

論文 97-34S-6-1

동적 신경망에 기초한 불확실한 로봇 시스템의 적응 최적 학습제어기

(DNN-Based Adaptive Optimal Learning Controller for Uncertain Robot Systems)

鄭 宰 旭 *, 鞠 兌 龍 *, 李 宅 鍾 *

(Chae-Wook Chung, Tae-Yong Kuc, and Taek-Jong Yi)

요 약

이 논문에서는 불확실한 로봇 시스템을 위한 적응 최적 학습제어기를 제시한다. 제안한 제어기는 단순한 형태의 DNN unit을 사용하여 불확실한 파라미터를 추정하고, 미지의 원하는 최적입력을 학습한다. 그리고 Lyapunov 함수를 이용하여 시스템내의 모든 오차신호들이 유계이며, 로봇의 실제궤적이 원하는 궤적으로 global하게 지수적으로 수렴하는 것을 보인다. 제안한 제어기의 효과를 2 자유도를 가진 로봇 머니플레이터에 적용하여 입증하고자 한다.

Abstract

This paper presents an adaptive optimal learning controller for uncertain robot systems which makes use of simple DNN(dynamic neural network) units to estimate uncertain parameters and learn the unknown desired optimal input. With the aid of a Lyapunov function, it is shown that all the error signals in the system are bounded and the robot trajectory converges to the desired one globally exponentially. The effectiveness of the proposed controller is shown by applying the controller to a 2-DOF robot manipulator.

I. 서 론

로봇 머니플레이터를 위한 제어기 설계시 등장하는 문제점으로는 제어에 이용되는 모든 신호들이 유계성(boundedness)을 가지면서 점근적인 궤적추적을 보장할 수 있는가 하는 것이다. 원하는 모든 궤적과 가능한 모든 초기조건에 대해 이러한 문제점을 해결할 수 있는 제어기는 global한 수렴성을 가졌다고 표현할 수 있다. 최근 적응제어를 이용한 로봇 머니플레이터 제어기 설계법들은 global한 수렴성에 대한 엄격한 증명들을 제시하고 있는데^[1, 2], 이들 제어기들은 global한 수렴성 판별을 만족시키지 못하는 근사화 모델, 공칭궤

적에 대한 선형화, 파라미터 변동의 서행성(slow parameter variation)에 기초한 기법들과는 달리 로봇 머니플레이터의 비선형 동력학모델에 기초하여 제어기를 설계하고 있다^[3].

적응제어 기법의 연구와 더불어 70년대말과 80년대 초에 들어, 로봇을 이용한 작업의 대부분은 반복적인 성질을 가진다는 점에 착안하여 주어진 작업궤적에 대해 반복적인 학습 훈련을 거치므로써 궤적에 대한 오차를 점진적으로 줄여가는 반복 학습제어 기법이 등장하였다^[4]. 이 기법은 학습하지 않은 궤적에 대해서도 추적 성능을 보장할 수 있는 능력, 즉 궤적의 일반화 능력에 대해서는 개선해야 될 여지가 있으나 기법자체가 가지는 단순성과 제한적이거나 수학적으로 안정하다는 점으로 많은 연구결과가 제시되고 있다^[5].

이 논문은 전술한 적응제어 기법과 후술한 학습제어

* 正會員, 成均館大學校 電子工學科

(Dept. of Electronic Eng., Sungkyunkwan Univ.)

接受日字: 1997年2月12日, 수정완료일: 1997年6月2日

기법을 이용하여 로봇의 동력학에 기초한 제어기를 설계하고, 이를 동적 신경망^[6, 7]을 도입하여 구성하고자 한다. 이 제어기는 파라미터와 외란의 불확실성을 극복할 뿐만 아니라 광역적인 수렴특성과 주어진 Lyapunov 함수를 최소화함으로써 점근적인 최적제어를 달성한다. 또한 가중치와 강인제어 입력을 적절히 조절함으로써 수렴속도를 임의로 설정할 수 있다. 이어질 절에서는 동적 신경망에 대해 기술한 뒤 로봇 제어 시스템의 설계문제를 정립하고, 적응 최적 학습제어기의 설계 및 로봇 제어 시스템에 적용을 통하여 제안한 제어기의 안정도와 수렴성을 분석하고자 한다. 또한 주어진 제어기를 동적 신경망을 이용하여 적절히 설계할 수 있음을 보이고 수렴특성 등 제어특성이 보존됨을 전체 제어구조를 고찰함으로써 밝힌다. 끝으로 설계한 제어기의 성능검증을 위해 2-DOF(degree of freedom)의 로봇 머니퓰레이터에 적용하고 그 모의실험 결과를 제시하여 제안한 제어기의 유용성을 입증하고자 한다.

II. 동적 신경망 요소

일반적으로 신경망 모델은 생물학적 신경구조를 모방하여 수학적 모델로 만든 것이다. 신경망 모델은 기본 소자인 뉴런(neuron)에 케환신호가 없는 정적 신경망(static neural network:SNN)과 케환신호를 가지는 동적 신경망(dynamic neural network:DNN)으로 나눌 수 있다. 그런데 DNN은 SNN보다 강인한 계산 특성 등 여러 장점들이 있을 뿐만 아니라 생물학적인 신경구조에 더 가까운 것으로 알려져 있다^[8]. 구조적으로 DNN은 뉴런의 출력이 케환되어 다시 뉴런에 입력되므로 원천적으로 동적 특성을 가지며, 구성 자체가 다층 구조인 SNN과는 달리 단층 구조를 가지지만 신경망에서 나타나는 정교한 동적 특성으로 인하여 우수한 표현능력을 가질 수 있다. 이들은 신호처리 분야에서 널리 쓰이는 FIR 필터(finite impulse response filter)와 IIR 필터(infinite impulse response filter)의 구조와 밀접한 관계가 있는데, 예를 들어 시간 지연이 필요한 시스템의 신호처리를 위해 필터를 설계할 때 이론적으로 무한차원의 FIR 필터와 단일 극점(pole)을 가지는 IIR 필터가 등가라는 점은^[8] 구조상으로 IIR 필터와 유사한 DNN이 SNN보다 계산적인 면이나 구현측면에서 상당한 이점을 가질 수 있는 명

백한 이유가 된다.

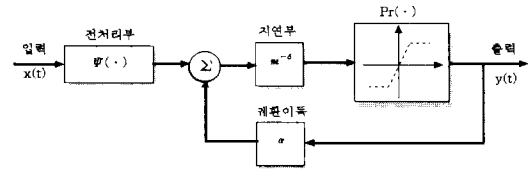


그림 1. DNN 부
Fig. 1. DNN Unit.

