

뉴로제어 및 반복학습제어 기법을 결합한 미지 비선형시스템의 적응학습제어

Adaptive Learning Control for Unknown Nonlinear Systems by Combining Neuro Control and Iterative Learning Control

최진영 · 박현주

Jin Young Choi and Hyun Joo Park

전기공학부, ERC-ACI, ASRI
서울대학교, Seoul 151-742, Korea
jychoi@ee.snu.ac.kr

요 약

본 논문은 뉴로제어 및 반복학습 제어기법에 기반한 미지의 비선형시스템의 적응학습제어 방법을 제안한다. 제안된 제어 시스템에서 반복학습제어기는 새로운 기준 궤적에 대해 시스템의 출력이 원하는 궤적으로 정확히 수렴하도록 하는 적응과 단기간 제어정보를 기억하는 기능을 수행한다. 상대차수만 알고 있는 미지 시스템에 대한 반복학습 법칙의 학습이득은 신경회로망을 이용하여 추정된다. 반복학습제어기에 의해 습득된 제어정보는 장기메모리에 기반한 앞먹임 뉴로제어기로 이전되어 누적 기억됨으로써 과거에 경험된 기준 궤적에 대해서는 신속하게 추종할 수 있도록한다. 2자유도 매니플레이터에 적용하여 제안된 기법의 타당성을 검증한다.

ABSTRACT

This paper presents an adaptive learning control method for unknown nonlinear systems by combining neuro control and iterative learning control techniques. In the present control system, an iterative learning controller (ILC) is used for a process of short term memory involved in a temporary adaptive and learning manipulation and a short term storage of a specific temporary action. The learning gain of the iterative learning law is estimated by using a neural network for an unknown system except relative degrees. The control informations obtained by ILC are transferred to a long term memory-based feedforward neuro controller (FNC) and accumulated in it in addition to the previously stored informations. This scheme is applied to a two link robot manipulator through simulations.

1. 서 론

시행착오에 의한 궤환오차(feedback error)는 생물체에서 제어 및 학습에 중요한 의미를 갖는다. 시행에 의한 오차를 인식하고 응답을 조절 제어하는 궤환제어, 반복시행에 의한 학습 등이 이러한 궤환 오차에 기반하고 있다. 궤환 오차에 의한 학습에 착안한 궤환오차 학습제어 기법이 제안되었다[1]. 이 연구에서는 앞먹임 제어기(feedforward controller)를 알고있는 함수들의 선형조합으로 구성하였기 때문에 일반적으로 사용하기에는 제약이 따른다. 이 앞먹임 제어기를 신경회로망[2]을 사용하여 학습하면 일반적으로 적용이 가능하나 원하는 출력으로의 수렴이 보장되지 않는다.

한편 반복학습에의해 유한 시간구간에서 원하는 출력으로 정확하게 수렴하는 것이 보장된 제어기법으로 반복학습제어(iterative learning control) 기법이 제안되었다[3~5]. 이 제어기법은 수렴성이 보장되나 수렴성을 보장하기 위해서는 입력갱신법칙의 학습이득이 충분조건을 만족해야만 한다[5]. 이를 알기 위해서는 제어대상 시스템의 상태 방정식을 알아야만 하는 문제가 있다. 반복학습제어의 또 다른 문제는 기준 입력 궤적이 바뀌면 학습된 제어 입력이 소용이 없다는 것이다. 즉, 새로운 기준 궤적에 대해서는 처음부터 다시 학습을 시작해야만 한다.

본 논문에서는 뉴로제어 기법과 반복학습제어 기법을 결합하여 상기의 문제들을 상호 보완하는 적응

* 서울대학교 발전기금 일반학술 연구비 지원에 의해 연구되었습니다.

학습제어 방법을 제안한다. 먼저 상대차수만 알고 있는 유사 비선형 시스템에 대하여 수렴조건을 거의 만족시키는 학습이득을 신경회로망을 사용하여 추정하는 방법을 제안한다. 추정된 학습이득을 사용하여 새로운 기준 궤적에 대해 반복학습하여 정확한 추정 능력을 갖도록 적응해 가기 위해 반복학습제어 기법을 사용한다. 새로운 기준 궤적에 대해 습득된 제어정보를 누적기억하기 위해 앞먹임 뉴로제어기를 도입하여 학습시킨다. 이 앞먹임 뉴로제어기는 다양한 기준 입력 궤적에 대한 제어입력들을 누적기억함으로써 반복학습제어기와는 달리 과거에 학습된 기준궤적에 대해서 신속히 기준출력을 추종할 수 있는 일반화 능력을 갖는다. 또한 순수 뉴로제어기와는 달리 정확한 기준 궤적으로의 수렴이 보장된다.

