

ニューラルネットワークの適応能力を利用した倒立振子のフィードバック制御

山海 陽一朗^{*} · 和久屋 寛^{**} · 森田 譲^{***}

Feedback Control of Inverted Pendulum Making Use of Adaptability of Neural Networks

By

Yoichiro SANKAI , Hiroshi WAKUYA and Yuzuru MORITA

Abstract : The final goal of control engineering is to develop a technique how we can operate an objective system desirably. At that time, identification of the dynamical system comes first, its modeling comes next, and finally derivation of control strategy follows. But the actual system is not so easy to handle, because external noise might disturb its dynamics even if the feedback control scheme known as its stability is adopted. In order to overcome such realistic problems, adaptability of neural networks is one of noticeable properties we can apply. Then, paying attention to such abilities, two different kinds of neuro-controllers are constructed in this article. One is a replacement-type of the conventional feedback controller (CFC) which is trained off-line in advance, and the other is a compensatory-type of the CFC under the operational mode of self-turning control. As a result of computer simulations, it is found that the proposed method has somewhat advantages over the conventional method.

Key words : neural network, neuro-controller, inverted pendulum, feedback control, self-turning

1 まえがき

制御工学の目的として、対象とするシステムに所望の動作をさせることが挙げられる。これを実現するためには、制御対象の数式モデル（物理的な動特性を微分方程式などで記述したもの）を求め、これを逆変換することで、所望の制御入力（操作量）を求める必要がある。ところが、実際の力学系の制御は、そのように単純なものではない。通常は外来信号（外乱）が入り込み、またそれが対象の動特性の変化をもたらす場合もあるため、徐々に適切な値に近づけていくフィードバック制御がよく用いられる。しかし一方では、制御対象の操作量が算出できてから修正動作を行うため、制御が後手に回りやすいなどの問題を含んでいる。そこで

本稿では、このような問題点を念頭において、ニューラルネットワークの適応能力に着目した制御手法について検討する。ここでは、非線形性を有する倒立振子の実機⁽¹⁾を用いた場合について焦点を絞り、提案法であるニューロ制御⁽²⁾⁻⁽⁴⁾の有効性について考える。

以下では、まず2.でニューラルネットワークを用いた倒立振子の制御モデルについて述べ、その後2つの観点からニューロ制御器の構築に関する実験を行う。すなわち、3.では置換型のニューロ制御器について、4.ではセルフチューニング方式⁽⁵⁾を志向した補償型ニューロ制御器についてである。そして5.でフィードバック誤差学習⁽⁶⁾との関連などについて考察し、最後に6.で結論を述べる。

2 倒立振子の制御モデル

2.1 制御系の基本構成

制御工学の分野において、制御系の解析・設計を行うときは、最も単純な線形フィードバック制御系を出

平成18年5月1日受理

* 大学院工学系研究科生体機能システム制御工学専攻

** 理工学部電気電子工学科

*** 文化教育学部

©佐賀大学理工学部

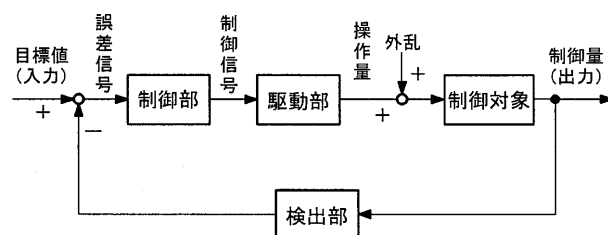


Fig.1 フィードバック制御系の基本構造

発点とすることが多い。Fig.1にフィードバック制御の基本構造を示す。まず目標となる量を目標値として与え、制御対象の出力を制御量とする場合を考える。このとき、制御量は入力側へフィードバックされて目標値と比較され、両者の間に差違があれば、それを誤差信号と定める。この誤差信号を基にして、調節部で適切な制御信号を作り出し、操作部で求められた操作量を制御対象に加えることになる。一般に、制御対象には未知の外乱が入り込むことが多く、仮に適切な操作量を加えたとしても、所望の制御量が得られるとは限らない。しかしフィードバック制御系では、常に制御量を検出して入力側へフィードバックし、目標値との差違がある限りは修正動作が繰り返されるため、上述のような外乱の影響が存在しても理想的には排除できると考えられる。

2.2 倒立振子の実験装置

本実験に用いる装置⁽¹⁾を Fig.2に示す。これは駆動部、制御対象、検出部、制御部という4つの部分に大別でき、その主な構成要素を Table 1にまとめる。

(a) 駆動部 :

サーボアンプとサーボモータ、プーリ、ワイヤー、リニアガイドから構成される。サーボアンプは、制御部であるコンピュータからの制御電圧を増幅し、サーボモータを駆動するための電流に変換するものである。サーボモータは、この電流の大きさに応じたトルクを発生させることで回転し、サーボモータの回転軸に取り付けられているプーリとワイヤーを利用して台車を水平方向に駆動させる。またサーボモータには、台車の位置を検出するためのオプティカルエンコーダ（後述）が取り付けられている。なおリニアガイドは、直線運動する台車を支持する機構である。

