'87 KACC 1987.10.16～17

ANALYSIS OF LEARNING CONTROL SYSTEMS WITH FEEDBACK
(Application to One Link Manipulators)

H.Hashimoto, Seongyun Kang, Jianxin XU, and F.Harashima

Institute of Industrial Science, University of Tokyo
7-22-1 Roppongi,Minato-ku,Tokyo 106 Japan
Tel. 03-402-6231 Ext.2334 Fax. 03-423-1484

Abstract: In this paper, we present a effective method to control robotic systems by an iterative learning algorithm. This method is based on the conceptsof the learning control law which is introduced in this paper, that is, avoidance of using derivative of system state and ignorance of high frequency influence in system performance. By means of the betterment of performance due to the improvement of estimated unknown information, the learning control algorithm compels the system to gradually approach in desired trajectory, and eventually the tracking error asymptotically converges upon zero. In order to verify its utility, one degree of freedom of manipulator has been used in the experiments and the results illustrate this control scheme is very effective.

1. Introduction

We have very many trials and errors from the cradle to the grave,and have a wide experience during our life. The experience to be obtained remains in memory and becomes the data to be referred on the behavior in future. When we try to take a action again, if it is experienced and can be pre-estimated there is no problem, but if it is not experienced and we cannot guess how to turn out, then we must solve it in proper sequence. At first we look back upon the past experience and select the optimal one in the informations of the experience to be accumulated in memory, then we go into action for the first time. In designing a control system, it is the same situation. If all priority information about the control environment is known and can be described deterministically, the optimal controller can be obtained by optimization techniques. In general cases, however, a prior information required is unknown or incompletely known. Hence in stead of using classical method of control design which is impossible to be applied, two diffe ent approaches have been proposed to solve this class of problems. One approach is to design a controller based only upon the amount of information available[1][2]. In this case, the unknown information is assumed to stay inside some known ranges and conservative design criterion is often used. Therefore the system is in general sub-optimal or even designed without considering optimization. The other approach is to design a controller which is capable of estimating the unknown information during its operation. If the estimated information gradually approaches the true information as time proceeds, then the controller thus designed will approach its goal in the sense of optimization. In other words, the performance of the designed controller eventually be as good as in the case where all priority information required is known. Because of the gradual improvement of performance due to the improvement of the estimated unknown information, this class of control systems may be called learning control systems. The controller learns the unknown information during operation and the learned information is in turn used as an experience for future decisions or controls. For example, taking into account the tracking control problem of a servo system whose dynamics is only described by a "fuzzy" model, that is, there are several unknown parameters and disturbances existed that make it impossible to calculate the inverse system to achieve the required control inputs according to the given trajectories. A learning control system is designed to learn to know the relationship between the required control inputs and desired trajectories, i.e. the control input is regarded as unknown information and recorded in memory as learned information. This kind of learned information is considered as an experience of the controller, and the experience will be used to improve the quality of control whenever similar control situations recur. The new information extracted from a recurred control situation (in this case, k+1th trial of tracking motion) is used to update the memory-stored information associated with former information (kth trial). Therefore, as the controller accumulate more information about the unknown relationship, the control law will be altered according to the updated information to get better system performance

From the point of view of Artificial Intelligence, "Learning" is one of the most important abilities that any intelligent machine should be incorporated into its basic functions. Learning Control could be divided into different levels: the first and second levels. The first level utilizes learned experience stored in memory as part of a priority to facilitate decision-making, pattern-recognition, environment perception and it is usually based on some conceptual and qualitative model, like fuzzy one[3], to describe characteristics of the whole control system and relationship between these characteristics[4]. On the other hand, the learning operation in the second level is essentially the numerical ones such as arithmetic, differential or integral ones, and it is typically based on some analytic, quantitative model such as differential or difference equations.[5][6].

The iterative learning method with feedback as being presented in this paper is involved in the learning control class of the second level, which is mainly towards to motion controls such as tracking control in servo systems. It is theoretically shown that the error decreases through repeating the operation of the system if some frequency condition is satisfied. Experiment results of a servo system confirm that this learning method is useful and effective.