2. 전체 시스템 구조

제안된 시스템은 시행착오에 의한 궤환제어, 반복 학습, 그리고 기억 메카니즘에 기반한다. 생물체의 뇌에는 두 가지 형태의 기억장치, 즉 단기메모리 (short term memory)와 장기메모리(long term memory)가 있다. 단기메모리는 임시적인 정보의 학습과 기억에 관련된 역할을 수행한다. 단기메모리에서 습득된 정보는 장기메모리로 전달되며 기존에 기억된 정보들에 추가되어 누적 저장된다. 이러한 장기메모리에 의해 기존에 경험된 상황에 신속히 응답할 수 있게 된다.

이러한 개념에 착안하여 그림 1과 같은 적응학습제어 시스템을 제안한다. 제안된 시스템은 학습이득 추정기(learning gain estimator: LGE)와 세 개의 제어기로 이루어지는데 제어기에는 되먹임 제어기(feed-

back controller: FBC), 단기메모리 기반의 반복학습 제어기(iterative learning controller: ILC), 장기메모리 기반의 앞먹임 뉴로 제어기(feedforward neuro controller: FNC)로 구성된다. LGE는 상대차수만 알고 있는 시스템의 입출력에 근거하여 반복학습 법칙의 학습 이득을 추정한다. FBC는 학습이 시작되기전의 초기상태에서 시스템을 안정화시킨다. 그러나 FBC는 원하는 출력에 정확하게 수렴하지 않을 수도 있다. ILC는 반복학습을 통해 시스템 출력이 원하는 출력에 정확히 수렴하도록 해준다. 반복학습이 끝난 후에 얻어진 제어정보는 새로운 기준 궤적에 대한 일반화 능력을 갖도록 하기 위하여 장기메모리 기반의 FNC에 이전되어 누적저장된다.

본 논문에서는 다음과 같은 유사 비선형 시스템을 대상 시스템으로 한다.

$$\begin{aligned} \dot{x}(t) &= f(x(t)) + G(x(t))u(t), \\ y(t) &= h(x(t)). \end{aligned} \quad (1)$$

이 때, $u \in R^m$ 는 시스템 입력, $x \in R^n$ 는 시스템 상태변수, $y \in R^m$ 는 시스템 출력이다. 또한, 사상 $f: R^n \rightarrow R^n$, $G: R^n \rightarrow R^{n \times m}$, $h: R^n \rightarrow R^m$ 는 모두 무한번 미분가능하다고 가정한다.

k 번째 반복학습에서 시스템의 제어입력은 다음과 같이 주어진다.

$$u^k(t) = u_{fb}^k(t) + u_{il}^k(t) + u_{fn}^k(t), \quad (2)$$

여기서 $u_{fn}^k(t)$, $u_{il}^k(t)$, $u_{fb}^k(t)$ 는 각각 FNC, ILC, FBC에 의해 생성된 입력이다. 제안된 제어 시스템의 전체적인 제어 과정은 다음과 같다:

1. 직관적으로 설계된 궤환제어기(FBC)에 의해 시

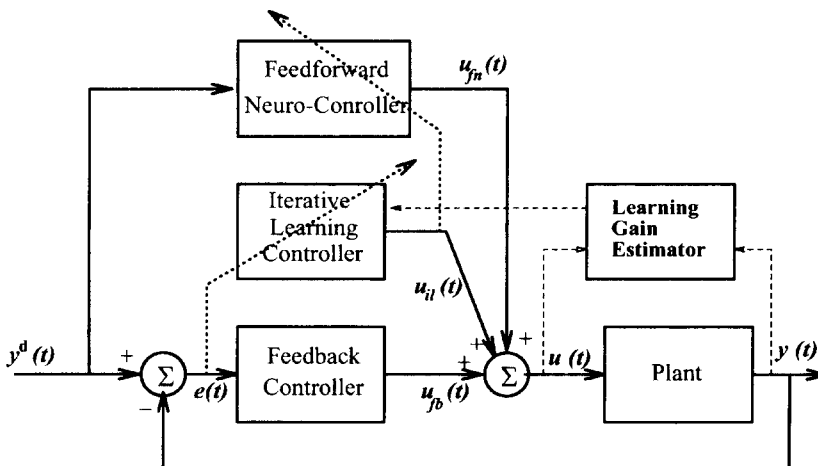


그림 1. 제안된 제어 시스템의 전체 구조.

시스템을 안정화시킨다.

2. 지속적인 여기 (persistently exciting) 입력을 시스템에 인가하여 신경회로망으로 시스템을 식별한다. 식별된 시스템으로부터 반복학습법칙의 학습 이득을 추정한다. 자세한 절차는 2.2.절에서 설명된다.

3. 유한시간구간 $t \in [0, T]$ 에서 특정의 기준 궤적 $y^d(t)$ 를 선정한다.