(b) 制御対象 :

台車と振り子から構成される。振り子は、重力を受けて倒れようとするのに対して、台車は左右に動くときの慣性力を利用して倒れないようにしている。なお、台車はヒノキの丸棒で作られた振り子を支持している。

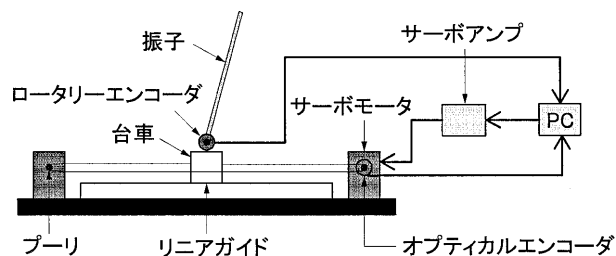


Fig.2 倒立振り子実験装置の構成

Table 1 倒立振り子実験装置を構成する各部の主な仕様

| 部位 | 構成要素 |
|------|------------------------------------|
| 駆動部 | サーボアンプ、サーボモータ、プーリ、ワイヤー、リニアガイド |
| 制御対象 | 振り子（角度制御）、台車（位置制御） |
| 検出部 | ロータリーエンコーダ（角度検出）、オプティカルエンコーダ（位置検出） |
| 制御部 | コンピュータ、入出力ボード |

(c) 検出部 :

ロータリーエンコーダとオプティカルエンコーダから構成される。両エンコーダから出力されるパルス信号をカウンターボードに取り込み、パルス数をカウントすることで、振り子の角度と台車の位置を検出している。

(d) 制御部 :

コンピュータと入出力ボードから構成される。入出力ボードは、センサからの角度と位置に関する情報をコンピュータに取り込むもので、コンピュータがこれらの情報に基づいてデジタル制御を行うために適切な制御電圧を作り出す。そして、入出力ボードを介してサーボアンプに送り出す。なお、制御プログラムを動作させるために必要な外部信号の取り込みも入出力ボードを通じて行われる。

2.3 ニューラルネットワークを用いた制御モデル

これまで、従来型のフィードバック制御の基本的な考え方、本実験で用いる倒立振り子の実機について述べた。次に、これをニューラルネットワークで制御する方法について説明する。駆動部と制御部をまとめた部分を「従来型フィードバック制御器 (CFC)」、それ以外の部分を実質的な「制御対象」とみなすと、Fig.1は両者の縦続接続によって構成されている。ここで、CFCと並列になるようにニューラルネットワークを配置し、必要に応じていくつかの信号経路を用意する。このよ

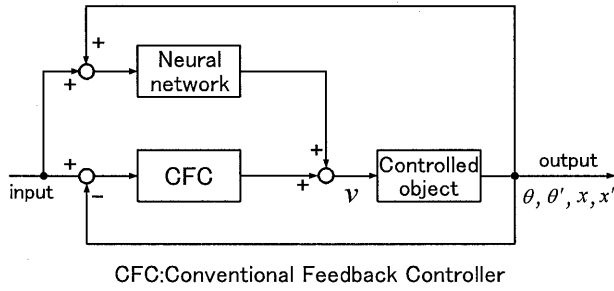


Fig.3 ニューラルネットワークを組み込んだ制御モデル

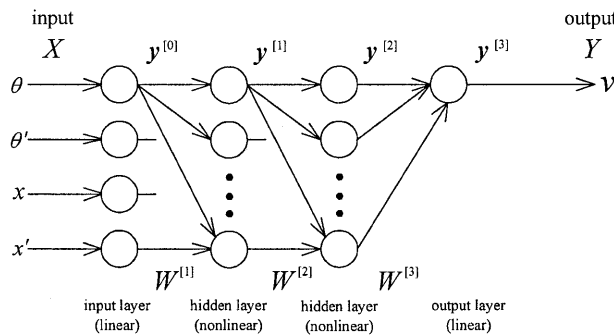


Fig.4 階層型ニューラルネットワーク

うにして構築した制御モデルをFig.3に示す。

また図中のニューラルネットワークは、Fig.4に示すように入力層のニューロンが4個、隠れ層のニューロンが各20個、出力層のニューロンが1個の階層構造型を用いている。ここで、各層の出力信号を $y^{[0]}$ 、 $y^{[1]}$ 、 $y^{[2]}$ 、 $y^{[3]}$ とおき、層間の結合荷重を $W^{[1]}$ 、 $W^{[2]}$ 、 $W^{[3]}$ として、その状態遷移則を次のように定義する。

$$y_i^{[0]} = X_i \quad (1)$$

$$y_j^{[1]} = f_1(u_j^{[y1]}) = f_1\left(\sum_i w_{ji}^{[1]} y_i^{[0]}\right) \quad (2)$$

$$y_k^{[2]} = f_2(u_k^{[y2]}) = f_2\left(\sum_j w_{kj}^{[2]} y_j^{[1]}\right) \quad (3)$$

$$Y = y^{[3]} = f_3(u^{[y3]}) = f_3\left(\sum_k w_k^{[3]} y_k^{[2]}\right) \quad (4)$$