2. Analytical Techniques of Learning Control

Several papers[6][7] proposed the iterative learning control scheme based on a modification of the present input by the derivative of the error vector of the previous trial. In case of controlling mechanical systems such as manipulators, this derivative corresponds to the use of angular acceleration, which may cause some problem with required differentiation of noisy measurement data obtained through tacho-generators. On the other hand, it is well known that in most situations the behavior of a mechanical system, especially with servoing mechanism, is similar to that of low pass filter, in other words, it is sufficient to consider the dynamics of such a system in a finite frequency domain without losing high accuracy. In brief, following two points are the basic considerations for a designer to choose a learning control law:

1) Avoidance of using derivative of system state.
2) Ignorance of high frequency influence in system performance.

A satisfactory learning method is to use a modification of present control input by combining the former control input, which is saved in memory, and real feedback which is based on track error between the desired output and the current system output, instead of previous system output.

Problem Formulation

To explain the essence of our proposed learning control, consider a nonlinear dynamic system whose mathematical expressions is described by

$$\dot{x} = f(x,t) + b(t)u \tag{1}$$

$$y(t) = c^t x(t) \tag{2}$$

$$x, f, b, c \in R^n \ , \ y, u \in R^1$$

where x is the system state vector and y and u are system output and input respectively.

Suppose that a desired output of the system is given over a fixed finite interval T:

$$y_d(t) = c^t x_d(t) \tag{3}$$

$$t \in [0,T] \tag{4}$$

The tracking error is defined between the desired and real system output

$$e_k(t) = y_d(t) - y_k(t)$$

$$= c^t(x_d - x_k) \tag{5}$$

where k means the number of operation trial.

Learning Control Input.

If there is an tracking error at the k th trial, the learning control at k+1th trial is constructed as

$$u_{k+1}(t) = u_k(t) + \lambda \rho(t) e_{k+1}(t) \tag{6}$$

where λ is a constant control parameter and ρ is a scalar function. The structure of this learning control scheme is illustrated in Fig.1.

It is assumed that all the system states are measurable and following conditions are known by priority knowledge.

(i) The sign of $c^t b(t)$ is known, and

$$\| c^t b(t) \| \geqq m_b > 0 \tag{7}$$

where $\| \cdot \|$ is some kinds of norm(in this paper,it is an operation of taking absolute value)

(ii) $c^t f(x,t)$ is lipshitz continuous, that is

$$\forall t \in [0,T] \ , \ x, x_k \in R^n$$

$$\exists M_f > 0$$

$$\| c^t f(x_k,t) - c^t f(x,t) \|$$

$$\leqq M_f \| c^t x_k - c^t x \| \tag{8}$$

According to condition (ii), the third condition can be stated as

(iii) System (1),(2) is a continuous mapping from u to y.

Error Model.

An error model is derived in order that control scheme can be analyzed rather easily. According to error definition (5), substitution of control input (6) into system equation (1) and (2) yields

$$\dot{e}_{k+1}(t) = c^t(\dot{x}_d - \dot{x}_{k+1})$$

$$= \dot{e}_k(t) + c^t(f_k - f_{k+1}) - \lambda c^t b(t) \rho(t) e_{k+1}(t) \tag{9}$$

where for convenience $f(x_k,t)$ is represented by f_k. Considering lipshitz condition (ii) following relation is obtained.