이 논문에서는 DNN 형태의 동적 신경망 구조를 가지는 적응 학습제어기를 설계하고자 하는데, 그림 1은 이 논문에서 사용할 DNN unit의 구조를 나타낸다. 그림 1에서 전처리부는 입력신호(후술할 학습신호들)를 학습규칙에 따라 시스템의 출력을 변환시키는 역할을 하며, 케환경로로부터 오는 출력신호와 학습 입력신호를 합산하여 저장하는 역할을 하는 지연부가 있다. 또한 시간지연 δ 는 임의의 양수이며, DNN unit의 출력은 지연부의 값을 지정된 영역으로 사상시키는 비선형 투영함수 $Pr(\cdot)$ 로부터 구해진다. $Pr(\cdot)$ 의 역할은 적용할 시스템의 물리적 특성을 고려하여 출력을 유계시킴과 동시에 과도한 적응신호의 변화를 방지하는데 있다. 케환이득인 α 는 망각인자(forgetting factor)로 작용하는데, 값의 범위는 $0 < \alpha \leq 1$ 을 만족하도록 선택한다. 이와 같은 기본뉴런을 결합하여 DNN을 구성하는데, 일반적인 신경망 응용의 경우 망의 출력이 원하는 출력과의 오차를 발생함에 따라 각 뉴런의 학습이 일어나도록 하여 궁극적으로 원하는 출력이 망의 출력단에서 나타나도록 설계하고 있다. 그러나 정적인 응용, 예를 들어 목표값이나 적용환경 자체가 비교적 정적인 패턴인식(pattern recognition) 분야에는 이러한 설계방법이 좋은 결과를 얻을 수 있지만, 목표값이 시변인 경우가 많고 적용환경 자체가 여러 가지 요인(잡음, 외란, 시스템이 가지는 불확실성 등)에 따라 동적인 요소가 많이 포함되는 제어기 설계분야의 경우는 시스템의 동특성을 고려하지 않고 단순히 출력오차를 학습에 이용하게 되면 적절한 학습이 일어나지 않을 뿐만 아니라 시스템의 안정도에도 심각한 영향을 줄 수 있다. 따라서 최근 연구자들은 제어시스템의 동특성 또는 물리적 특성을 고려하여 빠른 학습 특성과 시스템의 안정도를 보장하는 학습규칙을 개발하고 있는 추세이다^[9, 10].

Ⅲ. 로봇 제어시스템 설계문제와 전체 제어구조

1. 머니플레이터의 동역학 및 문제설정

N-DOF의 강체 로봇 머니플레이터의 동역학 방정식을 표현하면,

$$D(q)\ddot{q} + B(q, \dot{q})\dot{q} + g(q, \dot{q}) + d = \tau \quad (1)$$

로 나타낼 수 있다. 여기서, $q \in R^n$ 이며, 로봇 관절의 일반화 좌표벡터이다. 그리고 $D(q) \in R^{n \times n}$ 인 양의 한정(positive-definite) 관성행렬이고, $B(q, \dot{q})\dot{q}$ 는 구심력 및 코리올리(Coriolis)력을 나타내는 벡터이다. 또한 $g(q, \dot{q})$, d , τ 는 각각 중력과 마찰력, 확정외란, 입력토크의 벡터들을 표시한다. 제어기를 설계하기 위해 식(1)의 로봇시스템은 다음과 같은 성질을 만족한다고 가정한다^[2, 11].

성질1) 관성행렬 $D(q)$ 는 양의 한정성과 대칭성을 가지며 모든 q 에 대해 유계되어 있다.

성질2) 행렬 $\dot{D}(q) - 2B(q, \dot{q})$ 는 비대칭(skew-symmetric) 성질을 가진다.

성질3) $D(q)\ddot{q} + B(q, \dot{q})\dot{q} + g(q, \dot{q}) = Y(q, \dot{q}, \ddot{q})\theta$ 의 관계로 표현할 수 있다. 여기서 $Y(q, \dot{q}, \ddot{q}) \in R^{n \times l}$, $\theta \in R^l$ 은 회귀행렬과 미지의 파라미터 벡터를 나타낸다.

성질4) 미지의 확정외란 $d(t)$ 는 유계되어 있으며, $d(t) = d_1 + d_2(t)$ 로 기술할 수 있고, 부등식 $|d_2| \leq d_0$ 을 만족한다. 여기서 d_0 과 d_1 은 상수이다.

이러한 관계하에서 이 논문에서 다루고자하는 제어기 설계문제는 주어진 수렴속도를 가지면서 로봇시스템의 궤적이 원하는 궤적 q_d 를 추적하는 적응 최적 학습제어기를 구하는 것이다. 그러나 일반적으로 로봇의 동역학과 입력외란을 정확하게 기술하기란 어렵기 때문에 최적제어와 같은 고전적인 제어기법을 사용하여 지정된 수렴속도를 갖는 최적 제어입력을 구하기란 쉽지 않다. 따라서 이 논문에서는 최적입력과 미지의 파라미터들을 점근적으로 추정하면서 원하는 수렴속도를 반영한 성능지표를 최소화 시키는 안정적인 학습규칙을 가진 학습제어기를 설계하고자 한다. 설계된 적응 최적 학습제어기는 점근적인 최적입력과 파라미터 추정뿐만

아니라 시스템을 안정화 시키는 궤환제어 기능도 수행할 수 있도록 한다.

이러한 목적에 맞는 적응 최적 학습제어기는 다음과 같이 구성할 수 있다.

$$\tau(t) = \tau_{fb}(t) + \tau_{ce}(t) + \tau_o(t) \quad (2)$$

여기서 τ_{fb} , τ_{ce} , τ_o 는 각각 궤환 오차입력, 토크계산 오차입력(computed-torque-error input^[11]), 점근 최적 제어입력을 나타내며, 시스템을 안정화시키는 오차입력 τ_{fb} 와 τ_{ce} 는 다음과 같이 주어진다.

$$\tau_{fb}(t) = \Gamma z + \hat{d}_0 \text{sgn}(z) + \gamma(e, \dot{e}) \quad (3)$$

$$\tau_{ce}(t) = \hat{D}_e(q) \ddot{q}_d + \hat{B}_e(q, \dot{q}) \dot{q}_d + \hat{g}_e(q, \dot{q}) + a(\hat{D}(q) \dot{e} + \hat{B}(q, \dot{q})e) \quad (4)$$