4. 반복학습제어기 ILC에 의해 제어입력 궤적 $u_i(t)$, $t \in [0, T]$ 를 다음과 같이 갱신한다.

- (a) 최초 반복학습에서는 $u_{ii}^0(t)$ 를 0 으로 설정한다.
- (b) 입력 $u^k(t) = u_{ii}^k(t) + u_{ff}^k(t) + u_{fn}^k(t)$ 를 시스템에 인가한다.
- (c) ILC는 발생한 오차를 이용하여 $k+1$ 번째 입력 궤적 $u_{ii}^{k+1}(t)$ 를 다음과 같이 갱신한다.

$$u_{ii}^{k+1}(t) = u_{ii}^k(t) + u_{fn}^k(t) + S^k(t)e^k(t). \quad (3)$$

여기서 $e^k(t) = y^d(t) - y^k(t)$ 는 원하는 기준 궤적 $y^d(t)$ 와 시스템 출력 궤적과의 오차이다. 여기서 단계 2에서 추정된 학습이득 $S^k(t)$ 가 사용된다.

5. ILC 에서 학습된 제어정보 $u_{ii}^{k+1}(t)$ 를 다음의 절차에 의해 장기메모리 기반의 앞먹임 뉴로제어기 (FNC)로 전달한다.

- (a) FNC 학습을 위한 교사값(teacher value)은 다음과 같이 주어 진다.

$$desired_u_{fn}^{k+1}(t) := u_{fn}^k(t) + u_{ii}^{k+1}(t).$$

- (b) 다음의 조건을 만족할 때까지 FNC를 학습시킨다.

$$|desired_u_{fn}^{k+1}(t) - u_{fn}^{k+1}(t)| \leq \sum.$$

- (c) ILC로부터의 새로운 제어입력 $u_{ii}^{k+1}(t)$ 를 다음의 식에 의해 대체한다.

$$u_{ii}^{k+1}(t) \leftarrow u_{ii}^{k+1}(t) - u_{fn}^{k+1}(t) + u_{fn}^k(t).$$

이 입력값은 FNC 로의 제어정보의 이동후 남은 제어 정보로서 결국 $u_{ii}^{k+1}(t) + u_{fn}^{k+1}(t)$, 즉 앞먹임 제어정보의 총량은 변하지 않기 때문에 FNC 학습 후에도 출력 오차의 정확도는 나빠지지 않는다. 이 단계는 매 반복 학습마다 수행할 필요는 없다.

6. 시간구간 $t \in [0, T]$ 에서의 오차가 주어진 오차보다 작아질때까지 단계 4부터의 과정을 반복한다.

7. 새로운 기준 궤적이 주어지면, 단계 3부터의 과정을 반복한다.

2.1 반복학습 법칙

시스템 (1)에 대하여, ILC에 대한 입력갱신은 모든

시간구간 $t \in [0, T]$ 에서

$$u_{ff}^{k+1}(t) = u_{ff}^k(t) + u_{fb}^k(t) + S^k(t) \begin{bmatrix} e_1^{k(r_1)}(t) \\ \vdots \\ e_m^{k(r_m)}(t) \end{bmatrix} \quad (4)$$

와 같이 갱신된다. 여기서 $S^k: [0, T] \rightarrow R^{m \times m}$ 는 유계인 함수이고, $e^k(t) = [e_1^k(t), \dots, e_m^k(t)] \in R^m$ 는 원하는 출력 $y_i^d(t)$ 과 실제 출력 $y_i^k(t)$ 사이의 오차이다. 즉, $e_i^k(t) = y_i^d(t) - y_i^k(t)$ 이고 $e_i^{k(r_i)}(t)$ 는 $e_i^k(t)$ 를 j 번 미분한 값이다. 또한 r_i 는 i 번째 출력에 대한 상대차수이다. 다음의 정리는 ILC에 의해 시스템 출력이 원하는 출력으로 수렴함을 보인다.

정리 1 : 주어진 비선형 시스템 (1)에 대한 반복학습 제어에서 입력 갱신법칙이 (4)처럼 주어진다고 하자. 만약, $S^k(t)$ 가 모든 시간 $t \in [0, T]$ 에서 다음 부등식

$$\|I - S^k(t)J(x^k(t))\| \leq \rho < 1 \quad (5)$$

을 만족하면 $k \rightarrow \infty$ 일 때, 모든 시간 $t \in [0, T]$ 에서 출력 오차가 평등하게(uniformly) 0으로 수렴한다. 즉,

$$\lim_{k \rightarrow \infty} \|e^k(t)\| = 0, \quad \forall t \in [0, T] \quad (6)$$

여기서 $J(\cdot)$ 는 비간섭(decoupling) 행렬이다[5].