$$f_1(z) = f_2(z) = \tanh\left(\frac{1}{2}z\right) \quad (5)$$

$$f_3(z) = z \quad (6)$$

ここで式(6)を見ると、第4層は線形ニューロン層を構成しているが、これは出力値制限を除去するためである。またこのとき、

$$X_1 = \theta, \quad X_2 = \theta', \quad X_3 = x, \quad X_4 = x' \quad (7)$$

$$Y = v \quad (8)$$

とすれば、このネットワークは入力信号 \mathbf{X} が与えられたときの出力信号 Y を求める4入力1出力の構造となる。学習時における結合荷重 $w_k^{[3]}$ の更新則は、出力層のニューロンにおける誤差を e 、学習速度係数 η を小さな定数とおき、誤差逆伝搬学習法(バックプロパゲーション)より、

$$\begin{aligned} \Delta w_k^{[3]} &\equiv -\eta \frac{\partial e}{\partial w_k^{[3]}} = -\eta \frac{\partial e}{\partial y^{[3]}} \frac{\partial y^{[3]}}{\partial u^{[y3]}} \frac{\partial u^{[y3]}}{\partial w_k^{[3]}} \\ &= \frac{\partial e}{\partial y^{[3]}} f_3'(u^{[y3]}) y_k^{[2]} \end{aligned} \quad (9)$$

で表される。

3 置換型ニューロ制御器を用いた倒立振子のオンライン制御

ここではFig.5に示すように、まず従来型フィードバック制御器(CFC)とニューラルネットワークを並列に配置した制御系を考え、CFCをニューラルネットワークで置換できるかについて検討する⁽²⁾。もし、これが実現可能であれば、ニューロ制御器の学習能力を活用することで、さらなる性能改善が期待できると考える。

3.1 実験方法

本実験では、置換型ニューロ制御器の有効性について検討するに当たり、次の3つの段階に分けて考える。

第1段階: 倒立振子の実機による学習用データの取得 (Fig.5(a)参照)

第2段階: 階層型ネットワークによる制御器特性のオフライン学習 (Fig.5(b)参照)

第3段階: 階層型ネットワークを組み込んだ実機のオンライン制御 (Fig.5(c)参照)

まず第1段階では、Fig.5(a)に示すようにして、実際に倒立振子を制御して振子の角度 θ とその変化分 θ' 、台車の位置 x とその変化分 x' のデータを取得する。このとき、台車を制御するモータの電圧 v も同時に求めておく。なお、サンプリング間隔は25msに設定している。また、今回は倒立振子を持続的に制御している区間のみに着目するため、振子を振り上げる区間については考慮しないものとする。

次に第2段階では、上で調べたデータを利用して、制御器が具備すべき入出力特性

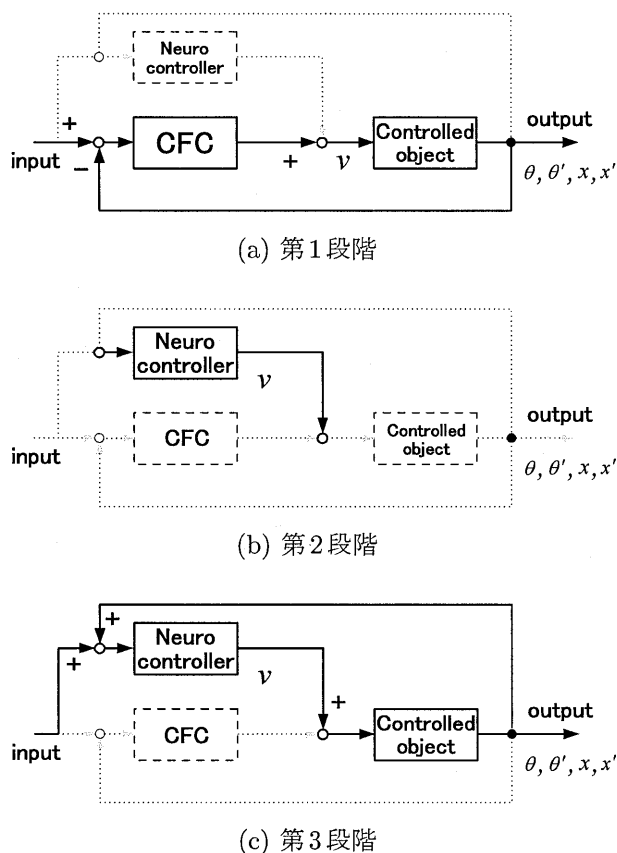


Fig.5 置換型ニューロ制御器を用いた倒立振子のオンライン制御実験

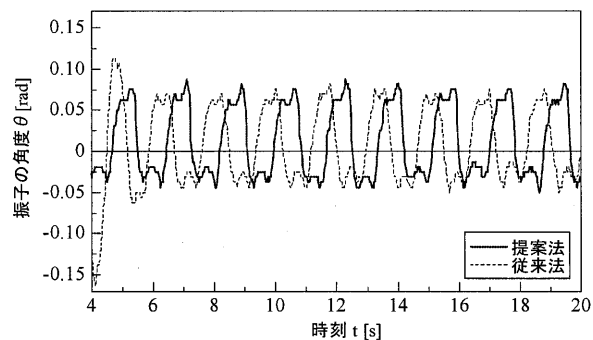
$$v = \mathcal{F}(\theta, \theta', x, x') \quad (10)$$

