

(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2010-20445

(P2010-20445A)

(43) 公開日 平成22年1月28日(2010.1.28)

(51) Int.Cl.

F I

テーマコード (参考)

G 0 6 N 3/08 (2006.01)

G 0 6 N 3/08

Z

G 0 6 N 3/04 (2006.01)

G 0 6 N 3/04

F

審査請求 有 請求項の数 10 O L (全 28 頁)

(21) 出願番号 特願2008-178806 (P2008-178806)
 (22) 出願日 平成20年7月9日 (2008.7.9)

(71) 出願人 000002185
 ソニー株式会社
 東京都港区港南1丁目7番1号
 (74) 代理人 100082131
 弁理士 稲本 義雄
 (74) 代理人 100121131
 弁理士 西川 孝
 (72) 発明者 伊藤 真人
 東京都港区港南1丁目7番1号 ソニー株
 式会社内
 (72) 発明者 青山 一美
 東京都港区港南1丁目7番1号 ソニー株
 式会社内
 (72) 発明者 野田 邦昭
 東京都港区港南1丁目7番1号 ソニー株
 式会社内

(54) 【発明の名称】 学習装置、学習方法、およびプログラム

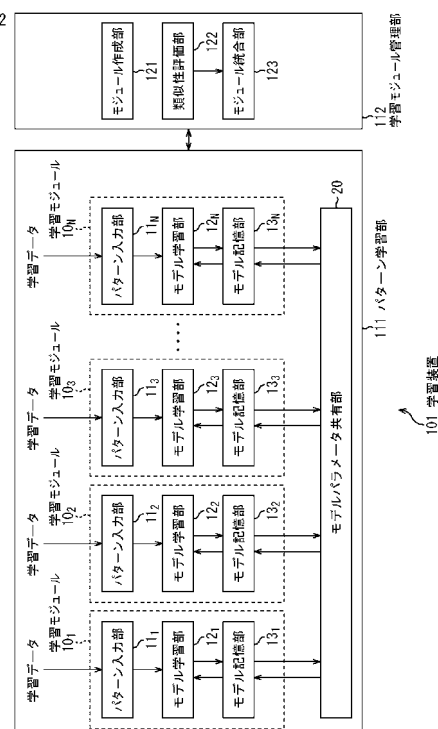
(57) 【要約】

【課題】規模拡張性があり、かつ、汎化特性を有するパターン学習モデルを得ることができるようにする。

【解決手段】学習モジュール10₁ないし10_Nは、入力データを用いて、パターン学習モデルの複数のモデルパラメータを更新する更新学習を行う。モデルパラメータ共有部20は、学習モジュール10₁ないし10_Nに、モデルパラメータを共有させる。新たな学習データが入力データとしてパターン学習部111に供給された場合、モジュール作成部121は、その新たな学習データに対応する新たな学習モジュールを作成する。そして、新たな学習モジュールを含む全ての学習モジュールで更新学習後、モジュール統合部123は、類似性評価部122による学習モジュール間の類似性に基づいて、学習モジュールの統合の可否を判定し、学習モジュールの統合を行う。本発明は、例えば、時系列パターン学習等に適用できる。

【選択図】図12

図12



【特許請求の範囲】**【請求項 1】**

パターンを学習するパターン学習モデルが有する複数のモデルパラメータを、入力データを用いて更新する更新学習を行う複数の学習モジュールと、

前記複数の学習モジュールの 2 以上の学習モジュールに、前記モデルパラメータを共有させるモデルパラメータ共有手段と、

前記パターンを学習させるための新たな学習データが前記入力データとして供給された場合に、前記新たな学習データに対応する新たな学習モジュールを作成するモジュール作成手段と、

前記新たな学習モジュールを含む全ての学習モジュールで前記更新学習が行われた後、学習モジュール間の類似性を評価する類似性評価手段と、

前記学習モジュール間の類似性に基づいて、前記学習モジュールの統合の可否を判定し、前記学習モジュールの統合を行うモジュール統合手段と

を備える学習装置。

【請求項 2】

前記モジュール作成手段は、既存の全学習モジュールのモデルパラメータの平均値を、前記新たな学習モジュールの前記複数のモデルパラメータの初期値として与える

請求項 1 に記載の学習装置。

【請求項 3】

前記モジュール統合手段は、統合対象の複数の前記学習モジュールのモデルパラメータの平均値を、統合後の前記学習モジュールのモデルパラメータに設定する

請求項 1 に記載の学習装置。

【請求項 4】

前記パターン学習モデルは、時系列パターン、又はダイナミクスを学習するモデルである

請求項 1 に記載の学習装置。

【請求項 5】

前記パターン学習モデルは、HMM、RNN、FNN、SVR、又はRNNPBである

請求項 1 に記載の学習装置。

【請求項 6】

前記モデルパラメータ共有手段は、複数の学習モジュールのすべて、又は一部に、前記モデルパラメータを共有させる

請求項 1 に記載の学習装置。

【請求項 7】

前記モデルパラメータ共有手段は、前記複数の学習モジュールの 2 以上の学習モジュールに、前記複数のモデルパラメータのすべて、又は一部を共有させる

請求項 1 に記載の学習装置。

【請求項 8】

前記モデルパラメータ共有手段は、前記 2 以上の学習モジュールの各学習モジュールが更新したモデルパラメータを、前記 2 以上の学習モジュールそれぞれが更新したモデルパラメータの重み付け平均値によって補正することにより、前記 2 以上の学習モジュールに、その 2 以上の学習モジュールそれぞれが更新したモデルパラメータを共有させる

請求項 1 に記載の学習装置。

【請求項 9】

複数の学習モジュールそれぞれにおいて、パターンを学習するパターン学習モデルが有する複数のモデルパラメータを、入力データを用いて更新する更新学習を行い、

前記複数の学習モジュールの 2 以上の学習モジュールに、前記モデルパラメータを共有させ、

前記パターンを学習させるための新たな学習データが前記入力データとして供給された場合に、前記新たな学習データに対応する新たな学習モジュールを作成し、

10

20

30

40

50

前記新たな学習モジュールを含む全ての学習モジュールで前記更新学習が行われた後、学習モジュール間の類似性を評価し、

前記学習モジュール間の類似性に基づいて、前記学習モジュールの統合の可否を判定し、前記学習モジュールの統合を行うステップを含む学習方法。

【請求項 10】

コンピュータを、

パターンを学習するパターン学習モデルが有する複数のモデルパラメータを、入力データを用いて更新する更新学習を行う複数の学習モジュールと、

前記複数の学習モジュールの 2 以上の学習モジュールに、前記モデルパラメータを共有させるモデルパラメータ共有手段と、

前記パターンを学習させるための新たな学習データが前記入力データとして供給された場合に、前記新たな学習データに対応する新たな学習モジュールを作成するモジュール作成手段と、

前記新たな学習モジュールを含む全ての学習モジュールで前記更新学習が行われた後、学習モジュール間の類似性を評価する類似性評価手段と、

前記学習モジュール間の類似性に基づいて、前記学習モジュールの統合の可否を判定し、前記学習モジュールの統合を行うモジュール統合手段と

して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、学習装置、学習方法、およびプログラムに関し、特に、規模拡張性があり、かつ、汎化特性を有するパターン学習モデルを得ることができるようにする学習装置、学習方法、およびプログラムに関する。

【背景技術】

【0002】

パターンを学習するパターン学習モデルとしては、例えば、RNN(Recurrent Neural Network)やRNNPB(Recurrent Neural Net with Parametric Bias)等がある。このようなパターン学習モデルの学習のスキームは、「局所表現」スキームと「分散表現」スキームとに分類される。

【0003】

「局所表現」スキームでは、複数のパターンが、パターン学習モデルの学習（パターン学習モデルが有するモデルパラメータの更新）を行う複数の学習モジュールにおいてそれぞれ学習され、これにより、1つの学習モジュールで、1つのパターンが記憶される。

【0004】

また、「分散表現」スキームでは、複数のパターンが、1つの学習モジュールにおいて学習され、これにより、1つの学習モジュールで、複数のパターンが同時に記憶される。

【0005】

「局所表現」スキームでは、1つの学習モジュールで、1つのパターンが記憶される、つまり、1つのパターン学習モデルが、1つのパターンを学習するので、ある学習モジュールと、他の学習モジュールとの間で、パターンの記憶の干渉が少なく、パターンの記憶の安定性が高い。そして、「局所表現」スキームは、学習モジュールを追加することにより、新たなパターンを学習することが容易に行うことができるという規模拡張性に優れる。

