

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2007-164704
(P2007-164704A)

(43) 公開日 平成19年6月28日(2007.6.28)

(51) Int. Cl. F I テーマコード(参考)
G05B 13/02 (2006.01) G05B 13/02 M 5H004
G06N 3/00 (2006.01) G06N 3/00 550E

審査請求 未請求 請求項の数 8 O L (全 15 頁)

(21) 出願番号 特願2005-363602(P2005-363602)
 (22) 出願日 平成17年12月16日(2005.12.16)

特許法第30条第1項適用申請有り 2005年6月16日 社団法人電子情報通信学会発行の「電子情報通信学会技術研究報告 信学技報Vol. 105 No. 130」にて発表

(71) 出願人 504174135
 国立大学法人九州工業大学
 福岡県北九州市戸畑区仙水町1番1号
 (74) 代理人 100099634
 弁理士 平井 安雄
 (72) 発明者 古川 徹生
 福岡県北九州市若松区ひびきの2-4 国立大学法人九州工業大学内
 Fターム(参考) 5H004 GA01 GB16 HA07 HB07 HB08
 JA03 JA07 JB01 JB21 KD35
 LA05 LA12

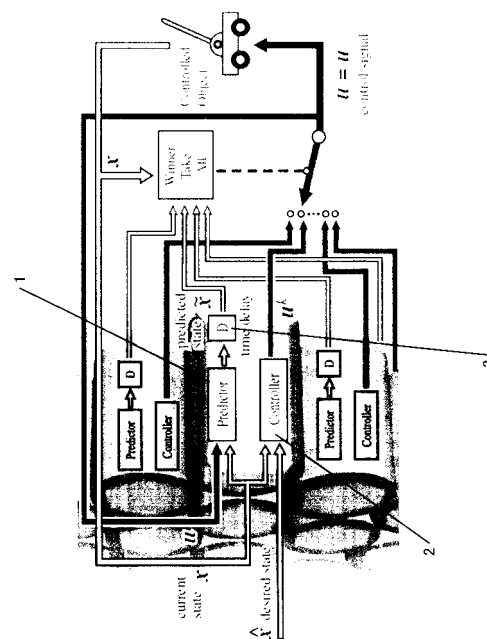
(54) 【発明の名称】 自己組織化マップを用いる装置、その方法及びプログラム

(57) 【要約】

【課題】 数少ないトレーニングケースから汎化性の高い制御を実現することができる装置を提供する。

【解決手段】 装置内に複数存在するニューラルネットワークのモジュールからなるユニットのうち、最も次時刻の制御対象の予測状態を正しく予測した予測器を含むユニットに係る制御器の制御信号を採用して制御対象を制御するので、即時性の高い制御を実現できると共に、自己組織化マップを形成することができる。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

ニューラルネットワークのモジュールからなるユニット間の競合的学習から実現される自己組織化マップを構築する装置であって、

当該ニューラルネットワークのモジュールは、制御対象を制御する制御器と制御対象の次時刻状態を予測する予測器を含み、

制御器が制御対象の理想状態と制御対象の現在状態を入力されることで制御信号を出力し、

予測器が制御対象の現時刻の制御信号と制御対象の現在状態を入力されることで次時刻の制御対象の予測状態を出力し、

制御対象の現在状態を最も近い予測を行った予測器を具備するユニットを最適ユニットとして特定し、

最適ユニットに係る制御器から出力された制御信号で制御対象を実際に制御し、

制御信号を採用されたユニットを最適ユニットとして自己組織化マップを更新する装置。

10

【請求項 2】

現在状態の最も近い予測を行った予測器を具備するユニットである候補ユニットと、前回の最適ユニットとなったユニットとが異なる場合に、

候補ユニットの現在状態の予測が前回の最適ユニットとなったユニットの現在状態の予測よりも所定以上制御対象の現在状態に近くないときには、前回の最適ユニットとなったユニットを最適ユニットとして維持する

前記請求項 1 に記載の装置。

20

【請求項 3】

制御対象の現在状態を入力されることで制御信号を出力すると共に制御器にも出力する制御器毎に用意された線形フィードバック制御器を新たに含む

前記請求項 1 に記載の装置。

【請求項 4】

予測器が出力した予測した制御対象の予測状態を少なくとも予測した時刻が到来するまで保持する予測毎に用意された遅延器を新たに含む

前記請求項 1 に記載の装置。

30

【請求項 5】

最適ユニットとして特定された対象のユニットが以前最適ユニットとして特定されたとき、又は、以前最適ユニットと特定されたユニットと自己組織化マップ上で近くにあったとき、そのときの制御対象の条件に基づき対象のユニットが対象としている制御対象の条件を推定する