증명 : 학습하는 동안에 $u_{fn}^{k+1}(t) = u_{fn}^k(t)$ 이므로 식 (4)의 좌변에 $u_{fn}^{k+1}(t)$ 를 더하고 우변에 $u_{fn}^k(t)$ 를 더하자. 그러면, 식 (4)은 다음과 같이 된다.

$$\begin{aligned} u_{ii}^{k+1}(t) + u_{fn}^{k+1}(t) \\ = u_{ii}^k(t) + u_{fn}^k(t) + u_{fb}^k(t) + S^k(t)e^k(t) \end{aligned} \quad (7)$$

앞먹임 입력을 $u_{ff}^k(t) := u_{ii}^k(t) + u_{fn}^k(t)$ 로 정의하자. 그러면 위의 정의와 식 (2)로부터 식 (7)은 다음과 같이 된다.

$$u_{ff}^k(t) = u^k(t) + S^k(t)e^k(t), \quad t \in [0, T] \quad (8)$$

만약, 식 (5)가 만족되면, [5]의 정리로부터 식 (8)에 의해 시스템 출력은 원하는 출력으로 수렴한다. (Q.E.D.)

부언 1 : 시스템의 출력이 원하는 출력으로 정확히 수렴하기 위해서는 플랜트와 제어기의 초기값이 매 번 반복학습시마다 같아야한다.

부언 2 : 학습 이득이 비간섭행렬의 역행렬과 같아질 때, 즉,

$$S^k(t) = J^{-1}(x^k(t)), \quad (9)$$

일때 식 (5)에 의해 출력의 수렴성이 보장될 뿐더러 수렴속도도 가장 빠르게 된다.

시스템의 상태방정식을 안다면, 비간섭행렬은 정의

로부터 바로 얻어질 수 있다. 그러나 시스템의 상태 방정식을 모른다면, 비간섭 행렬은 바로 얻을 수 없고, 따라서 빠른 수렴을 보장해주는 학습 이득을 추정할 수 없다. 그러므로 시스템의 상태방정식을 알 수 없는 경우에도 학습이득을 찾는 방법이 필요하며 이것을 다음 절에서 다루도록 한다.

2.2. 학습 이득의 추정

다음 특성을 가정하자.

가정1 : 전체 시스템을 안정화시키는 되먹임 제어기가 존재한다.

가정2 : 시스템의 상대차수는 모두 같다. 즉, $r_1=r_2=\dots=r_m=r$.

가정3 : $z=[y^T, \dot{y}^T, \dots, y^{(r-1)T}]^T$ 에서 x 로 가는 미분동형(diffeomorphism) K 가 존재한다. 즉,

$$x=K(y, \dot{y}, \dots, y^{(r-1)})=K(z). \quad (10)$$

시스템 출력의 r 차 미분은 Lie 미분 [7]에 의해 다음과 같이 주어진다.

$$y^{(r)}=L_f^r h(x)+L_g L_f^{r-1} h(x)u. \quad (11)$$

여기서 $f=[f_1, \dots, f_m]^T$, $g=[g_1, \dots, g_m]^T$ 이다. 이때

$$p(z(t))=L_f h(K(z)) \quad (12)$$

$$Q(z(t))=L_g L_f^{r-1} h(K(z)) \quad (13)$$

이라고 정의하면, 가정3 으로부터 식 (11)은 다음과 같이 표현된다.

$$\begin{aligned} y^{(r)}(t) &= p(z(t)) + Q(z(t))u(t) \\ &= p(y(t), \dot{y}(t), \dots, y^{(r-1)}(t)) \\ &\quad + Q(y(t), \dot{y}(t), \dots, y^{(r-1)}(t))u(t). \end{aligned} \quad (14)$$

식 (13)으로부터 비간섭행렬 $J(x)$ 는 다음의 관계를 갖는다[7].

$$J(x)=Q(z). \quad (15)$$

(14)를 샘플 데이터 시스템 (sampled data system)으로 근사화하기 위해 (14)에서 $y(t)$ 의 i 번째 미분에 Euler 방법 [8]에 의한 다음의 근사식을 대입한다.

$$y^{(i)}(t) \approx \frac{1}{(\Delta t)^i} \sum_{j=0}^i (-1)^{i-j} C_j y(t+j\Delta t). \quad (16)$$

여기서 C_j 는 $i!/j!(i-j)!$ 를 나타낸다. 그러면 다음의 샘플 데이터 시스템을 얻을 수 있다.

$$\begin{aligned} y(t+r\Delta t) &= \bar{p}(y(t), y(t+\Delta t), \dots, y(t+(r-1)\Delta t)) \\ &\quad + \bar{Q}(y(t), y(t+\Delta t), \dots, y(t+(r-1)\Delta t))u(t). \end{aligned} \quad (17)$$

