

직접 구동형 매니퓰레이터를 위한 학습 제어기의 실시간 구현에 관한 연구

전 종욱*, 안 현식**, 임 미섭***, 김 권호**, 김 광배**, 이 례희*

* 서강대학교 전자공학과, ** KIST 제어시스템 연구실

A Study on Implementation of a Real Time Learning Controller for Direct Drive Manipulator

Jong Wook Jeon*, Hyun Sik An**, Mee Seub Lim***, Kwon Ho Kim**, Kwang Bae Kim**, Kwae Hi Lee*

* Dept. of E.E. Sogang Univ., ** KIST Control Systems Lab.

Abstract

In this thesis, we consider an iterative learning controller to control the continuous trajectory of 2 links direct drive robot manipulator and process computer simulation and real-time experiment. To improve control performance, we adapt an iterative learning control algorithm, drive a sufficient condition for convergence from which is derived extended conventional control algorithm and get better performance by extended learning control algorithm than that by conventional algorithm from simulation results. Also, experimental results show that better performance is taken by extended learning algorithm.

I. 서 론

특정한 일을 주기적으로 수행하는 시스템에 적합한 반복 학습제어는 Arimoto[1,3], Craig[4], 및 Hideg[5] 등에 의하여 연구되었는데, Craig와 Hideg는 시스템을 주파수 영역에서 해석하기 위하여 비선형 시스템을 선형 시불변 시스템으로 근사화하여 학습제어 시스템의 수렴성을 해석하였다. 그러나 직접 구동형 로봇의 경우 각 링크의 상호 관계를 무시할 수 없기 때문에 선형 시불변 시스템으로 근사화 하는데 무리가 있다. 반면에 Arimoto는 로봇을 선형 시변시스템으로 근사화하여 학습제어 시스템을 해석하였다. 이 학습제어 방법은 각 링크를 독립적으로 제어할 수 있다는 가정이 필요 없으나 가속도 오차 성분만을 입력 토크 갱신에 사용하여 잡음의 영향을 크게 받으며, 위치 제어의 경우 초기 상태 오차가 존재할 때 안정도를 보장하지 못하는 단점을 가진다.

본 논문에서는 Arimoto의 학습제어 방법을 확장시킨 제어 방법을 고려하고 주어진 충분 조건하에서 출력 오차의 수렴성을 증명하였다. 이 방법은 학습 제어기의 이득상수의 수를 늘림으로써 제어입력이 효과적으로 생성되도록 하였다. 컴퓨터 시뮬레이션에서는 기존의 Arimoto 학습제어 방법과, 이를 확장시킨 제어방법을 수행하였고, 제안한 학습제어 방법을 이용하여 로봇을 실시간 제어하기 위하여 실시간 운영체제를 탑재한 단일 보드 컴퓨터 (Single board computer), 입출력 보드 (I/O board) 및 VME subrack등으로 이루어진 실시간 제어 시스템을 구성하고, 실시간 실험을 수행하였다.

II. 반복 학습제어 알고리즘

1. 기존의 반복 학습제어

일반적으로 선형 시불변 시스템은 주파수 영역에서 해석할 수 있기 때문에 학습제어에서도 주파수 영역에서 해석한 이론이 많이 나와 있으며, 이들은 선형 시변 시스템이나 비선형 시스템을 대상으로 해석한 이론에 비해 매우 간단하다.

Craig[4]는 계산 토크법을 사용하여 시스템을 선형화시킨 후, 다음과 같은 학습 제어를 제안하였다.

$$\tau_{k+1} = \tau_k + P * e_k \quad (* : convolution)$$

여기서,

τ_k : k번째 학습시 인가 토크

P : $P(s) = s^2 + (K_v - \mu)s + (K_p - \mu)$ 의 형태를 가지는 학습 필터 (learning filter)

e_k : k 번째 학습시 위치 오차.