前記請求項 1 に記載の装置。

【請求項 6】

現在の制御対象の条件が装置に入力された場合に、当該現在の制御対象の条件に対応する自己組織化マップ上の位置を特定する手段を新たに含む、

自己組織化マップ上の位置に対応するユニットを用いて現在の制御対象を制御する

前記請求項 1 に記載の装置。

40

【請求項 7】

ニューラルネットワークのモジュールからなるユニット間の競合的学習から実現される自己組織化マップを構築する装置を用いる方法であって、

ニューラルネットワークのモジュールに含まれる制御対象を制御する制御器が制御対象の理想状態と制御対象の現在状態を入力されることで制御信号を出力するステップと、

ニューラルネットワークのモジュールに含まれる制御対象の次時刻状態を予測する予測器が制御対象の現時刻の制御信号と制御対象の現在状態を入力されることで次時刻の制御対象の予測状態を出力するステップと、

制御対象の現在状態の最も近い予測を行った予測器を具備するユニットを最適ユニッ

50

トとして特定するステップと、

最適ユニットに係る制御器から出力された制御信号で制御対象を実際に制御するステップと、

制御信号を採用されたユニットを最適ユニットとして自己組織化マップを更新するステップを含む方法。

【請求項 8】

ニューラルネットワークのモジュールからなるユニット間の競合的学習から実現される自己組織化マップを構築するようにコンピュータを機能させるためのプログラムであって、

制御対象の理想状態と制御対象の現在状態を入力されることで制御信号を出力するニューラルネットワークのモジュールに含まれる制御対象を制御する制御器と、

制御対象の現時刻の制御信号と制御対象の現在状態を入力されることで次時刻の制御対象の予測状態を出力するニューラルネットワークのモジュールに含まれる制御対象の次時刻状態を予測する予測器と、

制御対象の現在状態の最も近い予測を行った予測器を具備するユニットを最適ユニットとして特定する手段と、

最適ユニットに係る制御器から出力された制御信号で制御対象を実際に制御する手段と、

制御信号を採用されたユニットを最適ユニットとして自己組織化マップを更新する手段としてコンピュータを機能させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は自己組織化マップを用いる装置であり、特に、即時性を要する制御に最適な装置に関する。

【背景技術】

【0002】

本発明の背景技術として、特開 2000 - 35956 号公報に開示されるエージェント学習装置がある。

【0003】

この背景技術のエージェント学習装置は、環境に対して働きかけ、その結果得られる報酬を最大化するための行動出力を決定する強化学習システムと、環境の変化を予測する環境予測システムとの組によりなる学習モジュールが複数備えられ、各学習モジュールの環境予測システムの予測誤差が少ないものほど大きな値を取る責任信号が求められ、この責任信号に比例して強化学習システムによる行動出力が重み付けされて、環境に対する行動が与えられる構成である。

【0004】

この背景技術のエージェント学習装置によれば、非線形性 / 非定常性を持つ制御対象やシステムなどの環境で、具体的な教師信号は与えられず、様々な環境の状態や動作モードに最適な行動の切り替えや組み合わせを行い、また先見知識を用いることなく柔軟に行動学習を行なうことができる。

【特許文献 1】特開 2000 - 35956 号公報

【非特許文献 1】Wolpert, D.M., Kawato, M.: Multiple paired forward and inverse models for motor control. Neural Networks 11, 1317-1329, 1998

【発明の開示】

【発明が解決しようとする課題】

【0005】

前記背景技術のエージェント学習装置は、教師付き学習の枠組みを適用することができない、何が正しい出力であるかは未知であるという実世界の問題を解決するものであった。

しかしながら、背景技術のエージェント学習装置であっても、トレーニングケースの少ない場合にあっては迅速に問題を適切に解決することができない場合があるという課題を有する。特に、環境1、環境2、環境3を仮定した場合に、環境2が環境1と環境3を折衷した環境である場合に、環境1を十分にトレーニングさせ、環境3を十分にトレーニングさせた後に、環境2をトレーニングさせた時には、環境1に対応した学習モジュールと、環境3に対応した学習モジュールとが環境2に歩み寄る形で対応するため、環境2に十分に対応するためには相当の時間がかかる。すなわち、既に学習した環境を利用して新規環境に対する即時性のある対応をできないという課題を有する。

【0006】

本発明は前記課題を解決するためになされたものであり、数少ないトレーニングケースから汎化性の高い制御を実現することができる装置を提供することを目的とする。 10

【課題を解決するための手段】

【0007】

すなわち、制御対象の特性の突然の変化に対応し、且つ、できるだけ少ない標本数から汎化的な制御能力を獲得することができる装置を実現することを目標としている。そこで自己組織化マップの考え方を導入した自己組織化適応制御器 (Self-Organizing Adaptive Controller : SOAC) を提案する。SOACの開発は制御工学における適応制御のテーマに関連するが、適応制御では基本的に、制御対象の特性が時間とともにゆっくり変化することを前提としており、その点相違する。

