

(19)日本国特許庁 (J P)

(12) 特 許 公 報 (B 1)

(11)特許番号

特許第2996952号
(P2996952)

(45)発行日 平成12年1月11日(2000.1.11)

(24)登録日 平成11年10月29日(1999.10.29)

(51)Int.Cl.⁷

識別記号

F I

G 0 5 B 13/02

G 0 5 B 13/02

L

請求項の数9(全10頁)

(21)出願番号	特願平10-203891	(73)特許権者	592179296 株式会社エイ・ティ・アール人間情報通信研究所 京都府相楽郡精華町大字乾谷小字三平谷5番地
(22)出願日	平成10年7月17日(1998.7.17)	(72)発明者	川人 光男 京都府相楽郡精華町大字乾谷小字三平谷5番地 株式会社エイ・ティ・アール人間情報通信研究所内
審査請求日	平成10年7月17日(1998.7.17)	(72)発明者	ダニエル ウォルバート 京都府相楽郡精華町大字乾谷小字三平谷5番地 株式会社エイ・ティ・アール人間情報通信研究所内
		(74)代理人	100064746 弁理士 深見 久郎 (外2名)
		審査官	仁科 雅弘

最終頁に続く

(54)【発明の名称】 学習制御方法および学習制御装置

3

(57)【特許請求の範囲】

【請求項1】 非線形性や非正常性などを有する制御対象を目標とする時間軌道に沿って制御するための学習制御方法であって、

運動指令が与えられたとき前記制御対象の状態を予測する順モデルと、前記目標とする時間軌道が与えられたとき制御対象への入力信号となる運動指令を推定する逆モデルが対として設けられた複数の制御モジュールを有し、

前記複数の制御モジュールの各順モデルによって推定された状態予測と、前記制御対象からの実際の状態とを比較して各々の順モデルの予測誤差を計算し、その予測誤差に基づいて前記複数の制御モジュールのうち予測の良好な順モデルを含む制御モジュールにより大きな責任信号を与え、この責任信号で重み付けして前記複数の逆モ

4

デルからの出力を線形加算した前向き運動信号とフィードバック運動指令信号とを加算した制御信号を出力し、前記責任信号は、前記各制御モジュールに与えられる0から1の間の数であって、前記各制御モジュールに与えられる数の和が1になるように正規化されていることを特徴とする、学習制御方法。

【請求項2】 非線形性や非正常性などを有する制御対象を目標とする時間軌道に沿って制御するための学習制御装置であって、

10 それぞれが運動指令の与えられたとき前記制御対象の状態を予測する順モデルと、目標軌道が与えられたとき前記制御対象への入力信号となる運動指令を推定する逆モデルとを含み、前記順モデルによって推定された状態予測と前記制御対象からの実際の状態とを比較して前記順モデルの予測誤差を計算する複数の制御モジュール、

前記各制御モジュールから出力された予測誤差に基づいて、前記複数の制御モジュールのうちの予測の良好な順モデルを含む制御モジュールにより大きな責任信号を与える責任信号推定手段、および前記責任信号推定手段からの責任信号で重み付けして、前記複数の逆モデルからの出力信号を線形加算した前向き運動信号とフィードバック運動指令信号とを加算した制御信号を前記制御対象に与える加算手段を備え、

前記責任信号は、前記各制御モジュールに与えられる 0 から 1 の間の数であって、前記各制御モジュールに与えられる数の和が 1 になるように正規化されていることを特徴とする、学習制御装置。

【請求項 3】 前記制御モジュールは、制御を開始する前にいずれかの制御モジュールが選択されるように前記責任信号推定手段からの責任信号を教師信号として学習する責任信号予測手段を含む、請求項 2 に記載の学習制御装置。

【請求項 4】 前記順モデルは、前記加算手段から出力された制御信号とフィードバック運動指令信号との和を入力として受ける入力部と、前記制御対象の状態を推定する出力部とを有する神経回路を含む、請求項 2 に記載の学習制御装置。

【請求項 5】 前記神経回路は、前記制御対象の実際の状態から前記制御対象の状態の推定を差し引いた信号と、前記神経回路が含まれるモジュール単位の責任信号との積を教師有り学習則の誤差信号としてパラメータを学習することを特徴とする、請求項 4 に記載の学習制御装置。

【請求項 6】 前記逆モデルは、前記目標とする時間軌道を入力として受ける入力部と、前記制御対象への入力信号となる運動指令を推定する出力部とを有する神経回路を含むことを特徴とする、請求項 2 に記載の学習制御装置。

【請求項 7】 前記神経回路は、フィードバック誤差学習、順逆モデリング、直接逆モデリングなど所定の学習制御教師有り学習則から得られる誤差信号と、前記逆モデルが含まれるモジュールの責任信号の積を誤差信号としてパラメータを学習することを特徴とする、請求項 6 に記載の学習制御装置。

【請求項 8】 前記責任信号推定手段は、前記各制御モジュール内の順モデルが出力する状態の予測値と前記制御対象の実際の状態との誤差を計算する入力部と、前記責任信号予測手段の予測と前記入力部の誤差の積を計算する中間部と、前記制御モジュールについて大きな誤差には小さな責任信号を単調減少飽和型関数で計算する出力部とを含む、請求項 3 に記載の学習制御装置。

【請求項 9】 前記責任信号予測手段は、前記責任信号推定手段によって計算される責任信号を教師信号として所定の教師有り学習則によりパラメータを学習することを特徴とする、請求項 3 に記載の学習制御装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は学習制御方法および学習制御装置に関し、特に、非線形性や非正常性などを有する制御対象を目標とする時間軌道に沿って制御する学習制御方法および学習制御装置であって、エレベータやロボットや自動車エンジンやブレーキなどの制御に適用される。

