学習方式によるマニピュレータの接触動作の改善[†]

山田 康二*・鈴木 達也*・大熊 繁*

Betterment Process of Contact Motion by Learning Method

Koji YAMADA*, Tatsuya SUZUKI* and Shigeru OKUMA*

A contact motion of a robotic manipulator to a workpiece is a fundamental operation which occurs frequently. The contact motion is a very short time phenomenon and a so-called impact force arises. If we use only a feedback control scheme, it is difficult to realize a desirable contact motion. In this paper, a new control method of the contact motion by a learning control is proposed. The learning control is based on an iterative operation, and is one of the feed-forward controls which do not need precise values of system parameters or a model of the impact force. Therefore it is useful for the contact motion control. Proposed learning controller has two features as follows. Firstly, it observes an input signals and output signals at the different period so as to improve a transient response of a system. Secondly, it can learn an initial state of the system, which includes an approaching velocity of the manipulator. By learning the approaching velocity, a fast contact motion can be achieved, because the magnitude of impact force depends on the approaching velocity. Simulation results are shown to verify an effectiveness of the proposed method.

Key Words: learning control, robotic manipulator, contact motion, approaching velocity

1. はじめに

近年、ロボットマニピュレータに対して、工具と作業対象とが、離れた状態から接触して一定の力を加える状態へ移行する動作（本論文ではこれを接触動作と

呼ぶ）の重要性が高まってきた。文献6)によれば、接触動作は、接触直後、対象物体とエンドエフェクタが独立に動く状態と、その後に現れる両者が一体化して動く状態とに分けて考えることができる。しかも前半の状態においては、対象物体とエンドエフェクタの間に衝突現象が頻繁に見られ、これに伴って発生する衝撃力の影響が接触動作の制御を困難なものとしている（本論文では、衝突現象によって発生する力を衝撃力、力センサが観測する力を接触力と呼ぶ）。接触動作において望まれることは、エンドエフェクタが対象に速やかに接近し、接触後はできるだけ短い時間で、かつ、オーバーシュートなしに接触力が目標値に到達することである。このような接触動作の制御に関する研究は最近になって行われるようになり^{1)~9)}、文献3)ではVSS理論を用いて衝突の際の衝撃力を抑制する手法が検討され、また文献8)では、環境の動特性を考慮にいてコンプライアントな動作を実現する制御系の設計手法がそれぞれ提案されている。しかし、衝突現象はきわめて短い時間の現象であり、しかも衝突時に急激な運動量変化に起因する衝撃力が発生するため、文献3)や、8)で示されているような衝突後のフィードバック制御のみで所望の接触動作を達成することは困難であるとする。この問題に対して、あらかじめ入手できるシステムのパラメータ、すなわち制御対象の物理パラメータやサーボ系の制御ゲインなどを用いて、接触後の力偏差の2乗積分値を最小とする接近速度を計算する手法が提案されており⁴⁾、この手法は、永田らによって多自由度系に拡張されている⁹⁾。これらの手法は、接触前の状態を操作する手法として興味深いが、求めた接近速度の最適性がシステムのパラメータ同定精度に大きく依存することや、文献6)で示されているような接触直後の力学モデルが考慮に入れられていないことから、その有効性に限界があると思われる。

そこで本論文では、接近速度の学習を含めた新しい学習コントローラを提案し、接触力の過渡応答が試行

† ロボティクス・メカトロニクス講演会 '90 で発表 (1990・6)

* 名古屋大学工学部 名古屋市千種区不老町

* Faculty of Engineering, Nagoya University, Nagoya

(Received August 17, 1990)

(Revised May 1, 1991)

ごとに改善されていくことを示す。本学習コントローラは過渡応答の改善を目的としており、従来軌道制御などで用いられてきた学習コントローラ¹³⁾と異なっており、システムの入力と出力を時間的にずらして考えている。また、本学習コントローラは衝撃力の影響をなくすことを目標にするのではなく、むしろある程度衝撃力の影響を利用して速い立ち上がりを得ようとするものである。したがって、最終的に獲得される接近速度は零とはならず、このことは接触動作に要する時間の短縮化という点で重要な意味をもつ。以下まず 2 章で、文献 6) に基づいた接触動作の力学モデルを示す。つぎに 3 章で、従来力制御によく用いられてきた比例、積分および速度フィードバックを行う力サーボ系に提案する学習コントローラを組み合せ、学習の収束条件について検討する。また、4 章で学習コントローラ的设计法、および衝撃力が学習の収束性に及ぼす影響について検討し、5 章でシミュレーション結果を示す。6 章はまとめである。