$$\| c^t f_k - c^t f_{k+1} \|$$

$$\leqq M_f \| c^t x_k - c^t x_{k+1} \|$$

$$\leqq M_f \| c^t(x_d - x_k) \| + M_f \| c^t(x_d - x_{k+1}) \|$$

$$\leqq M_f(\| e_K \| + \| e_{k+1} \|) \tag{10}$$

Hence one can find

$$\eta = \eta(x_k, x_{k+1}, x_d, t) \tag{11}$$

$$\xi = \xi(x_k, x_{k+1}, x_d, t) \tag{12}$$

that satisfy

$$c^t(f_k - f_{k+1}) = \eta e_{k+1} + \xi e_k \tag{13}$$

and error model (9) can be rewritten as follows

$$\dot{e}_{k+1} = \dot{e}_k + \xi e_k + (\eta - \lambda c^t b\rho)e_{k+1} \quad (14)$$

$$\| \xi \| \leqq M_f$$

$$\leqq M_f \| c^t x_k - c^t x \| \quad (8)$$

According to condition (ii), the third condition can be stated as

(iii) System (1),(2) is a continuous mapping from u to y.

<u>Error Model.</u>

An error model is derived in order that control scheme can be analyzed rather easily. According to error definition (5), substitution of control input (6) into system equation (1) and (2) yields

$$\dot{e}_{k+1}(t) = c^t(\dot{x}_d - \dot{x}_{k+1})$$

$$= \dot{e}_k(t) + c^t(f_k - f_{k+1}) - \lambda c^t b(t)\rho(t)e_{k+1}(t) \quad (9)$$

where for convenience $f(x_k, t)$ is represented by f_k. Considering lipshitz condition (ii) following relation is obtained.

$$\| c^t f_k - c^t f_{k+1} \|$$

$$\leqq M_f \| c^t x_k - c^t x_{k+1} \|$$

$$\leqq M_f \| c^t (x_d - x_k) \| + M_f \| c^t (x_d - x_{k+1}) \|$$

$$\leqq M_f (\| e_K \| + \| e_{k+1} \|) \quad (10)$$

Hence one can find

$$\eta = \eta(x_k, x_{k+1}, x_d, t) \quad (11)$$

$$\xi = \xi(x_k, x_{k+1}, x_d, t) \quad (12)$$

that satisfy

$$c^t(f_k - f_{k+1}) = \eta e_{k+1} + \xi e_k \quad (13)$$

and error model (9) can be rewritten as follows

$$\dot{e}_{k+1} = \dot{e}_k + \xi e_k + (\eta - \lambda c^t b\rho)e_{k+1} \quad (14)$$

$$\| \xi \| \leqq M_f$$

$$\| \eta \| \leqq M_f$$

Before going on next analysis, it is important to point out that, as long as desired trajectory (3) is continuous, tracking error $e_i(t)$ is then a continuous function over each operation interval owing to the continuity of the whole control system, as related in condition (iii).

ε <u>- Domain Analysis</u>

In order to investigate the convergence of the control sheme, suppose that in k th tracking trial

$$e_k(t) \not\equiv 0 \qquad t \in [0,T] \quad (15)$$

and at $t = \sigma_k$

$$e_k(\sigma_k) = \max_{t \in [0,T]} e_k(t) \quad (16)$$

Similarly, assume

$$e_{k+1}(\sigma_{k+1}) = \max_{t \in [0,T]} e_{k+1}(t) \quad (17)$$

and π_{k+1} is the most closest point left to σ_{k+1} where

$$e_{k+1}(\pi_{k+1}) = 0 \quad (18)$$

(Fig.2.). The asymptotic convergence can then be expressed as

$$e_{k+1}(\sigma_{k+1}) < e_k(\sigma_k) \quad (19)$$

Now considering the ε-domain of $e_{k+1}(t)$, by virtue of continuity of $e_{k+1}(t)$

$$\exists \varepsilon > 0, \quad t = \pi_{k+1} + \varepsilon, \quad \forall \tau \in [\pi_{k+1}, t]$$

$$sgn(e_{k+1}(\tau)) = sgn(e_{k+1}(t)) \quad (20)$$

Applying integration operation to both sides of equation(14) yields following relation due to the mean value theorem

$$e_{k+1}(t) + \int_{\pi_{k+1}}^{t} (\lambda c^t b\rho - \eta)(\tau)e_{k+1}(\tau)d\tau$$