【0002】

10 【従来の技術】従来の制御のほとんどは線形のフィードバック制御で行なわれているが、制御対象の特性に強い非線形性があったり、特性が時間的に急激に変動する場合には目標とする軌道に高い精度で追従することはできない。

【0003】

20 【ロボット制御の分野では、適応制御と呼ばれる一種の学習制御則が提案されてきた。これはロボットの逆ダイナミクスをラグランジュ方程式などで記述しておき、入出力データに基づいて未知のパラメータを推定する方法である。しかし、この方法は制御対象のダイナミクスがかなり正確に解析式で表現されていることを前提としているなどの理由で実用化に至っていない。

【0004】

一方、ニューラルネットワークの分野では、多層前向き神経回路やRadial basis function network などと学習則を組合せて、制御対象の逆ダイナミクスモデルを獲得する試みがなされてきた。そのような例としては、たとえば特開平 2 - 5 4 3 0 4 号公報や、特開平 4 - 1 0 2 1 0 2 号公報や、特開平 6 - 1 9 5 0 8 号公報において知られている。

【0005】

30 【発明が解決しようとする課題】しかし、非常に強い非線形性や非正常性がある場合に 1 個の神経回路を獲得することは非常に困難である。さらに、ニューラルネットワークの分野では、複数個のモジュールをgating-expert アーキテクチャという方式で同時に学習し、切り換える方式がたとえば R. A. Jacobs, et al: Adaptive mixtures of local experts. Neural Computation, Vol. 3, 79-78 (1990) で提案されている。これを学習制御に応用する試みもさなれている。しかし、このアーキテクチャでは、多数のモジュールを同時並列に学習し、かつ安定にスイッチングを起こすことが大変困難である。その理由は、主に以下の 2 つのためである。

【0006】

50 第 1 に、多数のモジュールのうち、どのモジュールがある時点で最も環境や動作点など制御対象の特性をよく再現しているのかを調べるためには、実際にあるモジュールを使って制御し、その制御性能を異なる時点で比較するしか方法がなかった。これはある時点では 1 つのモジュールの性能しか評価できないことを意味しているので、多数のモジュールのどれが多数の環境のうちのどれによく適応しているかを推定するには、モジュール数の指数関数的な時間が必要となって、モジュー

ルのスイッチングや学習が極端に困難になるという結果につながる。

【0007】第2にモジュールのスイッチングはgatingネットワークと呼ばれる単一の神経回路が行なうが、これは大変困難な課題となっている。なぜなら、少し複雑な問題に対して、スイッチングを正しく生じさせるためには、非常に多次元の空間を複雑で非線形な多数の境界で分割しなければならないが、これは神経回路の学習をもってしても困難である。

【0008】Narendraらは、予測器と制御器の対を複数用意しておき、最も予測器の性能の高い制御器を1つ選んで制御するスイッチング制御方式を提案している(K. Narendra, et al.: Adaptation and learning using multiple models, switching, and tuning. IEEE Control Systems Magazine, June, 37-51 (1995))。しかし、この方式では多数のモジュールの出力の組合せを臨機応変に選択して新しい環境に対応することが全くできない。また、制御器の学習を行なうこともできず、さらに制御を行なうまでには予めスイッチングを起こすことができない。

【0009】それゆえに、この発明の主たる目的は、制御対象の軌道追従制御を高い性能でかつ学習で行なうことができるような学習制御方法および学習制御装置を提供することである。

【0010】

【課題を解決するための手段】請求項1に係る発明は、非線形性や非正常性を有する制御対象を目標とする時間軌道に沿って制御するための学習制御方法であって、運動指令が与えられたとき制御対象の状態を予測する順モデルと、目標とする時間軌道が与えられたとき制御対象への入力信号となる運動指令を推定する逆モデルが対として設けられた複数の制御モジュールを有し、複数の制御モジュールの各モデルによって推定された状態予測と制御対象からの実際の状態とを比較して各々の順モデルの予測誤差を計算し、その予測誤差に基づいて複数の制御モジュールのうち予測の良好な順モデルを含む制御モジュールにより大きな責任信号を与え、この責任信号で重み付けして複数の逆モデルからの出力を線形加算した前向き運動信号とフィードバック運動指令信号とを加算して制御信号を出力し、責任信号は各制御モジュールに与えられる0から1の間の数字であって、各制御モジュールに与えられる数の和が1になるように正規化される。

【0011】請求項2に係る発明は、非線形性や非正常性を有する制御対象を目標とする時間軌道に沿って制御するための学習制御装置であって、それぞれが運動指令の与えられたとき制御対象の状態を予測する順モデルと、目標軌道が与えられたとき制御対象への入力信号となる運動指令を推定する逆モデルとを含み、順モデルによって推定された状態予測と制御対象からの実際の状

態とを比較して順モデルの予測誤差を計算する複数の制御モジュールと、各制御モジュールから出力された予測誤差に基づいて、複数の制御モジュールのうちの予測の良好な順モデルを含む制御モジュールにより大きな責任信号を与える責任信号推定手段と、責任信号推定手段からの責任信号で重み付けして、複数の逆モデルからの出力信号を線形加算した前向き運動信号とフィードバック運動指令信号とを加算した制御信号を制御対象に与える加算手段を備え、責任信号は各制御モジュールに与えられる0から1の間の数字であって、各制御モジュールに与えられる数の和が1になるように正規化される。