2. 接触動作のモデルおよび力サーボ系

本論文では、1 自由度マニピュレータの接触動作を考える。この章では、接触動作を以下に示す四つのフェイズに分類して、そのモデル化を行う。

第 1 フェイズ：エンドエフェクタと対象物体が離れている状態。この状態では、マニピュレータは位置制御系により制御される。

第 2 フェイズ：エンドエフェクタが対象物体に衝突した直後。両者間には衝突、跳ね返り、再衝突というチャタリング現象が発生し、急速な運動量変化に伴う衝撃力を生じながら、接触状態・非接触状態間の遷移を繰り返す。なお、このフェイズにおいて最初に衝撃力を検出した時点を時刻 $t=0$ とし、位置制御から力制御への切り換えは、時刻 $t=0$ に行われるものとする。

第 3 フェイズ：チャタリングが減衰した後、エンドエフェクタと対象物体が完全に一体となった状態。対象物体の粘性抵抗や、速度フィードバックの働きにより、やがて系は定常状態へと移行する。

第 4 フェイズ：接触力の応答が定常値に達した状態。このフェイズでは、力制御系の働きにより系は安定に力制御される。

従来、第 1 フェイズは軌道制御問題、第 4 フェイズは力制御問題として研究されてきた。本論文で取り扱う接触問題は、第 2 および第 3 フェイズにおける系の挙動を改善することを目的としている。そこで、以下、第 2、第 3 フェイズのモデルについて検討する。

2.1 第 2 フェイズの力学的モデル

第 2 フェイズにある系は Fig. 1 のようにモデル化され、マニピュレータ位置を X_a 、エンドエフェクタの位置を X_t 、対象物体表面位置を X_w とすると、その挙動は次式で表わされる。

$$\begin{aligned} m_a \ddot{X}_a &= U - f_s \\ m_t \ddot{X}_t &= f_s - f_c \\ m_w \ddot{X}_w &= f_c - k_w X_w - c_w \dot{X}_w \\ f_s &= k_s (X_a - X_t) + c_s (\dot{X}_a - \dot{X}_t) \end{aligned} \quad (1)$$

ただし、

f_s : 力センサが観測する接触力

f_c : エンドエフェクタと対象物体の衝突により生じる衝撃力

U : マニピュレータへの入力

である。(1)式において U は次節で示す力サーボ系により生成される。また、 f_c はエンドエフェクタと対象物体の衝突の度に断続的に系に加えられる衝撃力で、本論文では以下のようにモデル化する。エンドエフェクタと対象物体のはね返り係数を e とすると、衝突の際にエンドエフェクタと対象物体の間に働く力積 R は次式で表わされる。

$$R = (e+1) \frac{m_t m_w}{m_t + m_w} (\dot{X}_t - \dot{X}_w) \quad (2)$$

(2)式は運動量保存則とはね返り係数の定義より導くことができる。このとき、衝撃力と力積の関係は次式で表わされる。

$$R = \int_{t_h}^{t_h + \tau} f_c dt \quad (3)$$

ただし、 τ は衝撃力が作用する時間であり、 t_h は衝突の発生した時刻である。衝突する物体が理想的な剛体であった場合、 R が有限であるが、 τ は無限小、 f_c は無限大となる。しかし、実際には、物体表面の柔らかさのため τ 、 f_c は共に有限の値をとる。衝撃力の影響を検討する際に重要となるのは、衝撃力の形状そのものよりもむしろ、その積分値 R である。したがって、本論文では衝撃力 f_c のモデルとして次式で表わされる近似モデルを用いることにする。

Fig. 1 Model of contact motion in phase 2

$$f_c = \frac{R}{\delta} \quad (t_h \leq t \leq t_h + \delta) \quad (4)$$

ただし、 δ は衝突の際 f_c の作用する時間であり、この値は物体表面に依存する定数である。(4)式で表わされる f_c は、衝突直後より時間 δ の間一定値に維持され、その後次の衝突まで零に保たれるという方形波状のモデルを意味する。第2フェイズにおいて衝突と衝突の時間間隔が δ 以下になったとき、系は第3フェイズに移行するものとする。

2.2 第3フェイズの力学的モデルおよび

力サーボ系

第3フェイズの系は **Fig. 2** のようにモデル化され、その挙動は次式で表わされる。

$$\begin{aligned} m_a \ddot{X}_a &= U - f_s \\ (m_t + m_w) \ddot{X}_w &= f_s - k_w X_w - c_w \dot{X}_w \\ f_s &= k_s (X_a - X_w) + c_s (\dot{X}_a - \dot{X}_w) \end{aligned} \quad (5)$$

第2フェイズと第3フェイズのモデルにおいて、大きく異なる点は第2フェイズにおいては、エンドエフェクタと対象物体が独立に動くのに対し、第3フェイズでは両者が一体となって動くという点である。

第2, 第3, 第4フェイズに対して、次式で表わされる比例、積分動作、および速度フィードバックを行う連続時間系の力サーボ系を考える。

$$U = K_{fp}(f_d^* - f_s) + K_{fi} \int_0^t (f_d^* - f_s) d\xi - K_v \dot{X}_a \quad (6)$$