をニューラルネットワークに学習させる。時刻 t におけるネットワークの入力信号を $\mathbf{X}(t)$ 、出力信号を $Y(t)$ 、教師信号として $\{v(t) | (t = 1, 2, \dots, N)\}$ を与えると、誤差 e は、

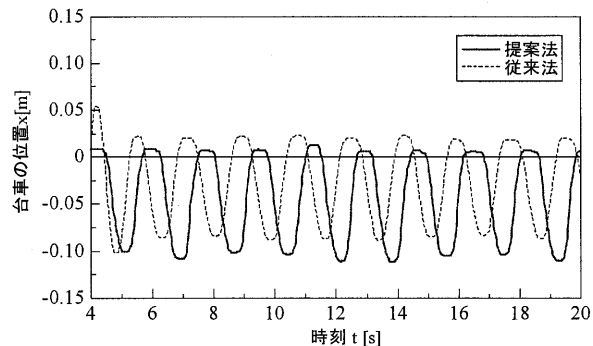
$$e = \frac{1}{2} \sum_t \{v(t) - y^{[3]}(t)\}^2 \quad (11)$$

となる。このとき、ニューラルネットワークは Fig.5(b) のように CFC から切り離れたオフライン状態とし、式 (11) に基づいて 30,000 回ほどの学習を行う。なおここでは、学習が容易となるように以下のような工夫を施している。

- 振子の角度 θ や台車の位置 x については、原点からの対称性を考慮して隠れ層の非線形変換関数を $(-1.0, 1.0)$ の範囲で変化する \tanh 関数へ変更している。
- 学習データのうち、 θ と x の絶対値が極めて小さいため、これを 10 倍に線形拡大することにより、



(a) 振子の角度 θ の時間変化



(b) 台車の位置 x の時間変化

Fig.6 置換型ニューロ制御器を用いた倒立振子のオンライン制御の実験結果

Table 2 従来法と提案法の制御能力の評価

| | 従来法 | 提案法 | 平均誤差 |
|----------------|-----------------------|-----------------------|-------|
| 振子の角度 θ | 7.00×10^{-2} | 7.25×10^{-2} | 3.57% |
| 台車の位置 x | 7.20×10^{-2} | 8.58×10^{-2} | 19.2% |

\tanh 関数の出力区間 $(-1.0, 1.0)$ に適切に分布するようにしている。

最後の第3段階では、制御系にニューラルネットワークを組み込んで実際に倒立振子のオンライン制御を行い、その制御能力を CFC の場合と比較する。Fig.5(c) に示すように、このときはフィードバック制御の要である CFC をニューラルネットワークで置換した構成 (実線部) となっている。

3.2 実験結果

Fig.6 は、第3段階における倒立振子の制御特性を表したもので、(a) が振子の角度 θ 、(b) が台車の位置 x について表したものである。これを見ると、従来法と提案法では振動の振幅がほぼ等しいものの、位相にずれのあることがわかる。これは、それぞれの実験において、時間軸の原点を振子の振り上げ開始時刻に設定し

たためであり、基準からのずれの存在は本質的な問題ではない。よって以下では、位相差よりも振幅に重点をおいた解析を試みる。

次に、ここで提案する置換型ニューロ制御と従来型フィードバック制御器の性能を比較するため、振動1周期あたりの誤差（基準からのずれ）に着目する。具体的には、振子の角度 θ であれば0radからのずれを表す領域面積であり、台車の位置 x であれば0mからのずれを表す領域面積である。ここでは、比較的安定している4周期を選び、そこから1周期あたりの平均誤差（平均面積）を求めることにした。Table 2は、従来法と提案法の比較結果である。なお表中の平均誤差とは、従来法に対する両手法の差の割合を示している。これによると、振子の角度 θ については従来法の1周期あたりの平均面積が 7.00×10^{-2} であったのに対して、提案法では 7.25×10^{-2} となり、約3.57%の差が認められた。また同様にして、台車の位置 x については約19.2%の差が生じた。台車の位置 x の誤差は角度 θ の誤差よりも明らかに大きくなっている。しかし、ここでの目標が倒立振子を制御することにあると考えると、上述のように θ の制御能力が良好であれば、他のパラメータについては多少の能力低下は許容するとの立場も容認できるであろう。またこのあとに、さらに学習を行うことで制御能力の改善も見込めるであろう。以上のことから、本実験では従来型フィードバック制御器をニューロ制御器で置換しても、ほぼ同程度の制御性能を示すことができたと判断する。

4 補償型ニューロ制御器を用いたセルフチューニング方式による倒立振子の制御

ここでは、3.とは異なる観点から倒立振子の制御実験を試みる。すなわち、従来型フィードバック制御器を基本に据え、これにニューラルネットワークを用いて補正を行うというセルフチューニング型のオンライン学習制御手法の確立を目指す⁽³⁾⁽⁴⁾。