【0012】

【0013】請求項3に係る発明では、請求項2の制御モジュールは、制御を開始する前にいずれかの制御モジュールが選択されるように責任信号推定手段からの責任信号を教師信号として学習する責任信号予測手段を含む。

【0014】請求項4に係る発明では、請求項2の順モデルは、最終的な制御信号を入力として受ける入力部と、制御対象の状態を推定する出力部とを有する神経回路を含む。

【0015】請求項5に係る発明では、請求項4の神経回路は、制御対象の実際の状態から制御対象の状態の推定を差し引いた信号と、神経回路が含まれるモジュール単位の責任信号との積を教師有り学習則の誤差信号としてパラメータを学習する。

【0016】請求項6に係る発明では、請求項2の逆モデルは、目標とする時間軌道を入力として受ける入力部と、制御対象への入力信号となる運動指令を推定する出力部とを有する神経回路を含む。

【0017】請求項7に係る発明では、請求項6の神経回路はフィードバック誤差学習、順逆モデリング、直接逆モデリングなど所定の学習制御教師有り学習則から得られる誤差信号と、逆モデルが含まれるモジュールの責任信号の積を誤差信号としてパラメータを学習する。

【0018】請求項8に係る発明では、請求項3の責任信号推定手段は、各制御モジュール内の順モデルが出力する状態の予測値と制御対象の実際の状態の誤差を計算する入力部と、責任信号予測手段の予測と入力層の誤差の積を計算する中間部と、制御モジュールについて大きな誤差に小さな責任信号を単調減少飽和型関数で計算する出力部とを含む。

【0019】請求項9に係る発明では、請求項3の責任信号予測手段は、責任信号推定手段によって計算される責任信号を教師信号として所定の教師有り学習則によりパラメータを学習する。

【0020】

【発明の実施の形態】図1はこの発明の一実施形態の全体の構成を示すブロック図である。図1において、順モデル2と逆モデル3と責任信号予測器4とを含む複数の

制御モジュール 1 1 , 1 2 ... 1 n が設けられる。各制御モジュール 1 1 , 1 2 ... 1 n には、文脈情報信号と運動指令信号と腕の目標軌道信号とが与えられる。ここで、文脈情報信号は、運動指令や制御対象の状態以外の信号で制御対象の動作点や環境、その種類などに関して得られる信号である。たとえば、後述の図 6 の物体操作の例では物体の視覚情報となる。

【 0 0 2 1 】 順モデル 2 と逆モデル 3 とは対になっており、各順モデル 2 は並列的に状態を予測して予測信号を減算器 2 1 , 2 2 ... 2 n に出力する。各減算器 2 1 , 2 2 ... 2 n には制御対象 9 としての腕の動きから検出された実現された腕軌道の信号 x_t が与えられており、各減算器 2 1 , 2 2 ... 2 n はこれらの信号と予測信号とを減算することによって、予測誤差信号を責任信号推定器 5 に出力する。各制御モジュール 1 1 , 1 2 ... 1 n の責任信号予測器 4 は、文脈情報信号に基づいて制御開始時に適当な制御モジュールが選択されるように学習を行なう。

【 0 0 2 2 】 責任信号推定器 5 は、各順モデル 2 が運動指令から制御対象 9 の状態を推定する誤差に基づいて予測の良好な順モデルを含む単位により大きな責任信号を与えるものである。この責任信号は、対応の逆モデル 3 からの運動指令信号と乗算器 6 1 によって乗算され、各乗算出力が加算器 7 によって加算される。そして、その総和が加算器 8 によってフィードバック制御器 3 0 からのフィードバック運動指令 u_{fb} と加算されて最終的な運動指令として制御対象 9 に与えられる。最終的な運動指令は、さらに順モデル 2 の入力層 2 0 1 に与えられる。制御対象 9 は上述の腕のみならず手や物体などであってもよい。

【 0 0 2 3 】 図 2 は順モデル 2 の動作を説明するために、図 1 の順モデル 2 に関連する部分を抽出して示した図である。図 2 において、順モデル 2 は運動軌道が与えられたとき、予測軌道を計算する。これに対して、逆モデル 3 は目標軌道が与えられたとき、それを実現する運動指令を計算する。これらのモデルは、いずれも少なくとも入力層と出力層とを含むニューラルネットワークによって構成される。順モデル 2 では、一般的にニューラルネットワークの入力層 2 0 1 に入力信号としての運動指令信号 u_t が与えられると、出力層 2 0 2 から制御対象状態を特定するための予測軌道 \hat{x}_{t+1}^i が出力される。

【 0 0 2 4 】 図 2 に示した例では、現在の状態信号 x_t と運動指令信号 u_t とが順モデル 2 の入力層 2 0 1 に与えられる。制御モジュール 1 1 , 1 2 ... 1 n のうちの i *

$$\Delta w_t^i = \varepsilon \lambda_t^i \frac{d\phi_i}{dw_t^i}(x_t - \hat{x}_t^i) = \varepsilon \frac{d\hat{x}_t^i}{dw_t^i} \lambda_t^i (x_t - \hat{x}_t^i) \quad (3)$$

【 0 0 3 1 】 図 3 は逆モデル 3 の動作を説明するために、図 1 の逆モデル 3 に関連する部分を抽出して示した

* 番目の順モデル 2 は時刻 t での次の状態の予測 \hat{x}_{t+1}^i を第 (1) 式より計算する。

【 0 0 2 5 】

【 数 1 】

$$\hat{x}_{t+1}^i = \phi(w_t^i, x_t, u_t) \quad (1)$$

【 0 0 2 6 】 ここで、 w_t^i は順ダイナミクスをモデル化するために用いられる関数近似器中のパラメータであり、より具体的には、ニューラルネットワークのシナプス荷重である。