【0006】

但し、「局所表現」スキームでは、1つのパターン学習モデルが、1つのパターンを学習する、つまり、パターンの記憶が、複数の学習モジュールでそれぞれ独立に行われるため、複数の学習モジュールそれぞれでのパターンの記憶間の関係性の構造化（共通化）による汎化特性を得ること、すなわち、例えば、ある学習モジュールに記憶されたパターン

10

20

30

40

50

とも異なり、他の学習モジュールに記憶されたパターンとも異なるが、それらの両方のパターンの、いわば中間のパターンを生成するようなこと等はできない。

【 0 0 0 7 】

一方、「分散表現」スキームでは、1つの学習モジュールで、複数のパターンが記憶される、つまり、1つのパターン学習モデルが、複数のパターンを学習するので、1つの学習モジュール内での複数のパターンの記憶間の干渉によって、その記憶間の共通化による汎化特性を得ることができる。

【 0 0 0 8 】

但し、「分散表現」スキームでは、パターンの記憶の安定性が低いため、規模拡張性がない。

10

【 0 0 0 9 】

ここで、特許文献1には、あるパターンを学習するRNNと、そのパターンと相関がある他のパターンを学習するRNNとの2つのRNNのコンテキストどうしの誤差に基づいて、その2つのRNNのコンテキストを変化させて、RNNの学習を行うこと、及び、その学習後の2つのRNNのうち的一方のコンテキストを、他方のRNNのコンテキストとして用いて、つまり、一方のRNNのコンテキストを、他方のRNNのコンテキストに影響させて、出力データを生成すること（RNNの入力層に入力データを入力し、その入力データに対応する出力データを、RNNの出力層から出力すること）が記載されている。

【 0 0 1 0 】

また、非特許文献1には、言語のパターンを学習するRNNPBと、動作のパターンを学習するRNNPBとの2つのRNNPBのPBどうしの差に基づいて、その2つのRNNPBをPBを変化させて、RNNPBの学習を行うこと、及び、その学習後の2つのRNNPBのうち一方のPBを、他方のPBに影響させて、出力データを生成することが記載されている。

20

【 0 0 1 1 】

【特許文献1】特開2002-024795号公報

【非特許文献1】Yuuya Sugita, Jun Tani, "Learning Semantic Combinatoriality from the Interaction between Linguistic and Behavioral Processes", Adaptive Behavior, Vol. 13, No. 1, 33-52 (2005)

【発明の開示】

【発明が解決しようとする課題】

30

【 0 0 1 2 】

以上のように、従来のパターン学習モデルの学習では、規模拡張性があるパターン学習モデルを得るか、又は汎化特性を有するパターン学習モデルを得ることはできるが、規模拡張性があり、同時に、汎化特性を有するパターン学習モデルを得ることは困難である。

【 0 0 1 3 】

本発明は、このような状況に鑑みてなされたものであり、規模拡張性があり、同時に、汎化特性を有するパターン学習モデルを得ることができるようにするものである。

【課題を解決するための手段】

【 0 0 1 4 】

本発明の一側面の学習装置は、パターンを学習するパターン学習モデルが有する複数のモデルパラメータを、入力データを用いて更新する更新学習を行う複数の学習モジュールと、前記複数の学習モジュールの2以上の学習モジュールに、前記モデルパラメータを共有させるモデルパラメータ共有手段と、パターンを学習させるための新たな学習データが前記入力データとして供給された場合に、前記新たな学習データに対応する新たな学習モジュールを作成するモジュール作成手段と、前記新たな学習モジュールを含む全ての学習モジュールで前記更新学習が行われた後、学習モジュール間の類似性を評価する類似性評価手段と、前記学習モジュール間の類似性に基づいて、前記学習モジュールの統合の可否を判定し、前記学習モジュールの統合を行うモジュール統合手段とを備える。

40

【 0 0 1 5 】

本発明の一側面の学習方法は、複数の学習モジュールそれぞれにおいて、パターンを学

50

習するパターン学習モデルが有する複数のモデルパラメータを、入力データを用いて更新する更新学習を行い、前記複数の学習モジュールの2以上の学習モジュールに、前記モデルパラメータを共有させ、パターンを学習させるための新たな学習データが前記入力データとして供給された場合に、前記新たな学習データに対応する新たな学習モジュールを作成し、前記新たな学習モジュールを含む全ての学習モジュールで前記更新学習が行われた後、学習モジュール間の類似性を評価し、前記学習モジュール間の類似性に基づいて、前記学習モジュールの統合の可否を判定し、前記学習モジュールの統合を行うステップを含む学習方法である。

【0016】

本発明の一側面のプログラムは、コンピュータを、パターンを学習するパターン学習モデルが有する複数のモデルパラメータを、入力データを用いて更新する更新学習を行う複数の学習モジュールと、前記複数の学習モジュールの2以上の学習モジュールに、前記モデルパラメータを共有させるモデルパラメータ共有手段と、パターンを学習させるための新たな学習データが前記入力データとして供給された場合に、前記新たな学習データに対応する新たな学習モジュールを作成するモジュール作成手段と、前記新たな学習モジュールを含む全ての学習モジュールで前記更新学習が行われた後、学習モジュール間の類似性を評価する類似性評価手段と、前記学習モジュール間の類似性に基づいて、前記学習モジュールの統合の可否を判定し、前記学習モジュールの統合を行うモジュール統合手段として機能させる。

【0017】

本発明の一側面においては、複数の学習モジュールそれぞれにおいて、パターンを学習するパターン学習モデルが有する複数のモデルパラメータを、入力データを用いて更新する更新学習が行われ、複数の学習モジュールの2以上の学習モジュールにおいて、モデルパラメータが共有される。また、パターンを学習させるための新たな学習データが入力データとして供給された場合に、新たな学習データに対応する新たな学習モジュールが作成され、新たな学習モジュールを含む全ての学習モジュールで更新学習が行われた後、学習モジュール間の類似性が評価され、学習モジュール間の類似性に基づいて、学習モジュールの統合の可否が判定され、学習モジュールの統合が行われる。

【発明の効果】

【0018】

本発明の一側面によれば、規模拡張性があり、同時に、汎化特性を有するパターン学習モデルを得ることができる。

【発明を実施するための最良の形態】

【0019】

図1は、本発明を適用する学習装置の基本となる学習装置の一実施の形態の構成例を示している。

【0020】

図1において、学習装置は、複数であるN個の学習モジュール 10_1 ないし 10_N と、モデルパラメータ共有部20とから構成される。

【0021】

学習モジュール 10_i ($i=1, 2, \dots, N$)は、パターン入力部 11_i 、モデル学習部 12_i 、及びモデル記憶部 13_i から構成され、入力データを用いて、パターン学習モデルの複数のモデルパラメータ(学習リソース)を更新する更新学習を行う。

【0022】

すなわち、パターン入力部 11_i には、モデル記憶部 13_i に記憶されたパターン学習モデルに獲得(学習)させるパターン(カテゴリ)の入力データが、パターン学習モデルの学習に用いる学習データとして供給される。

【0023】

パターン入力部 11_i は、そこに供給される学習データを、パターン学習モデルの学習に適切な形のデータにする処理をして、モデル学習部 12_i に供給する。すなわち、例え

10

20

30

40

50

ば、学習データが時系列のデータである場合には、パターン入力部 1 1_i は、例えば、その時系列のデータを、固定の長さに区切って、モデル学習部 1 2_i に供給する。

【 0 0 2 4 】

モデル学習部 1 2_i は、パターン入力部 1 1_i からの学習データを用いて、モデル記憶部 1 3_i に記憶されたパターン学習モデルの複数のモデルパラメータを更新する更新学習を行う。

【 0 0 2 5 】

モデル記憶部 1 3_i は、複数のモデルパラメータを有し、パターンを学習するパターン学習モデルを記憶する。すなわち、モデル記憶部 1 3_i は、パターン学習モデルの複数のモデルパラメータを記憶する。

10

【 0 0 2 6 】

ここで、パターン学習モデルとしては、例えば、時系列のパターンである時系列パターンや、時間変化する力学系を表すダイナミクスを学習（獲得）（記憶）するモデル等を採用することができる。

【 0 0 2 7 】

時系列パターンを学習するモデルとしては、例えば、HMM(Hidden Markov Model)等があり、ダイナミクスを学習するモデルとしては、例えば、RNN, FNN(Feed Forward Neural Network), RNNPB等のニューラルネットワークや、SVR(Support Vector Regression)等がある。