시간 영역에서 볼 때 위 학습 제어는 PID형태를 갖는다. Craig는 위 제어에 의해 수렴되는 학습 필터의 계수 범위를 얻었다. Hideg[5]는 선형 시불변 시스템에 대해 PD-type의 학습제어를 제안하였다. 그는 주파수 영역에서 외부 외란이나 부하의 변동이 주기적으로 나타날 경우 학습 제어에 의해 충분히 상쇄됨을 보였다. 이를 이용하여 로봇 시스템의 관성질량 행렬의 변하는 부분과 기타 비선형성을 갖는 부분들을 일종의 주기성을 갖는 잡음으로 간주하고 학습제어에 의해 이러한 비선형성을 충분히 극복할 수 있음을 보였다. Alessandrol[6]는 위 학자들과 마찬가지로 선형학습 필터를 사용하여 선형 시불변 시스템에 대해 시스템이 학습되어감을 보였다. 특히 안정성 (stability)에 대해 일종의 보상 필터를 제안하여 비안정성을 갖는 학습제어기를 안정화 시킬 수 있었다.

위와 같이 주파수 영역에서 전개한 학습제어 이론들의 공통된 가정은 단일 입출력 시스템이라는 것이다. 따라서 다중 입출력을 갖는 로봇, 특히 직접 구동형 로봇에 대해서는 이를 직접 적용하기 어렵다.

선형 시변 시스템에 대해서는, Arimoto가 Betterment process에 의해 시스템 (1)의 출력 오차를 0으로 수렴시킬 수 있음을 보였다[1].

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \\ y(t) &= C(t)x(t) \end{aligned} \quad (1)$$

여기서 x, u, y 는 각각 r의 차원을 갖는 벡터이다.

시스템 출력의 원하는 경로 $y_d(t)$ 가 유한시간구간 $t \in [0, T]$ 에서 지정되어 있고, $y_d(t)$, $u(t)$ 가 t 에 대해 연속적으로 미분가능하다고 가정하자.

이때 시스템 (1)에 대한 Betterment process를 다음과 같이 구성한다.

$$\begin{aligned} e_k(t) &= y_d(t) - y_k(t) \\ u_{k+1}(t) &= u_k(t) + \Gamma(t) \dot{e}_k(t) \end{aligned} \quad (2)$$

위에서 볼 수 있는 바와 같이 학습법칙은 출력오차의 변화율을 이용하여 다음번 시행시의 입력 토오르크를 결정한다. 즉 기존의 Arimoto에 의한 제어 법칙은 출력 오차의 미분값만을 사용하여 수렴성을 보였다. 여기서 제안된 제어법칙은 주파수 영역에서 해석한것과는 달리 선형시변 시스템으로 해석하였기 때문에 각 관절을 독립적으로 제어해도 된다는 가정이 필요없으므로 특히 직접구동형 로봇에 적용할 수 있는 장점이 있다.

2. 확장된 학습제어 알고리즘

본 절에서 제안한 학습제어 방법에서는 제어 입력의 결정시 출력 오차의 미분값 뿐만 아니라 출력 오차 자체도 이용함으로써 학습제어기의 형태를 확장시키고 이 확장된 제어 방법을 이용했을 때 시스템 출력오차의 수렴성을 보장할 수 있는 수렴조건을 유도하도록 한다. 또한 컴퓨터 시뮬레이션을 통하여 여기서 제시된 방법에 의해 출력오차가 더욱 빠르게 주어진 구간 전체에서 0으로 수렴함을 볼 수 있다. 즉 식 (2)에서 제어 입력의 생성식을 다음과 같이 수정한다.

$$\begin{aligned} e_{k+1}(t) &= y_d(t) - y_k(t) \\ u_{k+1}(t) &= u_k(t) + \Gamma(t) \dot{e}_k(t) + \phi(t) e_k(t) \end{aligned} \quad (3)$$

로봇트 제어의 경우에 출력은 일반적으로 상태변수들의 선형조합이므로 여기서는 시스템 (1)에서 $C(t) = C$ 인 경우만 다루기로 한다.

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \\ y(t) &= C \cdot x(t) \end{aligned} \quad (4)$$

이제 식 (3)과 같이 제안한 학습제어 방법을 선형시스템 (4)에 적용하였을 때, 시행을 반복함에 따라 출력 오차가 주어진 시구간 전체에서 0으로 수렴됨을 정리에서 보이기로 한다.