여기서,

$$\begin{aligned} \bar{p}(y(t), \dots, y(t+(r-1)\Delta t)) &= C_1 y(t+(r-1)\Delta t) + \dots - (-1)^r C_r y(t) \\ &\quad + (\Delta t)^r p(y(t), \dots, \frac{1}{(\Delta t)^{r-1}} \sum_{j=0}^{r-1} (-1)^{r-1-j} C_j y(t+j\Delta t)), \\ \bar{Q}(y(t), \dots, y(t+(r-1)\Delta t)) &= (\Delta t)^r Q(y(t), \dots, \frac{1}{(\Delta t)^{r-1}} \sum_{j=0}^{r-1} (-1)^{r-1-j} C_j y(t+j\Delta t)). \end{aligned} \quad (18)$$

따라서, 식 (15)와 (18)로부터 비간섭 행렬은 다음과 같이 표현된다.

$$\begin{aligned} J(x(t)) &= Q(y(t), \dot{y}(t), \dots, y^{(r-1)}(t)) \\ &\approx \bar{Q}(y(t), y(t+\Delta t), \dots, y(t+(r-1)\Delta t)) / (\Delta t)^r \\ &= \hat{Q}(\bar{x}) / (\Delta t)^r. \end{aligned} \quad (19)$$

여기서, $\bar{x}(t)=[y^T(t), y^T(t+\Delta t), \dots, y^T(t+(r-1)\Delta t)]^T$ 가정2와 가정3 으로부터 시스템 (1)은 샘플된 입출력 데이터의 관계식으로 표현되어질 수 있다. 그러므로 입출력식 (17)을 신경회로망을 이용하여 근사화할 수 있다. 학습이득을 얻기 위하여 실제 비간섭 행렬을 사용하는 대신에 다음과 같이 추정된 비간섭 행렬을 사용할 수 있다.

$$\begin{aligned} \hat{S}^k(t) &\approx \hat{J}^{-1}(x(t)) \\ &= (\Delta t)^r \hat{Q}^{-1}(\bar{x}) \end{aligned} \quad (20)$$

만약 신경회로망이 입출력 관계식 (17)을 매우 잘 근사한다면, 식 (5)의 높은 0에 가까워질 수 있고 수렴은 보장된다. 그러므로 수렴속도는 신경회로망이 시스템의 입출력 관계식을 얼마나 잘 근사하느냐에 달려있다.

$J(x)$ 를 추정하기 위해, 부분선형학습 신경망(Piecewise Linearly Trained Network: PLTN)을 사용한다[6]. PLTN은 범용 근사자이기 때문에 식 (17)에서 \bar{Q} 를 임의의 정확도를 가지고 근사할 수 있다. PLTN에 의해 근사화된 i 번째 출력은 다음과 같다.

$$\begin{aligned} \hat{y}_i(t+r\Delta t) &= \sum_{j=1}^M \bar{\mu}_{i,j}(\bar{x}(t)) (w_{i,j}^T \bar{x}(t) + \alpha_{i,1} \mu_1(t) + \dots + \alpha_{i,m} \mu_m(t) + \beta_{i,j}) \\ &= \sum_{j=1}^M \bar{\mu}_{i,j}(\bar{x}(t)) (w_{i,j}^T \bar{x}(t) + \beta_{i,j}) \\ &\quad + \sum_{j=1}^M \bar{\mu}_{i,j}(\bar{x}(t)) \alpha_{i,1} \mu_1(t) + \dots + \sum_{j=1}^M \bar{\mu}_{i,j}(\bar{x}(t)) \alpha_{i,m} \mu_m(t) \\ &= \hat{p}_i(\bar{x}(t)) + \hat{Q}_{i1}(\bar{x}(t)) \mu_1(t) + \dots + \hat{Q}_{im}(\bar{x}(t)) \mu_m(t). \end{aligned} \quad (21)$$

여기서

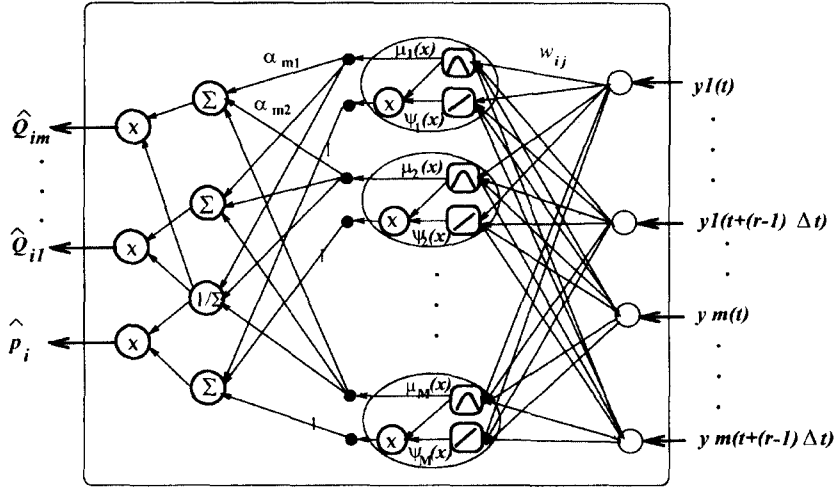


그림 2. 비간섭 행렬을 근사화하는 신경회로망.

