 있다. 전자계산기는 미리 잘 훈련된 사람에 의해 행해진 말들의 평균치를 취해 기준 Template를 만든다. 음성인식은 이 기준 Template와 주어진 음성신호를 비교하여 가장 가까운 기준 Template를 취함으로써 행해진다. 이상과 같은 방법을 사용하면 한 말을 인식하는데 1초의 샘플링시간을 합쳐 약 7초가량이 소요된다. 로봇트 팔의 동작을 위해 16개 단어를 각각 24번씩 반복했을때 약 88%의 인식의 정확성을 얻었다. 이들 16개 단어 및 이들의 조합이 표 1에 나타나 있다.

3.3 과제 조정계층(Task Coordination Level)

기능적으로 제어되는 로봇트의 기능은 그림 2에

표 1. User Commands and Their Meanings

1. TV	PLEASE	: turn on TV
	THANKS	: turn off TV
	MORE	: advance TV channel clockwise
	CHANNEL LESS	: advance TV channel ccw
	THANKS	: end of task
	MORE	: increase TV volume one unit
VOLUME	LESS	: decrease TV volume one unit
	THANKS	: end of task
	PLEASE	: turn on radio
2. RADIO	THANKS	: turn off radio
	MORE	: advance radio channel clockwise
	CHANNEL LESS	: advance radio channel ccw
	THANKS	: end of task
	MORE	: increase radio volume one unit
	VOLUME	LESS
THANKS		: end of task
3. DRINK		PLEASE
4. LIGHTS	THANKS	: replace cup
	PLEASE	: turn on lights
5. BOOK	THANKS	: turn off lights
	PLEASE	: fetch the book
6. FAN	THANKS	: replace the book
	PLEASE	: turn on fan
7. WINDOW	THANKS	: turn off fan
	PLEASE	: open drapes
8. BED	THANKS	: close drapes
	MORE	: raise bed one unit
	LESS	: lower bed one unit
	THANKS	: end of task
9. EAT	MORE	: get another bit to eat
	LESS	: get another bit to eat
	THANKS	: replace spoon, end of task
10. NURSE		: actuate nurse call button

서와 같이 크게 3개의 과제로 나뉘어 진다.

- 1) Sensory
- 2) Vision
- 3) Mechanical Motion

첫번째 과제는 근접센서, 압력 게이지 혹은 그 밖의 Sensor 들로부터 정보를 수집하는 일이다. 적당한 정보를 수집하고 데이터를 처리하여 필요한 Feedback 정보를 제공하는 일은 계층적 지능제어의 제 2계층인 조정자가 해야 할 일인 것이다.

두번째 과제는 한 개 혹은 그 이상의 카메라로부터의 시각정보를 처리하는 일이다. 이 작업은 물체의 인식, 분류 그리고 물체의 추적등과 동작제어를 위한 시각 Feedback 정보등을 포함하고 있다.

마지막 과제는 종단조건, 동작의 종류, 손의 방향 등에 관한 정보를 바탕으로 적당한 제어이득(Control Gain)을 결정하는 일을 하는 것이다.

### 3.4 조정된 준 최적제어

로봇트 방정식의 비선형성과 차수때문에 로봇트 제어를 위한 최적제어하는 구하기 힘이 든다. 그리고 또 최적해를 구하였다 하여도 최적제어기(Optimal Controller)를 real-time으로 실현시키는 거의 불가능하다. 따라서 로봇트 제어를 위하여 실현시키기 쉽고 구조가 간단한 준최적제어기(Suboptimal Controller)가 개발되어 왔다. 이 제어기의 구성은 선형 제어량과 비선형 제어량의 합으로 다음과 같이 표시된다.

$$u(x) = u_{NL}(x) + u_L(x) \quad (1)$$

또 비선형 최적제어의 근사이론(Approximation Theory)이 팔의 위치선정을 위한 효율적인 준최적 Feedback 제어기의 결정을 위해 연구, 개발되고 있다.

로봇트의 위의 3개의 접속부분을 위한 방정식은 Lagrange equation으로부터 다음과 같은 상태방정식으로 표시된다.

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ J^{-1}(x) N(x) \end{bmatrix} + \begin{bmatrix} 0 \\ J^{-1}(x) \end{bmatrix} u \quad (2)$$

여기서  $X(t)$ ; 시스템의 상태변수  $= (\theta_1, \theta_2, \theta_3, \dot{\theta}_1, \dot{\theta}_2, \dot{\theta}_3)^T$

$J^{-1}(x) N(x)$ ; Gravitational, Coriolis and

centrifugal torques and gear friction of motor  $u(x)$ ; 외부에서 가해준 torque이다.

그리고 주어진 평가함수는

$$IP(u) = \int_0^T [ (x - x^d)^T Q (x - x^d) + u^T u ] dt \quad (3)$$

로 주어진다. 그러면 문제는 주어진 시스템 상태 방정식(2)를 만족하면서 평가함수(3)을 최소로 하는 제어법칙  $u(x)$ 를 구하는 문제가 된다. 근사이론을 사용하여 준최적해를 구하는 문제는 현재도 연구가 진행중에 있다.