여기서 $r \times r$ 행렬 $G = (g_{ij})$ 의 노름 $\|G\|$ 은

$$\|G\| = \max_{1 \leq i \leq r} \left[\sum_{j=1}^r |g_{ij}| \right]$$

로 정의하며, 함수 $e(t)$ ($R: R \in r \times 1$ 벡터)의 λ -노름은

$$\|e\|_\lambda = \sup_{0 \leq t \leq T} \left\{ e^{-\lambda t} \max_{1 \leq i \leq r} |e_i(t)| \right\}$$

으로 정의한다.

정리

선형 시스템 (4)에 대하여 학습제어 방법 (3)을 적용하자. 이때 유한 시간 구간 $t \in [0, T]$ 내에서

$$\|I - \Gamma(t)CB(t)\|_\infty < 1 \quad (5)$$

이고, 초기 입력 $u_0(t)$ 가 연속이면 시행을 반복함에 따라 주어진 시간 구간에서 출력 오차는 0이 된다. 즉,

$$\lim_{k \rightarrow \infty} e_k(t) = 0$$

증명

원하는 출력 $y_d(t)$ 를 발생시킬 때의 상태 벡터를 $x_d(t)$, 제어 입력을 $u_d(t)$ 라고 하면 식 (3)으로부터,

$$\begin{aligned} e_k &= y_d(t) - y_k(t) \\ &= C(x_d(t) - x_k(t)) \end{aligned}$$

$$\begin{aligned} \dot{e}_k(t) &= C(\dot{x}_d(t) - \dot{x}_k(t)) \\ &= CA(t)(x_d(t) - x_k(t)) \\ &\quad + CB(t)(u_d(t) - u_k(t)) \end{aligned} \quad (6)$$

제어 입력 (3)과 식 (6)을 이용하여 입력 토오르크 값은 다음과 같이 나타낼 수 있다.

$$\begin{aligned} u_d(t) - u_{k+1}(t) &= [I - \Gamma(t)CB(t)] \\ &\quad - [\Gamma(t)CA(t) + \phi(t)C](x_d(t) - x_k(t)) \end{aligned}$$

위 양변에 노름을 취하면,

$$\|u_d(t) - u_{k+1}(t)\| \leq \|I - \Gamma(t)CB(t)\| \|u_d(t) - u_k(t)\| + \|\Gamma(t)CA(t) + \phi(t)C\| \|x_d(t) - x_k(t)\| \quad (7)$$

여기서

$$\begin{aligned} \|x_d(t) - x_k(t)\| &= \left\| \int_0^t [(A(\tau)x_d(\tau) + B(\tau)u_d(\tau)) - (A(\tau)x_k(\tau) + B(\tau)u_k(\tau))] d\tau \right\| \\ &\leq \int_0^t (\|A(\tau)\| \|x_d(\tau) - x_k(\tau)\| + \|B(\tau)\| \|u_d(\tau) - u_k(\tau)\|) d\tau \end{aligned} \quad (8)$$

식 (8)을 Bellman - Gronwall의 정리[7]를 이용하여 정리하고 식 (7)에 대입하면,

$$\begin{aligned} \|u_d(t) - u_{k+1}(t)\| &\leq \|I - \Gamma(t)CB(t)\| \|u_d(t) - u_k(t)\| \\ &\quad + \|\Gamma(t)CA(t) + \phi(t)C\| \int_0^t \|B(\tau)\| \|u_d(\tau) - u_k(\tau)\| e^{-a(t-\tau)} d\tau \end{aligned} \quad (9)$$

여기서 $a = \|A(t)\|$ 이다.

윗 식의 양변에 $e^{-\lambda t}$ ($\lambda > 0, \lambda \neq a$)를 곱하여 λ -노름을 취하면 다음과 같은 부등식을 얻을 수 있다.

$$\begin{aligned} \|\delta u_{k+1}(t)\|_\lambda &\leq [\|I - \Gamma(t)CB(t)\| + \|\Gamma(t)CA(t) \\ &\quad + \phi(t)C\| \|B(t)\| \frac{1 - e^{-(a-\lambda)T}}{\lambda - a}] \|\delta u_k(t)\|_\lambda \end{aligned}$$