【0008】

SOACは制御器を自己組織的に構成することを目的に考案されたものであり、大きく2つの特徴を持つ。第1の特徴は、自己組織化マップ (SOM) とモジュラーネットワークの双方の特徴を併せ持つ $m \times n$ SOM (modular network SOM) を元にしたアーキテクチャを用いていることである。すなわちSOACはニューラルネットの機能モジュールが多数集まった構造を持ち、それらはSOMのアルゴリズムに従って学習が行われる。第2の特徴は、SOACの各機能モジュールが制御器と予測器のペアから構成されることである。すなわちSOACは、特性の異なる制御器と予測器が多数並んだ構造をしている (図1)。ここで、図1は本発明のSOACの基本構成を示す図である。 20

【0009】

まず第1の特徴である $m \times n$ SOMであるが、これは従来型SOMの各ベクトルユニットをニューラルネットワークの機能モジュールに置き換えたものである。たとえば $m \times n$ SOMでは、MLP (Multi-Layer Perceptron) やRNN (Recurrent Neural Network) などのモジュールをSOMのベクトルユニットの代わりとして使うことができる。こうすることで従来型SOMでは扱えなかったデータベクトルの集合や、時系列データをマッピングすることができる。なお $m \times n$ SOMの機能モジュールとしてヘブ学習ニューロンを選べば $m \times n$ SOMは通常のSOMになることから、 $m \times n$ SOMはSOMの一般化とみることができる。 $m \times n$ SOMの機能モジュールはユーザーが自由にデザインできるので、SOMの応用範囲を大きく広げることができる。そこで発明者は鋭意努力の末、ニューラルネットワークを用いた制御器を機能モジュールとする $m \times n$ SOMを想到した。これがSOACの第1の特徴である。すなわちSOACは、ニューラルネットの制御器が多数集合したものであり、それらはSOMのアルゴリズムによって機能の分業・協調が行われる。ある制御対象が与えられたとき、制御対象をもっとも良く制御するモジュールが最適制御器 (Best Matching Controller : BMC) として選ばれ、BMCモジュールを用いて対象が制御される。もし制御対象の特性が突然変化した場合は、BMCも直ちに別のモジュールへと変わるため適応的に制御を行うことができる (図2)。ここで、図2は本発明のモジュール切り替えの説明図である。 30 40

【0010】

SOACの第2の特徴は、制御器と予測器がペアになったモジュール構造である。その必要性は制御タスクを実時間で行わなければならないことから生じる。制御対象の特性が変化したとき、それに対応して最適な制御器、すなわちBMCをただちに切り替えなけ 50

ればならない。そこでSOACでは、すべての制御器とペアになる予測器を用意しておき、制御対象の次時刻の状態をもっとも良く推定した予測器とペアの制御器が、その時刻におけるBMCであるとした。すなわち予測器はシステム同定器として働き、ペアとなる制御器は同定したシステムに対して最適な制御器になるよう事前に学習しておく。こうすることで、制御対象の突然の変化に対しても瞬時にBMCを切り替えることができるようになる。

より体系的には本発明は次のように説示できる。

【0011】

(1) 本発明に係る装置は、ニューラルネットワークのモジュールからなるユニット間の競合的学習から実現される自己組織化マップを構築する装置であって、当該ニューラルネットワークのモジュールは、制御対象を制御する制御器と制御対象の次時刻状態を予測する予測器を含み、制御器が制御対象の理想状態と制御対象の現在状態を入力されることで制御信号を出力し、予測器が制御対象の現時刻の制御信号と制御対象の現在状態を入力されることで次時刻の制御対象の予測状態を出力し、制御対象の現在状態を最も近い予測を行った予測器を具備するユニットを最適ユニットとして特定し、最適ユニットに係る制御器から出力された制御信号で制御対象を実際に制御し、制御信号を採用されたユニットを最適ユニットとして自己組織化マップを更新するものである。

10

【0012】

このように本発明によれば、装置内に複数存するニューラルネットワークのモジュールからなるユニットのうち、最も次時刻の制御対象の予測状態を正しく予測した予測器を含むユニットに係る制御器の制御信号を採用して制御対象を制御するので、即時性の高い制御を実現することができると共に、自己組織化マップを形成することができる。

20

なお、前記「ニューラルネットワークのモジュールからなるユニット間の競合的学習から実現される自己組織化マップを構築する装置」は、下位概念として「ニューラルネットワークのモジュールからなるユニット間の競合的学習と近傍関数による平滑化から実現される自己組織化マップを構築する装置」とすることもできる。