여기서 $e = q_d - q$, $z = \dot{e} + ae$ ($a > 0$), 궤환 이득 Γ 는 양의 대칭행렬이고, 비선형 함수 γ 는 머니플레이터의 동역학 방정식을 재편성하여 기술할 때 나타나는 회귀행렬(regression matrix)과 파라미터 행렬의 상한경계를 위해 도입된 함수이다. \hat{d}_0 는 d_0 의 추정값이고, $z=0$ 일 때 $\text{sgn}(z)=0$, $|z| \neq 0$ 일 때 $\text{sgn}(z) = \frac{z}{|z|}$ 이다. ($\hat{\cdot}$)는 추정된 파라미터로 구성된 추정시스템이며 다음과 같이 정의한다.

$$\begin{aligned} \hat{D}_e(q) &\equiv \hat{D}(q) - \hat{D}(q_d) \\ \hat{B}_e(q, \dot{q}) &\equiv \hat{B}(q, \dot{q}) - \hat{B}(q_d, \dot{q}_d) \\ \hat{g}_e(q, \dot{q}) &\equiv \hat{g}(q, \dot{q}) - \hat{g}(q_d, \dot{q}_d) \end{aligned}$$

식(2)를 식(1)에 대입하고 추정시스템의 정의와 유사한 다음의 정의식들을 이용하여 정리하면 식(5)와 같은 오차시스템을 구할 수 있다.

$$\begin{aligned} D_e(q) &\equiv D(q) - D(q_d) \\ B_e(q, \dot{q}) &\equiv B(q, \dot{q}) - B(q_d, \dot{q}_d) \\ g_e(q, \dot{q}) &\equiv g(q, \dot{q}) - g(q_d, \dot{q}_d) \\ \hat{D}_e(q) &\equiv \hat{D}_e(q) - \hat{D}_e(q_d) \\ \hat{B}_e(q, \dot{q}) &\equiv \hat{B}_e(q, \dot{q}) - \hat{B}_e(q_d, \dot{q}_d) \\ \hat{g}_e(q, \dot{q}) &\equiv \hat{g}_e(q, \dot{q}) - \hat{g}_e(q_d, \dot{q}_d) \end{aligned}$$

$$D(q) \dot{z} + B(q, \dot{q})z + \Gamma z = d_2 - \hat{d}_0 \text{sgn}(z) + Y_e \tilde{\theta} + \tilde{\tau}_o - \gamma(e, \dot{e}) \quad (5)$$

여기서 파라미터 오차벡터와 점근적인 최적입력 오차 벡터는 각각 $\tilde{\theta} = \theta - \hat{\theta} \in R^l$ 과 $\tilde{\tau}_o = \tau_d - \tau_o$ 로 정한다. 그리고 회귀행렬 $Y_e \in R^{n \times l}$ 은 파라미터 재편성

(reparameterization)에 따라 $Y_e \tilde{\theta} = \tau_{ce}^* - \tau_{ce}$ 의 방정식으로 부터 유도되며, 여기의 τ_{ce}^* 는 추정 파라미터 벡터 $\tilde{\theta}$ 를 시스템 파라미터 벡터 θ 로 대체한 τ_{ce} 를 나타낸다. 한편, 원하는 궤적 q_d 에 대한 바람직한 제어입력을 $\tau_d = D(q_d) \dot{q}_d + B(q_d, \dot{q}_d) \dot{q}_d + g(q_d, \dot{q}_d) + d_1$ 로 정의하고, $|\tau_d| < \tau_{\max}$ 로 유계되어 있다고 가정한다. 그리고 비선형 함수 γ 는 다음의 proposition 1을 사용하여 다음 절에서 정의할 것이다.

Proposition 1^[12]

계산토크 오차입력 τ_{ce}^* 는 다음의 관계에 따라 유계된다. $|Y_e \theta| \leq \xi_0 |e| + \xi_1 | \dot{e}| + \xi_2 |e|^2 + \xi_3 |e| | \dot{e}| + \xi_4 (\dot{e})$

여기서 ξ_i ($i=0,1,2,3,4$)는 양수이며, $\xi_4(0)=0$ 이다.

proposition 1에 대한 증명은 [12]에 있으며, 참고로 유사한 proposition과 증명이 [13]에 제시되어 있다. 이러한 설정을 근간으로 적응 최적제어 설계문제는 오차시스템 (5)가 지정된 수렴속도로 수렴하도록 만드는 $\tau_o, \hat{\theta}, \hat{d}_0$ 의 학습규칙을 찾는 문제로 바뀌며, 다음 절에서는 변환된 문제가 원하는 수렴속도를 반영한 성능지수를 사용하여 쉽게 풀 수 있음을 보일 것이다.

2. 적응 최적 regulation문제

전 절의 내용을 토대로 식(5)의 오차시스템을 수렴시킬 수 있는 regulation 문제에 대한 학습규칙을 유도하고자 한다. 우선 성능지표로서 다음과 같이 지수적으로 가중된 함수들의 범함수(functional)인 $V(t)$ 를 정의하자.

$$V(t) = \frac{1}{2\beta} \int_{t-\delta}^t W(\eta) d\eta + \frac{1}{2} w^T(t-\delta) D w(t-\delta) \quad (6)$$

여기서 $w(t) = e^{\lambda t} z(t)$ ($\lambda \geq 0$), $\beta > 0$ 이고 δ 는 적응신호의 수렴구간을 나타내며,

$$W(t) = \tilde{\tau}_{ow}^T L^{-1} \tilde{\tau}_{ow} + \tilde{\theta}_w^T S \tilde{\theta}_w + \hat{d}_{ow}^2,$$

$$\tilde{\tau}_{ow} = \tau_d - e^{\lambda t} \tau_o, \quad \tilde{\theta}_w = \theta - e^{\lambda t} \hat{\theta}, \quad \hat{d}_{ow} = d_0 - e^{\lambda t} \hat{d}_0$$

이고, L, S 는 양의 대각행렬이다. 식 (6)에서 우변의 첫항은 입력오차, 파라미터 추정오차, 외란 추정오차로 구성되어 있고, 둘째항은 속도와 위치에 관련된 궤환오차를 포함한다.