여기서 λ 를 충분히 크게 하면,

$$\begin{aligned} \rho &= \|I - \Gamma(t)CB(t)\| \\ &\quad + \|\Gamma(t)CA(t) + \phi(t)C\| \|B(t)\| \frac{1 - e^{-(a-\lambda)T}}{\lambda - a} \\ &\approx \|I - \Gamma(t)CB(t)\|, \end{aligned}$$

이 되고, 조건식 (5)에 의해,

$$\|\delta u_{k+1}(t)\|_\lambda \leq \rho \|\delta u_k(t)\|_\lambda, \quad 0 \leq \rho < 1$$

이므로,

$$\|\delta u_k\|_\lambda \rightarrow 0 \quad \text{as } k \rightarrow \infty, \quad t \in [0, T]$$

즉,

$$u_k(t) \rightarrow u_d(t) \quad \text{as } k \rightarrow \infty, \quad t \in [0, T]$$

또 식 (9)에 의해,

$$x_k(t) \rightarrow x_d(t) \quad \text{as } k \rightarrow \infty, \quad t \in [0, T]$$

$$y_k(t) \rightarrow y_d(t) \quad \text{as } k \rightarrow \infty, \quad t \in [0, T]$$

즉, (3)의 제어 법칙에 의해, 시스템이 학습을 반복함에 따라 경로 $t \in [0, T]$ 에서 $y_k(t)$ 는 $y_d(t)$ 에 접근하여, $e_k(t) \rightarrow 0$ 임을 알 수 있다.

3. 매니퓰레이터에의 응용

로봇트 매니퓰레이터 동특성은 다음과 같이 주어진다.

$$R(\theta)\ddot{\theta} + f(\dot{\theta}, \theta) + g(\theta) = \tau \quad (10)$$

여기서

$R(\theta)$: 관성 질량 행렬

$f(\dot{\theta}, \theta)$: 코리올리력, 원심력, 마찰력.

$g(\theta)$: 중력

위와 같이 로봇의 모델링은 비선형 시스템이므로 본 논문에서 제시한 학습제어 방법을 적용시키기 위해서는 선형화가 필요하다. 또, 본 논문의 실험에 사용한 로봇은 수평 다관절형이므로 $g(\theta) = 0$ 이다.

다음과 같은 제어 법칙을 사용하면,

$$\tau = K_p(\theta_d - \theta) + K_v(\dot{\theta}_d - \dot{\theta})$$

아래와 같이 선형 시변 시스템으로 근사화 된다.

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \\ y(t) &= Cx(t) \\ A(t) &= \begin{bmatrix} -\ddot{\theta}_d & 0 \\ 0 & -\ddot{\theta}_d \end{bmatrix} & B(t) &= R^{-1}(t) \\ C &= [0 \quad I] \end{aligned}$$

수렴 조건식에 위 값을 대입하면 다음과 같은 관계를 얻을 수 있다.

$$\|I_n - \Gamma(t)C(t)B(t)\| = \|I_n - \Gamma(t)R^{-1}(t)\| < 1 \quad (11)$$

III. 제어 시스템의 구성

전체 제어시스템은 크게 단일보드 컴퓨터인 제이거 보드 MVME147과 모터 구동기를 주 프로세서와 인터페이스하기 위한 입출력 보드로 구성되어 있다.

실제 프로그래밍하는 환경은 SUN 워크스테이션이며 여기서 모든 프로그래밍과 실행 코드를 만든다. 따라서 이 실행코드를 MVME147로 전송하기 위해, Ethernet을 사용하여 통신 모듈인 MVME712를 통해 전송한다. 또 토오크 출력을 위한 A/O(아날로그 출력) 모듈은 Macro 6781를 사용하였다. 레졸버로부터 받는 펄스(위치 각)는 펄스 입력 모듈인 HIMV-606A를 사용하였다. 이와 같은 입출력 모듈중 HIMV-606A는 주 프로세서인 MVME147에서 사용하는 VME bus를 사용하지만, Macro 6781은 자체 I/O bus를 사용하므로 VME bus와 연결하기 위해 버스 변환기인 MVME316을 사용하였다. 그림 1은 전체 제어 시스템에 대한 구성도이다.