【 0 0 2 8 】

20

例えば、HMMについては、HMMにおいて状態が遷移する確率を表す状態遷移確率や、状態が遷移するときに、HMMからある観測値が出力される確率を表す出力確率、又は確率密度を表す出力確率密度関数が、HMMのモデルパラメータである。

【 0 0 2 9 】

また、例えば、ニューラルネットワークについては、ニューロンに相当するユニット（ノード）において、他のユニットからの入力に付されるウエイト（重み）が、ニューラルネットワークのモデルパラメータである。

【 0 0 3 0 】

なお、HMMの状態遷移確率や、出力確率、又は出力確率密度関数、ニューラルネットワークのウエイトは、いずれも複数存在する。

30

【 0 0 3 1 】

モデルパラメータ共有部 2 0 は、N個の学習モジュール 1 0₁ないし 1 0_Nのうちの、2以上の学習モジュールに、モデルパラメータを共有させる共有処理を行う。モデルパラメータ共有部 2 0 が共有処理を行うことにより、N個の学習モジュール 1 0₁ないし 1 0_Nのうちの、2以上の学習モジュールは、モデルパラメータを共有する。

【 0 0 3 2 】

なお、以下では、説明を簡単にするため、モデルパラメータ共有部 2 0 は、N個の学習モジュール 1 0₁ないし 1 0_Nのすべてに、モデルパラメータを共有させる共有処理を行うこととする。

【 0 0 3 3 】

40

次に、図 2 のフローチャートを参照して、図 1 の学習装置が行う、パターン学習モデルを学習する学習処理について説明する。

【 0 0 3 4 】

ステップ S 1 1 において、学習モジュール 1 0_i のモデル学習部 1 2_i は、モデル記憶部 1 3_i に記憶されたモデルパラメータを、例えば、乱数等によって初期化して、処理は、ステップ S 1 2 に進む。

【 0 0 3 5 】

ステップ S 1 2 では、学習モジュール 1 0_i が、その学習モジュール 1 0_i で学習すべき学習データが供給（入力）されるのを待って、その学習データを用いて、モデルパラメータを更新する更新学習を行う。

50

【 0 0 3 6 】

すなわち、ステップ S 1 2 では、学習モジュール 1 0_i において、パターン入力部 1 1_i が、学習モジュール 1 0_i に供給された学習データを、必要に応じて処理し、モデル学習部 1 2_i に供給する。

【 0 0 3 7 】

さらに、ステップ S 1 2 では、モデル学習部 1 2_i が、パターン入力部 1 1_i からの学習データを用いて、モデル記憶部 1 3_i に記憶されたパターン学習モデルの複数のモデルパラメータを更新する更新学習を行い、その更新学習によって得られた新たな複数のモデルパラメータによって、モデル記憶部 1 3_i の記憶内容を更新する（上書きする）。

【 0 0 3 8 】

ここで、ステップ S 1 1 及び S 1 2 の処理は、N 個の学習モジュール 1 0₁ ないし 1 0_N のすべてで行われる。

【 0 0 3 9 】

ステップ S 1 2 の後、処理は、ステップ S 1 3 に進み、モデルパラメータ共有部 2 0 は、N 個の学習モジュール 1 0₁ ないし 1 0_N のすべてに、モデルパラメータを共有させる共有処理を行う。

【 0 0 4 0 】

すなわち、学習モジュール 1 0_i が有する複数のモデルパラメータのうちの、例えば、m 番目のモデルパラメータに注目すると、モデルパラメータ共有部 2 0 は、N 個の学習モジュール 1 0₁ ないし 1 0_N それぞれの m 番目のモデルパラメータに基づいて、学習モジュール 1 0₁ の m 番目のモデルパラメータを補正する。

【 0 0 4 1 】

さらに、モデルパラメータ共有部 2 0 は、N 個の学習モジュール 1 0₁ ないし 1 0_N それぞれの m 番目のモデルパラメータに基づいて、学習モジュール 1 0₂ の m 番目のモデルパラメータを補正し、以下、同様にして、学習モジュール 1 0₃ ないし 1 0_N それぞれの m 番目のモデルパラメータを補正する。

【 0 0 4 2 】

以上のように、モデルパラメータ共有部 2 0 が、学習モジュール 1 0_i の m 番目のモデルパラメータを、N 個の学習モジュール 1 0₁ ないし 1 0_N それぞれの m 番目のモデルパラメータに基づいて補正することで、N 個の学習モジュール 1 0₁ ないし 1 0_N の m 番目のモデルパラメータのそれぞれは、N 個の学習モジュール 1 0₁ ないし 1 0_N の m 番目のモデルパラメータのすべての影響を受ける（N 個の学習モジュール 1 0₁ ないし 1 0_N の m 番目のモデルパラメータのそれぞれに、N 個の学習モジュール 1 0₁ ないし 1 0_N の m 番目のモデルパラメータのすべてに影響させる）。

【 0 0 4 3 】

このように、複数の学習モジュールのモデルパラメータすべてを、その複数の学習モジュールのモデルパラメータのそれぞれに影響させること（複数の学習モジュールのモデルパラメータのそれぞれが、その複数の学習モジュールのモデルパラメータすべての影響を受けること）が、複数の学習モジュールによるモデルパラメータの共有である。

【 0 0 4 4 】

モデルパラメータ共有部 2 0 は、ステップ S 1 3 において、学習モジュール 1 0_i のモデル記憶部 1 3_i に記憶された複数のモデルパラメータのすべてを対象に、共有処理を行い、その共有処理によって得られたモデルパラメータによって、モデル記憶部 1 3₁ ないし 1 3_N の記憶内容を更新する。

【 0 0 4 5 】

ステップ S 1 3 の後、処理は、ステップ S 1 4 に進み、図 1 の学習装置は、学習の終了条件が満たされているかどうかを判定する。

【 0 0 4 6 】

ここで、ステップ S 1 4 での学習の終了条件としては、例えば、学習の回数、つまり、ステップ S 1 2 及び S 1 3 が繰り返された回数が、あらかじめ定められた所定の回数とな

10

20

30

40

50

ったことや、あらかじめ用意された学習データのすべてを用いて、ステップS 1 2の更新学習が行われたこと、あるいは、ある入力データに対して出力されるべき出力データの真値が分かっている場合に、その入力データに対してパターン学習モデルから出力される出力データの、真値に対する誤差が所定値以下であること、等を採用することができる。

【0047】

ステップS 1 4において、学習の終了条件が満たされていないと判定された場合、処理は、ステップS 1 2に戻り、以下、同様の処理が繰り返される。

【0048】

また、ステップS 1 4において、学習の終了条件が満たされていると判定された場合、処理は終了する。

【0049】

なお、ステップS 1 2とステップS 1 3は、その処理の順番を逆にすることも可能である。即ち、N個の学習モジュール 10_1 ないし 10_N のすべてに、モデルパラメータを共有させる共有処理を行った後、モデルパラメータを更新する更新学習を行うようにすることもできる。

【0050】

次に、図3は、パターン学習モデルとして、RNNPBを採用した場合の、図1の学習装置の構成例を示している。

【0051】

なお、図3においては、学習モジュール 10_i のパターン入力部 11_i 及びモデル学習部 12_i の図示を省略してある。

【0052】

モデル記憶部 13_i には、RNNPB（を定義するモデルパラメータ）が記憶されている。ここで、モデル記憶部 13_i に記憶されたRNNPBを、以下、適宜、RNNPB#iとも記載する。

【0053】

RNNPBは、入力層、隠れ層（中間層）、及び出力層により構成されている。入力層、隠れ層、及び出力層は、それぞれ任意の数の、ニューロンに相当するユニットにより構成されている。

【0054】

RNNPBでは、入力層の一部のユニットである入力ユニットに、時系列データ等の入力データ x_t が入力（供給）される。ここで、入力データ x_t としては、例えば、画像や音声の特徴量や、ロボットの手や足に相当する部分の動きの軌道等を採用することができる。

【0055】

また、入力層の、入力データ x_t が入力される入力ユニット以外のユニットの一部であるPBユニットには、PB(Parametric Bias)が入力される。PBによれば、同一の状態のRNNPBに対して、同一の入力データ x_t が入力されても、PBを変更することにより、異なる出力データ x_{t+1}^* を得ることができる。

【0056】

入力層の、入力データ x_t が入力される入力ユニット以外のユニットの残りであるコンテキストユニットには、出力層の一部のユニットより出力される出力データが、内部状態を表すコンテキストとしてフィードバックされる。