Fig. 2 Model of contact motion in phase 3

ここで、

f_d^* : 力指令値

である。ただし、 f_d^* は力の目標値と次章で説明する学習コントローラの出力するフィードフォワード指令値の和により構成される。(6)式は積分動作を含んでいるため、力サーボ系は1型の閉ループ系となりステップ指令に対して定常偏差の生じないシステムとなる。今、(5), (6)式で表わされる力サーボ系において状態変数 x , 入力 u , 出力 y を以下のようにとる。

$$\begin{aligned} x &= \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} X_a \\ \dot{X}_a \\ X_w \\ \dot{X}_w \\ \int_0^t (f_d^* - f_s) d\xi \end{pmatrix} \\ u &= f_d^*, \quad y = f_s \end{aligned} \quad (7)$$

このとき、第3フェイズにおける系の状態変数表示は次式で与えられる。

$$\begin{aligned} \dot{x} &= Ax + bu \\ y &= Cx \end{aligned} \quad (8)$$

ただし、

$$\begin{aligned} A &= \begin{pmatrix} 0 & 1 & 0 \\ \frac{k_s(K_{fp}+1)}{m_a} - \frac{K_v - c_s(K_{fp}+1)}{m_a} & \frac{k_s(K_{fp}+1)}{m_a} & 0 \\ 0 & 0 & 0 \\ \frac{k_s}{m_t+m_w} & \frac{c_s}{m_t+m_w} & -\frac{k_s+k_w}{m_t+m_w} \\ -k_s & -c_s & k_s \\ 0 & 0 & 0 \\ \frac{c_s(K_{fp}+1)}{m_a} & \frac{K_{fi}}{m_a} & 0 \\ 1 & 0 & 0 \\ -\frac{c_s+c_w}{m_t+m_w} & 0 & 0 \\ c_s & 0 & 0 \end{pmatrix} \\ b &= \begin{pmatrix} 0 \\ \frac{K_{fp}}{m_a} & 0 & 0 & 1 \end{pmatrix}^T \end{aligned}$$

Fig. 3 System configuration

$$\mathbf{C} = (k, c_s - k, -c, 0) \quad (9)$$

である。(8), (9)式で表わされる力サーボ系は、可制御、可観測な5次の系である。また、得られた力サーボ系のブロック図を、Fig. 3 の上の破線部に示す。

この章では、接触動作の力学モデルを二つのフェイズに分けて示した。一般に学習コントローラを設計する際には、制御対象のモデルが必要になるが、接触動作における系の基本的な挙動は、第3フェイズによって支配される。したがって、本論文では以後、第3フェイズのモデルをもとに学習コントローラを設計し、第2フェイズの影響、つまり第3フェイズと比べた場合の次数の違いや(4)式で表わされる衝撃力の影響はすべて外乱とみなすことにする。

3. 学習コントローラ

この章では、提案する学習コントローラの具体的動作を示すと共に、 $t=0$ で第3フェイズが始まると想定して(つまり、外乱の影響を無視して)学習の収束条件を示す。これは2章の終わりで述べたように、学習コントローラを第3フェイズのモデルに基づいて設計するため(4.1節参照)、そのために必要となる条件を示すものである。また、 $t=0$ で第2フェイズが始まるとした場合の(つまり、外乱の影響を考慮にいた)収束条件、および上述のように設計された学習コントローラを用いた場合の、第2フェイズが収束性に与える影響については4.2節で検討する。

3.1 学習コントローラの動作

2章で述べた力サーボ系に学習コントローラを加えたものを Fig. 3 に示す。Fig. 3 において f_d は力の目標を表わし、添字の $k, k+1$ は学習の試行回数を表わす。力サーボ系は連続時間系で、学習コントローラは離散時間系で構成されるものとする。今、力サーボ系において、(10), (11)式で表わされるように k 回目の試行における時刻0の状態 $\mathbf{x}(0)$ と、時刻 t_0 までのフィードフォワード力指令サンプル列 $f^*(i)$ ($i=1, 2, 3, \dots, n_c$) (サンプル数 n_c , サンプル周期 T_c で零次ホールドを通して力サーボ系に加えられる)を (n_c+5) 次元の入力系列ベクトル $\mathbf{u}_k$ とし、時刻 t_0 以降の接触力の応答サンプル列 $y(i)$ ($i=1, 2, 3, \dots, n_s$) (サンプル数 n_s , サンプル周期 T_s)を n_s 次元の出力系列ベクトル $\mathbf{y}_k$ とする。

$$\mathbf{u}_k = (\mathbf{x}_k(0)^T f^*(1) \cdots f^*(n_c))^T \quad (10)$$

$$\mathbf{y}_k = (y_k(1) y_k(2) \cdots y_k(n_s))^T \quad (11)$$

また、力サーボ系の n_s 次元の目標出力系列ベクトル $\mathbf{f}_d$ は次式で与えられるものとする。

$$\mathbf{f}_d = (f_d \cdots f_d)^T \quad (12)$$

このとき、学習コントローラは(13)式で与えられる誤差系列ベクトル $\mathbf{e}_k$ を基に、

$$\mathbf{e}_k = (e_k(1) \cdots e_k(n_s))^T = \mathbf{f}_d - \mathbf{y}_k \quad (13)$$