$$\begin{aligned} \hat{p}(\bar{x}(t)) &= \sum_{j=1}^M \bar{\mu}_{i,j}(\bar{x}(t)) (w_{ij}^T \bar{x}(t) + \beta_{ij}) \\ \hat{Q}_{iq}(\bar{x}(t)) &= \sum_{j=1}^M \bar{\mu}_{i,j}(\bar{x}(t)) \alpha_{i,qj}, \quad q=1, \dots, m. \end{aligned} \quad (22)$$

여기서 \hat{Q}_{iq} 는 식 (20)에서 \bar{Q} 의 i 행 q 열 원소의 추정치가 된다. \bar{w}_{ij} 는 $(m \times r)$ -차원의 매개변수 벡터이고 β_{ij} 는 바이어스이다. $\bar{\mu}_{i,j}(\bar{x})$ 는 다음과 같은 방사형 기저함수에 기반한 국부화함수이다.

$$\bar{\mu}_{i,j}(\bar{x}) = \exp(-\|\bar{x} - \bar{c}_{i,j}\|^2 / \gamma), \quad j=1, \dots, M, \quad (23)$$

$$\bar{\mu}_{i,j}(\bar{x}) = \mu_{i,j}(\bar{x}) / \sum_{k=1}^M \mu_{i,k}(\bar{x}). \quad (24)$$

여기서 $\|\bar{x} - \bar{c}_{i,j}\|$ 는 입력 \bar{x} 와 노드의 중심 $\bar{c}_{i,j}$ 사이의 유클리드 거리이고 M 은 생성된 전체 유닛의 개수이다. 경쟁변수 γ 는 각 국부영역의 크기를 결정한다.

비간섭 행렬을 근사화하는 신경망의 구조는 그림 2와 같다. 유사 비선형 플랜트 (17)을 식별하는 PLTN의 학습과정에는 앞서 발표한 논문[6]에서와 같이 자기구성과 순환최소자승추정법을 사용하는 학습 방법을 사용하였다.

3. 시뮬레이션

제안된 방법을 2축 로봇 매니퓰레이터에 적용했다. 로봇 매니퓰레이터의 시스템 동역학은 다음과 같이 표현된다.

$$M(\theta(t))\ddot{\theta}(t) + c(\theta(t), \dot{\theta}(t)) + g(\theta(t)) = u(t). \quad (25)$$

여기서, $\theta(t)$, $M(\theta)$, $c(\theta, \dot{\theta})$, $g(\theta)$, $u(t)$ 는 각각 관절의

각도, 관성행렬(inertial matrix), 코리올리(Coriolis)와 구심력, 중력에 의한 힘, 제어입력이다. 이 시스템의 상대차수는 $r_1, \dots, r_m=2, \dots, 2$ 이며, 비간섭 행렬은 $J(x(t)) = M^{-1}(\theta(t))$ 이다[5]. 그러므로 구하고자 하는 학습이득은 관성행렬 $M(\theta(t))$ 가 되어야 한다.

시스템을 안정화시키는 되먹임 제어기로는 비례이득과 미분이득이 다음의 K_p , K_d 와 같은 제어기를 사용했다.

$$K_p = \begin{bmatrix} 60 & 0 \\ 0 & 10 \end{bmatrix}, K_d = \begin{bmatrix} 8 & 0 \\ 0 & 2.5 \end{bmatrix}. \quad (26)$$

이 경우 가정 1~가정 3이 만족되는지는 쉽게 확인할 수 있다. 컴퓨터 시뮬레이션을 위해 링크의 길이와 무게는 $l_1=l_2=0.5$ m, $m_1=m_2=2$ kg으로 했다. 그림 3은 $\hat{l}_1=\hat{l}_2=0.6$ m, $\hat{m}_1=\hat{m}_2=0.6$ kg으로 링크의 길이와 무게에 각각 20%와 25%의 모델링 오차가 존재하는 관성행렬 $\hat{M}(\theta(t))$ 을 학습이득으로 하고, 반복학습 제어를 실행한 결과이다. 결과는 느린 수렴속도를 보여주고 있다.