【0013】

(2) 本発明に係る装置は必要に応じて、現在状態の最も近い予測を行った予測器を具備するユニットである候補ユニットと、前回の最適ユニットとなったユニットとが異なる場合に、候補ユニットの現在状態の予測が前回の最適ユニットとなったユニットの現在状態の予測よりも所定以上制御対象の現在状態に近くないときには、前回の最適ユニットとなったユニットを最適ユニットとして維持するものである。

30

【0014】

このように本発明によれば、最適ユニットが他のユニットに移行する可能性がある場合に、前回の最適ユニットの予測状態と制御対象の現在状態との差が、候補ユニットの予測状態と制御対象の現在状態との差よりも所定閾値よりも大きくない場合に、前回の最適ユニットを継続して採用するので、最適ユニットの交代が頻繁になされることがなく、安定した制御対象の制御を実現することができる。すなわち、候補ユニットがいくらか正確に制御対象の予測状態を予測していた場合であっても現状の最適ユニットでもさほど変わらず十分制御することができる場合にはユニットの切換を抑制している。ユニットの切換の乱発で不安定な制御系にならないように配慮している。

40

【0015】

(3) 本発明に係る装置は必要に応じて、制御対象の現在状態を入力されることで制御信号を出力すると共に制御器にも出力する制御器毎に用意された線形フィードバック制御器を新たに含むものである。

このように本発明によれば、各制御器に線形フィードバック制御器を配しているので、制御器であるニューラルネットワークが十分に学習していない場合であっても線形フィードバックが制御信号を補間して適切な制御を実現できると共に、線形フィードバック制御器が出力する信号が制御器にも入力されて学習も補間することができる。

【0016】

50

(4) 本発明に係る装置は必要に応じて、予測器が出力した予測した制御対象の予測状態を少なくとも予測した時刻が到来するまで保持する予測毎に用意された遅延器を新たに含むものである。

このように本発明によれば、予測器が予測状態を出力するタイミングが予測した時刻でなかった場合であっても遅延器が調整し、適切に最適ユニットを特定することができる。

【0017】

(5) 本発明に係る装置は必要に応じて、最適ユニットとして特定された対象のユニットが以前最適ユニットとして特定されたとき、又は、以前最適ユニットと特定されたユニットと自己組織化マップ上で近くにあったとき、そのときの制御対象の条件に基づき対象のユニットが対象としている制御対象の条件を推定するものである。

10

このように本発明によれば、同一制御対象の条件は自己組織化マップの同位置若しくは周辺位置に配置されるため、ユニットが最適ユニットとなったときに以前最適ユニットとなったときの制御対象の条件からおおよそその対象の制御対象の条件を把握することができる。

【0018】

(6) 本発明に係る装置は必要に応じて、現在の制御対象の条件が装置に入力された場合に、当該現在の制御対象の条件に対応する自己組織化マップ上の位置を特定する手段を新たに含み、自己組織化マップ上の位置に対応するユニットを用いて現在の制御対象を制御するものである。

20

【0019】

このように本発明によれば、入力された制御対象の条件に対応する自己組織化マップ上の位置を特定し、この特定した位置に対応するユニットに係る制御器を用いて制御するので、入力された制御対象の条件に関して直接的に学習がなされていない場合であっても、比較的安定して制御することができる。例えば、制御対象の条件Aと制御対象の条件Bがあり、これらの条件に関しては学習がなされている場合に、制御対象の条件Aと制御対象の条件Bの中間の制御対象の条件Cが入力されたとき、自己組織化マップ上で制御対象の条件Aと制御対象の条件Bとの間の位置に対応するユニットを用いて制御対象を制御することでかかる条件に関して学習がない場合であっても当初からおおよそ適切に制御を行うことができる。

30

【0020】

(7) 本発明に係る方法は、ニューラルネットワークのモジュールからなるユニット間の競合的学習から実現される自己組織化マップを構築する装置を用いる方法であって、ニューラルネットワークのモジュールに含まれる制御対象を制御する制御器が制御対象の理想状態と制御対象の現在状態を入力されることで制御信号を出力するステップと、ニューラルネットワークのモジュールに含まれる制御対象の次時刻状態を予測する予測器が制御対象の現時刻の制御信号と制御対象の現在状態を入力されることで次時刻の制御対象の予測状態を出力するステップと、制御対象の現在状態の最も近い予測を行った予測器を具備するユニットを最適ユニットとして特定するステップと、最適ユニットに係る制御器から出力された制御信号で制御対象を実際に制御するステップと、制御信号を採用されたユニットを最適ユニットとして自己組織化マップを更新するステップを含むものである。前記装置は方法としても把握することができる。

