接触瞬間での制御則切替を必要としない 繰り返し学習型接触動作制御*

板橋 界 児^{*1}, 山田 康 二^{*1}
鈴木 達 也^{*1}, 大熊 繁^{*2}

Contact Force Control Based on Learning Operation without Switching the Servo Mode at the Instant of the Impact

Kaiji ITABASHI, Koji YAMADA,
Tatsuya SUZUKI and Shigeru OKUMA

We propose a new method of controlling the contact force based on the iterative learning operation. Most of the conventional contact force control techniques must change the servo mode from "the position control mode" to "the force control mode" at the instant of the impact. However, if there exists a time delay during the change of the servo mode, an unstable phenomenon arises. On the contrary, the proposed method is based on the damping control and change only the force command signal at a certain time T_0 after the impact. The learning controller optimizes the force command signal before T_0 and realizes the dead-beat contact force response after the impact. This results in the smooth contact operation. We show some simulation and experimental results to verify the effectiveness of the proposed method.

Key Words: Contact Force Control, Learning Control, Impact Phenomenon

1. はじめに

ロボットが、環境との接触を伴いながら行なう作業として、バリ取り、穴開け、ボンディング等の作業が挙げられる。これらの作業をスムーズに行なうためには、ロボットが自由空間を運動している状態から、環境との接触を伴う状態への移行をスムーズに行なう必要がある。この問題は接触動作制御問題として、これまでさまざまな研究がなされてきた⁽³⁾⁻⁽¹⁰⁾。中でも、筆者らは繰り返し学習アルゴリズムを用いた接触動作の制御アルゴリズムを提案してきた⁽⁸⁾⁻⁽¹⁰⁾。そこで提案されたアルゴリズムを用いれば、例えば環境に関する正確な情報がなくても、繰り返し動作を行なわせることによって、最終的に接触力の有限整定応答が得られ、しかも望ましい接近速度を獲得することができた。しかしながら、この手法においては、接触瞬間に位置制御モードから力制御モードへ切替えることを前提としており (Fig.1), 従って、切替に無視できないほどの時間的遅れがある場合、大きな衝撃力が発生し、学習の安定性が確保できないという問題点があった。

以上の点を踏まえて本論文では、制御則の切替を接触直前の任意の時点に設定することができ、接触瞬間での制御則の切替を必要としない (Fig.1(b)), 繰り返し学習に基づいた新しい接触動作制御手法を提案する。提案する手

法においては、力の比例制御、フィードフォワード項、速度ダンピング項からなる制御則を採用し、接触後のある時刻 (T_0) で力指令値を変化させるという手法を用いる。そして、 T_0 以前の力指令値を学習により最適化していく。この制御則を用いる場合、環境との接触がない状態においては、力指令値がマニピュレータの速度の定常値に比例するので、 T_0 以前の力指令値を学習するということは、接近速度を学習することを意味している。

以下、本論文では、一自由度マニピュレータを対象とし、環境がバネ定数のみでモデル化できる場合と、バネ定数と粘性抵抗でモデル化できる場合とに分けて、力指令値を変化させるタイミングを示し、学習アルゴリズムを示す。また、提案手法の有効性を、シミュレーション、実験を用いて検証する。

Fig. 1. Switching of force command signal and servo mode

* 原稿受付 1995年7月3日.

^{*1} 名古屋大学大学院 (〒464-01 名古屋市千種区不老町).

^{*2} 正員, 名古屋大学工学部.

2. 接触動作モデルと力サーボ系

本章では、環境がバネ要素のみからなる場合と、バネと粘性抵抗要素からなる場合に分けて、接触動作のモデルと制御則を示す。

2.1 環境がバネ要素のみからなる場合 環境がバネ要素のみからなる場合の接触動作モデルを Fig.2 に示す。

Fig. 2. Physical model of 1-DOF manipulator I

このとき、系のダイナミクスは、式(1)(2)で表される。

$$m\ddot{x} = u - f_s \quad (1)$$

$$f_s = kx \quad (2)$$

ここで、 m はマニピュレータの質量、 u はマニピュレータへの入力、 f_s は力センサにより観測される接触力、 k は環境のバネ定数である。また x はマニピュレータの手先位置を表しており、接触の瞬間を時刻 $t = 0$ とする。この系を、接触の前後を通して式(3)で表されるサーボ系で制御する。

$$u = K_p(f_u - f_s) + f_u - K_v\dot{x} \quad (3)$$

ここで、 K_p は比例ゲイン、 K_v は速度フィードバックゲインである。 f_u は式(4)で与えられる力目標値であり、Fig.3のような波形となる。

$$f_u = \begin{cases} F_{d0} & (t < T_0) \\ F_d & (t \geq T_0) \end{cases} \quad (4)$$