이제, 식 (5)의 양변에 $e^{\lambda t}$ 를 곱하면 지수적으로 가중된 새로운 오차시스템을 얻는다.

$$D(q) \dot{w} + B(q, \dot{q}) w + \Gamma_0 w = e^{\lambda t} \{ d_2 - \hat{d}_0 \text{sgn}(z) + Y_e \tilde{\theta} + \tilde{\tau}_o - \gamma(e, \dot{e}) \}. \quad (7)$$

여기서 $\Gamma_0 = \Gamma - \lambda D$ 이다. 제어식 (2), (3), (4)에 대하여 최적 제어입력 τ_o , 파라미터 추정값 $\hat{\theta}$, 외란 경계 상한의 추정값 \hat{d}_0 에 대한 학습규칙을 구하기 위하여 오차시스템 방정식 (7)의 구속조건으로 주어진 성능지수 $V(t)$ 를 점근적으로 최소화 시키도록 한다. 즉, Lyapunov 안정도 판별법을 이용하여 성능지수인 $V(t)$ 를 최소화 시키므로써 변환된 제어문제가 해를 가짐을 보이고자 한다. 오차시스템 (7)에 대하여 성능지수 $V(t)$ 를 미분하면 다음과 같다.

$$\begin{aligned} \dot{V}(t) &= w^T (D \dot{w} + \frac{1}{2} \dot{D} w) + \Delta W(t) \\ &= -w^T(t-\delta) \Gamma_0 w(t-\delta) + w^T(t-\delta) e^{\lambda(t-\delta)} \\ &\quad \{ Y_e \tilde{\theta} + \tilde{\tau}_o - \gamma(e, \dot{e}) \} + w^T(t-\delta) e^{\lambda(t-\delta)} [d_2 - \hat{d}_0 \text{sgn}(z(t-\delta))] \\ &\quad + \Delta W(t). \end{aligned}$$

여기서 $\Delta W(t) \equiv \frac{1}{2\beta} \{ W(t) - W(t-\delta) \}$ 이다. 정의식

$\|(\cdot)\|_M^2 \equiv (\cdot)^T M (\cdot)$ 을 이용하여 $\Delta W(t)$ 를 정리하면 다음과 같다.

$$\begin{aligned} \Delta W(t) &= \frac{1}{2\beta} \{ W(t) - W(t-\delta) \} \\ &= \frac{1}{2\beta} \{ e^{2\lambda t} |\tau_o(t)|_{L^{-1}}^2 - e^{2\lambda(t-\delta)} |\tau_o(t-\delta)|_{L^{-1}}^2 \} \\ &\quad + \frac{1}{2\beta} \{ e^{2\lambda t} |\hat{\theta}(t)|_S^2 - e^{2\lambda(t-\delta)} |\hat{\theta}(t-\delta)|_S^2 \} \\ &\quad + \frac{1}{2\beta} \{ e^{2\lambda t} |\hat{d}_0(t)|^2 - e^{2\lambda(t-\delta)} |\hat{d}_0(t-\delta)|^2 \} \\ &\quad - \frac{1}{\beta} \tau_o^T L^{-1} \{ e^{\lambda t} \tau_o(t) - e^{\lambda(t-\delta)} \tau_o(t-\delta) \} \\ &\quad - \frac{1}{\beta} \theta^T S \{ e^{\lambda t} \hat{\theta}(t) - e^{\lambda(t-\delta)} \hat{\theta}(t-\delta) \} \\ &\quad - \frac{1}{\beta} d_0^T \{ e^{\lambda t} \hat{d}_0(t) - e^{\lambda(t-\delta)} \hat{d}_0(t-\delta) \}. \end{aligned}$$

이 결과를 원래의 식 $\dot{V}(t)$ 대입하여 정리하면 다음을 얻는다.

$$\begin{aligned} \dot{V}(t) &= -|w(t-\delta)|_{\Gamma_0}^2 + \frac{\beta}{2} |w(t-\delta)|_{L_0}^2 + w^T(t-\delta) e^{\lambda(t-\delta)} \\ &\quad \{ Y_e \tilde{\theta} + \tau_d + d_0 - \gamma(e, \dot{e}) \} - w^T(t-\delta) \{ Y_e \tilde{\theta} + \tau_d + d_0 \} \\ &\quad + E(t) + \Delta d. \end{aligned}$$

여기서

$$\begin{aligned} E(t) &\equiv \frac{1}{2\beta} e^{2\lambda t} \{ |\tau_o(t)|_{L^{-1}}^2 + |\hat{\theta}(t)|_S^2 + |\hat{d}_0(t)|^2 \} \\ &\quad - \frac{1}{2\beta} |e^{\lambda(t-\delta)} \tau_o(t-\delta) + \beta L w(t-\delta)|_{L^{-1}}^2, \\ &\quad - \frac{1}{2\beta} |e^{\lambda(t-\delta)} \hat{\theta}(t-\delta) + \beta S^{-1} Y_e w(t-\delta)|_S^2 \\ &\quad - \frac{1}{2\beta} |e^{\lambda(t-\delta)} \hat{d}_0(t-\delta) + \beta w(t-\delta)|^2 \end{aligned}$$

$$\begin{aligned}
 & -\frac{1}{\beta} \tau_d^T L^{-1} \{ e^{\lambda t} \tau_o(t) - e^{\lambda(t-\delta)} \tau_o(t-\delta) - \beta L w(t-\delta) \} \\
 & -\frac{1}{\beta} \theta^T S \{ e^{\lambda t} \theta(t) - e^{\lambda(t-\delta)} \theta(t-\delta) - \beta S^{-1} Y_e^T w(t-\delta) \} \\
 & -\frac{1}{\beta} d_0^T L^{-1} \{ e^{\lambda t} \hat{d}_0(t) - e^{\lambda(t-\delta)} \hat{d}_0(t-\delta) - \beta w(t-\delta) \},
 \end{aligned}$$

$$\begin{aligned}
 \Delta d & \equiv d_0^T w(t-\delta) - d_0^T |w(t-\delta)| + d_2^T w(t-\delta) e^{\lambda(t-\delta)} \\
 & - d_0^T w(t-\delta) e^{\lambda(t-\delta)}
 \end{aligned}$$

이다. 안정도 판별을 위해 $|d_2| \leq d_0$ 관계를 이용하여 Δd 의 부호를 조사해 보면

$$\begin{aligned}
 \Delta d & \leq d_0^T |w(t-\delta)| - d_0^T |w(t-\delta)| + (d_2^T - d_0^T) w(t-\delta) e^{\lambda(t-\delta)} \\
 & \leq (d_2^T - d_0^T) w(t-\delta) e^{\lambda(t-\delta)} \\
 & \leq \{ |d_2^T| |w(t-\delta)| - d_0^T |w(t-\delta)| \} e^{\lambda(t-\delta)} \\
 & \leq \{ d_0^T |w(t-\delta)| - d_0^T |w(t-\delta)| \} e^{\lambda(t-\delta)} \\
 & \leq 0.
 \end{aligned}$$

따라서 $\dot{V}(t)$ 는 다음과 같이 정리할 수 있다.

$$\dot{V}(t) \leq -|w(t-\delta)|_{L_0}^2 + \frac{\beta}{2} |w(t-\delta)|_{L_0} + w^T(t-\delta) e^{\lambda(t-\delta)} \bar{\gamma} + E(t),$$

여기서

$$\begin{aligned}
 L_0 & \equiv I + L + Y_e S^{-1} Y_e^T, \\
 \bar{\gamma} & \equiv (Y_e \theta + \tau_d + d_0)(1 - e^{-\lambda(t-\delta)}) - \gamma(e, \dot{e}).
 \end{aligned}$$