40

【0021】

(8) 本発明に係るプログラムは、ニューラルネットワークのモジュールからなるユニット間の競合的学習から実現される自己組織化マップを構築するようにコンピュータを機能させるためのプログラムであって、制御対象の理想状態と制御対象の現在状態を入力されることで制御信号を出力するニューラルネットワークのモジュールに含まれる制御対象を制御する制御器と、制御対象の現時刻の制御信号と制御対象の現在状態を入力されることで次時刻の制御対象の予測状態を出力するニューラルネットワークのモジュールに含まれる制御対象の次時刻状態を予測する予測器と、制御対象の現在状態の最も近い予測を

50

行った予測器を具備するユニットを最適ユニットとして特定する手段と、最適ユニットに係る制御器から出力された制御信号で制御対象を実際に制御する手段と、制御信号を採用されたユニットを最適ユニットとして自己組織化マップを更新する手段としてコンピュータを機能させるためのものである。前記装置はプログラムとしても把握することができる。

これら前記の発明の概要は、本発明に必須となる特徴を列挙したのではなく、これら複数の特徴のサブコンビネーションも発明となり得る。

【発明を実施するための最良の形態】

【0022】

(本発明の第1の実施形態)

10

[1. 基本構成]

SOACの構成を図1に示す。基本的なSOACのアーキテクチャはmnSOMと同じであり、mnSOMの機能モジュールが予測器ブロックと制御器ブロックから構成されたものである。

k - thモジュールの制御器ブロックは、制御対象の現在の状態 $x(t)$ と目標状態 $\hat{x}(t)$ を入力とし、制御信号 $u^k(t)$ を出力とする。すなわち

【0023】

【数1】

$$u^k(t) = cf^k(x(t), \hat{x}(t)) \quad (1)$$

20

【0024】

と表されとする。一方、k - thモジュールの予測器ブロックは、制御対象の現在の状態 $x(t)$ と制御信号 $u(t)$ を入力とし、 t 秒後の制御対象の状態の予測値 $\tilde{x}^k(t + \Delta t)$ を出力する。すなわち

【0025】

【数2】

$$\tilde{x}^k(t + \Delta t) = pf^k(x(t), u(t)) \quad (2)$$

30

【0026】

と表されとする。

SOACは、学習モードと実行モードの2つのモードを持つ。学習モードでは、全モジュールの予測器と制御器をmnSOMのアルゴリズムに従って学習する。実行モードでは学習の完了したモジュールを用いて実際に制御対象を制御する。

【0027】

[2. 実行モード]

学習モードについて説明する前に、まずSOACを実際に運用する実行モードについて説明する。制御対象の挙動と予測器が予想した挙動との誤差を次式で定義する。

【0028】

40

【数3】

$$pe^k(t) = (1 - \varepsilon)pe^k(t - \Delta t) + \varepsilon\|x(t) - \tilde{x}^k(t)\|^2 \quad (3)$$

【0029】

$pe^k(t)$ は予測誤差の指数減衰平均である。すなわち実行モードにおいては、ごく近い過去から現在までの予測誤差の時間平均を取る。時間平均を取る区間の長さは Δt で決まり、 ε が小さいほど時間平均の区間は長くなり、逆に $\varepsilon=1$ のときはその瞬間の予測誤差のみで $pe^k(t)$ が決まる。そして $pe^k(t)$ をもっとも小さくするモジュールが時刻 t におけるBMCとなる。 Δt の値は制御対象に加わる外乱やノイズの大きさによって決まり、一般に

50

外乱やノイズが大きいほど は小さくとる方がよい。BMCの添字を*とすれば、

【0030】

【数4】

$$*(t) = \arg \min_k p e^k(t) \quad (4)$$

【0031】

【数5】

$$u(t) = u^*(t) = c f^*(x(t), \hat{x}(t)) \quad (5)$$

10

となり、BMCの出力が実際の制御信号となり制御対象へ入力される。

【0032】

[3. 学習モード(予測器ブロック)]

本節では、予測器ブロックの学習について説明し、制御器ブロックの学習は次節で説明する。今、事前にI個の既知な制御対象があり、これらを学習に使用するものとする。従って、これらを制御する制御器もI個用意する。よってI個の時系列データ[(x_i(t), u_i(t))](i=1, ..., I)が得られる。

【0033】

予測器ブロックの学習アルゴリズムはmnSOMのアルゴリズムと等しい。したがって、予測器ブロックのアルゴリズムはmnSOMと同様に(1)評価過程(2)競合過程(3)協調過程(4)適応過程の4過程から成る。ここで、予測器はMLPモジュールであると仮定し、重みベクトルを^pw^kとする。

20

【0034】

[3.1 評価過程]

まず、教師パターンとの予測誤差をI個全てに対して求める。

【0035】

【数6】

$${}^p E_i^k = \frac{1}{T} \int_0^T \|x_i(t) - \tilde{x}_i^k(t)\|^2 dt \quad (6)$$