($k+1$) 回目の試行における入力系列ベクトル $\mathbf{u}_{k+1}$ を(14)式で表わされるアルゴリズムによって生成する($n_c=3, n_s=5$ の場合を Fig. 4 に示す)。

$$\mathbf{u}_{k+1} = \mathbf{u}_k + \mathbf{H} \cdot \mathbf{e}_k \quad (14)$$

(14)式において、 $\mathbf{H}$ は $(n_c+5) \times n_s$ 次元の入力修正行列である。ただし、1回目の試行における $f^*(i)$ はすべて零とする。この繰り返しアルゴリズムによって、時刻 t_0 以降の接触力の応答が試行ごとに目標値に近づいていく。

ここで、従来軌道制御で用いられてきた学習コントローラ¹³⁾と比べたときの本学習コントローラの特徴について述べておこう。従来軌道制御に用いられてきた学習コントローラをそのまま用いようとするならば、設計者は目標過渡波形なるものを用意しなければならず、接触動作のように過渡時に外乱が作用するシステムにおいては、これは非常に困難なことである。これに対し、本手法は Fig. 3 のようにシステムの入力系列と出力系列を時間的にずらすことによって、目標過渡波形を必要としない学習法となっている(必要な目標は指令値 f_d のみである)。また、本学習コントローラのもう一つの特徴は、入力系列の中に初期状態を含めていることである。これは、接触前の状態(接近速度 $x_2(0)$ と積分器の初期値 $x_5(0)$)を学習するためである。前章で述べたように、接触問題においては第2フェイズにおいて発生する衝撃力が大きな問題となる。この衝撃力の大きさは第1フェイズにおける接近速度に依存しているため、制御の立場から見た場合の接触問題は、初期状態に依存した外乱を含む系に対して、いかに過渡特性を向上させるかという問題とみなせる。つまり接触問題においては、接近速度が過渡応答を決める一つの大きなパラメータとなるので、本学

Fig. 4 Operation of learning controller

習コントローラは学習する量の中に初期状態を含めている。

第3フェイズを想定した場合、出力系列から時刻 t_0 における状態を知り、かつ入力系列によってその状態が任意に操作できるためには、入力系列、出力系列のそれぞれのサンプル数 n_c , n_s がつぎの条件を満たす必要がある。

$$n_c \geq 3, \quad n_s \geq 5 \quad (15)$$

(15)式において、入力系列の中には初期状態が含まれており、 $x_2(0)$ と $x_5(0)$ の二つの量が操作可能であるため、 $n_c \geq 5-2=3$ で良いことに注意されたい。

3.2 学習の収束条件

この節では $t=0$ で第3フェイズが始まると想定して、前節で示した学習アルゴリズムを用いた場合の学習の収束条件について検討する。本節で示される結果は、4.1節において収束速度の速い学習コントローラ的设计指針を与える。

今、 $f_k^*(i)$ を与えるサンプル周期は T_c であるので、第3フェイズを想定した場合、 $f_k^*(i)$ の出力終了直後、すなわち時刻 $t_0 = T_c \times n_c$ での状態変数 x_k は u_k を用いて次式のように表わされる。

$$x_k = G_c \cdot u_k + g_c \quad (16)$$

ただし、

$$\begin{aligned} G_c &= (A_c^{n_c} \ A_c^{n_c-1} b_c \ A_c^{n_c-2} b_c \cdots b_c) \\ g_c &= A_c^{n_c-1} b_c f_d + A_c^{n_c-2} b_c f_d + \cdots + b_c f_d \\ G_c &: 5 \times (n_c + 5), \quad g_c: 5 \times 1 \end{aligned} \quad (17)$$

である。ここで、 A_c , b_c は、(9)式で表わされる A , b をサンプル周期 T_c で離散化して得られる力サーボ系の離散時間状態推移行列および入力行列である。今、 (A_c, b_c) が可制御であるように T_c が選ばれているとすると、 G_c のランクは5となる。

同様に、観測のサンプル周期は T_s であるから、時刻 t_0 以降 $t_0 + t_s \times (n_s - 1)$ までの出力系列ベクトル y_k は x_k を用いて次式のように表わすことができる。

$$y_k = G_s \cdot x_k + g_s \quad (18)$$

ただし、

$$\begin{aligned} G_s &= \begin{pmatrix} C \\ CA_s \\ CA_s^2 \\ \vdots \\ CA_s^{n_s-1} \end{pmatrix} \\ g_s &= \begin{pmatrix} 0 \\ Cb_s f_d \\ CA_s b_s f_d + Cb_s f_d \\ \vdots \\ CA_s^{n_s-2} b_s f_d + \cdots + Cb_s f_d \end{pmatrix} \\ G_s &: n_s \times 5, \quad g_s: n_s \times 1 \end{aligned} \quad (19)$$