그림 4는 신경회로망에 의해 추정된 학습이득을 사용한 경우이다. 입출력 관계식 (17)을 근사화하기 위해 PLTN을 샘플링 주기 5 ms로 60분동안 학습시켰다. PLTN이 원하는 학습이득을 잘 근사화했기 때문에 (그림 6), 제안된 방법은 출력 오차가 0.001 rad보다 작아지기 위해서 단지 3번의 학습이면 충분했다. 반면, 20~25% 모델링 오차를 가진 학습이득의 경우 57번의 학습이 필요했다.

그림 5는 FNC를 10가지 다른 궤적에 대하여 학습시킨후 새로운 기준궤적에 대한 추종성능을 보여준다. 그림에서 보듯이 10개의 기준궤적에 대한 반복학

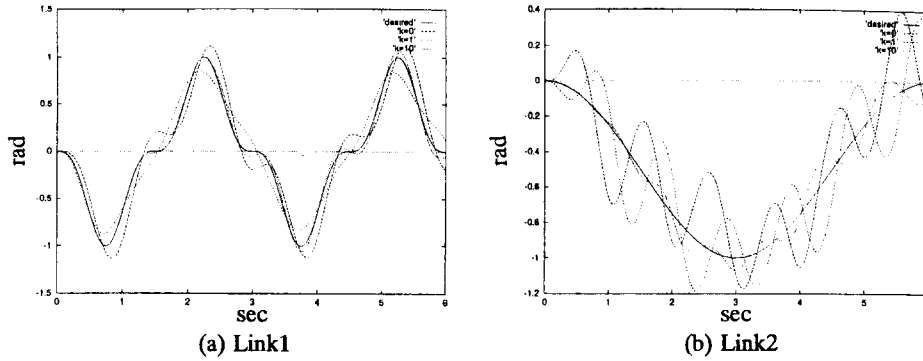


그림 3. 모델링 오차가 존재하는 관성행렬을 학습이득으로 사용했을 때의 수렴성능.

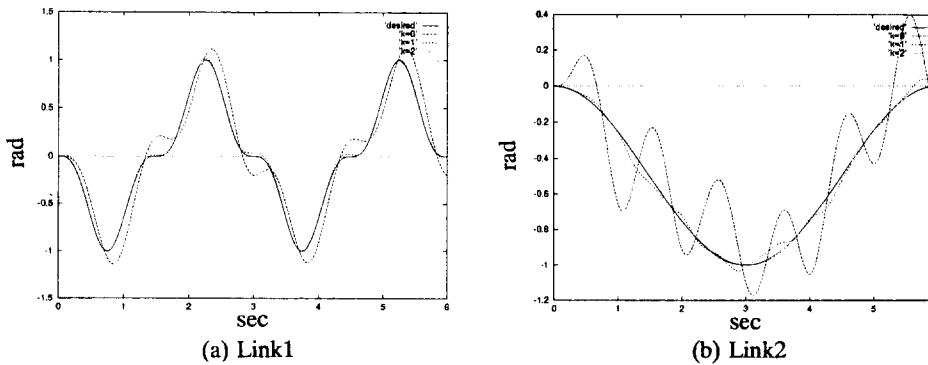


그림 4. 신경회로망으로 추정된 관성행렬을 학습이득으로 사용했을 때의 수렴성능.

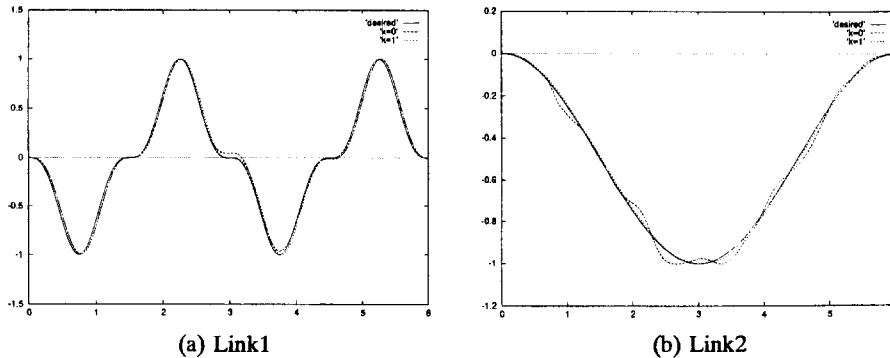


그림 5. 10개의 서로다른 기준체적을 FNC에 학습한 후의 추종성능.

습을 거친후 추종성능은 되먹임 제어기만 있는 경우 보다 현격히 향상되었음을 알 수 있다. 이는 과거 반복학습에 의해 적응되었던 제어 정보들이 FNC에 누적 기억되고 있음을 보여 준다. 이는 FNC가 일반화 능력이 있음을 의미한다. 제안된 방법의 경우 초기 상태의 오차가 작기 때문에 시스템의 출력이 원하는 출력으로 수렴하기 위해서 되먹임 제어기만 사용한 경우보다 반복학습 제어의 시행횟수가 더 적게 필요함

을 알 수 있다.