30

【0036】

ここで、 $x_i^k(t)$ と^pE_i^kはそれぞれi番目の教師に対するk-t h予測器の出力と平均予測誤差である。また、Tは時系列の長さを表す。

【0037】

[3.2 競合過程]

予測誤差を求めた後、全ての教師パターンについてBMCを決める。BMCは次式に示すように平均予測誤差を最小としたモジュールにより決定される。

【0038】

【数7】

$$*_i = \arg \min_k {}^p E_i^k \quad (7)$$

40

【0039】

[3.3 協調過程]

近傍関数を用いて学習分配率を決定する。

【0040】

【数 8】

$$\psi_i^k = \frac{\exp[-\|\xi^k - \xi_i^*\|^2/2\sigma]}{\sum_{i'=1}^I \exp[-\|\xi^k - \xi_{i'}^*\|^2/2\sigma]} \quad (8)$$

ここで、 ξ^k 、 ξ_i^* は k - t h モジュールと BMC のマップ空間における座標を表す。

【0041】

[3.4 適応過程]

予測器の重みベクトルは学習分配率 $\{\psi_i^k\}$ を用いて次式のように表される。

【0042】

10

【数 9】

$$\Delta^p w^k = -\eta \sum_{i=1}^I \psi_i^k \frac{\partial PE_i^k}{\partial w^k} \quad (9)$$

【0043】

これら 4 過程をネットワークが定常状態になるまで繰り返す。その結果、近い性質を持つモジュールはマップ空間上の近い位置に配置される。

なお、この学習分配率に応じてマップが更新される。

20

【0044】

[3.5 学習モード(制御器ブロック)]

SOAC の制御器としてフィードバック誤差学習を用いる。フィードバック誤差学習を用いることの利点は、制御器として従来型の線形フィードバック制御器を用いて訓練することができ、事前に最適な制御器を決定する必要がないこと、従って追加学習も可能になることである。

【0045】

SOAC の 1 モジュールのブロック線図を図 3 に示す。閉ループ適応制御系にフィードバック誤差学習を導入したモデルである。このモデルは、制御器が従来型の線形フィードバック制御器 (Conventional Feedback Controller : CFC) とニューラルネット制御器 (Neural Network Controller : NNC) から構成される。CFC と NNC を並列にすることで、単に NNC を CFC で学習できることのみならず、CFC によって制御系を安定させられること、NNC が非線形な補償を実現することができることなどの長所が生じる。今、多入出力系のフィードバック係数行列を ^{cfc}W とし、 k - t h モジュールのフィードバック係数行列を添字をつけて $^{cfc}W^k$ とする。このとき、SOAC の制御則は以下のように表される。

30

【0046】

【数 10】

$$^{nnc}u^k(t) = ^{nnc}f^k(\hat{x}(t) - x(t)) \quad (10)$$

40

【0047】

【数 11】

$$^{cfc}u^k(t) = ^{cfc}W^k(\hat{x}(t) - x(t)) \quad (11)$$

【0048】

【数 12】

$$u^k(t) = ^{nnc}u^k(t) + ^{cfc}u^k(t) \quad (12)$$

50

【 0 0 4 9 】

また、NNCの誤差信号 ^{nnc}E はCFCの出力を用いて次式で定義する。

【 数 1 3 】

$$^{nnc}E^k = \frac{1}{T} \int_0^T \|^{cfc}u^k\|^2 dt \quad (13)$$

【 0 0 5 0 】

最後に、フィードバック $^{cfc}W^k$ と重みベクトル $^{nnc}W^k$ は以下の式に従って更新される。ここで、学習分配率 $^k\eta_i$ は予測器の学習で得られた値を用いる。すなわち制御器はペアになっている予測器が同定したシステムを正しく制御するように学習する。 10

【 0 0 5 1 】

【 数 1 4 】

$$^{cfc}W^k := \sum_{i=1}^I \psi_i^k \ ^{cfc}W_i \quad (14)$$

【 0 0 5 2 】

【 数 1 5 】

$$\Delta^{nnc}w^k = -\eta \frac{\partial^{nnc}E^k}{\partial^{nnc}w^k} \quad (15)$$

20

以上がSOACのアーキテクチャと学習アルゴリズムである。

【 0 0 5 3 】

[4 . ハードウェア構成図]

図4は本実施形態に係る装置の構成要素のハードウェア構成図である。本装置は汎用的なコンピュータを用いることができる。ハードウェアの構成としてはCPU(Central Processing Unit)11、DRAM(Dynamic Random Access Memory)12等のメインメモリ、外部記憶装置であるHD(hard disk)13、表示装置であるディスプレイ14、入力装置であるキーボード15及びマウス16、ネットワークに接続するための拡張カードであるLANカード17、CD-ROMドライブ18等からなる。 30