Fig. 3. Force command signal

ここで、 T_0 は力センサにより観測される接触力が初めてピークに達する時刻とする。 F_d は最終的な接触力目標値、 F_{d0} は接触力応答改善のため $t < T_0$ において与えられる仮

の力目標値である。ここで注意すべき点は、本手法では接触の前後を通して式(3)で制御を行なうので、Fig.3に示すように接触前の時刻零以前から力目標値が与えられている点である。サーボ系のブロック図を Fig.4 に示す。また、式(3)(4)で表される制御則を用いた場合の系の挙動は以下の通りである。

1. 接触前 ($t < 0$)

式(4)より、 $t < 0$ なので $f_u = F_{d0}$ である。また、マニピュレータは自由空間にあるので、 $f_s = 0$ である。ここで、マニピュレータは定常状態、つまり一定速度で環境に接近する状態にあるとすると、 $\ddot{x} = 0$ を式(1)(3)に代入することによって式(5)を得る。

$$\dot{x} = \frac{K_p + 1}{K_v} F_{d0} = v_0 \quad (5)$$

従って、マニピュレータは式(5)で与えられる接近速度 v_0 で環境に接近することがわかる。

2. 接触後 I ($0 \leq t < T_0$)

式(4)より、 $f_u = F_{d0}$ である。このとき、式(1)(2)(3)より、式(6)の運動方程式を得る。

$$\ddot{x} + \frac{K_v}{m}\dot{x} + \frac{(K_p + 1)k}{m}x = \frac{K_p + 1}{m}F_{d0} \quad (6)$$

ここで、 $n = \frac{K_v}{2m}$ 、 $p = \sqrt{\frac{(K_p + 1)k}{m}}$ とすると系の固有振動数 $p_1 = \sqrt{p^2 - n^2}$ となり、式(6)の解は式(7)で与えられる。

$$x(t) = \left\{ \exp^{-nt} (A_1 \sin p_1 t + A_2 \cos p_1 t) + \frac{1}{k} \right\} F_{d0} \quad (7)$$

ただし、 A_1, A_2 は初期条件によって定められる定数である。式(7)に、初期条件 $x(0) = 0$ 、 $\dot{x}(0) = v_0$ を代入し、式(8)を得る。

$$\begin{aligned} A_1 &= \frac{\frac{K_p + 1}{K_v} - \frac{n}{k}}{p_1} \\ A_2 &= -\frac{1}{k} \end{aligned} \quad (8)$$

$f_s(t)$ の応答は Fig.5 の実線 ($0 \leq t < T_0$ の領域) のようになる。また、力センサにより観測される接触力が初めてピークとなる時刻 T_0 は、 $\dot{f}_s(t) = k\dot{x}(t) = 0$ を満たす t であるから、式(7)を微分することによって求められ、式(9)で与えられる。

$$T_0 = \frac{1}{p_1} \arctan\left(\frac{A_1 p_1 - A_2 n}{A_1 n + A_2 p_1}\right) \quad (9)$$

ここで、 T_0 が F_{d0} に依存しない点に注意されたい。これより、時刻 $t = T_0$ における接触力 $f_s(T_0)$ は式(10)で与えられる。

$$\begin{aligned} f_s(T_0) &= kx(T_0) \\ &= k\{\exp^{-nT_0}(A_1 \sin p_1 T_0 + A_2 \cos p_1 T_0) + \frac{1}{k}\}F_{d0} \\ &= \alpha \cdot F_{d0} \end{aligned} \quad (10)$$

ここで,

$$\alpha = k \cdot \exp^{-nT_0}(A_1 \sin p_1 T_0 + A_2 \cos p_1 T_0) + 1 \quad (11)$$

である. 式 (11) より, α は F_{d0} に依存しないので, $f_s(T_0)$ は F_{d0} に比例することがわかる.