$$= e_k(t) - e_k(\pi_{k+1}) + \xi(t')e_k(t')\varepsilon \quad (21)$$

where $t' \in [\pi_{k+1}, t]$ .From equation(16),we have

$$\forall \tau \in [\pi_{k+1}, t]$$

$$\| e_k(\tau) \| \leqq \| e_k(\sigma_k) \| \quad (22)$$

further from condition (iii) and (14), control parameter λ and $\rho(t)$ can be chosen in such way

$$\lambda c^t b(\tau)\rho(\tau) - \eta(\tau) > 0 \quad (23)$$

that

$$\| e_{k+1}(t) \| + \int_{\pi_{k+1}}^{t} (\lambda c^t b\rho - \eta)(\tau) \| e_{k+1}(\tau) \| d\tau$$

$$\leqq \| e_k(t) \| + \| e_k(\pi_{k+1}) \| + \| \xi(t')e_k(t')\varepsilon \|$$

$$\leqq (2 + M_f\varepsilon) \| e_k(\sigma_k) \| \quad (24)$$

or

$$\| e_{k+1}(t) \| \leqq (2 + M_f\varepsilon) \| e_k(\sigma_k) \|$$

$$- (\lambda m_b - M_f)\int_{\pi_{k+1}}^{t} \| e_{k+1}(\tau) \| d\tau \qquad (25)$$

This relation holds at moment $t = \sigma_{k+1}$, and it is easy to deduce

$$\| e_{k+1}(\sigma_{k+1}) \|$$

$$\leqq (2 + M_f\varepsilon_{k+1})\exp((M_f - \lambda_{mb})\varepsilon_{k+1}) \| e_k(\sigma_k) \| \qquad (26)$$

where $\varepsilon_{k+1} \triangleq \sigma_{k+1} - \pi_{k+1}$

It is obvious that to guarantee the inequality(19) is equal to find a control parameter λ such that

$$(2 + M_f\varepsilon_{k+1})\exp((M_f - \lambda m_b)\varepsilon_{k+1}) < 1 \qquad (27)$$

Let

$$\lambda = (M_f + \lambda_0) / m_b \qquad (28)$$

and assume ε_{k+1} is sufficient small in contrast to $1/M_f$, then relation(27) can be expressed approximately as

$$\lambda_0 > 1 / \varepsilon_{k+1} ln2 \qquad (29)$$

This denotes that learning control scheme(6) is valid in the case all the frequencies tracking error contains are lower than $\sqrt{2}\lambda_0$

<u>Low Pass Filter</u>

Although the effectiveness of this learning scheme is verified through amount of simulations and experiments in the case of applying it to the tracking control of manipulators, one can still observe the accumulation of tracking error resulted from uncontrollable high frequencies due to the nonlinearity of dynamics and the shape of given trajectories. To overcome this drawback, a low pass filter is introduced to suppress the undesirable influence, that is , to make the learning mechanism insensive to those high frequencies whose influence to tracking accuracy in feedback loop is essentially ignorable, meantimes to make the learning mechanism sensitive to those low frequencies which dominate the system performance. To achive this, a low pass filter is applied to the memory-stored information, hence control scheme should be rewritten as

$$u_{k+1}(t) = L(u_k(t)) + \lambda\rho(t)e_{k+1}(t) \qquad (30)$$

where "L" denotes such operators as low pass filters.

Experiment and Results

<u>Description of Experimental System.</u>

One-degree of freedom of robotic manipulator is used to examine the validity of our learning control method. the manipulator is shown in Fig.3. A DC Servo motor (17W) is coupled to the arm through a harmonic drive. The DC Servo motor is fed by a DC-DC amplifier operating at 17 KHz. A 1000 pulses/rev shift-encoder is used to sense the output position and the tacho-generator coupled directly to the servo motor provides an analog signal for the output speed. In addition, a current loop is used to decrease the electrical time delay. The microcomputer system used in this experiment is a INTEL-8086 based system with 10 MHz clocks. these are illustrated in Fig.4.