である。ここで、 A_s , b_s は、(9)式で表わされる A ,

b をサンプル周期 T_s で離散化して得られる状態推移行列および入力行列である。 (C, A_s) が可観測であるように T_s が選ばれているとすると、 G_s のランクは5となる。

以上をもとに、学習の収束条件について検討する。(14)式の両辺に G_c , G_s を順に左からかけ、(13), (16), (18)式を用いると次式を得る。

$$\begin{aligned} e_{k+1} &= (I - G_s \cdot G_c \cdot H) e_k \\ &= S \cdot e_k \quad (S: n_s \times n_s) \end{aligned} \quad (20)$$

(20)式は試行ごとの出力誤差系列の遷移を表わし、 S を誤差遷移行列と呼ぶ。ここで、学習が収束するための条件として、つぎのような条件を考える。

$$\|e_{k+1}\|_2 < \|e_k\|_2 \quad (21)$$

ただし、 $\|\cdot\|_2$ はベクトルの l^2 ノルムを表わす。(20), (21)式より学習が収束するための十分条件は次式で与えられる。

$$\|S\|_2 = \sqrt{\lambda_{\max}(S^T \cdot S)} < 1 \quad (22)$$

ただし、 $\lambda_{\max}(\cdot)$ は行列の最大固有値を表わし、 $\|\cdot\|_2$ は、ベクトルの l^2 ノルムから誘導された行列ノルムを表わす。

$n_s = 5$ の場合は(22)式を使って、学習の収束性を検討できる。しかし、 $n_s > 5$ のとき、つまり観測する出力数がシステムの次数より多い場合、 G_s は縦長の行列となる。この場合、学習が収束するかどうかにかかわらず S は固有値1をもつことになり、(22)式の条件は満たされない。したがって、 n_s を5より大きく取る場合、(22)式の条件に代わって、新たな収束条件が必要となり、つぎのように与えられる¹²⁾。(13), (18)式より、

$$e_k = f_d - y_k = f_d - g_s - G_s \cdot x_k \quad (23)$$

となる。また、時刻 t_0 における力サーボ系の目標状態を x_d とすると

$$f_d - g_s = G_s \cdot x_d \quad (24)$$

となることは明らかである。したがって、(23), (24)式より $e_k = G_s \cdot (x_d - x_k) \in \text{Im } G_s$ となることがわかる。つまり、 $n_s > 5$ であっても、出力誤差系列ベクトル e_k の張る空間は R^{n_s} における5次元部分空間 $\text{Im } G_s$ 内に制限されることになる。したがって、(20)式を S による線形変換とみなしたときに、 S の代わりに定義域を $\text{Im } G_s$ に制限した変換 $S|_{\text{Im } G_s}$ を $\text{Im } G_s$ 内の正規直交基底を用いて表現し (S_1 とする)、(22)式のかわりに S_1 の誘導 l^2 ノルムを用いて学習の収束性を評価すればよい。このとき S_1 は次式で与えられる。

$$S_1 = P^{-1} \cdot (I - G_c \cdot H \cdot G_s) \cdot P \quad (25)$$

$$(S_1: 5 \times 5), \quad (P: 5 \times 5)$$

ただし、 P は G_s の5本の列ベクトルよりなる基底か

ら正規直交基底への変換行列を表わす¹²⁾. $n_s=5$ ならば, $\mathbf{P}=\mathbf{G}_s^{-1}$ と選ぶことによって(25)式と(22)式は等価となるので, 結局, 学習の収束条件は次式で与えられることになる.

$$\|\mathbf{S}_1\|_2 = \sqrt{\lambda_{\max}(\mathbf{S}_1^T \cdot \mathbf{S}_1)} < 1 \quad (26)$$

4. 学習コントローラ的设计および 第2フェイズの影響

4.1 H的设计法

学習コントローラを設計する際には学習の収束速度を考慮に入れる必要がある. 学習の収束を速めるには, (25)式の値をできるだけ小さくするように $\mathbf{H}$ を決めればよい. そこで, あらかじめ入手できる物理パラメータを用いて $\mathbf{G}_c$, $\mathbf{G}_s$ を計算し(これを $\hat{\mathbf{G}}_c$, $\hat{\mathbf{G}}_s$ とする), $\hat{\mathbf{G}}_c \cdot \mathbf{H} \cdot \hat{\mathbf{G}}_s = \mathbf{I}$ となるように $\mathbf{H}$ を決める. このとき, パラメータ同定誤差および第2フェイズの影響がなければ(25)式左辺の値は零となり, 2回の試行で誤差は零に収束する. $\hat{\mathbf{G}}_c \cdot \mathbf{H} \cdot \hat{\mathbf{G}}_s = \mathbf{I}$ を満たす $\mathbf{H}$ はいろいろ考えられるが, ここでは $\mathbf{H}=\mathbf{H}_c \cdot \mathbf{H}_s$ と分解し, 次式で与えられる擬逆行列を用いることにする.