【 0 0 2 7 】 これらの次の状態の予測 \hat{x}_{t+1}^i は、実際の次の状態と減算器 2 0 3 で比較され、各順モデル 2 がどの程度現在の制御対象の振る舞いを説明しているかを表わす責任信号を計算するのに使われる。すなわち、順モデル 2 の出力層 2 0 2 は制御対象の実際の状態 x_t から遅延回路 (DL) 2 0 4 の出力である制御対象の推定を差し引いた信号 $x - \hat{x}_t$ と、そのニューラルネットワークが含まれるモジュール単位の責任信号 λ_t^i との積を所定の教師有り学習則の誤差信号としてパラメータを学習する。ここで、 i 番目の責任信号は、責任信号推定器 5 によって次の第 (2) 式に示すソフトマックス関数で計算される。

【 0 0 2 8 】

【 数 2 】

$$\lambda_t^i = \frac{e^{-|x_t - \hat{x}_t^i|^2 / 2\sigma^2}}{\sum_{j=1}^n e^{-|x_t - \hat{x}_t^j|^2 / 2\sigma^2}} \quad (2)$$

【 0 0 2 9 】 ここで、 x_t は真の状態、 λ_t^i はスケール定数である。ソフトマックスは、指数関数と正規化によって責任信号が「 0 」と「 1 」の間に入って、かつすべての単位モジュールで足し合わせると 1 になるようにしている。現在の振る舞いをよく再現し、したがって予測誤差の小さな単位が大きな責任信号を持つことになる。責任信号は図 1 と次の第 (3) 式に示すように、順モデル 2 の学習を競合的に制御するためにも使われる。つまり、責任信号の大きい単位モジュールはそれに比例して誤差信号を多く受取る。したがって、責任信号の値が「 1 」に近い程その制御モジュールは責任が大きいことになり、制御にもそれだけ大きく貢献し、順モデルと逆モデルの学習もそれに比例して速く進むことになる。

【 0 0 3 0 】

【 数 3 】

図である。図 3 において、逆モデル 3 は、前述のごとく目標とする時間軌道を入力信号として受ける入力層 3 0

1 と、制御対象への入力信号となる運動指令を特定する出力層 3 0 2 とを含むニューラルネットワークで構成されている。

【 0 0 3 2 】より具体的に説明すると、逆モデル 3 の入力層 3 0 1 には、次の目標状態 x_{t+1}^* と現在の状態信号 x_t とが入力され、第 i 番目の逆モデル 3 は運動指令信号 u_{t+1}^i を次の第 (4) 式のように計算する。な *

$$u_{t+1}^i = \Psi(\alpha_t^i, x_{t+1}^*, x_t) \tag{4}$$

【 0 0 3 4 】ここで、 Ψ はある関数近似器のパラメータ、具体的にはニューラルネットワークのシナプス荷重である。最終的な運動指令は、加算器 7 によって第 (5) 式に示すように n 個の逆モデル 3 の出力の責任信

$$u_t = \sum_{i=1}^n \lambda_t^i u_{t+1}^i = \sum_{i=1}^n \lambda_t^i \Psi(\alpha_t^i, x_{t+1}^*, x_t) \tag{5}$$

【 0 0 3 6 】図 3 において、責任信号 λ_t^i は乗算器 3 0 3 によって運動指令 u_{t+1}^i と乗算される。同様に責任信号 λ_t^i は乗算器 3 0 4 によって運動指令の誤差 u_{fb}^i と乗算され、逆モデル 3 の学習に重みが付けられる。これによって、対となる順モデル 2 が正しい予測をしているときだけ逆モデル 3 が制御し、学習することが保証される。

【 0 0 3 7 】なお、教師付学習の場合には、目標とする

$$\Delta \alpha_t^i = \varepsilon \lambda_t^i \frac{d\Psi_i}{d\alpha_t^i}(u_t^* - u_t^i) = \varepsilon \frac{du_t^i}{d\alpha_t^i} \lambda_t^i (u_t^* - u_t^i) \cong \varepsilon \frac{du_t^i}{d\alpha_t^i} \lambda_t^i u_{fb}^i \tag{6}$$

【 0 0 3 9 】ここで、責任信号 λ_t^i は 3 通りに使われている。第 1 に第 (3) 式で明らかなように順モデル 2 の学習を制御し、第 2 に第 (6) 式で明らかなように逆モデル 3 の学習を制御し、第 3 に第 (5) 式で明らかなように逆モデル 3 の最終運動指令への寄与を制御する。

【 0 0 4 0 】図 4 は責任信号予測器の動作を説明するために、図 1 の責任信号予測器に関連する部分を抽出して示した図である。責任信号予測器 4 も入力層 4 0 1 と出力層 4 0 2 とを含むニューラルネットワークによって構成されていて、入力層 4 0 1 には文脈情報信号 y_t が入力され、出力層 4 0 2 はそれぞれの単位に対して責任信号推定器 5 によって計算される責任信号の予測値 $\hat{\lambda}_t^i$ を次の第 (7) 式に基づいて計算する。

【 0 0 4 1 】

【数 7】

$$\hat{\lambda}_t^i = \eta(y_t^i, y_t) \tag{7}$$

【 0 0 4 2 】これらの責任信号予測値 $\hat{\lambda}_t^i$ は責任信

* お、出力層 3 0 2 には、フィードバック誤差学習、順逆モデリング、直接逆モデリングなど所定の学習制御教師有り学習則から得られる誤差信号を入力して学習を行なう。