<u>Desired Path.</u>

During the experiments, the whole system is forced to track the desired trajectory shown in Fig.5, which is a polynominal of time t represented by equation.

$$x_{1d} = at^3(1-t/T)^3 \qquad [rad] \qquad (31)$$

$$x_{2d} = \dot{x}_{1d} \qquad [rad/sec] \qquad (32)$$

$$y_d = x_{1d} + 4x_{2d} \qquad (33)$$

Here,the operation period T is 2 seconds and the maximum joint displacement is one radian and maximum joint velocity is about 1.7 radian per second.

<u>Decision of Learning Control.</u>

In the design procedure, only partially known information is available, which gives a rough mathematical description of the robotic manipulator dynamics

$$J\ddot{\theta}(t) + \alpha\dot{\theta}(t) + \beta sin(\theta) + d(t) = u(t) \qquad (34)$$

where θ and u are joint angle and torque exerted on the joint respectively. α is the unmeasurable coefficient of viscous friction, J represents the total inertia of manipulator joint. It should be noted that both J and β change corresponding to the payload the manipulator picks up. d(t) denotes unknown disturbance and noise. In such a case the proposed control method is especially effective because the information it needs to know is no more than the sign of J so as to regulate control parameters. From the theoretical analysis, control parameter λ should be as possible large as real situation permits, in our experiments λ=64, and $\rho(t)$ is selected unit for simplicity. During the repeating learning operations, a third order filter is used as low pass filter to suppress effects of high frequencies and its cut-off frequency ω_c is set to be 20 Hz. The flowchart of the whole learning process is shown in Fig.6, where δ is a small constant used to judge whether the required tracking precision is achieved.

<u>Experimental Results.</u>

Experimental results are obtained in two cases. One is the case without payload as shown in Fig.7, the other is the case with payload of 850 g located at the end of the arm as shown in Fig.8. In both cases we can observe the learning effects from these experimental results, that is, although the tracking error is quite large in the initial trial it nevertheless decreases significantly once the learning operations begin, hence accurate tracking performances are obtained, which are illustrated in the 1st trial, 2nd trial, and 10th trial respectively.

Conclusion

The learning algorithm presented in this paper is suited for motion control of mechanical or servo systems such as robotic manipulators. According to this control scheme, the control input which achieves desired system performance is constructed through iterative learning operations. To implement this scheme, it is necessary to memorize only previous data information of measurable system states to update present control input with current feedback. This kind of controllers is not only easy to be realized from the point of view of hardware, but also effective in real applications, which is supported by several experimental results.

Reference

[1]. V.I.Utkin "variable Structure System With Sliding Mode", IEEE on AC, Vol.AC-22, pp.212-222, April 1977.

[2]. F.Harashima, J.Xu, H.Hashimoto et al, "Tracking Control of Robot Manipulator Using Sliding Mode", Proceedings of 15th International Symposium on Industrial Robot, Vol.2. pp.657-664, September,1985.

[3]. Braae.M, and Rutherford.D.A., "Theoretical and Linguistic aspects of the Fuzzy Logic Controller", Automatica, Vol.15, pp.553-577, 1979.

[4]. P.H.Winston "Artifitial Intelligence", Addison-Weslay.1984.

[5]. Y.Z.Tsypkin "Adaptive and Learning in Automatic Systems", Academic Press,1971.

[6]. S.Arimoto, S.kawamura and I.Miyazaki "Bettering Operation of Robots by Learning", J. of Robotic Systems, 1-2, pp.-- 1984.

[7]. S.Arimoto, S.kawamura and I.Miyazaki "Betterment Process: A New Control Theory for Servomechanism and Mechatronics Systems", 13th SICE Symposium on Control Theory, pp.15-20, 1984.