例えば、CD-ROMに格納されているプログラムがHD13上に複製(インストール)され、必要に応じてプログラムがメインメモリ12に読み出され、CPU11がかかるプログラムを実行することで装置を構成する。

【 0 0 5 4 】

[5 . 動作]

図5は本実施形態に係る装置の実行モードの動作のフローチャートの一例である。

CPU11(予測器2)は制御対象の現時刻の制御信号と制御対象の現在状態を用いて次時刻の制御対象の状態を予測する(ステップ100)。ここでの予測は全ユニットの予測器2が実施する。 40

CPU11(遅延器4)は予測した次時刻の制御対象の実際の状態が入力されるまで待機させる(ステップ200)。ここでの待機は全ユニットの遅延器4が実施する。

CPU11は予測した次時刻の制御対象の実際の状態と予測された状態とを比較し、最も近い予測を行った予測器2を特定する(ステップ300)。ここで特定された予測器2を含むユニットが最適ユニットとなる。

【 0 0 5 5 】

CPU11(制御器1)は制御対象の理想状態と制御対象の現在状態を用いて制御信号を求め、モータ等の駆動源に制御信号を出力する(ステップ400)。 50

C P U 1 1 は最適ユニットに基づき自己組織化マップを更新する (ステップ 5 0 0)

ここで、最適ユニットと判明した後に最適ユニットに係る制御器 1 が制御信号を求めているが、最適ユニットと判明する前に全ユニットの制御器 1 が制御信号を求める構成にすることもできる。

【 0 0 5 6 】

(その他の実施形態)

[自己組織化マップを用いた制御対象の条件推定]

例えば、ある制御対象の条件を制御した場合に最適ユニットが図 6 (b) に示すマップ上で真中に位置する場合には、制御対象の条件は振り子の重心までの距離が「 L o n g 」で振り子の質量が「 H e a v y 」であることが推測される。

すなわち、制御対象を制御することで自己組織化マップ上での位置を特定し、かかる位置に対応している従前の制御対象の条件から現制御対象の条件を推定することができる。

【 0 0 5 7 】

[自己組織化マップを用いたユニットの特定]

ある制御対象の条件が入力された場合に、かかる制御対象の条件と同一条件の学習がなされていなかった場合であっても、類似する制御対象の条件に対応している自己組織化マップ上で近い位置と対応しているユニットを用いることで初動時の制御の乱れを抑制することができる。

【 0 0 5 8 】

また、制御対象の条件により自己組織化マップ上で画定することも可能であり、各画定領域の中で新たに入力された制御対象の条件と最も合致する画定領域内のユニットを用いることで同様に初動時の制御の乱れを抑制することができる。また、選択可能なユニットを制限することで、制御に用いるユニットの選択にかかる労力を低減することができる。

【 0 0 5 9 】

[その他]

もし制御対象のダイナミクスが隠れパラメータによって連続的に変化するような場合、それに対応したマップができると考えられ、かように生成された特徴マップを有効に活用することができる。

また、実行モードにおいても未学習の対象を制御しながら各モジュールの特性を修正する、追加学習の機能を加えることもできる。

以上の前記各実施形態により本発明を説明したが、本発明の技術的範囲は実施形態に記載の範囲には限定されず、これら各実施形態に多様な変更又は改良を加えることが可能である。そして、かような変更又は改良を加えた実施の形態も本発明の技術的範囲に含まれる。このことは、特許請求の範囲及び課題を解決する手段からも明らかなことである。

【実施例】

【 0 0 6 0 】

[S O A C を倒立振子の制御に適用した例]

S O A C の性能を調べるために倒立振子系を用いてシミュレーション実験を行った。実験で使用した振子のパラメータを表 1 に示す (図 6 (a) 参照) 。

【 0 0 6 1 】

10

20

30

40

【表 1】

l	振子の重心までの距離	0.2 [m] (Short) 0.5 [m] (Half) 0.8 [m] (Long)
m	振子の質量	0.1 [kg] (Light) 0.35 [kg] (Middle) 0.6 [kg] (Heavy)
M	台車の質量	5.0 [kg]
C	振子の粘性摩擦係数	4.0×10^{-4} [kgm ² /s]
f	台車の粘性摩擦係数	10.0 [kg/s]
g	重力加速度	9.8 [m/s ²]

10

20

【0062】

SOACのモジュールとしては図3と同様のものを用いた。シミュレーションでは振子の長さおよび重さは可変であるものとし、パラメータの異なる9組のパラメータセットを学習用パターンとして用意した。学習用パターンは表1に示すように振子の長さが“Long”、“Half”、“Short”の3種類と、振子の重さが“Heavy”、“Middle”、“Light”の3種類の組み合わせで生成した。CFCのフィードバック係数行列（ベクトル）は状態フィードバック制御法により求めた。NNCは3層MLPを用い、予測器は線形ニューラルネットを用いた。また、台車には外乱としてガウス白色ノイズを与えた。