【 0 0 3 3 】

【数 4】

号 λ_t^i による重み付け加算して得られる。

【 0 0 3 5 】

【数 5】

20 運動指令信号 u_t^* が必要となって、一般にそれは得られないが、フィードバック誤差学習に基づいて ($u_t^* - u_t^i$) をフィードバック制御器 3 0 から与えられるフィードバック運動指令 u_{fb}^i で置換えて第 (6) 式のように学習できる。このような教師付学習については、特許第 2 6 7 6 3 9 7 号において知られている。

【 0 0 3 8 】

【数 6】

号推定器 5 によって計算された実際の責任信号 λ_t^i と減算器 4 0 3 によって比較される。これらの誤差は、責任信号予測器 4 のパラメータを教師有り学習によって更新するために使用される。

【 0 0 4 3 】なお、責任信号推定器 5 は、各制御モジュール 1 1 , 1 2 ... 1 n が出力する状態の予測値と制御対象の状態との誤差を計算する入力部 4 0 4 と、責任信号予測器 4 による予測と入力部 4 0 4 の誤差の積を計算する中間層 4 0 5 と、制御モジュール 1 1 , 1 2 ... 1 n のそれぞれに関して大きな誤差には小さな責任信号を単調減少飽和型関数で計算する出力部 4 0 6 とから構成されている。

【 0 0 4 4 】図 5 は図 2 ~ 図 4 で説明した順モデル 2 と逆モデル 3 と責任信号予測器 4 との関連を理解しやすいように系統的に示した図である。

【 0 0 4 5 】順モデル 2 は運動指令信号に基づいて、フィードフォワード的に次の状態を予測し、減算器 2 1 によって予測値とフィードバックされた現在の状態とを減算して予測誤差が求められる。この予測誤差は尤度モデ

ルとされる。

【0046】図4で説明したように、責任信号予測器4によって文脈情報信号に基づいて演算された予測値 \hat{x}_t^i は事前確率とみなされる。この事前確率は乗算器101によって前述の尤度モデルと乗算されて正規化され責任信号推定器5に入力され、責任信号推定器5からフ*

$$\lambda_t^i = \hat{\lambda}_t^i e^{-|x_t - \hat{x}_t^i|^2 / 2\sigma^2} / \sum_{j=1}^n \hat{\lambda}_t^j e^{-|x_t - \hat{x}_t^j|^2 / 2\sigma^2} \quad (8)$$

【0048】逆モデル3は目標軌道から運動指令を演算し、乗算器61で運動指令と責任信号推定値とが乗算され、加算器7で他の制御モジュールの運動指令が加算されてフィードフォワード運動指令となる。また、フィードバック運動指令は乗算器62によって責任信号推定値と乗算され、逆モデル3の学習の誤差信号として用いられる。

【0049】図6~図8はこの発明の効果を説明するための図であって、特に、図6は操作される物体を手で保持している状態を示し、図7はこの発明の学習の過程を示す図であり、図8は操作物体を制御するためのスイッチングがこの発明の学習の結果得られた3つの制御モジュールによって高い精度で行なわれている様子を示す図である。

【0050】図6において、操作される物体の質量をMとし、弾性をKとし、粘性をBとする。

【0051】図7において、目標として与えられる3種類の物体 α, β, γ の特性を推定できれば、目標を達成したことになる。学習であるので、最初の3つのモジュールA, B, Cの物理的特性はランダムに選ばれ、異なる点の初期状態から出発し、学習していくうちにそれぞれのモジュールA, B, Cの3つの順モデルがモデル化する操作対象物体 α, β, γ の特性に学習で収束していくことがわかる。

【0052】図8(a), (b), (c)は学習初期に3つのモジュールA, B, Cが図8(d)に示す目標軌道を追従する場合の責任信号のスイッチング状態を示している。学習初期では、責任信号は皆同じく1/3の値を保っていてスイッチングできず、図8(d)に示すように、目標軌道aと実現した軌道bとが大きく異なっている。

【0053】これに対して、図8(e), (f), (g)はモジュールA, B, Cの責任信号の一例を示しており、この発明による学習をした結果5000ステップごとに正しくスイッチングが生じており、図8(h)

	M_j^F	B_j^F	K_j^F	M_j^I	B_j^I	K_j^I
A	1.0020	2.0080	8.0000	1.0711	2.0080	8.0000
B	5.0071	7.0040	4.0000	5.0102	6.9554	4.0089
C	8.0029	3.0010	0.9999	7.8675	3.0467	0.9527

*フィードバック的に責任信号推定値つまり、事後確率 λ_t^i が計算されて出力される。責任信号推定器5によって最終的に計算される責任信号 λ_t^i は第(8)式のようにBayesの法則によって決定される。

【0047】

【数8】

10 に示すように、目標軌道と実現した軌道とをほぼ一致させることができた。すなわち、時間 $0 \sim 0.5 \times 10^4$, 1.5×10^4 , 2.0×10^4 の2つの時間帯では、物体Aを操作するためのモジュール α が正しくスイッチオンされていることがわかる。一方、時間 $0.5 \times 10^4 \sim 1.0 \times 10^4 \sim 2.0 \times 10^4 \sim 2.5 \times 10^4$ では、物体BのためにモジュールB β が、時間 $1.0 \times 10^4 \sim 1.5 \times 10^4$, $2.5 \times 10^4 \sim 3.0 \times 10^4$ では物体Cのためにモジュール γ が正しくスイッチオンされている。

20 【0054】全体として非常に短い時間の誤動作(縦方向の線分として 0.9×10^4 , 2.2×10^4 , 2.8×10^4 あたりに3回出現している)を除いて、ほぼ100%の精度で、新しい物体が把持されるたびに古いモジュールがスイッチオフされ、正しいモジュールがスイッチオンされている。