$$\mathbf{H}_s = (\hat{\mathbf{G}}_s^T \cdot \hat{\mathbf{G}}_s)^{-1} \cdot \hat{\mathbf{G}}_s^T \quad (27)$$

$$\mathbf{H}_c = \mathbf{W} \cdot \hat{\mathbf{G}}_c^T \cdot (\hat{\mathbf{G}}_c \cdot \mathbf{W} \cdot \hat{\mathbf{G}}_c^T)^{-1} \quad (28)$$

$$\mathbf{H}_s: 5 \times n_s, \quad \mathbf{H}_c: (n_c+5) \times 5$$

$$\mathbf{W}: (n_c+5) \times (n_c+5)$$

ここで, (27), (28)式の物理的意味についてであるが, (27)式の $\mathbf{H}_s$ は観測した出力誤差系列を時刻 t_0 での状態誤差に変換する役割をもち, また, (28)式の $\mathbf{H}_c$ は時刻 t_0 での状態誤差を時刻 t_0 までの入力系列ベクトルの誤差に変換する役割をもち, また, $\mathbf{W}$ は入力系列の修正量に自由度をもたせるための行列である.

なお, $\mathbf{H}$ の選択にあたっては以下のことを考慮する必要がある. (8), (9)式で表わされる系の場合, マニピュレータと対象物体の初期(接触時の)位置 $x_1(0)$, $x_3(0)$, および対象物体の初期速度 $x_4(0)$ は修正することができないので(つまり試行ごとに同じ値をとる), $x_1(0)$, $x_3(0)$ および, $x_4(0)$ の修正を担う $\mathbf{H}$ の第1, 3, 4行は零にする必要がある.

4.2 第2フェイズの影響

前節では, 第3フェイズのモデルを基に入力修正行列 $\mathbf{H}$ を決めた. この節では, 前節で設計した $\mathbf{H}$ を用いた場合, 外乱とみなした第2フェイズの影響をどう受けるかについて検討する. なお, 議論を簡単にするため第3フェイズに関しては正確なモデルが得られているものと仮定する. つまり,

$$\hat{\mathbf{G}}_c = \mathbf{G}_c, \quad \hat{\mathbf{G}}_s = \mathbf{G}_s \quad (29)$$

であるとする.

今, $t=0$ で第2フェイズが始まるとすると, (16)式は次式のように書き改められる.

$$\mathbf{x}_k = (\mathbf{G}_c + \Delta\mathbf{G}_c) \cdot \mathbf{u}_k + \mathbf{g}_c \quad (30)$$

$$(\Delta\mathbf{G}_c: 5 \times (n_c+5))$$

ただし, $\Delta\mathbf{G}_c$ は第2フェイズと第3フェイズの次数の相違や第2フェイズにおける衝撃力の影響を表わす行列である. また, 整定時間 t_0 までに系は第3フェイズに移行しているものとする(つまり, (18)式中の $\mathbf{G}_s$ は第2フェイズの影響を受けないものとする). このとき, 学習の収束条件は(26)式から次式に書き改められる.

$$\|\mathbf{P}^{-1} \cdot (\mathbf{I} - (\mathbf{G}_c + \Delta\mathbf{G}_c) \cdot \mathbf{H} \cdot \mathbf{G}_s) \cdot \mathbf{P}\|_2 < 1 \quad (31)$$

(31)式に(29), (28), (27)式を代入すると, $\mathbf{G}_c \cdot \mathbf{H}_c = \mathbf{H}_s \cdot \mathbf{G}_s = \mathbf{I}$ より次式を得る.

$$\begin{aligned} & \|\mathbf{P}^{-1} \cdot (\mathbf{I} - (\mathbf{G}_c + \Delta\mathbf{G}_c) \cdot \mathbf{H} \cdot \mathbf{G}_s) \cdot \mathbf{P}\|_2 \\ &= \|\mathbf{P}^{-1} \cdot (\mathbf{I} - (\mathbf{G}_c + \Delta\mathbf{G}_c) \cdot \mathbf{H}_c \cdot \mathbf{H}_s \cdot \mathbf{G}_s) \cdot \mathbf{P}\|_2 \\ &= \|\mathbf{P}^{-1} \cdot \Delta\mathbf{G}_c \cdot \mathbf{H}_c \cdot \mathbf{P}\|_2 \end{aligned} \quad (32)$$

(32)式において, $\mathbf{P}$ は $\mathbf{G}_s$ に依存する行列である. したがって, 第2フェイズの影響がない場合(この場合, (29)式の仮定より(26)式の左辺は零となる)と比べて, $\|\mathbf{P}^{-1} \cdot \Delta\mathbf{G}_c \cdot \mathbf{H}_c \cdot \mathbf{P}\|_2$ の分だけ学習の収束速度が遅くなることがわかる. 特に, エンドエフェクタの質量の値やはね返り係数の値が大きい場合には $\|\Delta\mathbf{G}_c\|_2$ の値が大きくなるので学習の収束性は劣化する(場合によっては発散する). したがって, このような場合には収束速度のみではなく, $\|\Delta\mathbf{G}_c \cdot \mathbf{H}_c\|_2$ の値をも小さくするように $\mathbf{H}_c$ を決める必要がある. このような $\mathbf{H}_c$ の具体的な設計法については別の機会に報告したい.