4.1 実験方法

上述のように、制御系にニューラルネットワークの適応能力を援用すると、ノイズに強くかつ汎化性に豊かな非線形システムのモデル化が期待できる。このときの各信号の流れをFig.7に示す。基本的な構成はFig.5と同一であるが、CFCの出力信号（誤差を減少させるための制御信号）が零となるようにニューラルネットワークを誤差逆伝搬法で学習する点が異なっている。ここでは、この補償型ニューロ制御器の初期結合荷重として以下の2とおりを考える。

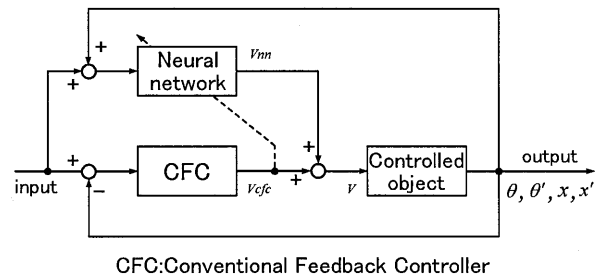


Fig.7 補償型ニューロ制御器を用いたセルフチューニング方式による倒立振子の制御実験

- i) 制御器特性を事前にオフライン学習したもの
- ii) 未学習のもの

前者の制御器特性を事前にオフライン学習したものは、3.で獲得した結合荷重を初期値として用いる。これに対して、後者の未学習のものは $[-1.0, 1.0]$ の範囲で発生した乱数を結合荷重の初期値として用いる。

このようにして各結合荷重の初期値を決定し、セルフチューニング方式を用いた場合の実際の倒立振子の制御能力について調べてみる。なお、サンプリング間隔は15msに設定してデータの取得 $(\theta, \theta', x, x', v)$ と学習を繰り返し、振子の角度 θ と台車の位置 x が収束するまで継続する。

4.2 実験結果

Fig.8は、初期値に制御器特性のオフライン学習時のデータを用いた場合の一例で、倒立振子の制御特性を振子の角度 θ 、台車の位置 x について表したものである。ここで横軸は学習回数（時刻）、縦軸はそれぞれ振子の角度 θ 、台車の位置 x を表している。Fig.8を観察すると、主観的ではあるが学習回数に応じて補償型ニューロ制御器の状態をいくつかに分類できそうである。これをまとめたものがTable 3である。これによると、学習回数が0回から20,000回付近までは、初期の過渡的な部分を除くと学習しているものの結果として表れていない状態と考えられる。次の20,000回から28,000回付近で、ようやく学習の効果が表れ始めて少しずつ振動の振幅が減少してくる。そして、29,000回付近で振動の振幅が最小となり、その後、増加に転じて不安定な振動現象が出現した。同様の実験を数回試みたものの、どの試行も類似した傾向を示したため、この振動現象の出現には再現性があるものと考えられる。

次に、初期値に乱数を用いた場合をFig.9に示す。Fig.8と同様の傾向を示すが、横軸のスケールが大きく異なっている。具体的には収束に要する学習回数が

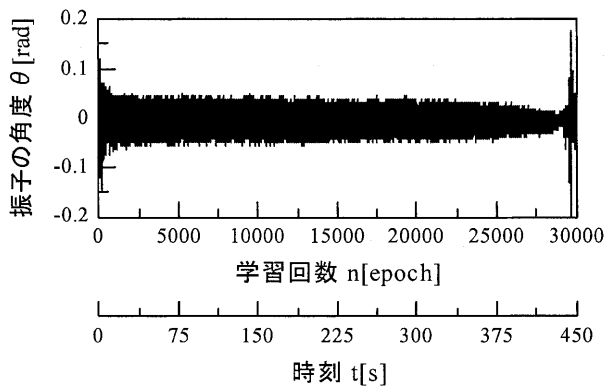
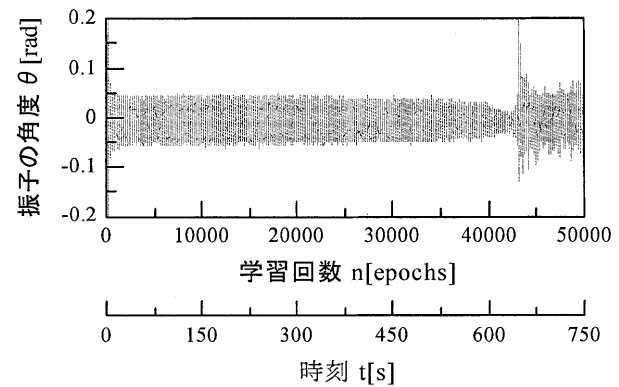
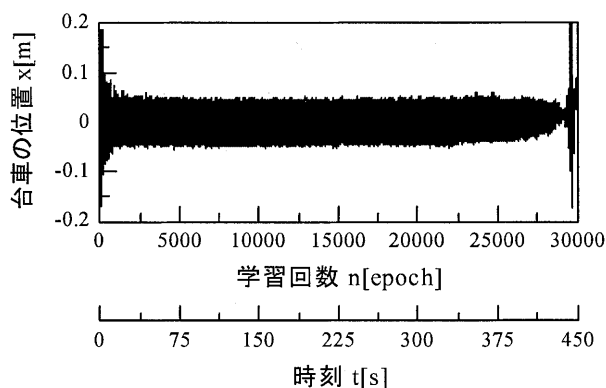
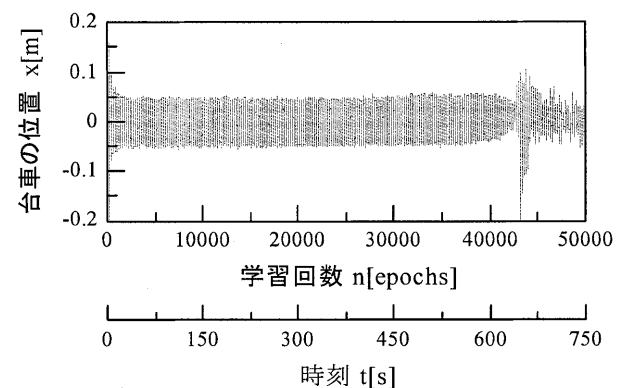
(a) 振子の角度 θ の時間変化(a) 振子の角度 θ の時間変化(b) 台車の位置 x の時間変化(b) 台車の位置 x の時間変化