이제, $\dot{V}(t)$ 식의 E(t)항에서

$$\begin{aligned}
 e^{\lambda t} \tau_o(t) & = e^{\lambda(t-\delta)} \tau_o(t-\delta) + \beta L w(t-\delta), \\
 e^{\lambda t} \theta(t) & = e^{\lambda(t-\delta)} \theta(t-\delta) + \beta S^{-1} Y_e^T w(t-\delta), \quad (8) \\
 e^{\lambda t} \hat{d}_0(t) & = e^{\lambda(t-\delta)} \hat{d}_0(t-\delta) + \beta w(t-\delta)
 \end{aligned}$$

와 같은 학습신호들간의 관계를 설정하면 E(t)는 명백히 영이 된다. 이제 $\Gamma = \beta L_0$ 로 두고, $\lambda \neq 0$ 일 때 식(9)와 같은 포화형(saturation-type) 제어입력 γ 를 다음과 같이 정의하자.

$$\gamma = \begin{cases} 0, & \lambda = 0 \\ \text{sgn}(z) \{ \alpha(e, \dot{e}) + \sigma_0 \}, & \lambda \neq 0. \end{cases} \quad (9)$$

여기서 $\alpha(e, \dot{e})$ 와 σ_0 는 각각 $|Y_e \theta| \leq \alpha(e, \dot{e})$, $\tau_d + d_0 \leq \sigma_0$ 를 만족한다.

이러한 관계로부터 proposition 1은 e 와 \dot{e} 가 유계이면 항상 γ 도 유계임을 보장하므로

$$\alpha(e, \dot{e}) = \psi_0 |e| + \psi_1 |\dot{e}| + \psi_2 |e|^2 + \psi_3 |e| |\dot{e}| + \psi_4 (\dot{e})$$

로 두기로 한다. 여기서 $\psi_i \geq \zeta_i$ ($i=0, 1, 2, 3, 4$)이고 $\psi_4(\dot{e})$ 는 마찰력 보상을 위한 비선형항이다.

선택한 케환이득 Γ 와 비선형 함수 γ 를 $\dot{V}(t)$ 에 대입하여 정리하면 다음과 같다.

$$\begin{aligned}
 \dot{V}(t) & \leq -\frac{\beta}{2} w(t-\delta) \bar{L} w^T(t-\delta) \\
 & \leq 0.
 \end{aligned} \quad (10)$$

여기서 행렬 $\bar{L} = \bar{L}^T > 0$ 은 $\bar{L} = I + L + Y_e S^{-1} Y_e^T - \frac{2\lambda}{\beta} D > 0$ 으로 정의된다.

이러한 관계를 만족하도록 선정된 학습규칙과 케환 입력들은 시간이 경과함에 따라 Lyapunov 안정도 판별법을 위한 성능지표 V(t)를 점근적으로 최소화 시키면서 적응 최적제어 문제를 해결하게 된다. 아울러 학습신호 $\tau_o, \theta, \hat{d}_0$ 에 다음과 같은 투영함수 $(\cdot)^+$ 를 도입하자.

$$x_i^+ = \text{Pr}(x_i) = \begin{cases} \bar{x}_i, & x_i > \bar{x}_i \\ x_i, & x_i \leq x_i \leq \bar{x}_i \\ \underline{x}_i, & \text{나머지구간.} \end{cases}$$

여기서 x_i 는 \hat{d}_0 와 τ_o, θ 의 i번째 요소를 나타내며, \bar{x}_i 는 상한을 \underline{x}_i 는 하한을 뜻한다. 예를 들어 관절 1에 대한 입력토크의 경우 [표1]을 참조하면 250Nm, -250Nm가 각각 상하한을 나타낸다.

그러면 식(6)의 W(t)에 대한 투영함수 적용결과는 다음으로 주어진다.

$$W^+(t) = \tau_{ow}^+ T L^{-1} \tau_{ow}^+ + \theta_w^+ T S \theta_w^+ + \hat{d}_{0w}^{+2}.$$

여기서 $\tau_{ow}^+ = \tau_d - e^{\lambda t} \tau_o^+$, $\theta^+ = \theta - e^{\lambda t} \theta^+$, $\hat{d}_0^+ = d_0 - e^{\lambda t} \hat{d}_0^+$. $W^+(t) \leq W(t)$ 이므로 W(t)대신 $W^+(t)$ 를 대입하여 최종적인 학습규칙을 구하면 다음과 같다.

$$\tau_o(t) = \alpha \tau_o^+(t-\delta) + \alpha \beta L z(t-\delta), \quad (11)$$

$$\hat{\theta}(t) = \alpha \hat{\theta}^+(t-\delta) + \alpha \beta S^{-1} Y_e^T(t-\delta) z(t-\delta), \quad (12)$$

$$\hat{d}_0(t) = \alpha \hat{d}_0^+(t-\delta) + \alpha \beta z(t-\delta). \quad (13)$$

여기서 $\alpha = e^{-\lambda \delta}$ ($0 < \alpha \leq 1$)로서 망각인자로 간주할 수 있다. 그리고 t=0에서 초기조건들은 $\tau_o(0) = \hat{\tau}_d(0)$, $\theta(0) = \theta_0$, $\hat{d}_0(0) = \bar{d}$ 로 설정하며 $\hat{\tau}_d(0), \theta_0, \bar{d}$ 는 각각 τ_d , 공칭 파라미터 벡터, 공칭 외란 경계의 초기 추정값들을 나타낸다.

그러면 투영함수를 포함한 학습규칙 (11), (12), (13)과 제어입력 (2)를 이용한 적응 학습제어기의 수렴특성

은 proposition 2로 요약할 수 있다.

Proposition 2

적응 최적 제어시스템이 방정식 (1), 제어입력 (2), (3), (4)와 학습규칙 (11), (12), (13)처럼 구성되고, β , L , S 가 $\bar{L} \equiv (I+L+Y_e S^{-1} Y_e^T - \frac{2\lambda}{\beta} D) > 0$ 을 만족한다면, 주어진 적응 최적 제어시스템은 다음과 같이 수렴한다.

- i) $V_0(t) \in L_\infty$,
- ii) $\lambda=0$ 이면, global하게 점근적으로 $\lim_{t \rightarrow \infty} z(t) = 0$ 으로 수렴하며,
- iii) $\lambda > 0$ 이면, 수렴속도 $e^{-\lambda t}$ 로써 global하게 지수적으로 $\lim_{t \rightarrow \infty} z(t) = 0$ 에 수렴한다.