3. 接触後 II ($t \geq T_0$)

接触後 I の状態を保つと, Fig.5 の 1 点鎖線のような力応答となる. そこで, 力センサにより観測される接触力が初めてピークに達した時点 ($t = T_0$) で, 式 (4) に従って力目標値 f_u を F_d に変化させる. $t = T_0$ において $\dot{f}_s(T_0) = \dot{x}(T_0) = 0$ となるため, F_{d0} を適切に選ぶことによって, $f_s(T_0) = F_d$ とすることができれば, 式 (3) の第 1 項および第 3 項は零となり, 全ての状態変化が零となる有限整定が実現できる. 次章で示す学習コントローラはこの関係を満たすように F_{d0} を修正していくものである.

Fig. 4. System configuration

Fig. 5. Desired output

2.2 環境に粘性抵抗を含む場合 前節では, 環境がバネ要素のみからなるとしてモデルと指令値切替時刻を示した. 本節では, 環境の粘性抵抗を考慮した Fig.6 で表される系について考える.

Fig.6 より, 力センサで観測される接触力 f_s は, 式 (12) で与えられる.

$$f_s = kx + c\dot{x} \quad (12)$$

前節では力目標値を変化させる時刻 T_0 を, 力センサにより観測される接触力が初めてピークに達する時刻, つまり

Fig. 6. Physical model of 1-DOF manipulator II

$\dot{f}_s = 0$ となる時刻とした. これは, F_{d0} を適当に選ぶことによって $f_s(T_0) = F_d$ とすることができれば接触力の有限整定応答が実現できるからである. しかしながら, 環境に粘性抵抗が存在する場合は, $\dot{x} = 0$ となる時刻を T_0 としないと有限整定応答は実現できない.(環境がバネ要素のみの場合は, $\dot{f}_s = 0$ と $\dot{x} = 0$ が等価である点に注意されたい.) 従って, 本節では, T_0 をマニピュレータの速度が初めて零となる時刻に選ぶものとし, そのときの系の挙動について以下に示す. ただし, 接触前 ($t < 0$) と接触後 II ($t \geq T_0$) については, 前節と同じなのでここでは省略する.

● 接触後 I ($0 \leq t < T_0$)

式 (4) より $f_u = F_{d0}$ である. 式 (1)(3)(12) より, 運動方程式は式 (13) で与えられる.

$$\ddot{x} + \frac{(K_p + 1)c + K_v}{m}\dot{x} + \frac{(K_p + 1)k}{m}x = \frac{K_p + 1}{m}F_{d0} \quad (13)$$

ここで, $l = \frac{(K_p + 1)c + K_v}{2m}$, $p = \sqrt{\frac{(K_p + 1)k}{m}}$ とすると系の固有振動数 $p_2 = \sqrt{p^2 - l^2}$ となり, 式 (13) の解は式 (14) で与えられる.

$$x(t) = \{\exp^{-lt}(B_1 \sin p_2 t + B_2 \cos p_2 t) + \frac{1}{k}\}F_{d0} \quad (14)$$

ただし, B_1, B_2 は初期条件によって定められる定数である. 式 (14) に, 前節と同じ初期条件 $x(0) = 0$, $\dot{x}(0) = v_0$ を代入すると, 式 (15) を得る.

$$\begin{aligned} B_1 &= \frac{\frac{K_p + 1}{K_v} - \frac{1}{k}}{p_2} \\ B_2 &= -\frac{1}{k} \end{aligned} \quad (15)$$

また, マニピュレータの速度が初めて零になる時刻 T_0 は $\dot{x}(t) = 0$ を満たす t であるから, 式 (14) を微分することによって求められ, 式 (16) で与えられる.

$$T_0 = \frac{1}{p_2} \arctan\left(\frac{B_1 p_2 - B_2 l}{B_1 l + B_2 p_2}\right) \quad (16)$$