【0057】

ここで、時刻tの入力データ x_t が入力層の入力ユニットに入力されるときに入力層のPBユニットとコンテキストユニットに入力される時刻tのPBとコンテキストを、それぞれ、 P_{B_t} と c_t と記載する。

【0058】

隠れ層のユニットは、入力層に入力される入力データ x_t 、 P_{B_t} 、コンテキスト c_t を対象として、所定のウェイト（重み）を用いた重み付け加算を行い、その重み付け加算の結果を引数とする非線形関数の演算を行って、その演算結果を、出力層のユニットに出力する。

。

10

20

30

40

50

【 0 0 5 9 】

出力層の一部のユニットからは、上述したように、次の時刻 $t+1$ のコンテキスト c_{t+1} となる出力データが出力され、入力層にフィードバックされる。また、出力層の残りのユニットからは、例えば、入力データ x_t に対応する出力データとして、その入力データ x_t の次の時刻 $t+1$ の入力データ x_{t+1} の予測値 x^*_{t+1} が出力される。

【 0 0 6 0 】

ここで、RNNPBでは、ユニットへの入力に重み付け加算されるが、この重み付け加算に用いられるウェイト（重み）が、RNNPBのモデルパラメータである。RNNPBのモデルパラメータとしてのウェイトには、入力ユニットから隠れ層のユニットへのウェイト、PBユニットから隠れ層のユニットへのウェイト、コンテキストユニットから隠れ層のユニットへのウェイト、隠れ層のユニットから出力層のユニットへのウェイト、及び、隠れ層のユニットからコンテキストユニットへのウェイトの5種類がある。

【 0 0 6 1 】

パターン学習モデルとして、以上のようなRNNPBを採用した場合、モデルパラメータ共有部20には、RNNPBのモデルパラメータとしてのウェイトを、学習モジュール10₁ないし10_Nに共有させるウェイトマトリクス共有部21が設けられる。

【 0 0 6 2 】

ここで、RNNPBのモデルパラメータとしてのウェイトは、複数あるが、その複数のウェイトをコンポーネントとするマトリクスを、ウェイトマトリクスという。

【 0 0 6 3 】

ウェイトマトリクス共有部21は、モデル記憶部13₁ないし13_Nに記憶されたRNNPB#1ないしRNNPB#Nの複数のモデルパラメータとしてのウェイトマトリクスすべてを、学習モジュール10₁ないし10_Nのそれぞれに共有させる。

【 0 0 6 4 】

すなわち、RNNPB#iのウェイトマトリクスを w_i と表すこととすると、ウェイトマトリクス共有部21は、ウェイトマトリクス w_i を、N個の学習モジュール10₁ないし10_Nそれぞれのウェイトマトリクス w_1 ないし w_N のすべてに基づいて補正することで、ウェイトマトリクス w_i に、ウェイトマトリクス w_1 ないし w_N のすべてを影響させる共有処理を行う。

【 0 0 6 5 】

具体的には、ウェイトマトリクス共有部21は、例えば、次式(1)に従い、RNNPB#iのウェイトマトリクス w_i を補正する。

【 0 0 6 6 】

【数1】

$$w_i = w_i + \Delta w_i \quad \dots (1)$$

【 0 0 6 7 】

ここで、式(1)において、 w_i は、ウェイトマトリクス w_i を補正する補正成分であり、例えば、式(2)に従って求められる。

【 0 0 6 8 】

【数2】

$$\Delta w_i = \alpha_i \sum_{j=1}^N \beta_{ij} (w_j - w_i) \quad \dots (2)$$

【 0 0 6 9 】

式(2)において、 β_{ij} は、RNNPB#iのウェイトマトリクス w_i に、RNNPB#j ($j=1, 2, \dots, N$)のウェイトマトリクス w_j に影響させる度合いを表す係数である。

【 0 0 7 0 】

したがって、式(2)の右辺のサメーション $\sum \beta_{ij} (w_j - w_i)$ は、係数 β_{ij} を重みとした、RNNPB#iのウェイトマトリクス w_i に対するRNNPB#1ないしRNNPB#Nのウェイトマトリクス w_1 ないし w_N それぞれの偏差（差分）の重み付け平均値を表し、 α_i は、その重み付け平均値

10

20

30

40

50

$\sum \beta_{ij}(w_j - w_i)$ を、ウエイトマトリクス w_i に影響させる度合いを表す係数である。

【0071】

係数 α_i 及び β_{ij} としては、例えば、0.0より大で1.0より小の値を採用することができる。

【0072】

式(2)によれば、係数 α_i が小であるほど、いわば共有が弱くなり(ウエイトマトリクス w_i が受ける重み付け平均値 $\sum \beta_{ij}(w_j - w_i)$ の影響が小さくなり)、係数 α_i が大であるほど、いわば共有が強まる。

【0073】

なお、ウエイトマトリクス w_i の補正の方法は、式(1)に限定されるものではなく、例えば、式(3)に従って行うことが可能である。

【0074】

【数3】

$$w_i = \alpha'_i \cdot w_i + (1 - \alpha'_i) \cdot \sum_{j=1}^N \beta'_{ij} \cdot w_j \quad \cdots (3)$$

【0075】

ここで、式(3)において、 β'_{ij} は、RNNPB#iのウエイトマトリクス w_i に、RNNPB#j(j=1, 2, ..., N)のウエイトマトリクス w_j に影響させる度合いを表す係数である。

【0076】

したがって、式(3)の右辺の第2項におけるサメーション $\sum \beta'_{ij} \cdot w_j$ は、係数 β'_{ij} を重みとした、RNNPB#1ないしRNNPB#Nのウエイトマトリクス w_1 ないし w_N の重み付け平均値を表し、 α'_i は、その重み付け平均値 $\sum \beta'_{ij} \cdot w_j$ を、ウエイトマトリクス w_i に影響させる度合いを表す係数である。

【0077】

係数 α'_i 及び β'_{ij} としては、例えば、0.0より大で1.0より小の値を採用することができる。

【0078】

式(3)によれば、係数 α'_i が大であるほど、共有が弱くなり(ウエイトマトリクス w_i が受ける重み付け平均値 $\sum \beta'_{ij} \cdot w_j$ の影響が小さくなり)、係数 α'_i が小であるほど、共有が強まる。

【0079】

次に、図4のフローチャートを参照して、パターン学習モデルとして、RNNPBを採用した場合の、図1の学習装置の学習処理について説明する。

【0080】

ステップS21において、学習モジュール10_iのモデル学習部12_iは、モデル記憶部13_iに記憶されたRNNPB#iのモデルパラメータであるウエイトマトリクス w_i を、例えば、乱数等によって初期化して、処理は、ステップS22に進む。

【0081】

ステップS22では、学習モジュール10_iが、その学習モジュール10_iで学習すべき学習データ x_t が入力されるのを待って、その学習データ x_t を用いて、モデルパラメータを更新する更新学習を行う。

【0082】

すなわち、ステップS22では、学習モジュール10_iにおいて、パターン入力部11_iが、学習モジュール10_iに供給された学習データ x_t を、必要に応じて処理し、モデル学習部12_iに供給する。

【0083】

さらに、ステップS22では、モデル学習部12_iが、パターン入力部11_iからの学習データ x_t を用いて、モデル記憶部13_iに記憶されたRNNPB#iのウエイトマトリクス w_i を更新する更新学習を、例えば、BPTT(Back-Propagation Through Time)法により行い、その

10

20

30

40

50

更新学習によって得られた新たなモデルパラメータとしてのウエイトマトリクス w_i によって、モデル記憶部 13_iの記憶内容を更新する。

【0084】

ここで、ステップS21及びS22の処理は、N個の学習モジュール10₁ないし10_Nのすべてで行われる。

【0085】

また、BPTT法については、例えば、特開2002-236904号公報等に記載されている。

【0086】

ステップS22の後、処理は、ステップS23に進み、モデルパラメータ共有部20のウエイトマトリクス共有部21は、N個の学習モジュール10₁ないし10_Nのすべてに、ウエイトマトリクス w_1 ないし w_N のすべてを共有させる共有処理を行う。

10

【0087】

すなわち、ステップS23において、ウエイトマトリクス共有部21は、例えば、式(2)に従い、モデル記憶部13₁ないし13_Nに記憶されたウエイトマトリクス w_1 ないし w_N を用いて補正成分 w_1 ないし w_N をそれぞれ求め、その補正成分 w_1 ないし w_N により、モデル記憶部13₁ないし13_Nに記憶されたウエイトマトリクス w_1 ないし w_N を、式(1)に従ってそれぞれ補正する。