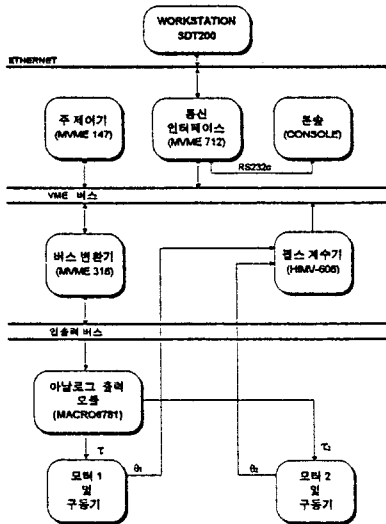


그림 1 전체 제어 시스템 구성

IV. 시뮬레이션 및 실험 결과

시뮬레이션 및 실험에서는 다음과 같은 경로를 사용하였다.

$$\begin{aligned} \theta_{d1} &= t^3(4-3 \cdot t) \\ \dot{\theta}_{d1} &= 12 \cdot t^2(1-t) \\ \ddot{\theta}_{d1} &= 12 \cdot t(2-3 \cdot t) \quad t \in [0, 1] \end{aligned}$$

시뮬레이션은 출력오차의 미분값만을 학습 입력 토오크 갱신에 사용할 경우와 출력 오차 자체를 학습 입력 토오크 갱신에 추가한 2가지로 나누어서 실행하였다.

표 1은 출력오차의 미분값만을 사용하였을 때 설정한 매개 변수이다.

표 1 Arimoto에 의한 학습제어 매개변수

Γ_{11}	Γ_{12}	Γ_{22}
$3.31+0.20 \cdot \cos(t)$	$0.10+0.165 \cdot \cos(t)$	0.165

여기서,

$$\Gamma = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{12} & \Gamma_{22} \end{bmatrix} \quad \Phi = \begin{bmatrix} \phi_1 & 0 \\ 0 & \phi_2 \end{bmatrix}$$

위의 값들은 모델링한 관성 질량 행렬과 동일하게 선택하여 얻은 값들이다.

기존의 출력 오차의 미분값 뿐만 아니라, 출력 오차 자체를 학습제어의 입력 토오크에 반영하였을 때 설정한 매개변수는 표 2와 같다.

표 2 확장된 학습제어 매개 변수

ϕ_1	ϕ_2
12	10

표에 나타난 바와 같이 $\Gamma(t)$ 는 모두 동일하게 사용하여, Φ 에 의해 학습제어의 수렴성을 비교해 보고자 하였다. 위와 같은 매개 변수에 의한 시뮬레이션 결과를 그림 2, 3에 나타내었으며, 이는 각 링크에 대해 2가지 방법에 의한 수렴 속도를 비교한 것이다. 실제 실험의 결과는 그림 4에 나타내었다.

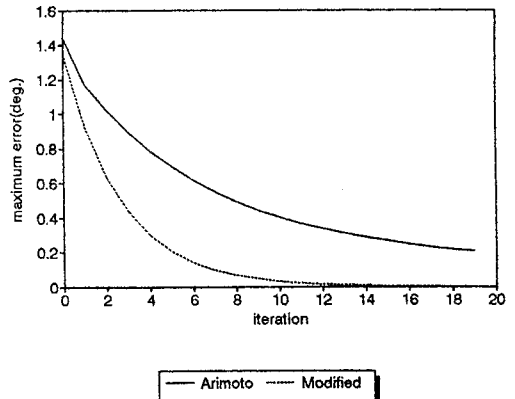


그림 2. 링크 1의 최대 오차 비교 (시뮬레이션).

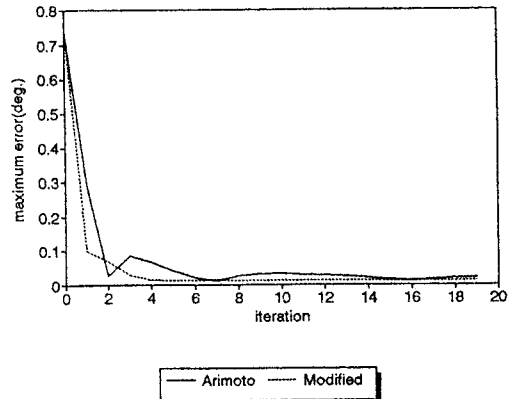


그림 3. 링크 2의 최대 오차 비교 (시뮬레이션).