このように、環境が粘性抵抗を含む場合でも、 T_0 は F_{d0} に依存しない。また、時刻 $t = T_0$ における接触力 $f_s(T_0)$ は、式(17)で与えられる。

$$\begin{aligned} f_s(T_0) &= kx(T_0) \\ &= k\left\{\exp^{-\beta T_0}(B_1 \sin p_2 T_0 + B_2 \cos p_1 T_0) + \frac{1}{k}\right\}F_{d0} \\ &= \beta \cdot F_{d0} \end{aligned} \quad (17)$$

ここで、 β は

$$\beta = k \cdot \exp^{-\beta T_0}(B_1 \sin p_2 T_0 + B_2 \cos p_2 T_0) + 1 \quad (18)$$

である。

3. 学習コントローラ

環境が粘性抵抗を含むと含まざるとに関わらず、前章で定義した時刻 T_0 において、接触力が最終的な力目標値と一致していなければならない。すなわち、

$$f_s(T_0) = F_d \quad (19)$$

を満たす必要がある。また前章で述べた通り、粘性抵抗を含む場合も、含まない場合も $f_s(T_0)$ は仮の力目標値 F_{d0} に比例するため、適切な F_{d0} を与えれば式(19)の条件を満たすことが可能となる。しかし、摩擦等、非線形な要素のために F_{d0} を解析的に求めることは不可能であること、パラメータ推定誤差があった場合に条件を満たすことが保証できないことなどから、ここでは繰り返し試行に基づく学習により試行ごとに時刻 T_0 での力応答誤差 e から F_{d0} を修正していき、最終的に有限整定を実現する最適な F_{d0} を得ることを考える。

Fig. 7. Relation between F_{d0} and force response

3.1 学習則 学習則を式(20)に示す。

$$F_{d0,k+1} = F_{d0,k} + h \cdot e_k \quad (20)$$

ここで、添字 $k, k+1$ は試行回数を表す。また、 h は学習係数、 e_k は k 回目の試行における時刻 T_0 での力応答誤差

で、式(21)で与えられる (Fig.7)。

$$e_k = F_d - f_{s,k}(T_0) \quad (21)$$

式(20)(21)は、観測された誤差に基づき F_{d0} を試行ごとに修正していき、最終的に誤差を零とする F_{d0} を得ようというものである。このアルゴリズムは文献(8)で示されているアルゴリズムと比べて、非常に簡易なアルゴリズムとなっている。学習アルゴリズムを付加したブロック図を、Fig.8に示す。

Fig. 8. System configuration with learning controller

3.2 学習の収束条件 本節では学習の収束条件を、環境が粘性抵抗を含む場合と、含まない場合とに分けて示す。まず、粘性抵抗を含まない場合であるが、式(10)(11)(21)より、次式を得る。

$$\begin{aligned} e_{k+1} &= F_d - f_{s,k+1}(T_0) \\ &= e_k + f_{s,k}(T_0) - f_{s,k+1}(T_0) \\ &= e_k - \alpha(F_{d0,k+1} - F_{d0,k}) \\ &= (1 - \alpha h)e_k \end{aligned} \quad (22)$$

$k+1$ 回目における力応答誤差が k 回目よりも小さければ学習は収束するので、学習の収束条件は式(23)で与えられる。

$$\begin{aligned} |1 - \alpha h| &< 1 \\ 0 &< h < \frac{2}{\alpha} \end{aligned} \quad (23)$$

この式において h が $2/\alpha$ に近いほど学習の収束は速くなる(つまり、試行回数が少なくて済む)が、一般には、 α の正確な値は入手不可能であるため、多少収束速度を犠牲にしても、 h は小さめに設定しておく必要がある。

次に、環境が粘性抵抗を含む場合であるが、この場合は、上述の展開において α を β に変更するだけで良いので詳細は省略する。

4. シミュレーション結果

4.1 環境がバネ要素のみからなる場合 環境がバネ要素のみからなる場合のシミュレーション結果を Fig.9に示す。シミュレーションに用いたパラメータは Table 1の通りである。

Fig. 9. Simulation result I

Table 1. Parameters used for simulation I

nomenclature	item	value	unit
m	manipulator mass	0.3	kg
k	workpiece stiffness	40000	N/m
K_p	proportional feedback gain	1	—
K_v	velocity feedback gain	64	—
h	learning controller gain	0.2	—