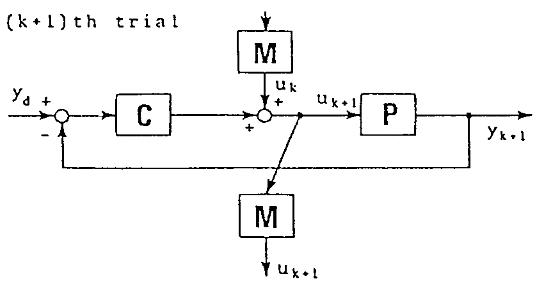


Fig. 1. Learning control system with feedback loop.

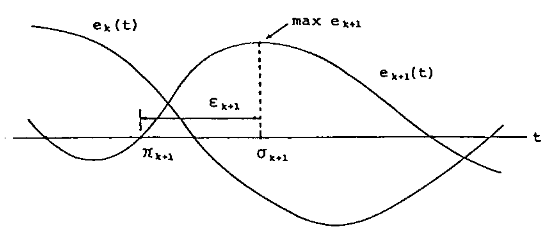


Fig. 2. ε-domain of k+1th trial.

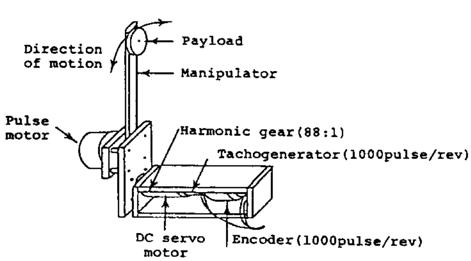


Fig. 3. Manipulator used in this experiment.

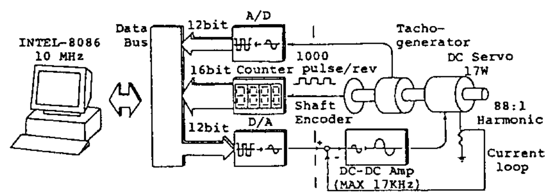


Fig. 4. Block diagram of control system.

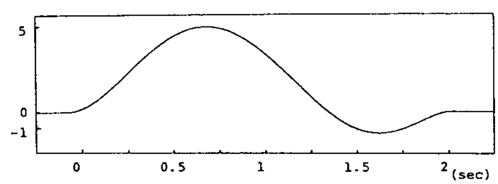


Fig. 5. Desired trajectory.

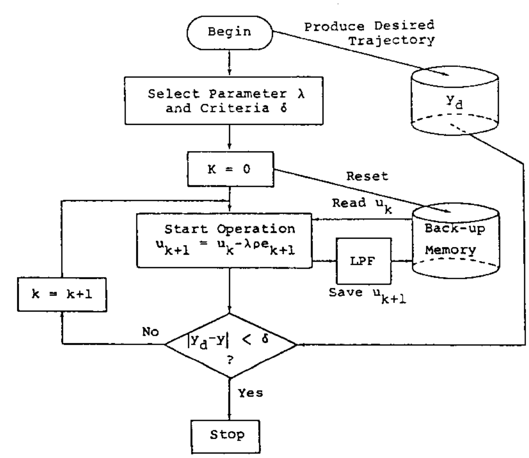
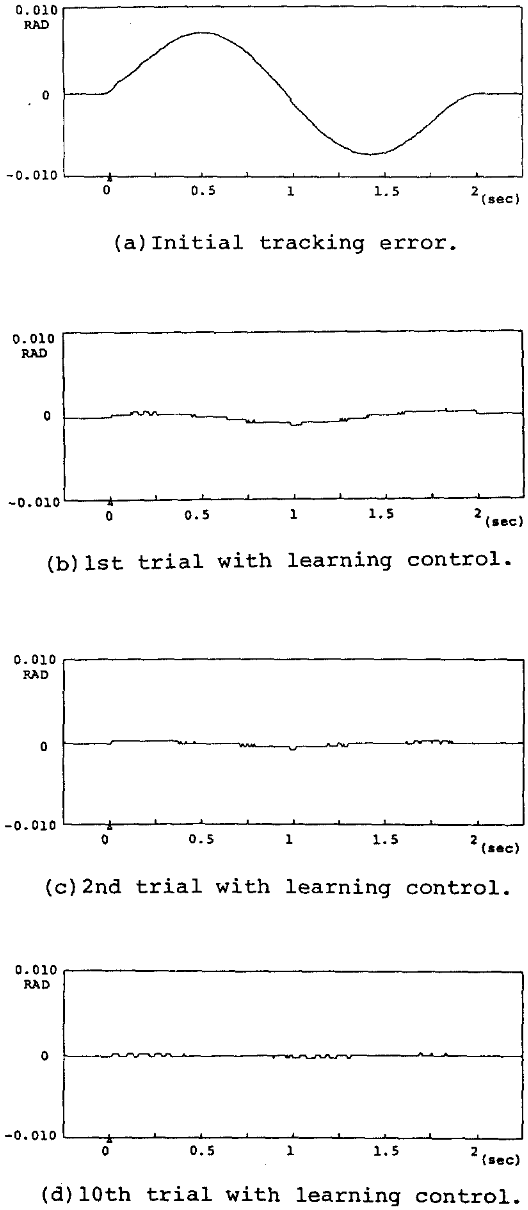