### 3.5 시각시스템

본 로봇트 시스템의 중요한 부분의 하나는 로봇트를 위한 시각조정시스템(Vision Coordinating System)이다. 시각시스템은 영상 데이터의 분석을 위해 고정된 문턱값을 사용했고, 인식알고리즘을 위해 면적과 둘레를 특성치로 사용하였다. 이를 이용하여 이 시스템은 주위환경에서 물체를 식별하고 그 위치를 계산하는 것이다. TV 카메라는 작업공간을 128 gray level을 갖는  $256 \times 256$  frame으로 나누고 그것을 저장한다. 저장된 데이터로부터 Fixed Threshold Segmentation Method에 의해 물체를 주위 배경으로부터 분리한다. 그후 면적( $y_1$ )과 둘레( $y_2$ )를 계산하여 특성벡터  $y^T = [y_1, y_2]$ 를 만들고 물체를 식별하기 위해 선형 식별함수  $g(y)$ 를 사용한다.

$$g(y) = w_1 y_1 + w_2 y_2 = w^T y \quad (4)$$

이 되고 여기서  $w$ 는 Weighting Coefficient이다.

식(4)로 계산된 값을 원래 주어진 물체에 대한 값들  $g(y^*)$ 와 비교하여 가장 근사한 것으로 그 물체가 무엇인지를 결정한다.

### 3.6 실험결과

음성입력, 시각을 이용한 물체인식, 그리고 준최적제어를 사용하여 다음과 같은 실험을 한 결과 성공적으로 이루어졌다.

- 1) 물이 들어 있는 컵을 집어 환자에게 마시게 가져다 준다.
- 2) TV를 키고 채널을 바꿔준다.

3) 간호원을 부르는 단추, 혹은 전등을 키는 단추등을 누른다.

#### 4) 결 론

이 글을 쓰는 목적은 두가지이다. 하나는 익숙하지 않은 환경에서 사람과의 관계를 최소로 하면서 거의 독립적으로 일을 수행하는 로봇의 제어를 위한 계층적 지능제어 방법을 소개하는 것이고 다른 하나는 그 방법을 불구자를 위한 보조기구의 개발에 적용할 수 있는 예를 보이는 것이다. 시각압력을 갖고 음성으로 제어되는 로봇 시스템이 Purdue University의 AARL에서 개발되었고 그 시스템에 계층적 지능제어방법을 효율적으로 적용하고 있다. 그외에도 이 방법은 독립적인 로봇의 지능을 요구하는 시스템 즉, 교통제어시스템, 원자로 제어로봇, 우주개발로봇 혹은 자동 생산라인에 투입된 로봇의 제어에 까지 적용될 수 있을 것이다.<sup>12)</sup>

그러나 이와같은 로봇들은 비교적 간단한 Computer 구조 혹은 간단한 Hardware에도 불구하고 실제로 생산되기 위해서는 인공지능(Artificial Intelligence)때문에 상당히 많은 비용이 요구된다. 그러나 병원 등에서 음성으로 제어되는 로봇를 시분할(Time Sharing)등의 방법을 사용하여 응급실 등에서 여러환자가 공동으로 이용할 수 있을 것이다. 그리고 이 분야의 계속적인 인구가 병원 혹은 재활센터(Rehabilitation Center)등과의 긴밀한 협조하에 이루어져야 될 것으로 생각된다.

#### 참 고 문 헌

1) Alderson, Samuel, The Electric Arm, Chapter 13 in Klopsteg and Wilson's, "Human Limbs and Their Substitutes", McGraw-Hill, 1954, reprinted by Hafner press, 1969.  
 2) Bejczy, A. K.; "Remote Manipulator System

Technology Review", Tech. Rep. No. 760-77, Jet Propulsion Lab., Pasadena, CA, July 1972.  
 3) Freedy, A., Hull, F. C., Lucaccini, L. F., and Lyman, L.; "A Computer-Based Learning System for Remote Manipulator Control". IEEE Trans. Systems, Man, Cybern., SMC-1, (4), 356-364, 1971.  
 4) Fu, K. S.; "Learning Control Systems and Intelligent Control Systems : An Intersection of Artificial Intelligence and Automatic Control", IEEE Trans. Automatic Control, AC-16, (1), 70-72, 1971.  
 5) Fu, K. S. and Fung, L. W.; "Decision Making in a Fuzzy Environment", Tech. Rept. TR-EE 73-22, Purdue University, West Lafayette, IN, May, 1973.  
 6) Graham, J. H., Saridis, G. N.; "Linguistic Methods for Hierarchically Intelligent Control", TR-EE 80-34. Report, Purdue University, Dec. 1980.  
 7) Graupe, D. et al; "A Microprocessor System for Multifunctional Control of Upper Limb Prostheses via EMG Identification", IEEE Trans. on Aut. Control, Vol. AC-23, pp.538-544, Aug. 1978.  
 8) Lee, C. S. G., Saridis, G. N.; "Computer Control of a Trainable Manipulator", Technical Report TR-EE 78-42, Purdue University, December, 1978.  
 9) McGee, R. B.; Control of Legged Locomotion Systems, Proceedings 1978, JACC, San Francisco, June 1977.  
 10) Rosen, C. A. And Nilsson, N. J.; "An Intelligent Automation", 1967 IEEE International Convention Record, Part 9, New York, March 1967.  
 11) Saridis, G. N., "Toward the Realization of Intelligent Controls", IEEE Proceedings, Vol. 67, No. 8, Aug. 1979.