Fig.8 セルフチューニング方式による制御結果(1) — 初期結合荷重を事前にオフライン学習したもの —

Fig.9 セルフチューニング方式による制御結果(2) — 初期結合荷重が未学習のもの —

Table 3 学習回数による振動の振幅の変化

| 学習回数 | 振動の振幅 |
|---------------|-------|
| 0~3,000 | 減少 |
| 3,000~20,000 | 一定 |
| 20,000~28,000 | 減少 |
| 28,000~29,000 | 収束 |
| 29,000~ | 振動 |

42,000回付近となり、事前にオフライン学習したものより14,000回ほど遅くなることがわかる。なお、この学習回数の大きな差は事前に制御器特性をオフライン学習した効果と考えられるが、実験環境の気温の高低に応じて大幅に変動するという実験結果も得られつつあり、詳細についてはさらなる検討が必要である。

ここで、CFCと補償型ニューロ制御器がサーボモータを駆動するために出力する信号(電圧)について調べてみよう。Fig.10とFig.11はCFCの出力 v_{cfc} とニューロ制御器の出力 v_{nn} の時間変化を表したものである。こ

れらは全体像を把握するためのものであり、これを拡大して時々刻々と変化していく波形 v_{cfc} と v_{nn} を再描画したものが、Fig.12とFig.13である。これらは一例であるが、本稿で重要となる3つの区間について取り上げている。横軸は学習回数、縦軸は台車を制御するモータの電圧 $v(=v_{cfc}+v_{nn})$ を個別に表したものである。初期結合荷重にCFCの特性を事前にオフライン学習したものを用いた場合、500回付近では同相、その後無関係となり、27,000回付近では逆相となっている。同様の実験を数回試みたものの、いずれの試行も類似した傾向を示したため、この現象の出現には再現性があるものと考えられる。

また、初期結合荷重を乱数とした場合も、同様の傾向(同相→無関係→逆相)を示していることが確認できた。ただし、相違点はそれぞれが出現するまでに要する学習回数であり、例えば逆相となるのは40,000回付近となって、事前に学習したものより13,000回程度遅れることを確認した。なお、このような現象の出現はセルフチューニング方式を採用したオンライン学習

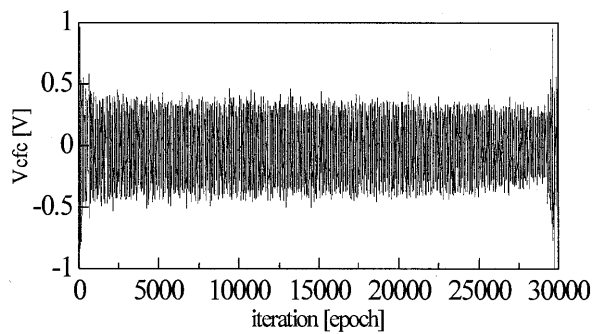
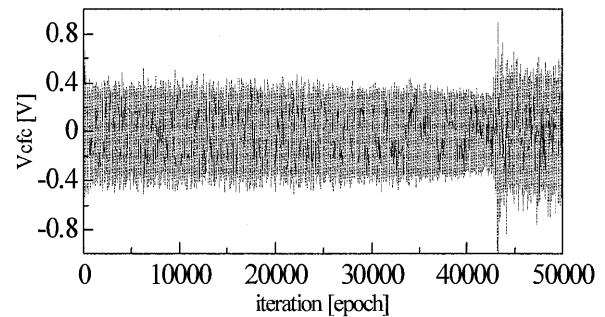
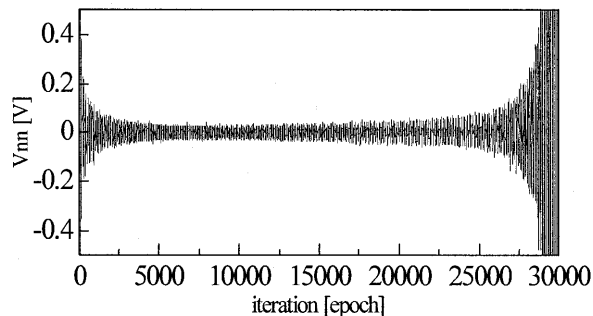
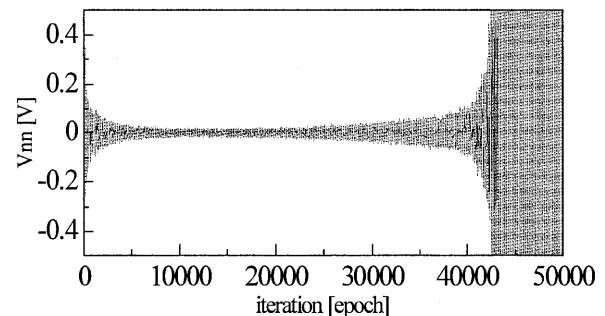
(a) CFCの出力 v_{cfc} の時間変化(a) CFCの出力 v_{cfc} の時間変化(b) ニューロ制御器の出力 v_{nn} の時間変化(b) ニューロ制御器の出力 v_{nn} の時間変化