Fig. 6. Flowchart of proposed learning control system.(LPF: Low-Pass Filter)



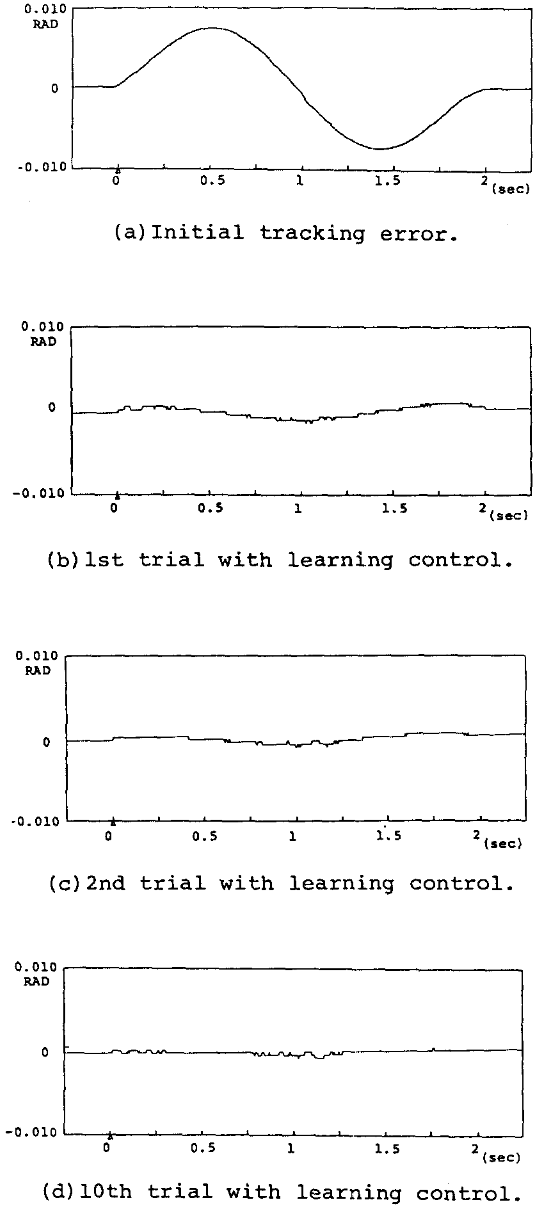
(a) Initial tracking error.

(b) 1st trial with learning control.

(c) 2nd trial with learning control.

(d) 10th trial with learning control.

Fig. 7. Experimental results of tracking error without payload in the system response (3rd order filter).



(a) Initial tracking error.

(b) 1st trial with learning control.

(c) 2nd trial with learning control.

(d) 10th trial with learning control.

Fig. 8. Experimental results of tracking error with payload (850g) in the system response (3rd order filter).