4. 결 론

본 논문에서는 뉴로제어와 반복학습제어를 결합하여 미지의 비선형 시스템에 대한 적응학습제어 기법을 제시하였다. 적응학습제어는 새로운 명령에 대해서는 적응하여 제어정보를 습득하고 이를 누적기억

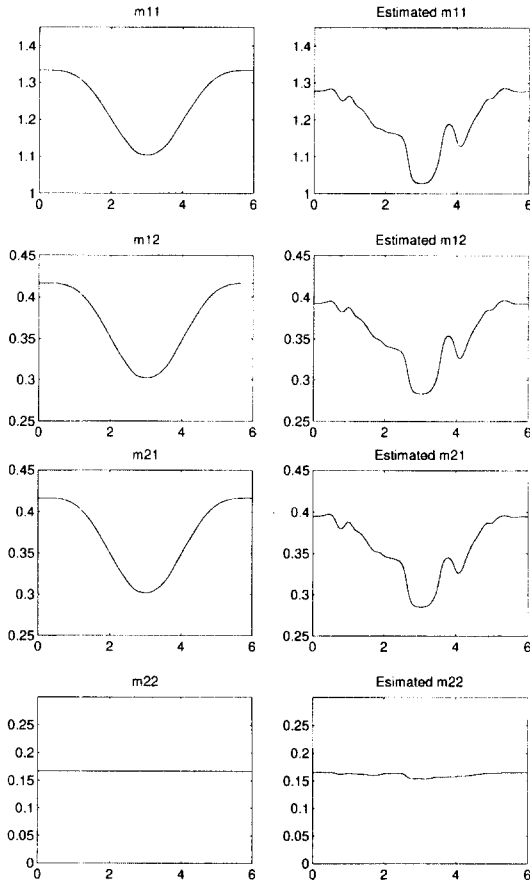


그림 6. 학습이득의 참값과 신경회로망에 의한 추정값 비교.

함으로써 과거에 경험된 명령에 대해서는 신속하고 정확히 응답하는 것을 목표로 한다. 본 논문에서 제안된 방법의 특징을 요약하면 다음과 같다. 첫째, 상대 차수만 알고 있는 임의의 비선형 플랜트에 적용될 수 있다. 이는 신경회로망을 이용하여 시스템의 동역학과 학습이득을 추정함으로써 가능하다. 둘째, 기존의 순수 뉴로제어 기법과는 달리 원하는 출력으로의 수렴이 보장될 뿐 아니라 수렴속도도 빠르게 된다. 끝으로, 제안된 제어방법은 순수 반복학습제어와는 달리 과거에 학습된 제어정보를 누적기억함으로써, 많은

기준입력케속에 대하여 학습이 진행될수록 점진적으로 일반화 능력이 증가한다.

참고문헌

- [1] T. Setoyama, R. Suzuki, H. Miyamoto and M. Kawato., Feedback-error-learning Neural Network For Trajectory Control of a Robotic Manipulator, *Neural Networks*, **1**, pp. 251-265, 1988.
- [2] D. Sbarbaro-Hofer and K. J. Hunt., Studies in Neural Network Based Control, In Warwick K. et al., editor, *Neural Networks for Control and Systems*, pages 94-122. *IEE Cont. Eng. Series* **46**, 1992.
- [3] S. Arimoto., Learning Control Theory for Robotic Motion, *Int. J. Adaptive Contr. Sig. Processing*, **4**, pp. 543-564, 1990.
- [4] L. Gambardella, P. Bondi and G. Casalino., On the Iterative Learning Control Theory for Robot Manipulators, *IEEE Trans. on Robotics and Automation*, **4**(1), pp. 14-22, 1988.
- [5] C.-H. Choi, T.-J. Jang and H.-S. Ahn., Iterative Learning Control in Feedback Systems. *Automatica*, **31**(2), pp. 243-248, 1995.
- [6] J. Y. Choi and H. J. Park., Adaptive Learning Control of Affine Nonlinear Systems Using Piecewise Linearly Trained Networks, *IEEE International Conference on Neural Networks (ICNN)*, pp. 2338-2341, November 1995.
- [7] A. Isidori. *Nonlinear Control Systems*, Springer-Verlag, 1985.
- [8] G. J. Borse., *Numerical Methods with MATLAB*, PWS publishing Company, 1997.

최진영(Jin Young Choi)

제 7권 제 4호 참조
현재 : 서울대학교 전기공학부 조교수

박현주(Hyun Joo Park)

1995년 : 서울대학교 제어계측 공학과 학사
1997년 : 서울대학교 제어계측 공학과 석사
1998년 ~ 현재 : 태국 Naresuan 대학 강사