【0088】

ステップS23の後、処理は、ステップS24に進み、図1の学習装置は、学習の終了条件が満たされているかどうかを判定する。

20

【0089】

ここで、ステップS24での学習の終了条件としては、例えば、学習の回数、つまり、ステップS22及びS23が繰り返された回数が、あらかじめ定められた所定の回数となったことや、ある入力データ x_t に対してRNNPB#iが出力する出力データ x^*_{t+1} 、すなわち、入力データ x_{t+1} の予測値 x^*_{t+1} の、入力データ x_{t+1} に対する誤差が所定値以下であること、等を採用することができる。

【0090】

ステップS24において、学習の終了条件が満たされていないと判定された場合、処理は、ステップS22に戻り、以下、同様の処理、すなわち、ウエイトマトリクス w_i の更新学習と、共有処理とが交互に繰り返される。

30

【0091】

また、ステップS24において、学習の終了条件が満たされていると判定された場合、処理は終了する。

【0092】

なお、図4においても、ステップS22とステップS23は、その処理の順番を逆にすることが可能である。

【0093】

以上のように、規模拡張性に優れた複数の学習モジュール10₁ないし10_Nそれぞれにおいて、モデルパラメータを共有しながら、その複数の学習モジュール10₁ないし10_Nそれぞれのモデルパラメータを更新する更新学習を行うことにより、1つの学習モジュールだけで行われる学習で得られる汎化特性が、複数の学習モジュール10₁ないし10_Nの全体で得ることができ、その結果、規模拡張性があり、同時に、汎化特性を有するパターン学習モデルを得ることができる。

40

【0094】

すなわち、多くのパターンを獲得(記憶)することができ、かつ、複数のパターンの共通性を獲得することができる。さらに、複数のパターンの共通性を獲得することで、その共通性に基づいて、未学習のパターンの認識や生成を行うことが可能となる。

【0095】

具体的には、学習データとして、例えば、N種類の音韻の音声データを、N個の学習モジュール10₁ないし10_Nにそれぞれ与えて、パターン学習モデルの学習を行った場合に、

50

そのパターン学習モデルによれば、学習に用いられていない時系列パターンの音声データの認識や生成を行うことができる。さらに、例えば、学習データとして、例えば、ロボットのアームを駆動するための、N種類の駆動データを、N個の学習モジュール10₁ないし10_Nにそれぞれ与えて、パターン学習モデルの学習を行った場合に、そのパターン学習モデルによれば、学習に用いられていない時系列パターンの駆動データの生成を行うことができ、その結果、ロボットは、教えられていないアームの動きをすることが可能となる。

【0096】

また、学習後のパターン学習モデルによれば、パターン学習モデル間のモデルパラメータ（リソース）どうしの距離に基づいて、パターン学習モデルどうしの類似性を評価し、類似性が高いパターン学習モデルどうしをクラスタとして、パターンのクラスタリングを行うことができる。

10

【0097】

次に、図5ないし図9を参照して、本件発明者が行った、図1の学習装置による学習処理（以下、適宜、共有学習処理という）のシミュレーションの結果について説明する。

【0098】

図5は、共有学習処理で学習を行ったパターン学習モデルについての各データを示している。

【0099】

なお、シミュレーションでは、パターン学習モデルとして、2つのPBが入力層に入力され、3つのコンテキストが入力層にフィードバックされる9個のRNNPB#1ないしRNNPB#9を採用し、学習データとして、3つのパターンP#1,P#2,P#3の時系列データのそれぞれに、異なる3つのノイズN#1,N#2,N#3のそれぞれを重畳して得られる9個の時系列データを用いた。

20

【0100】

また、RNNPB#1には、パターンP#1の時系列データにノイズN#1を重畳して得られる時系列データを、RNNPB#2には、パターンP#1の時系列データにノイズN#2を重畳して得られる時系列データを、RNNPB#3には、パターンP#1の時系列データにノイズN#3を重畳して得られる時系列データを、それぞれ、学習データとして与えた。

【0101】

同様に、RNNPB#4には、パターンP#2の時系列データにノイズN#1を重畳して得られる時系列データを、RNNPB#5には、パターンP#2の時系列データにノイズN#2を重畳して得られる時系列データを、RNNPB#6には、パターンP#2の時系列データにノイズN#3を重畳して得られる時系列データを、それぞれ、学習データとして与え、RNNPB#7には、パターンP#3の時系列データにノイズN#1を重畳して得られる時系列データを、RNNPB#8には、パターンP#3の時系列データにノイズN#2を重畳して得られる時系列データを、RNNPB#9には、パターンP#3の時系列データにノイズN#3を重畳して得られる時系列データを、それぞれ、学習データとして与えた。

30

【0102】

なお、更新学習は、入力データ x_t に対してRNNPBが出力する出力データとしての、入力データ x_{t+1} の予測値 x_{t+1}^* の、入力データ x_{t+1} に対する誤差（予測誤差）を小さくするように行った。

40

【0103】

図5上から1番目は、学習後のRNNPB#1ないしRNNPB#9に、学習時に与えられた学習データを入力データとして与えたときに、RNNPB#1ないしRNNPB#9それぞれが出力する出力データ(output)と、その出力データの予測誤差(error)とを示している。

【0104】

図5上から1番目において、予測誤差は、ほぼ0になっており、したがって、RNNPB#1ないしRNNPB#9は、入力データ、つまり、学習時に与えられた学習データとほぼ一致する出力データを出力する。

50

【 0 1 0 5 】

図 5 上から 2 番目は、学習後のRNNPB#1ないしRNNPB#9が、図 5 上から 1 番目に示した出力データを出力するときの、3 つのコンテキストの時間変化を示している。

【 0 1 0 6 】

また、図 5 上から 3 番目は、学習後のRNNPB#1ないしRNNPB#9が、図 5 上から 1 番目に示した出力データを出力するときの、2 つのPB（以下、適宜、2 つのPBそれぞれを、PB#1，PB#2と記載する）の時間変化を示している。

【 0 1 0 7 】

図 6 は、学習後のRNNPB#1ないしRNNPB#9のうちの、例えば、5 番目のRNNPB#5が、各値のPB#1，PB#2に対して出力する出力データを示している。

10

【 0 1 0 8 】

なお、図 6 では、横軸がPB#1を表し、縦軸がPB#2を表している。

【 0 1 0 9 】

図 6 によれば、RNNPB#5は、PB#1が0.6程度のときに、学習時に与えられた学習データとほぼ一致する出力データを出力しており、これにより、RNNPB#5は、学習時に与えられた学習データのパターンP#2を獲得していることが分かる。

【 0 1 1 0 】

また、RNNPB#5は、PB#1が0.6より小さいときに、RNNPB#1ないしRNNPB#3に学習させたパターンP#1や、RNNPB#7ないしRNNPB#9に学習させたパターンP#3に類似する時系列データを出力しており、これにより、RNNPB#5が、RNNPB#1ないしRNNPB#3が獲得したパターンP#1や、RNNPB#7ないしRNNPB#9が獲得したパターンP#3の影響を受け、いわば、RNNPB#5に対して学習時に与えられた学習データのパターンP#2が、RNNPB#1ないしRNNPB#3が獲得したパターンP#1や、RNNPB#7ないしRNNPB#9が獲得したパターンP#3の方向に変形していくときに現れる中間のパターンをも獲得していることが分かる。

20

【 0 1 1 1 】

さらに、RNNPB#5は、PB#1が0.6より大であるときに、9 個のRNNPB#1ないしRNNPB#9のいずれにも学習させていないパターンの時系列データを出力しており、これにより、RNNPB#5が、RNNPB#1ないしRNNPB#3が獲得したパターンP#1や、RNNPB#7ないしRNNPB#9が獲得したパターンP#3の影響を受け、いわば、RNNPB#5に対して学習時に与えられた学習データのパターンP#2が、RNNPB#1ないしRNNPB#3が獲得したパターンP#1や、RNNPB#7ないしRNNPB#9が獲得したパターンP#3の方向とは逆方向に変形していくときに現れるパターンをも獲得していることが分かる。

30

【 0 1 1 2 】

次に、図 7 は、9 個のRNNPB#1ないしRNNPB#9それぞれのウェイトマトリクスどうしの相関としての距離、すなわち、例えば、ウェイトマトリクスを構成する各ウェイトをコンポーネントとするベクトルの空間における、そのベクトルどうしの距離を表す長方形のマップを示している。