Fig.9より、試行を繰り返すごとに T_0 以前の力目標値 F_{d0} が修正されていき、最終的に有限整定を実現していることがわかる。ここで、 T_0 は一回目の試行において、接触力応答の微分値が最初に零になる時刻に設定してある。

4.2 環境が粘性抵抗を含む場合 環境が粘性要素を含む場合のシミュレーション結果を Fig.10に示す。シミュレーションに用いたパラメータを Table 2に示す。

Table 2. Parameters used for simulation II

nomenclature	item	value	unit
m	manipulator mass	0.3	kg
k	workpiece stiffness	40000	N/m
c	workpiece damping	10	N/m/s
K_p	proportional feedback gain	1	—
K_v	velocity feedback gain	64	—
h	learning controller gain	0.2	—

このシミュレーションにおいては1回目の試行において最初に速度が零となる時刻を T_0 としている。この場合も

Fig. 10. Simulation result II

同様に有限整定応答が実現できていることがわかる。ここで注意すべき点は、Fig.10を見てもわかる通り、 T_0 と力応答のピーク時刻の差が非常に少ないという点である。通常の工業材料においては、粘性抵抗は少なからず存在すると考えられるため、理論的には時刻 T_0 としてマニピュレータの速度が零になる時刻を用いる必要がある。しかしながら、そのためには精度の良い速度センサを新たに付加する必要があり、望ましくない(式(3)の $K_v \dot{x}$ の演算の際には非常に簡易な速度センサを用いている)。そこで、環境が粘性抵抗を含む場合でも接触力の応答の微分値が零となる時刻を T_0 として採用することを考える。この時のシミュレーション結果を Fig.11, Fig.12に示す。Fig.11では通常の工業材料を想定しているのに対し、Fig.12では粘性抵抗定数の非常に大きな工業材料を想定してシミュレーションを行なった。具体的には、Fig.11のシミュレーションに用いたパラメータは Table 2と同じであるが、Fig.12のシミュレーションでは粘性抵抗定数 c のみを3倍の 30 N/m/s に変更した。まず Fig.11のシミュレーション結果より、完全な有限整定応答は実現できないものの、通常の工業材料に対しては実用上問題の無いと思われる応答が実現されていることがわかる。これに対し、Fig.12のように粘性抵抗定数の大きな工業材料を扱う場合には、接触力応答の微分値を元に T_0 を決める方法では良い応答が得られない。この場合は、前述のようにマニピュレータの速度が初めて零になる時刻を検出して T_0 を設定し、有限整定を実現する必要がある。次章では、対象とする環境の粘性抵抗定数が通常の工業材料の範囲内であると想定して、 T_0 を接触力応答の微分値が零となる時刻にとった実験結果を示す。

Fig. 11. Simulation result III-I ($c = 10$) Fig. 12. Simulation result III-II ($c = 30$)

5. 実験結果

前章で述べたように、本章で示す実験においては、 $f_s = 0$ となる時刻を T_0 として採用する。

5.1 実験装置 本研究で用いた実験装置の概略を Fig.13に示す。

1 自由度マニピュレータとしては VCM(Voice Coil Motor) を、対象物体としては両端支持したプラスチック板を用いた。支持位置によってコンプライアンスを変化させることも可能である。

力サーボはアセンブラで記述されており、制御のサンプリングタイムは $0.1[ms]$ である。また、学習コントローラはオフラインで実行されるため、高速な処理は必要ないのて C++ で記述されている。

実験環境にまつわる各パラメータの値を Table 3に示す。

Fig. 13. Experimental setup

Table 3. Parameters used for experiment

nomenclature	item	value	unit
m	manipulator mass	0.4	kg
K_p	proportional feedback gain	1	—
K_v	velocity feedback gain	48	—
T_s	sampling time	0.1	ms
h	learning controller gain	0.2	—