Fig.10 セルフチューニング方式を採用したときの制御器出力(1) —初期結合荷重を事前にオフライン学習したもの—

Fig.11 セルフチューニング方式を採用したときの制御器出力(2) —初期結合荷重に乱数を用いたもの—

時の補償型ニューロ制御器とCFCの関係に起因する問題ではないかと考えている。

5 考察

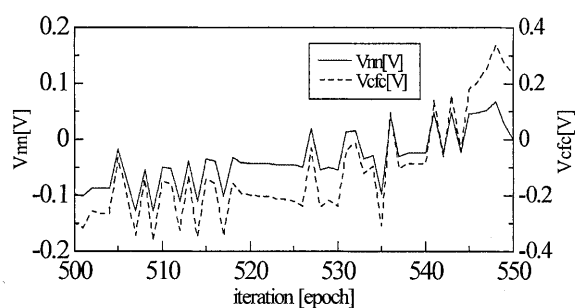
本研究では、ニューラルネットワークを用いた倒立振子の制御モデルに関する検討をFig.3に示すような構成で行ったが、この分野の研究では川人らによるフィードバック誤差学習という手法 (Fig.14) が有名である⁽⁶⁾。これは学習の初期段階ではフィードバック制御が働くものの、制御器の出力 (誤差信号) が小さくなるように並列配置したニューラルネットワークの学習を行うことで、学習の進行とともに制御方式の主体がフィードフォワード方式に移行していくというものである。これは生体のモデルともなっており、我々が新しい動作をマスターするとき、最初は身体の動作に注意しながら、しかし、一度身に付けてしまえば特に意識することなく、素早く実行できることを模倣したものになっている。このうち、従来型フィードバック制御器をニューロ制御器で置換しようという発想は、3.で行った実験と共通である。また、ニューラルネットワークの学習

法 (特に教師信号の設定法) については、4.で行った実験と共通である。ここでは、純粋に工学的な観点からフィードバック制御方式の方が安定した動作が可能ではないかとの方針に基づき、このような構成を採用したが、必ずしもこれに固執するものではない。

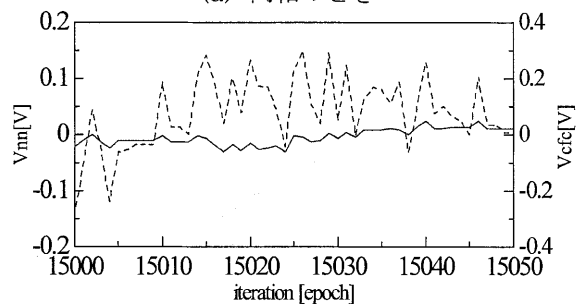
ところで、今回は倒立振子の制御に当たり、目標値として振子の角度 $\theta=0$ [rad]、台車の位置 $x=0$ [m]とした。このことは、Fig.3においてそもそも入力信号が存在しないことを意味し、この信号経路の重要性を検証できていないことを意味する。したがって、この点については倒立振子実験装置を傾けたりすることで $\theta \neq 0$ [rad]としたり、台車をリニアガイド上の一方に偏らせて $x \neq 0$ [m]として実験を行うことも必要と考える。なお、角度 θ の時間変化分 θ' と位置 x の時間変化分 x' については、倒立振子の制御という性格上、0以外の値にすることは厳しいようである。

6 むすび

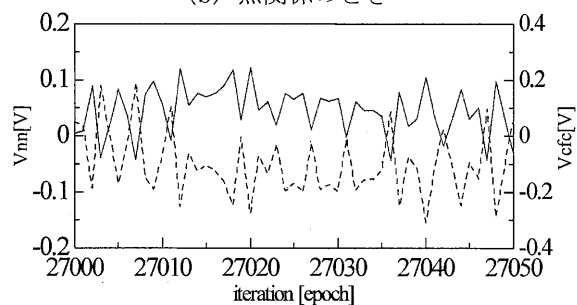
本稿では、非線形性を有する倒立振子の実機を取り上げ、ニューラルネットワークを用いた制御モデルに