【 0 1 1 3 】

なお、ウェイトマトリクスどうしの距離が小であるほど、その 2 つのウェイトマトリクスどうしの相関が高い。

40

【 0 1 1 4 】

図 7 のマップでは、横軸と縦軸のそれぞれに、9 個のRNNPB#1ないしRNNPB#9それぞれのウェイトマトリクスをとって、その横軸のウェイトマトリクスと、縦軸のウェイトマトリクスとの距離が、濃淡で示されており、濃い（黒い）部分ほど、距離が小であることを表す（淡い（白い）部分ほど、距離が大であることを表す）。

【 0 1 1 5 】

図 7 において、横×縦が5×3個のマップのうちの、左上のマップは、学習の回数が0回目であるときのウェイトマトリクスどうしの距離、すなわち、初期化がされたウェイトマトリクスどうしの距離を示しており、マップにおいて、対角線上に並ぶ、同一のRNNPB#iのウェイトマトリクスどうしの距離だけが小になっている。

50

【 0 1 1 6 】

以下、図 7 では、右に行くほど、そして、下に行くほど、学習が進行したときのマップが示されており、右下のマップが、学習の回数が1400回目であるときのウエイトマトリクスどうしの距離を示している。

【 0 1 1 7 】

図 7 によれば、学習が進行するにつれ、同一のパターンP#1の時系列データを学習したRNNPB#1ないしRNNPB#3のウエイトマトリクスどうしの距離、同一のパターンP#2の時系列データを学習したRNNPB#4ないしRNNPB#6のウエイトマトリクスどうしの距離、及び、同一のパターンP#3の時系列データを学習したRNNPB#7ないしRNNPB#9のウエイトマトリクスどうしの距離が小さくなることが分かる。

10

【 0 1 1 8 】

図 8 は、図 5 ないし図 7 の場合とは別の時系列データを用いて学習を行ったRNNPBのウエイトマトリクスどうしの相関としての距離を表す、図 7 と同様のマップを示している。

【 0 1 1 9 】

なお、図 8 のマップを作成するシミュレーションでは、図 9 に示す 5 種類のパターンP#1,P#2,P#3,P#4,P#5の時系列データのそれぞれに、異なる 4 つのノイズN#1,N#2,N#3,N#4のそれぞれを重畳して得られる 2 0 個の時系列データを用意し、各時系列データを、1 つのRNNPBに学習させた。したがって、図 8 のマップを作成するシミュレーションで用いたRNNPBは、2 0 個のRNNPB#1ないしRNNPB#20である。

【 0 1 2 0 】

20

また、学習では、RNNPB#1ないしRNNPB#4には、パターンP#1の時系列データを、RNNPB#5ないしRNNPB#8には、パターンP#2の時系列データを、RNNPB#9ないしRNNPB#12には、パターンP#3の時系列データを、RNNPB#13ないしRNNPB#16には、パターンP#4の時系列データを、RNNPB#17ないしRNNPB#20には、パターンP#5の時系列データを、それぞれ与えた。

【 0 1 2 1 】

図 8 左の 5×3 個のマップは、共有が弱い場合、すなわち、20個のRNNPB#1ないしRNNPB#20のウエイトマトリクス w_1 ないし w_{20} それぞれに、その 2 0 個のウエイトマトリクス w_1 ないし w_{20} のすべてを影響させる度合いが小さい場合、具体的には、式 (2) の係数 α_i が小である場合 (α_i がほぼ0である場合) のマップを示している。

【 0 1 2 2 】

30

また、図 8 右の 5×3 個のマップは、共有が強い場合、すなわち、20個のRNNPB#1ないしRNNPB#20のウエイトマトリクス w_1 ないし w_{20} それぞれに、その 2 0 個のウエイトマトリクス w_1 ないし w_{20} のすべてを影響させる度合いが大きい場合、具体的には、式 (1) の係数 α_i が小でない場合のマップを示している。

【 0 1 2 3 】

共有が弱い場合も強い場合も、学習の回数が0回目であるときの左上のマップでは、対角線上に並ぶ、同一のRNNPB#iのウエイトマトリクスどうしの距離だけが小になっている。

【 0 1 2 4 】

40

そして、共有が弱い場合には、図 8 左に示すように、学習が進行しても、ウエイトマトリクスどうしの距離に、特に傾向は現れないが、共有が強い場合には、図 8 右に示すように、同一のパターンの時系列データを学習したRNNPBの間で、ウエイトマトリクスどうしの距離が小さくなることが分かる。

【 0 1 2 5 】

したがって、共有処理によって、複数の学習モジュールをまたいで分散表現が形成され、複数のRNNPBが汎化特性を有するようになっていくことが分かる。

【 0 1 2 6 】

なお、モデル学習部 1 2_iによるモデルパラメータの更新学習の方法、及び、モデルパラメータ共有部 2 0 による共有処理の方法は、上述した方法に限定されるものではない。

【 0 1 2 7 】

50

また、本実施の形態では、モデルパラメータ共有部 20 による共有処理において、N個の学習モジュール 10_1 ないし 10_N のすべてに、モデルパラメータとしてのウエイトマトリクスを共有させるようにしたが、その他、例えば、N個の学習モジュール 10_1 ないし 10_N のうちの一部だけに、モデルパラメータとしてのウエイトマトリクスを共有させることが可能である。

【0128】

さらに、本実施の形態では、モデルパラメータ共有部 20 による共有処理において、学習モジュール 10_i に、複数のモデルパラメータとしての、ウエイトマトリクスを構成する複数のウエイトすべてを共有させるようにしたが、共有処理では、ウエイトマトリクスを構成する複数のウエイトすべてではなく、そのうちの一部のウエイトだけを共有させるようにすることが可能である。

10

【0129】

また、N個の学習モジュール 10_1 ないし 10_N のうちの一部だけに、ウエイトマトリクスを構成する複数のウエイトのうちの一部のウエイトだけを共有させることも可能である。

【0130】

なお、図 1 の学習装置は、モデルパラメータ共有部 20 が、複数の学習モジュール 10_1 ないし 10_N に、モデルパラメータを共有させる、つまり、各学習モジュール 10_i におけるパターン学習モデルとしての RNNPB#i のモデルパラメータであるウエイトマトリクス w_i に、学習モジュール 10_1 ないし 10_N それぞれにおける RNNPB#1 ないし RNNPB#N のウエイトマトリクス w_1 ないし w_N を影響させる点で、RNN の学習時に、2 つの RNN のコンテキストどうしの誤差に基づいて、その 2 つの RNN のコンテキストを変化させる、つまり、各 RNN のコンテキストに、2 つの RNN のコンテキストに影響させる特許文献 1 に記載の技術と共通する。

20

【0131】

しかしながら、図 1 の学習装置では、影響を受けるのが、モデルパラメータであるウエイトマトリクスである点で、モデルパラメータではなく、内部状態であるコンテキストが影響を受ける特許文献 1 に記載の技術と相違する。

【0132】

すなわち、例えば、関数で表現されるパターン学習モデルを例にすれば、パターン学習モデルのモデルパラメータは、そのパターン学習モデルを表現する関数を定義する、学習によって求められる定数（例えば、入力が u と、出力が y と、内部状態が x と、それぞれ表されるシステムをモデル化する状態方程式 $y=Cx+Du$ 、及び $x'=Ax+Bu$ （ x' は x の微分を表す）を例にすれば、 A, B, C, D にあたる）であり、そもそも定数ではない内部状態（状態方程式の例では、内部状態 x ）とは異なる。

30

【0133】

同様に、図 1 の学習装置は、各学習モジュール 10_i におけるパターン学習モデルとしての RNNPB#i のモデルパラメータであるウエイトマトリクス w_i に、学習モジュール 10_1 ないし 10_N それぞれにおける RNNPB#1 ないし RNNPB#N のウエイトマトリクス w_1 ないし w_N を影響させる点で、RNNPB の学習時に、2 つの RNNPB の PB どうしの差に基づいて、その 2 つの RNNPB の PB を変化させる、つまり、各 RNNPB の PB に、2 つの RNNPB の PB を影響させる非特許文献 1 に記載の技術と共通する。

40

【0134】

しかしながら、図 1 の学習装置では、影響を受けるのが、モデルパラメータであるウエイトマトリクスである点で、モデルパラメータではなく、内部状態である（あるいは、内部状態に相当する）PB が影響を受ける非特許文献 1 に記載の技術と相違する。