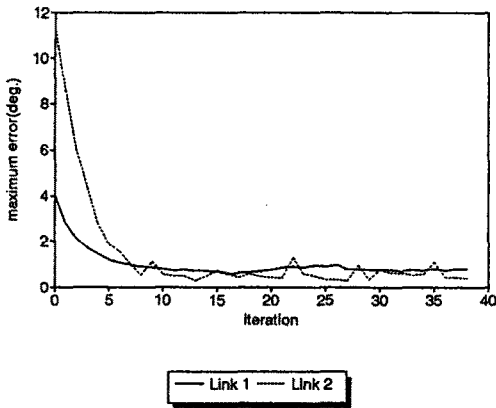


그림 4. 각 링크의 최대 오차 (실험).

VI. 결론

본 논문에서는 비선형이고 각 링크간의 영향이 큰 2축 직접 구동형 로봇을 실시간 제어하기 위하여 실시간 운영체제 Vxworks 및 단일 보드 컴퓨터, 입출력 board 등으로 구성된 제어시스템을 제작하였다.

제어법칙은 제어정밀도를 향상시키기 위해 반복 학습제어 알고리즘을 적용하였으며 기존의 Arimoto에 의한 학습제어의 수렴속도를 보완하기 위해 새로운 학습제어를 제안하였다. 여기서 제안한 제어법칙은 출력오차 자체를 추가한 것으로서, 선형 시변 시스템에서 오차가 0으로 수렴함을 유도하였고 시뮬레이션 결과 기존의 Arimoto의 방법에 비해 수렴속도가 빨라짐을 볼 수 있었다.

실제 실험에서는 제안한 제어법칙을 사용하였으며 실험 결과에서 링크 1의 경우 약 0.7 - 0.8 도의 정상 상태 오차가 생긴 것은 초기 상태를 항상 일정하게 유지시켜 주지 못한 것이라 생각한다. 특히 잔류 모터 인덕턴스가 매우 큰 관계로 초기 토크 및 속도를 정확히 0으로 맞출 수 없었다.

이러한 하드웨어적인 제한으로 시뮬레이션에서 보인 만큼의 오차로 수렴할 수 없었으며 보다 정밀한 제어를 위해 모터의 전기적인 특성을 이해하는 것이 필요하다. 또한 본 논문에서 사용한 반복 학습제어는 출력의 미분 성분을 사용하여 잡음의 영향을 많이 받는다. 따라서 출력의 미분 성분을 필요로 하지 않고 수렴할 수 있는 제어가 필요하며, 매 학습시 초기 상태가 변화여도 이를 보상할 수 있는 알고리즘이 필요하다.

참 고 문 헌

[1] S. Arimoto, S. Kawamura, and F. Miyazaki, "Bettering operation of dynamic systems by learning: a new control theory for servomechanism and mechatronics systems," in *Proc. 23rd Conf. Dec. and Contr.*, pp. 1064-1069, 1984.

[2] S. Arimoto, S. Kawamura, and F. Miyazaki, "Can mechanical robots learn by themselves?," *Robotics Research: The Second International Symposium*, H. Hanafusa and H. Inoue, Eds, Cambridge, MA: MIT Press, pp. 127-134, 1984.

[3] S. Arimoto, "Mathematical Theory of Learning with Applications to Robot Control," *Adaptive and Learning Systems*, K. S. Narendra, Ed., New York: Plenum, pp. 379-388, 1986.

[4] J. J. Craig, "Adaptive Control of Manipulators through Repeated Trials," in *Proc. IEEE American Control Conference*, pp. 1566-1573, June 1984.

[5] L. Hideg, and R. Judd, "Frequency Domain Analysis of Learning Systems," *Conf. on Dec. and Contr.*, pp. 586-591, Dec. 1988.

[6] A. De Luca, and G. Ulivi, "A Frequency-Domain Approach to Learning Control: Implementation for a Robot Manipulator," *IEEE Trans. Industrial Electronics*, vol. 39, no. 1, pp. 1-10, 1992.

[7] M. Vidyasagar, *Nonlinear Systems Analysis*, Prentice-Hall, 1978.