여기서 $(n-1)\delta \leq t \leq n\delta$ ($n=1, 2, 3, \dots$) 에 대해 $V_0(t) = \frac{1}{2\beta} \int_0^t W^*(\eta) d\eta + \frac{1}{2} \sum_{k=1}^n w^T(t-k\delta) D w(t-k\delta)$ 로 정의한다. (증명은 부록 A 참조)

proposition 2가 의미하는 바는 λ 를 설정함에 따라 수렴속도를 조절할 수 있다는 것이다. 더욱이 λ 가 영이 아닐 때 나타나는 비선형 케환 함수 γ 의 크기는 λ 의 크기와 무관하고, 케환이득 L , 적응이득 β , S 가 부등식 $\bar{L} > 0$ (여기서 βL 과 βS^{-1} 의 크기는 λD 에 비례함) 을 만족하는 한 지수적 수렴속도 λ 를 임의대로 정할 수 있음을 의미한다. 한편, 연속시간 제어를 가정한 다면 적응구간 δ 또한 학습규칙에서 임의대로 정할 수 있다. 즉, 이것은 연속시간 적응 알고리즘들^[14, 15] 에 비해 학습신호 τ_d , $\hat{\theta}$, \hat{d}_0 에 대한 계산적 부담을 줄일 수 있음을 의미한다.

부연

- i) 비선형 케환 함수 γ 는 가중합수를 사용하지 않으면 ($\lambda=0$) 필요하지 않다.
- ii) $\text{sgn}(\cdot)$ 대신 $\tanh(\cdot)$ 와 같은 유연한 함수를 사용하여 global 수렴성을 가지면서 $z=0$ 의 근방에서 유연한 천이특성을 얻을 수 있다^[13, 16].

3. 적응 최적 추적문제(Tracking)

적응구간 δ 가 충분히 작거나 $\tau_d(t)$ 가 시간적으로 서행하여 원하는 입력 $\tau_d(t)$ 가 $\tau_d(t-\delta)$ 에 충분히 근접하게 되면, 앞절에서 유도한 학습규칙을 적응 추적문제에도 적용할 수 있다. 또한 로봇의 반복작업과 같이

작업에 필요한 케적 $q_d \in C^2$ (적어도 두 번이상 미분이 가능한 함수의 집합)와 τ_d 가 주기 T 를 가질 경우에도 이미 유도한 학습규칙에서 $\delta=T$ 로 두어 최적 추적문제에 이들 학습규칙을 바로 적용할 수 있다. 이러한 관계로부터 나온 결과를 proposition 3으로 요약하였다.

Proposition 3

다음의 조건중 하나를 만족하면 앞의 학습규칙들은 전절의 결과 proposition 2와 같이 적응 최적 추적제어 시스템을 수렴시킨다.

- i) $\tau_d(t)$ 가 시간적으로 서행하거나 $\delta \ll 1$ 을 만족하여, $\tau_d(t)$ 가 $\tau_d(t-\delta)$ 에 충분히 근접한다.
- ii) $\tau_d(t)$ 가 주기 T 를 가져서 $\delta=T$ 일 때 $\tau_d(t) = \tau_d(t-\delta)$ 를 만족한다.

부연

원하는 케적이 주기 T 를 가지는 경우, 원하는 입력 $\tau_d(t)$ 에 포함된 d_1 은 주기적인 외란을 나타낸다.

proposition 3의 증명은 proposition 2와 비슷한 절차로 쉽게 유도되므로 여기서는 생략한다.

4. 전체 제어구조

전절에서 유도한 적응 학습제어기를 II절에서 기술한 DNN unit을 이용하여 구성하고자 한다. 우선 그림 1의 전처리부 입력($\psi(\cdot)$)은 학습신호에 따라 학습규칙 (11), (12), (13)에서 우변의 두 번째 항이 되고, 케환이득은 α 이다. 지연부 $m^{-\delta}$ 의 δ 는 신호의 갱신주기로 됨으로써 DNN unit으로 학습규칙을 구현할 수 있다. 이러한 학습규칙을 가진 DNN을 이용한 로봇 제어 시스템의 전체 구조를 그림 2와 같이 나타낼 수 있다.

제어기의 목표출력은 식 (2)에서 나타낸 로봇 입력토크를 생성하는데 있다. 이러한 목적을 위해 이 논문에서 사용하고 있는 제어 전략은 과도상태의 안정화와 외란의 경계값을 추정해 내기 위한 케환 제어, DNN을 이용한 최적 입력 학습제어, 로봇 파라미터 추정을 이용한 적응제어으로써 로봇에 필요한 입력토크를 만드는 것이다. 위와 같은 제어기 구성에서 한가지 유의해야 될 사항은 보조 제어입력 계산식 (4)의 관계를 직접 이용하지 않고 식 (1)로 표현되는 로봇 방정식의 성질

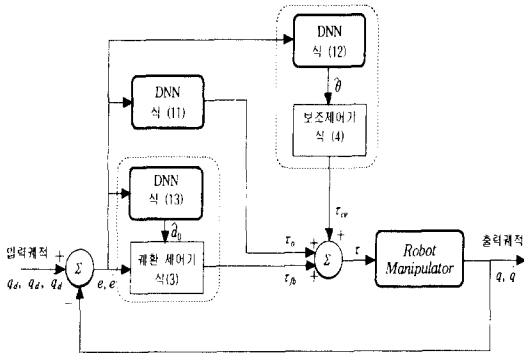


그림 2. DNN을 이용한 적응 최적 학습제어 구조
Fig. 2. Adaptive optimal learning control scheme using the DNN units.

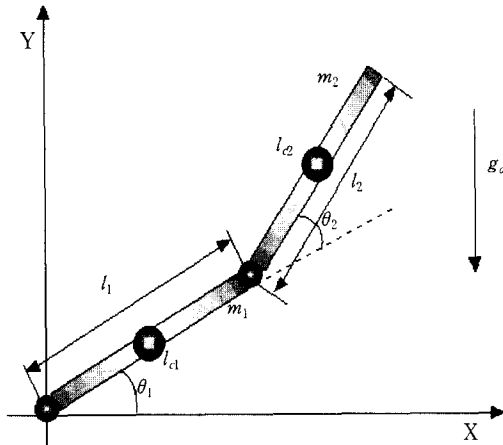


그림 3. 2-DOF 로봇 매니퓰레이터
Fig. 3. 2-DOF robot manipulator used in simulation.