【0135】

すなわち、上述したように、パターン学習モデルのモデルパラメータとは、パターン学習モデルを表現する関数を定義する、学習によって求められる定数であり、定数ではない内部状態とは異なる。

50

【0136】

そして、モデルパラメータは、パターン学習モデルを表現する関数を定義する、学習によって求められる定数であるがゆえに、学習時は、学習をしようとするパターンに対応する値になるように更新（変更）されるが、出力データを生成するとき（パターン学習モデルとしてのRNNPBの入力層に入力データを入力し、その入力データに対応する出力データを、RNNPBの出力層から出力するとき）には、変更されない。

【0137】

一方、特許文献1に記載の技術が対象としているコンテキスト、及び、非特許文献1に記載の技術が対象としているPBは、モデルパラメータとは異なる内部状態であるがゆえに、学習時は勿論、出力データを生成するときも変更される。

10

【0138】

以上のように、図1の学習装置は、特許文献1及び非特許文献1に記載の技術のいずれとも相違し、その結果、規模拡張性があり、同時に、汎化特性を有するパターン学習モデルを得ることができる。

【0139】

すなわち、図1の学習装置では、例えば、図10に示すように、RNNPBなどのパターン学習モデルのモデルパラメータを共有させる。

【0140】

その結果、図1の学習装置によれば、図11に示すように、規模拡張性に優れるが、汎化特性に欠ける「局所表現」スキームと、汎化特性があるが、規模拡張性に欠ける「分散表現」スキームとの2種類の学習の両方の長所を有する、いわば「中間表現」スキームの学習を行い、規模拡張性があり、同時に、汎化特性を有するパターン学習モデルを得ることができる。

20

【0141】

ところで、学習装置がN個の学習モジュール 10_1 ないし 10_N で構成されている場合に、新たな学習モジュールを追加する必要があると判断するかが難しいという問題がある。

【0142】

即ち、新しい学習データ（学習サンプル）が供給された場合に、その供給された学習データが、既存のパターン学習モデルで学習済みの時系列パターンに類似するものであるのか、既存のN個の学習モジュール 10_1 ないし 10_N のいずれかの学習モジュールに更新学習させれば済むものであるのか、あるいは、N個の学習モジュール 10_1 ないし 10_N で学習済みの時系列パターンのいずれにも類似するものではないのか、新たな学習モジュールを追加して学習させるべきものであるのかを判断することは難しい。

30

【0143】

規模拡張性に優れ、必要に応じて学習モジュールを追加して学習するようにしたモジュール型学習においては、例えば、既存の学習モジュールに対する新しい学習モジュールの新規性を数値で表し、その数値で表された新しい学習モジュールの新規性が所定の閾値を超えた場合に、新しい学習モジュールを追加するという手法が採用されている。

【0144】

しかしながら、上述した手法では、学習モジュールの追加を決定する基準となる閾値の設定が困難であり、設定を間違えると、本来は別々の学習モジュールに学習させるべき学習データを1つの学習モジュールが学習してしまったり、反対に、1つの学習モジュールに学習させるべき学習データを別々の学習モジュールが学習してしまうという問題があった。

40

【0145】

そこで、以下では、上述したモデルパラメータの共有処理によって、類似するモデルパラメータは、その距離が小さくなるという特性を生かして、学習モジュール追加の可否を判断する必要がなく、学習モジュールを追加しての学習（追加学習）ができるとともに、学習モジュール数の不要な増加も抑制することができるようにした実施の形態について説明する。

50

【 0 1 4 6 】

図 1 2 は、本発明を適用した学習装置の一実施の形態の構成例を示している。

【 0 1 4 7 】

図 1 2 において、上述した図 1 の学習装置と対応する部分については同一の符号を付し
てあり、その説明は省略する。

【 0 1 4 8 】

即ち、図 1 2 の学習装置 1 0 1 は、図 1 の学習装置と同様の構成を有するパターン学習
部 1 1 1 と、学習モジュールの管理を行う学習モジュール管理部 1 1 2 とにより構成され
ている。

【 0 1 4 9 】

パターン学習部 1 1 1 は、学習モジュール管理部 1 1 2 により制御された学習モジュ
ール数である N 個の学習モジュール 1 0₁ ないし 1 0_N により、パターン学習モデルの複数の
モデルパラメータ（学習リソース）を学習（更新）する更新学習を行う。

【 0 1 5 0 】

学習モジュール管理部 1 1 2 は、モジュール作成部 1 2 1、類似性評価部 1 2 2、およ
びモジュール統合部 1 2 3 から構成され、パターン学習部 1 1 1 の学習モジュール 1 0₁
ないし 1 0_N の個数（N）を制御する。

【 0 1 5 1 】

モジュール作成部 1 2 1 は、学習装置 1 0 1 のパターン学習部 1 1 1 に新しい学習デー
タが供給された場合に、その新しい学習データに対応する新たな学習モジュールを無条件
にパターン学習部 1 1 1 に作成（追加）する。

【 0 1 5 2 】

類似性評価部 1 2 2 は、パターン学習部 1 1 1 の各学習モジュール間の類似性を評価す
る。学習モジュール間の類似性の評価としては、例えば、学習モジュール間のモデルパラ
メータのユークリッド距離（以下、パラメータ距離と称する）を採用することができる。

【 0 1 5 3 】

具体的には、学習モジュール 1 0₁ と学習モジュール 1 0₂ のパラメータ距離 $D_{parameter}(1,2)$
は、式（4）で計算することができる。なお、式（4）における k は、学習モジ
ュール 1 0₁ および 1 0₂ のモデルパラメータを識別する変数であり、例えば、 $p_{1,k}$ は、
学習モジュール 1 0₁ の k 番目（k = 1, 2, ..., Q）のモデルパラメータを表す。

【 0 1 5 4 】

【数 4】

$$D_{parameter}(1,2) = \sqrt{\sum_{k=1}^Q (p_{1,k} - p_{2,k})^2} \quad \cdots (4)$$

【 0 1 5 5 】

時系列パターンに対してパターン学習モデルのモデルパラメータが一意に決まるような
冗長性の低いパターン学習モデルであれば、パターン学習モデルの類似性の評価にパラメ
ータ距離を用いることは容易に想像できるが、例えば、ニューラルネット（RNN）などの
ように冗長性の高いパターン学習モデルにおいて、パラメータ距離をパターン学習モデル
の類似性の評価に用いることができるのは、上述した共有学習処理により、類似した時系
列パターンを学習する学習モジュール間では、パラメータ距離が小さくなるという特性を
有しているからである。

【 0 1 5 6 】

モジュール統合部 1 2 3 は、類似性評価部 1 2 2 が求めた学習モジュール間の類似性に
基づいて、学習モジュールの統合の可否を判定する。そして、統合することができると判
定した学習モジュールがある場合には、モジュール統合部 1 2 3 は、それらの統合を行う
。

【 0 1 5 7 】

次に、図 1 3 のフローチャートを参照して、図 1 2 の学習装置 1 0 1 による、学習モジ

10

20

30

40

50

ジュールの追加をとまなう学習である追加学習処理について説明する。

【0158】

パターン学習部111に新しい学習データが供給されると、ステップS41において、モジュール作成部121は、新しい学習データに対して新たな学習モジュールをパターン学習部111に作成する。以下では、新たな学習モジュールを追加した後の学習モジュールの個数がN個であるとする。

【0159】

ステップS42において、パターン学習部111は、ステップS41の処理により追加された新しい学習モジュールを含めて、学習処理を行う。この学習処理は、図2を参照して説明した学習処理と同様であるので、その詳細な説明は省略する。

10

【0160】

ステップS43において、学習モジュール管理部112は、学習モジュール間の類似性に基づいて、学習モジュールを統合する統合処理を行う。統合処理の詳細は、図14を参照して後述する。

【0161】

ステップS44において、新しい学習データがあるか、即ち、パターン学習部111に供給された学習データのなかで、まだ学習処理に含めていない学習データがあるかを判定する。新しい学習データがあると判定された場合、処理はステップS41に戻り、ステップS41乃至S44の処理が繰り返される。一方、新しい学習データがないと判定された場合、追加学習処理は終了する。

20

【0162】

次に、図14のフローチャートを参照して、図13のステップS43における統合処理の詳細について説明する。

【0163】

統合処理では、初めに、ステップS61において、類似性評価部122が、学習モジュール間の類似性を評価する。即ち、類似性評価部122は、各学習モジュール間のパラメータ距離を、N個の学習モジュール 10_1 ないし 10_N のすべての組み合わせについて求める。