5.2 実験結果 実験結果を Fig.14, Fig.15, Fig.16に示す。各実験結果は、パラメータは Table 3のままで、環境のプラスチック板のみ変化させて実験を行なった結果である。Fig.14は $2[mm]$ 厚のプラスチック板を $5[cm]$ 幅で両端支持したものを環境として実験したものであり、これを標準状態とおいた場合、Fig.15は $3[mm]$ 厚のプラスチック板を $5[cm]$ 幅で両端支持したものを環境としたものであり、《硬い環境》とすることができる。また、Fig.16は $2[mm]$ 厚のプラスチック板を $8[cm]$ 幅で両端支持したものを環境としているので、《軟らかい環境》とすることができる。いずれの実験結果においても、全て最終的に $f_s(T_0) = F_d$ となっているために学習は収束している。また、試行を繰り返すごとにオーバーシュートが低減してゆき、最終的にほぼ有限整定応答を実現していることがわかる。

これらの実験結果からわかるように、本手法では学習定数を適切に選ぶことによって、学習を収束させるための試行の回数は変わるものの、環境が多少変化しても有限整定を実現することが可能である。

6. まとめ

本論文では、試行を繰り返すごとに力目標値を修正する繰り返し学習制御を用いて接触動作を改善する手法を提案し、シミュレーションと実機による検証を行なった。本論文で提案した手法の特徴を以下に述べる。

Fig. 14. Experimental result (normal environment)

Fig. 15. Experimental result (high stiffness environment)

Fig. 16. Experimental result (low stiffness environment)

1. 一般的な接触動作制御法が、接触瞬間に位置制御モードから力制御モードへの制御モードの切替を必要とするのに対し、本手法では接触直前の任意の時点で制御モードの切替を行なうことによって、接触の前後を通して同一の制御モードで制御を行なうことができる。従って制御モードの切替遅れに起因する不安定さを招く心配がない。
2. 1回目の力応答の最初のピークを与える時刻 T_0 とし、その時刻以前の力指令値を学習することで、ほぼ有限整定に近い応答を得ることが可能となる。
3. 学習係数は対象物体をもとに決定する必要があるが、対象物体が未知の時は学習係数を十分小さくすることによって安定した学習を得ることができる。ただし、試行回数は多少増えることになる。

また今後の課題としては、粘性抵抗による影響が無視できないほど大きい工業材料に対して、速度センサを取り付けられない環境下における T_0 の決定法があげられる。これについても学習を導入するのが望ましいと思われる。また、接触直後の衝撃力が非常に大きい場合の学習安定性についても検討する必要がある。これら全てを考慮した学習アルゴリズムを構築することが最終的に必要である。

文 献

- (1) R.P.Paul: Problems and research issues associated with the hybrid control of force and displacement, Proc. of IEEE Int. Conf. on Robo. and Auto., pp.1966-1971, 1987
- (2) O.Khatib and J.Burdick: Motion and Force Control of Robot Manipulators, Proc. of IEEE Int. Conf. on Robo. and Auto., pp.1381-1386, 1986
- (3) T.Yamada and T.Yabuta: Changing Method of Control Mode for Manipulators at Impact of Objects, Proc. of 20th ISIR, pp.593-600, 1989
- (4) 正司康雅, 稲葉忠, 福田敏男: 衝突を含むロボットマニピュレータの統一のアプローチに基づく安定制御, 日本ロボット学会誌, 第11巻, 第5号, pp.677-692, 1993
- (5) 北垣高成, 内山勝: 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, 第8巻, 第4号, pp.413-420, 1990
- (6) 永田和之, 小笠原司, 小俣透: 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 Vol.26, No.4, pp.435-442, 1990
- (7) 羅志偉, 伊藤正美: 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 Vol.26, No.4, pp.427-434, 1990
- (8) 山田康二, 鈴木達也, 大熊繁: 学習方式によるマニピュレータの接触動作の改善, 計測自動制御学会論文集, Vol.27, No.9, pp.1009-1017, 1991
- (9) 鈴木達也, 山田康二, 山本邦利, 近藤真実, 大熊繁: 学習に基づいた2自由度マニピュレータの接触動作の制御, 電気学会論文誌 C, 113巻, 8号, pp.603-612, 1993
- (10) 山田康二, 鈴木達也, 大熊繁: 学習によるマニピュレータの接触動作制御-実機による検証-, 日本ロボット学会誌, 第12巻, 第8号, pp.120-127, 1994