표 1. 로봇 매니퓰레이터의 제원
Table 1. Specification of robot manipulator used in simulation.

	link 1	link 2
m: 질량[kg]	10	5
l: 길이[m]	1	1
lc: 질량중심위치[m]	0.5	0.5
I: link 관성[kgm ²]	0.833	0.417
τ: 최대입력토크[Nm]	250	220
k: 점성 마찰계수	1.0	1.0
p: 쿨롱 마찰계수	4.0	4.0

3)을 이용하여 파라미터 재편성을 한 후 DNN에서 추정된 파라미터 $\hat{\theta}$ 를 입력으로 하여 토크를 계산하는 것이 전체 계산량을 줄이는데 도움이 된다는 것이다^[11].

IV. 모의 실험(Simulation)

제안한 적응 최적 학습제어기의 성능검증을 위해 그림 3과 같은 2-DOF의 로봇 매니퓰레이터에 적용하였다. 이 로봇의 각 제원은 표 1과 같다.

또한, 방정식 (1)로 표현되는 로봇 매니퓰레이터의 동역학 방정식의 요소와 로봇 시스템의 성질 3)에 따른 파라미터 벡터와 오차 방정식의 회귀행렬은 문헌 [11]의 규칙과 파라미터 선정기법을 이용하였다. 아울러 모의실험에 사용된 기준궤적과 외란모델은 다음과 같다.

$$q_{d1}(t) = 2.0 \sin(t) \quad d^1(t) = 10 + 3 \sin(3000t), \text{ (link 1)}$$

$$q_{d2}(t) = 1.5 \cos(t) \quad d^2(t) = 10 + 2 \sin(3000t), \text{ (link 2)}$$

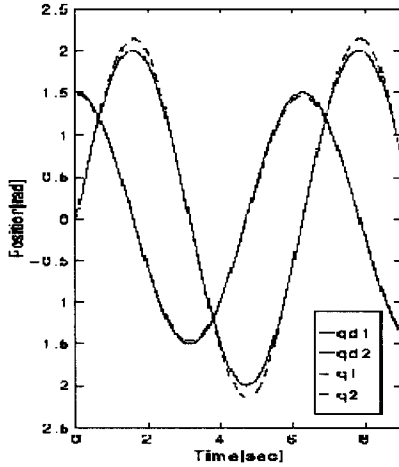
표 1에서 주어진 로봇의 제원에 따라 다음과 같은 3 종류의 실험을 선형 PD 제어기와 제안한 학습제어기에 대해 수행하였다.

- i) 무부하, 무외란시의 성능비교-그림 4
- ii) 무부하, 외란존재시의 성능비교-그림 5
- iii) 7.5kg의 부하와 외란존재시의 성능비교-그림 6

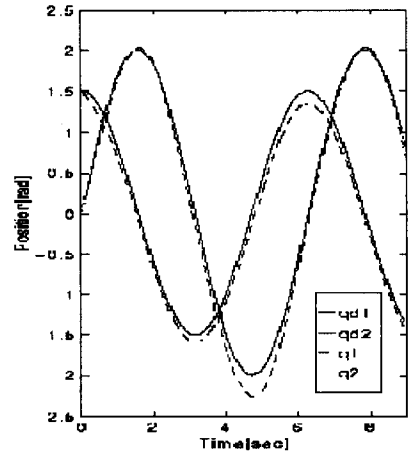
작업 i)과 같은 이상적인 환경에서는 비교적 두 제어기 모두 기준궤적을 잘 추적하였지만 실제 작업환경과 같이 외란, 부하변동등을 고려한 두 실험 ii), iii)에서는 선형 PD 제어기의 경우 큰 오차가 발생한 반면, 제안한 학습제어기는 외란보상 및 부하운반 능력이 우수하게 나타났다. 이것은 제안한 학습제어기가 주어진 궤적에 대해 최적의 입력토크를 잘 학습할 뿐만 아니라 외란에 대한 대처 능력 및 부하 인가에 따른 파라미터 학습능력이 우수하며 강인한 성질을 보여주고 있다 하겠다.

V. 결론

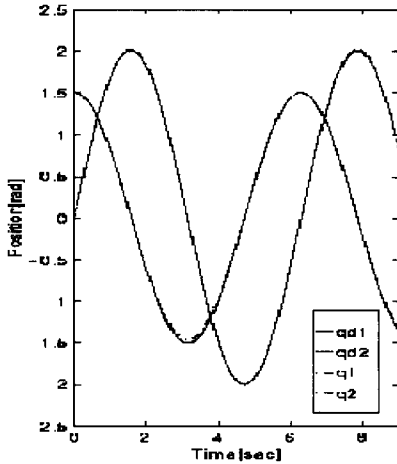
이 논문에서는 새로운 구조의 적응 최적 학습제어기를 제안하였고, 이를 동적 신경망을 이용하여 구현하고, 로봇 제어시스템에 적용하여 그 성능의 우수성을 입증하였다.



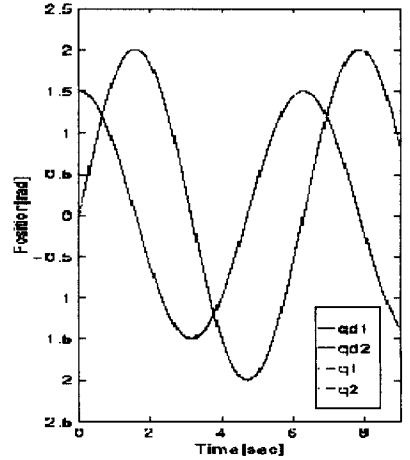
(a)



(a)



(b)



(b)

그림 4. 무부하, 무외란시의 성능비교 (a) PD 제어기 (b) 학습제어기

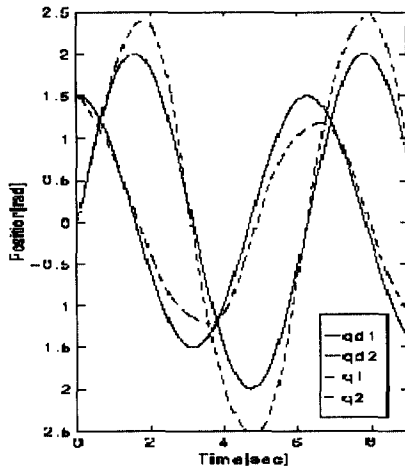
Fig. 4. Trajectories of the system without payload and disturbance (a) PD controller (b) Proposed learning controller.

그림 5. 무부하, 외란존재시의 성능비교 (a) PD 제어기 (b) 학습제어기

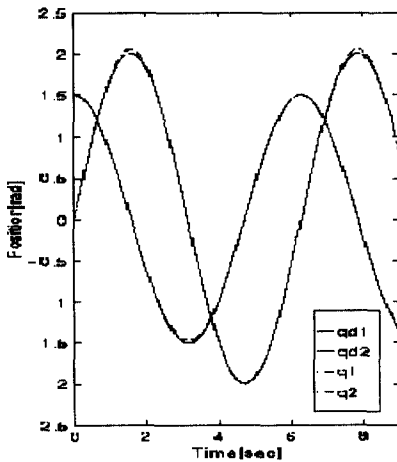
Fig. 5. Trajectories of the system with disturbances without payload (a) PD controller (b) Proposed learning controller.