5. シミュレーション結果および考察

5.1 接近速度の学習効果

本論文では, パラメータ同定誤差として, 一般に同定が困難であると思われる作業対象のバネ定数 k_w を取り上げる. 対象としたシステムのパラメータをTable 1に示す. (28)式中の重み行列 $\mathbf{W}$ は, $\mathbf{H}$ の第1, 3, 4行を零にする必要があるため次式のように選択した.

$$\mathbf{W} = \text{diag}(0 \ 1 \ 0 \ 0 \ 1 \ 1 \dots 1) \quad (33)$$

シミュレーションにおいては, $t=0$ で第2フェイズが始まると想定しており, 第2フェイズにおける f_c の値は衝突の度(2), (4)式から計算する. Fig. 5に k_w の同定誤差が50%のときの学習のようすを示す. 上図が接触力 f_c の応答を表わし, 下図がフィードフォワード指令値を表わす. 図中のパラメータ1, 4,

Table 1 System parameters

Nomenclature	Item	Actual value	Estimated value	Unit
m_a	Manipulator mass	15.0	15.0	kg
k_s	Sensor stiffness	20,000.0	20,000.0	N/m
c_s	Sensor damping	10.0	10.0	N/m/s
m_t	End-effector mass	0.6	0.6	kg
m_w	Workpiece mass	2.0	2.0	kg
k_w	Workpiece stiffness	20,000.0	30,000.0	N/m
c_w	Workpiece damping	50.0	50.0	N/m/s
K_{fp}	Proportional feedback gain	10.0	—	—
K_{fi}	Integral feedback gain	8.0	—	—
K_v	Velocity feedback gain	3,000.0	—	N/m/s
n_c	—	6	—	—
n_s	—	12	—	—
T_c	—	8.0	—	ms
T_s	—	8.0	—	ms
e	Elastic modulus	0.8	—	—

7は試行回数を表わし、下図における V_0 , S_0 はそれぞれ各試行における接近速度 ($x_2(0)$) と積分器の初期値 ($x_3(0)$) を表わす。なお、下図において1回目の試行におけるフィードフォワード指令値はすべて零であり、上図においてはすべての試行において、 $t=0$ で $f_s=0$ である。また、これらのシミュレーションにおいて5回目以降の試行では、フィードフォワード力指令値、システム初期値共にほとんど変化が見られないため、学習は5回で打ち切った。これを見ると、試行を繰り返すごとに、接触力の応答波形が改善されていることがわかる。ここで注目すべき点は、本学習コン

トローラは衝撃力の影響が零となるように接近速度を修正してはいない点である。つまり、本学習コントローラは衝撃力の影響をなくしてしまうのではなく、むしろある程度衝撃力の影響を利用した立ち上がりの早い応答を実現する接近速度を、学習によって得ていると見なせる。比較のため、Fig. 6に接近速度を0とし、かつ、それを学習していない場合のシミュレーション結果を示す。この場合は衝撃力の効果を利用できないためフィードフォワード指令値に大きな負担がかかっており、操作量の飽和につながる可能性が十分にある。

Fig. 5 Simulation results 1
(Estimation error of $k_w=50\%$)
(with approaching velocity learning)
(Elastic modulus $e=0.8$)

Fig. 6 Simulation results 2
(Estimation error of $k_w=50\%$)
($V_0=0$, without approaching velocity learning)
(Elastic modulus $e=0.8$)

Fig. 7 Simulation results 3
(Estimation error of $k_w=50\%$)
(with approaching velocity learning)
(Elastic modulus $e=0.9$)

5.2 はね返り係数 e の影響

第2フェイズにおける衝突の際生じる衝撃力ははね返り係数 e の定数に従って減衰する。このため e が大きい物質間の衝突の場合、衝撃力の減衰が小さく第2フェイズの影響が大きく表われる。4章で述べたように、本学習コントローラは第3フェイズのモデルを基に構成されるので、第2フェイズの影響が大きいほど(たとえば、 e の値が大きいほど)、学習の収束性は劣化する。Fig. 7 に $e=0.9$ とした場合のシミュレーション結果を示す。Fig. 5と比較すると学習の収束速度がやや遅いことがわかる。