30

【0063】

学習終了後、ネットワークを固定して制御実験を行った。図6(b)にSOACの学習により得られたマップを示す。マップは“Short”、“Middle”、“Long”の3つのクラスターを形成した。図6(b)中は倒立振子制御モジュールのマップを示し、マップ中のグレースケールは予測器とその近傍の距離を示す。

40

【0064】

SOACの適応能力を調べるために、30秒間隔で振子の長さおよび重さを変化させた。ただし最初の30秒は学習のときに与えたパラメータ(Half-Middle)を用い、その後は未学習のパラメータとした。また制御器の切り替えを行う場合と行わない場合、使用する制御器としてCFCのみを用いた場合、NNCのみを用いた場合の4通りについてそれぞれ実験を行った。実験結果を図7に示す。図をみてわかるように、学習したパラメータについては、すべてのケースにおいて振子を倒すことなく制御できた。しかし、未学習のパラメータを与えた場合で、制御器の切り替えをおこなわなかった場合(non-adaptive)に予告なく変えると、CFC、NNCどちらの場合も開始から75秒付近で振子は倒れてしまった。一方、制御器の切り替えをおこなった場合(adaptive)、CFC、NNCどちらの場合も振子を倒すことなく制御可能であった。特にNNCを用いた場合はCFC

50

の場合よりも振動の少ない制御が可能であった。

【0065】

シミュレーションの結果、SOACは倒立振子のパラメータ変化に対し適応的にBMCを切り替え、安定した制御が可能であった。このことからSOACは高い適応能力を持つことが分かった。さらに、SOACは適応制御器としての機能のみならず、制御対象の特徴マップを自己組織的に獲得することもできた。

【図面の簡単な説明】

【0066】

【図1】本発明のSOACの基本構成を示す図である。

【図2】本発明のモジュール切り替えの説明図である。

10

【図3】本発明の実施形態に係るSOACの1モジュールのブロック線図である。

【図4】本発明の実施形態に係る装置で用いるハードウェア構成図である。

【図5】本発明の実施形態に係る装置の実行モードの動作のフローチャートの一例である。

【図6】実施例に係るSOACの学習により得られたマップの例である。

【図7】実施例に係る実験結果である。

【符号の説明】

【0067】

1 制御器

2 予測器

20

3 CFC

4 遅延器

10 コンピュータ

11 CPU

12 DRAM

13 HD

14 ディスプレイ

15 キーボード

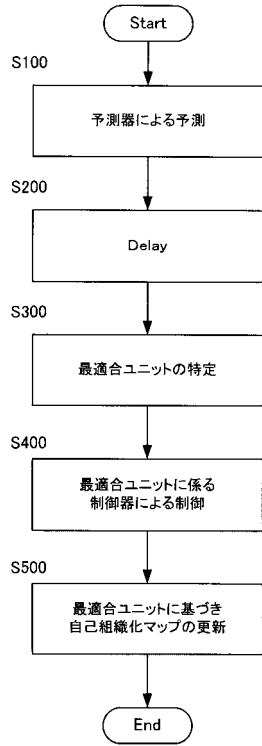
16 マウス

17 LANカード

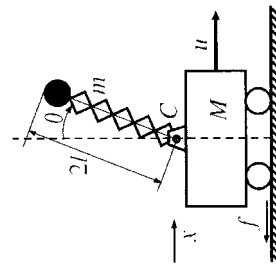
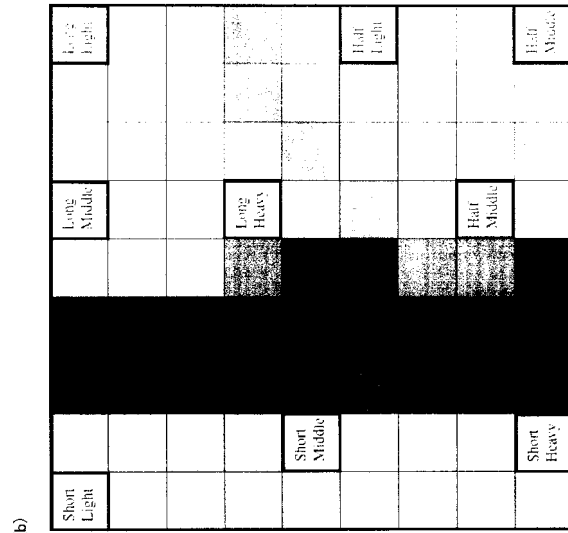
30

18 CD-ROMドライブ

【 図 5 】



【 図 6 】



【 図 7 】