【0055】これらの結果、本願発明は前述の特開平6-19508号公報で提案された学習制御装置に比べて圧倒的に優れていることがわかる。

30 【0056】表1はこの発明によって学習される4つの物体 $\alpha, \beta, \gamma, \delta$ のパラメータとしての質量M, 粘性B, 弾性Kである。

【0057】

【表1】

	M (Kg)	B (Nm ⁻¹ s)	K (Nm ⁻¹)
α	1.0	2.0	8.0
β	5.0	7.0	4.0
γ	8.0	3.0	1.0
δ	2.0	10.0	1.0

【0058】表2はこの発明の学習の結果得られた順モデルと逆モデルの物理的特性を示している。

【0059】

【表2】

【0060】表2は表1に記載された4つの物体のうち、～の3つの物体についてのみ学習を行なった結果であり、 M^F_j, B^F_j, K^F_j は順モデルによるモジュールの物理特性のパラメータであり、 M^I_j, B^I_j, K^I_j は逆モデルのモジュールによる物理特性のパラメータである。表2の物体と表1の最初の3つの物体とを対比すれば明らかのように、表2の順モデルおよび逆モデルで学習されたパラメータは、表1のパラメータにかなり近い値となっていることがわかり、この発明の優位性がわかる。

【0061】次に、物体の数が単位の数に比べて大きい場合もテストした。まず、表1に示した4つの物体、 $(M, B, K) = (2.02, 3.23, 4.47)$ を学習しておき、次に新しい物体 $(M, B, K) = (2.02, 3.23, 4.47)$ を与えた。 (M, B, K) の空間での線形和のベクトルは、理論的には $(0.15, 0.20, 0.35, 0.30)$ である。この理論的な値は、次の2つの式を解くことによって一意に解かれた答である。

$$\begin{aligned} \text{【0062】} & \quad = a_1 + a_2 + a_3 + a_4 \\ a_1 + a_2 + a_3 + a_4 & = 1 \end{aligned}$$

これに対して、この発明では、 $(0.14, 0.24, 0.37, 0.26)$ となり、全く学習していない物体も最初から学習なしで常に学習済の単位をうまく組合せて制御できることがわかった。これは、1時点では1つの制御器しか使えない Narendra とは全く異なる望ましい特性である。

【0063】なお、この発明は、物体の物理特性のみならず、物体の視覚形状と物理的特性の組合せを、責任信号予測器によって学習するようにしてもよい。その場合には、前述の特開平6-19508号公報に記載されている視覚パターンと同じ視覚刺激を用いて形状と物体を形状1 - , 形状2 - , 形状3 - のように組合せて学習すればよい。

【0064】

【発明の効果】以上のように、この発明によれば、複数の制御モジュールの各順モデルによって推定された状態予測と制御対象からの実際の状態とを比較して各々の順モデルの予測誤差を計算し、その予測誤差に基づいて複数の制御モジュールのうちの予測のよい順モデルを含む制御モジュールにより大きな責任信号を与え、この責任信号に基づいて複数の逆モデルからの出力を線形加算して制御信号を出力するようにしたので、たとえば制御対象であるロボットが把持している物体の種類や力制御を行なっている環境の違いによらず、常に所望の軌道追従性能を発揮することができる。

【図面の簡単な説明】

【図1】この発明の一実施形態の全体の構成を示すブロック図である。

【図2】順モデルの動作を説明するために図1の順モデルに関連する部分を抽出して示した図である。

【図3】逆モデルの動作を説明するために、図1の逆モデルに関連する部分を抽出して示した図である。

【図4】責任信号予測器の動作を説明するために、図1の責任信号予測器に関連する部分を抽出して示した図である。

【図5】責任信号予測器によってフィードフォワード的に推定された事前確率と順モデルの予測に基づいてフィードバック的に計算された予測誤差を組合せる機構を示す図である。