(a) 同相のとき



(b) 無関係のとき



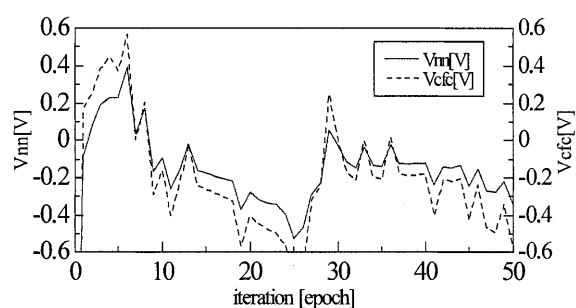
(c) 逆相のとき

Fig.12 セルフチューニング方式を採用したときのオンライン学習時の制御器出力(1) —初期結合荷重を事前にオフライン学習したもの—

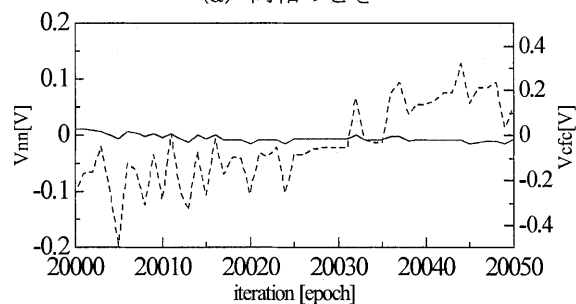
ついて検討した。まず、従来型フィードバック制御器(CFC)の特性を事前にオフラインで学習し、その後、制御系に組み込んでオンライン制御を行ったところ、置換型ニューロ制御器はCFCと同等の能力を示すことを確認した。次に、セルフチューニング方式を採用しオンラインで学習しながら制御を行ったところ、この補償型ニューロ制御器の制御能力向上の可能性が認められた。ただし現段階では、この能力向上は短時間だけとなっており、さらに不安定な振動現象が出現するという問題点も残っている。今後は、このような現象の出現メカニズムの解明とともに、さらなる制御能力の向上について検討する予定である。

謝辞

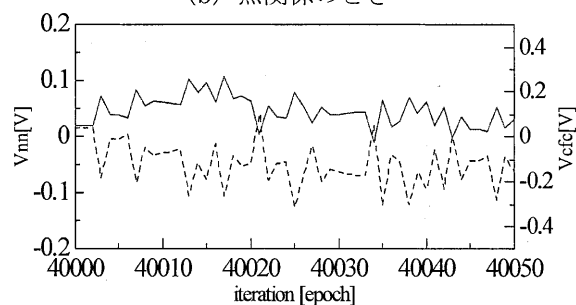
本研究の一部は、平成17年度佐賀大学学長経費(中期計画実行経費)「人工頭脳構築のための生体における



(a) 同相のとき



(b) 無関係のとき



(c) 逆相のとき

Fig.13 セルフチューニング方式を採用したときのオンライン学習時の制御器出力(2) —初期結合荷重が未学習のもの—

情報処理様式の解明とその応用」(研究代表者:野口義夫教授)の補助を受けた。記して感謝します。

参考文献

- [1] 陣内健一: “二重倒立振子に関する研究”, 平成13年度佐賀大学大学院教育学研究科教科教育専攻技術教育専修修士学位論文, 2002
- [2] 山海陽一朗, 和久屋 寛, 森田 譲: “ニューロコントローラを用いた倒立振子の制御”, 第1回人工頭脳工学シンポジウム講演論文集, O-3, pp.22-23, 2005
- [3] 山海陽一朗, 和久屋 寛, 森田 譲: “セルフチューニングを用いた適応的な手法による倒立振子の制御”, 平成17年度(第58回)電気関係学会九州支部連合大会講演論文集, 13-1A-07, p.522, 2005
- [4] 山海陽一朗, 和久屋 寛, 森田 譲: “セルフチューニング型オンライン学習制御方式における指令信号波形の反

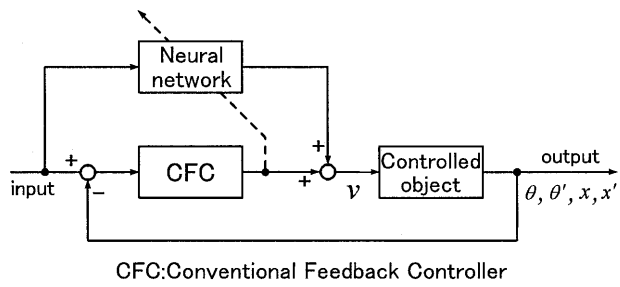


Fig.14 フィードバック誤差学習

転現象”，2006年総合大会講演論文集（情報・システム講演論文集1），D-2-9, p.12, 2006

- [5] 大松 繁, 山本 透 編：“セルフチューニングコントロール”，コロナ社, 1996
- [6] 川人光男：“脳の計算理論”，産業図書, 1998