5.3 考察

5.1節でも述べたように、本学習コントローラによって獲得される接近速度が零とはならないことは、接触動作に要する時間の短縮化という点で大きな意味をもつ。これに対し、Fig. 6のように衝撃力の影響がほとんど現れない接触動作を行う場合には、接触前に十分に速度を落とす必要があり、接触動作に要する時間が大幅に長くなると思われる。このことから、本学習方式が速やかな接触動作を実現するものであることがわかる。また、本学習コントローラによって獲得される接近速度を用いた場合、極端に n_c が小さいか大きい場合を除き、ほとんどの場合オーバーシュートのない波形が得られることをシミュレーションによって確認している。したがって、本学習コントローラはある意味で衝撃力によるオーバーシュートを抑制する能力があるともいえる。なお、本論文では、時刻0で第2フ

ェイズが始まると仮定しているが、実際のシステムでは、検出遅れによるサーボ切換えの遅れのため、学習の収束性の悪化が予想される。しかし、本章で示したシミュレーションのパラメータの場合、検出遅れが2ms程度ならいずれも収束することを確認しており、実機に適用した場合でも検出遅れの影響は小さいと思われる。

6. まとめ

本論文では、繰返し動作に基づいた学習制御を用いて、マニピュレータと作業対象との接触動作を試行ごとに改善する手法を提案し、学習コントローラ的设计法と収束性について論じた。本論文で示した学習制御法の特徴を以下に述べる。

1. 本学習コントローラの特徴は、システムの入力系列と出力系列を時間的にずらして考えているため目標過渡波形を用意する必要がないこと、および、学習する量の中に接近速度を含んでいることである。

2. 本学習コントローラは、学習する量の中に接近速度を含んでいるため、衝撃力を利用した立ち上がりの速い応答が実現できる。また、このことは、接触動作に要する時間の短縮(接近速度が必要以上に小さくなりすぎないため)という点で重要な意味をもつ。

3. 学習コントローラを設計する段階で衝撃力のモデルに関する情報は用いない。したがって、本学習コントローラは衝突時に発生する衝撃力の影響を外乱と見なすことになり、衝撃力の影響が強くなればなるほど学習の収束速度は遅くなる。

なお、ガタなどの不確定要素が存在し、接触現象の再現性が保証されない場合、学習の収束性が保証されない可能性がある。これについての実機による確認、および本学習法を用いた場合に、学習によって得られる波形がオーバーシュートを含むかどうかについての理論的検討が、今後の課題である。また衝撃力が非常に大きく学習が発散するような場合には、衝撃力の影響に対してロバストな学習アルゴリズムを構築する必要があると思われるが、これらについては別の機会に報告したい。

参考文献

- 1) O. Khatib and J. Burdich: Motion and Force Control of Robot Manipulators, Proc. of IEEE Int. Conf. on Robo. and Auto., 1381/1386 (1986)
- 2) T. Yamada and T. Yabuta: Changing Method of Control Mode for Manipulators at Impact of Objects, Proc. of 20th ISIR, 593/600 (1989)
- 3) K. Bo Sim, H. Hashimoto and F. Harashima: Contact Force Control of Robot Hand using VSS, Proc.

- of 4th Korean Auto. Cont. Conf., 1080/1084 (1989)
- 4) 北垣, 内山: 外部環境に対するマニピュレータの最適接近速度, 日本ロボット学会誌, **8-4**, 413/420 (1990)
 - 5) J. Kahng and F.M.L. Amirouche: Impact Force Analysis, Proc. of IEEE Int. Conf. on Robo. and Auto., 2061/2068 (1987)
 - 6) K. Youcef-Toumi and D. A. Gutz: Impact and Force Control, Proc. of IEEE Int. Conf. on Robo. and Auto., 410/415 (1989)
 - 7) Y. Shoji, M. Inaba, T. Fukuda and H. Wada: Contact/Impact Position and Force Hybrid Control, Proc. of IEEE Int. Workshop on Adv. Mot. Cont., 164/172 (1990)
 - 8) 羅, 伊藤: コンプライアントマニピュレーションのためのロボットの制御設計, 計測自動制御学会論文集, **26-4**, 427/434 (1990)
 - 9) 永田, 小笠原, 小俣: 多関節ロボットの柔軟な突き当てのための最適速度ベクトル, 計測自動制御学会論文集, **26-4**, 435/442 (1990)
 - 10) 鈴木, 山田, 大熊: 学習に基づいた外部環境に対するマニピュレータの最適接触法, 第7回日本ロボット学会学術講演会予稿集, No. 1503, 193/196 (1989)
 - 11) 井上, 越智, 斎藤: 繰り返し制御による過渡応答の改善, 計測自動制御学会論文集, **22-1**, 30/35 (1986)
 - 12) 鈴木, 大熊, 内川: 入力数が出力数より少ない線形系に対する学習安定条件, 第28回 SICE 学術講演会予稿集, 179/180 (1989)
 - 13) 鈴木, 安江, 大熊, 内川: ロボットマニピュレータの離散時間学習制御, 日本ロボット学会誌, **8-2**, 179/187 (1990)
-