【図6】操作される物体を手で保持している状態を示す図である。

【図7】操作対象物体に対応して複数のモジュールの特性がこの発明の学習で収束する様子を示す図である。

【図8】この発明による学習前と学習後の目標軌道を追従するための効果を説明するための図である。

【符号の説明】

2 順モデル

3 逆モデル

4 責任信号予測器

5 責任信号推定器

7, 21 ~ 2n, 203, 403 減算器

8, 61 ~ 63, 303, 304 乗算器

9 制御対象

11 ~ 1n 制御モジュール

30 フィードバック制御器

201, 301, 401, 404 入力層

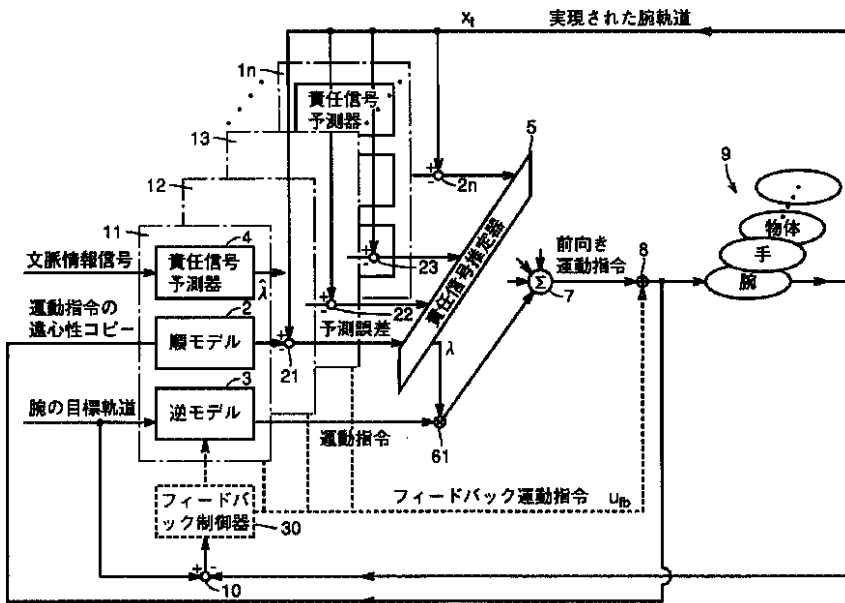
202, 302, 402, 406 出力層

【要約】

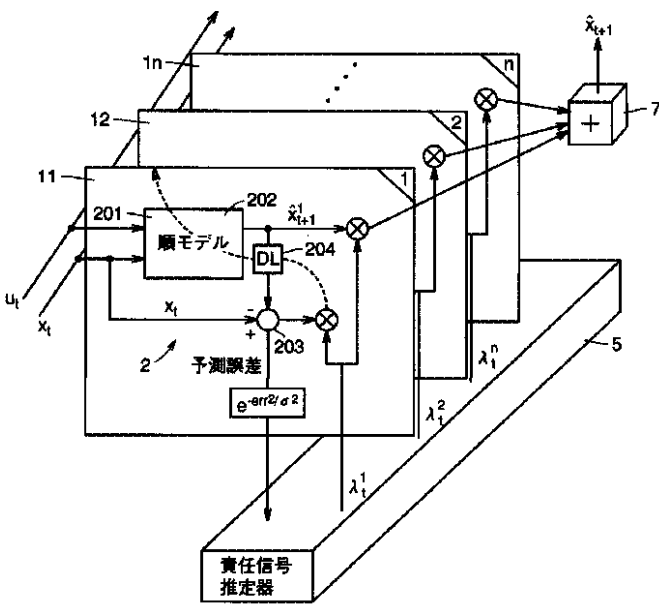
【課題】 制御対象の軌道追従制御を高い性能でかつ学習で行なうことのできるような学習制御方法および学習制御装置を提供する。

【解決手段】 順モデル2と逆モデル3と責任信号予測器4とを有する多数の制御モジュール11, 12...1nが設けられ、各順モデル2の状態予測とフィードバックされた実際の状態とを比較して各順モデルの予測誤差を計算し、この予測誤差と責任信号予測器4の出力とに基づいて、責任信号推定器5が責任信号を計算し、この責任信号で各逆モデル3の出力を加算して最終的な運動指令を計算し、制御対象9に与える。