제안한 제어기는 적응 최적 학습 제어시스템의 모든 오차신호들을 유계시킬 뿐만 아니라 로봇의 움직임을 원하는 궤적으로 수렴시킬 수 있었다. 더욱이 외란과 부하의 유무에 상관없이 강인한 제어특성을 보여 주고 있으며, 아울러 주어진 궤적의 주기성에 관계없이 적용이 가능하다. 또한 로봇제어를 위해 관절의 각기속도 (angular acceleration) \ddot{q} 를 측정하거나 추정할 필요가 없고, 아울러 추정 관성행렬의 역행렬 존재도 가정

할 필요가 없다. 그리고 주어진 궤적에 대해 로봇이 필요한 최적 입력토크를 학습할 수 있고, 원하는 수렴속도로써 로봇의 운동을 수렴시키는 장점이 있다. 특히 종래의 신경망을 이용한 제어기법에서 설계된 제어기의 안정도 또는 수렴성의 검증없이 출력오차 수렴만을 보인것과는 달리 동적 신경망 구조로써 로봇 동력학에 근거하여 구현된 제어기에 대하여 수학적으로 엄격한 수렴성을 제시하고 있다.



(a)



(b)

그림 6. 부하인가 및 외란존재시의 성능비교 (a) PD 제어기 (b) 학습제어기

Fig. 6. Trajectories of the system with payload and disturbances (a) PD controller (b) Proposed learning controller.

부록 A. Proposition 2의 증명

proposition 2에서 정의한 $V_0(t)$ 는 다음의 부등식을 만족한다.

$$V_0(t) \leq V_n(t) = \frac{1}{2\beta} \sum_{k=1}^n \int_{t-(n-1-k)\delta}^{t-(n-k)\delta} W^T(\eta) d\eta + \frac{1}{2} \sum_{k=1}^n w^T(t-k\delta) D w(t-k\delta) \quad (A-1)$$

각각 시구간을 가지는 n 개의 적분항에 대해

$W^+(t) \leq W(t)$ 의 관계와 식 (10)를 유도한 과정을 이용하면

$$\dot{V}_n(t) \leq -\frac{\beta}{2} \sum_{k=1}^n w^T(t-k\delta) \bar{L} w(t-k\delta) \leq 0 \quad (A-2)$$

이 되고 여기서 \bar{L} 은 식 (10)에서 정의한 것과 같다. 상기 부등식은 $V_0(t) \leq V_n(t) < V_1(0)$ 과 $w(t) \in L_2 \cap L_\infty$ 가 됨을 의미한다. 그리고 오차시스템 (5)에서 학습신호들 $\tau_o^+, \theta^+, \hat{d}_0^+$ 가 유계되어 있으므로, $z \in L_\infty$ 가 된다. 이것은 z 가 균일(uniform)하게 유계됨을 입증하는 것이 되며, 따라서 Barbalat의 Lemma^[13]에 의해 global하게 점근적으로 $\lim_{t \rightarrow \infty} z(t) = 0$ 에 수렴하는 것을 의미한다. 더욱이 $\lambda > 0$ 인 경우, $w(t) \in L_\infty$ 이므로 iii)의 성질을 얻을 수 있다. ■

참 고 문 헌

- [1] Craig, J.J., *Adaptive Control of Mechanical Manipulators*, Addison-Wesley, 1988.
- [2] Slotine, J.J., W. Li, "On the Adaptive Control of Robot Manipulators," *Int. J. Robotics Res.*, Vol.6, No.3, pp.49-59, 1987.
- [3] Ortega, R., M. Spong, "Adaptive Motion Control of Rigid Robots: A Tutorial," *Automatica*, pp.877-888, 1989.
- [4] Arimoto, S. et al., "Bettering Operation of Robots by Learning," *J. Robotic Syst.*, pp.123-140, 1984.
- [5] Moore, K.L., *Iterative Learning Control for Deterministic Systems*, Springer-Verlag, 1993.
- [6] Chung, Chae-Wook, Tae-Yong Kuc, Taek-Chong Yi, "A Dynamic Neural Network for Adaptive Optimal Learning of Robot Motion with Guaranteed Convergence Rate," *IEEE Conf. on SMC*, Beijing, Vol.2, pp.1315-1320, Oct. 1996.
- [7] 정재욱, 국태용, 이택중, "로봇 시스템의 동적 신경망과 적응 최적 학습제어기," *대한전자공학회 하계종합학술대회 논문집*, 제 19 권, 제 1 호, pp.179-182, 6월 1996
- [8] Gupta, M.M., D.H. Rao, "Neuro-Control Systems: A Tutorial," *Neuro-Control Systems-Theory and Applications*, IEEE

- Press, pp.1-43, 1994.
- [9] Pham, D.T., Liu Xing, *Neural Networks for Identification, Prediction and Control*, Springer-Verlag, 1995.
- [10] Jagannathan, S., F.L. Lewis, "Multilayer Discrete-Time Neural-Net Controller with Guaranteed Performance," *IEEE Trans. on NN*, Vol.7, No.1, pp.107-130, 1996.
- [11] Kuc, Tae-Yong, J.S. Lee, "An Adaptive Learning Control of Robot Manipulators," *IEEE Conf. Dec. & Control*, pp.1206-1211, 1991.
- [12] Kuc, Tae-Yong, *A Learning Control Theory with Application to Robotic Systems*, Ph.D. Thesis, Dept. of EE, POSTECH, Feb. 1993.
- [13] Lewis, F.L. et al., *Control of Robot Manipulators*, Macmillan Pub., NY, 1993.
- [14] Craig, J.J. et al., "Adaptive Control of Mechanical Manipulators," *IEEE Conf. on R&A, CA*, pp.190-195, 1986.
- [15] Slotine, J.J., W. Li, "Composite Adaptive Control of Robot Manipulators," *Automatica*, Vol.25, No.4, pp.509-519, 1989.
- [16] Kuc, Tae-Yong et al., "Tuning Convergence Rate of a Robust Learning Controller for Robot Manipulators," *IEEE Conf. Dec. & Control*, pp.1714-1719, 1995.

 저 자 소 개

鄭 宰 旭(正會員) 電子工學會誌 第 22卷 第 11號
參照

鞠 兌 龍(正會員) 電子工學會誌 第 22卷 第 11號
參照

李 宅 鍾(正會員) 電子工學會誌 第 22卷 第 11號
參照