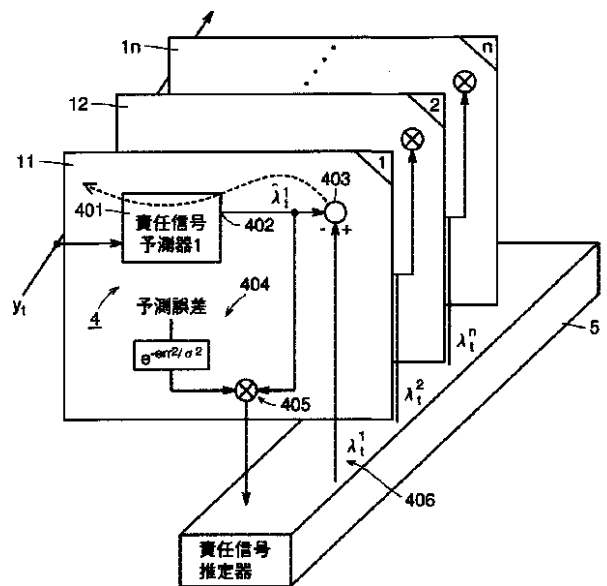
【図1】



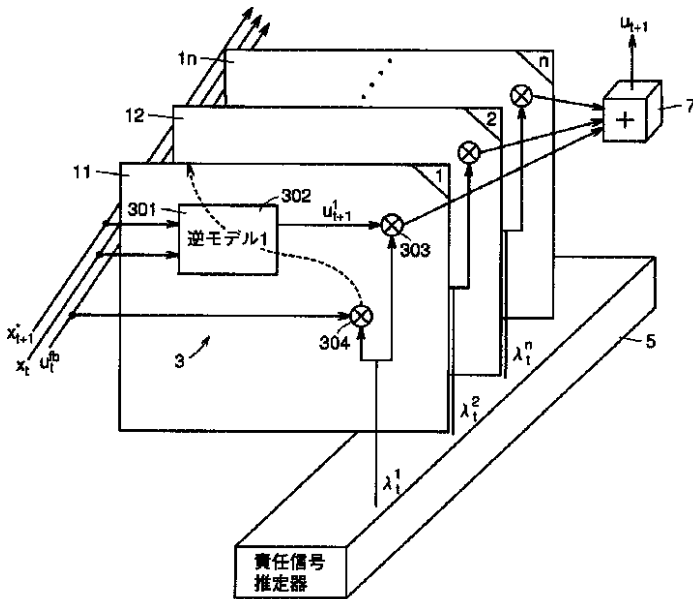
【図2】



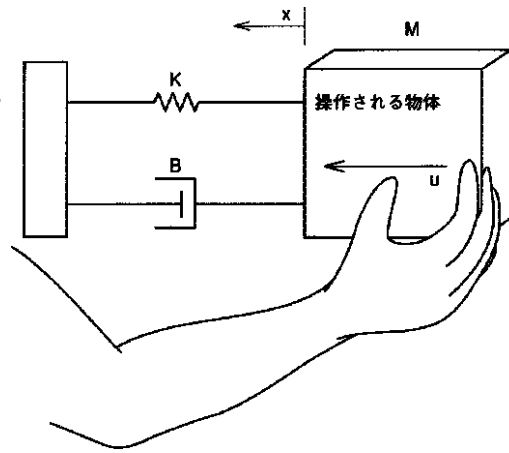
【図4】



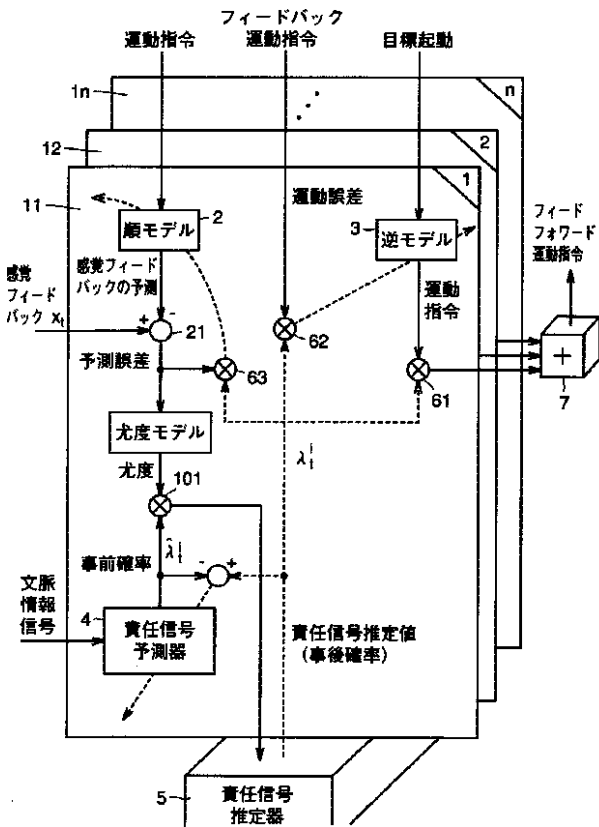
【図3】



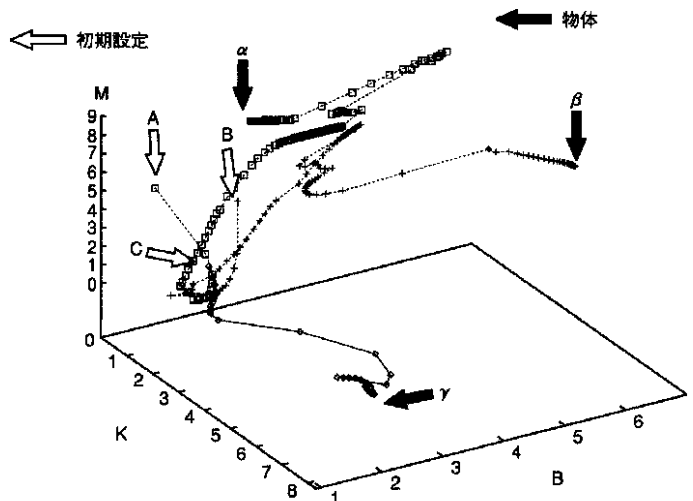
【図6】



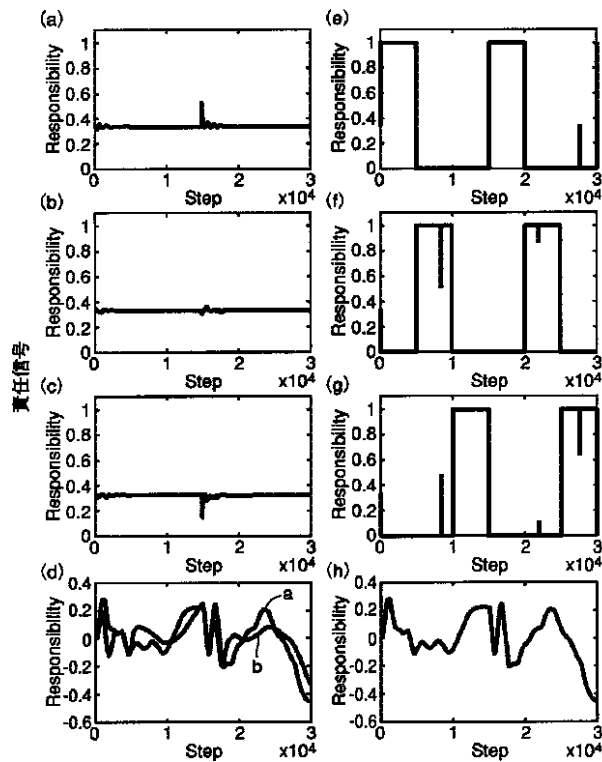
【図5】



【図7】



【図 8】



フロントページの続き

(72)発明者 春野 雅彦
 京都府相楽郡精華町大字乾谷小字三平谷
 5番地 株式会社エイ・ティ・アル人
 間情報通信研究所内

(56)参考文献 特開 平 8 - 152902 (J P , A)
 特開 平 6 - 19508 (J P , A)
 特開 平 6 - 222839 (J P , A)
 特開 平 3 - 123902 (J P , A)
 田中健一、外 1 名、「運動系ニューラ
 ルネットワークとロボット制御」、シス
 テム / 制御 / 情報、システム制御情報学
 会、平成 4 年 10 月 15 日、第 36 巻、第 10
 号、P . 653 - 660

(58)調査した分野(Int.Cl.⁷, D B 名)
 G05B 13/00 - 13/04
 J I C S T ファイル (J O I S)