【0164】

ステップS62において、モジュール統合部123は、類似性評価部122が求めた各学習モジュール間の類似性（学習モジュール間のパラメータ距離）に基づいて、統合すべき学習モジュールがあるかを判定する。具体的には、モジュール統合部123は、ステップS61で求めたパラメータ距離が所定の閾値 $D_{threshold}$ より小さいものがある場合に、そのパラメータ距離が得られた2つの学習モジュールは、統合すべき学習モジュールであるとして、統合すべき学習モジュールがあると判定する。

30

【0165】

ステップS62で、統合すべき学習モジュールがあると判定された場合、処理はステップS63に進み、モジュール統合部123は、統合すべきと判定された学習モジュールを統合する。具体的には、モジュール統合部123は、統合対象の2つの学習モジュールのモデルパラメータの平均値を計算し、その値を、統合後に存続させる一方の学習モジュールのモデルパラメータに設定し、他方の学習モジュールをパターン学習部111から削除する。

40

【0166】

なお、学習が十分に進んでいない学習モジュールを統合するのは適当ではないため、統合すべきと判定された各学習モジュールが、十分学習された状態であることを確認した上で、学習モジュールを統合する必要がある。統合対象の2つの学習モジュールが十分学習された状態であるかを判断するには、統合すべきと判定された2つの学習モジュールの学習スコアが十分学習された状態であることを表す所定の閾値以上であることを確認するか、または、各学習モジュールの学習スコアが所定の閾値以上であることを確認した上で、学習モジュール間の類似性を判断すればよい。

50

【 0 1 6 7 】

一方、ステップ S 6 2 で、統合すべき学習モジュールがないと判定された場合、ステップ S 6 3 の処理はスキップして統合処理は終了する（図 1 3 の追加学習処理に戻る）。

【 0 1 6 8 】

次に、パターン学習モデルとして RNN を採用した場合について説明する。RNN は、入力層に PB ユニットを有していない点のみが RNNPB と異なり、それ以外の更新学習等は RNNPB と同様に行うことができる。

【 0 1 6 9 】

パターン学習モデルとして RNN を採用した場合、図 1 2 の学習装置 1 0 1 の構成例を示すブロック図は、図 1 2 のパターン学習部 1 1 1 が図 3 に示したようになる。但し、図 3 の RNNPB#i が、PB ユニットを持たない RNN#i に置き換わる。

10

【 0 1 7 0 】

また、パターン学習モデルとして RNN を採用した場合の追加学習処理のフローチャートは、図 1 3 のステップ S 4 1 で作成する学習モジュールが RNN となるので、ステップ S 4 2 の学習処理が RNNPB#i を RNN#i に置き換えた図 4 の学習処理となり、ステップ S 4 3 の統合処理が図 1 5 に示す処理となる。

【 0 1 7 1 】

そこで、図 1 5 のフローチャートを参照して、パターン学習モデルとして RNN を採用した場合の、図 1 3 のステップ S 4 3 における統合処理について説明する。

【 0 1 7 2 】

20

ステップ S 8 1 において、類似性評価部 1 2 2 は、学習モジュール間の類似性を評価する。RNN では、ウエイトがモデルパラメータであるので、類似性評価部 1 2 2 は、ウエイトマトリクスのユークリッド距離（以下、ウエイト距離と称する）を、RNN 間の類似性の評価として採用する。

【 0 1 7 3 】

例えば、RNN#1 のウエイトマトリクス w_1 の各ウエイトを $w_{1,k,l}$ ($1 \leq k \leq Q, 1 \leq l \leq R$)、RNN#2 のウエイトマトリクス w_2 の各ウエイトを $w_{2,k,l}$ としたとき、RNN#1 と RNN#2 のウエイト距離 $D_{weight}(1,2)$ は、式 (5) で表すことができる。

【 0 1 7 4 】

【 数 5 】

30

$$D_{weight}(1,2) = \sqrt{\sum_{k=1}^Q \sum_{l=1}^R (w_{1,k,l} - w_{2,k,l})^2} \quad \cdots (5)$$

【 0 1 7 5 】

類似性評価部 1 2 2 は、各 RNN 間のウエイト距離を、N 個の学習モジュール 1 0₁ ないし 1 0_N (RNN#1 ないし RNN#N) のすべての組み合わせについて求める。

【 0 1 7 6 】

ステップ S 8 2 において、モジュール統合部 1 2 3 は、類似性評価部 1 2 2 が求めた各 RNN 間の類似性に基づいて、統合すべき学習モジュールがあるかを判定する。すなわち、モジュール統合部 1 2 3 は、ステップ S 8 1 で求めたウエイト距離が所定の閾値 $D_{threshold}$ より小さいものがある場合に、そのウエイト距離が得られた 2 つの学習モジュールは、統合すべき学習モジュールであるとして、統合すべき学習モジュールがあると判定する。

40

【 0 1 7 7 】

ステップ S 8 2 で、統合すべき学習モジュールがあると判定された場合、処理はステップ S 8 3 に進み、モジュール統合部 1 2 3 は、統合すべきと判定された学習モジュール (RNN) を統合する。具体的には、モジュール統合部 1 2 3 は、統合対象の 2 つの RNN のウエイトマトリクスの平均値を計算し、その値を、統合後に存続させる一方の RNN のウエイトマトリクスに設定し、他方の RNN をパターン学習部 1 1 1 から削除する。

【 0 1 7 8 】

パターン学習モデルが RNN である場合においても、統合対象の 2 つの RNN が十分学習され

50

た状態であることを確認する必要がある。RNNでは、例えば、学習誤差が所定の閾値以下であるかを判断することで、RNNが十分学習された状態であることを確認してから、統合対象の2つのRNNを統合する。

【0179】

一方、ステップS82で、統合すべき学習モジュールがないと判定された場合、ステップS83の処理はスキップして統合処理は終了する(図13の追加学習処理に戻る)。

【0180】

図16および図17は、学習装置101による追加学習処理を概念的に示した図である。

【0181】

図16は、1回の追加学習処理に対して1つの新しい学習データが供給され、図13の追加学習処理を1回行うごとに、1つの新たな学習モジュールをモジュール作成部121が追加していく処理を概念的に示した図である。

【0182】

新しい学習データ DAT_1 がパターン学習部111に供給されると、第1回目の追加学習処理が実行され、学習モジュール作成部121は、その学習データ DAT_1 に対する学習モジュール 10_1 を新たに作成する。

【0183】

次に、新しい学習データ DAT_2 がパターン学習部111に供給されると、第2回目の追加学習処理が実行され、学習モジュール作成部121は、その学習データ DAT_2 に対する学習モジュール 10_2 を新たに作成する。さらに、新しい学習データ DAT_3 がパターン学習部111に供給されると、第3回目の追加学習処理が実行され、学習モジュール作成部121は、その学習データ DAT_3 に対する学習モジュール 10_3 を新たに作成する。以下同様にして、新しい学習データ DAT_5 がパターン学習部111に供給されると、第5回目の追加学習処理が実行され、学習モジュール作成部121は、その学習データ DAT_5 に対する学習モジュール 10_5 を新たに作成する。

【0184】

第1回目ないし第5回目の各追加学習処理では、図13を参照して説明したように、追加された学習モジュールを含めて学習処理(ステップS42の処理)が行われ、続いて、統合処理(ステップS43の処理)が行われる。

【0185】

そして、第1回目ないし第4回目の各追加学習処理においては、統合すべき学習モジュールがないと判定され、第5回目の追加学習処理において、学習モジュール 10_1 と学習モジュール 10_5 を統合することができると判定されたとする。

【0186】

図17は、学習モジュール 10_1 と学習モジュール 10_5 を統合する場合の処理を概念的に示した図である。

【0187】

第5回目の追加学習処理において、学習処理終了後、モジュール統合部123が、類似性評価部122が求めた各学習モジュール間の類似性に基づいて、統合すべき学習モジュールがあるかを判定したところ、学習モジュール 10_1 と学習モジュール 10_5 は統合することができるという判定結果が得られたとする。すなわち、学習モジュール 10_1 と学習モジュール 10_5 のパラメータ距離 $D_{parameter}(1,5)$ が、閾値 $D_{threshold}$ より小さいという結果が得られたとする。

【0188】

この場合、モジュール統合部123は、学習モジュール 10_1 のモデルパラメータ P_1 と、学習モジュール 10_5 のモデルパラメータ P_5 の平均値を計算し、その平均値を統合後の学習モジュール 10_1 のモデルパラメータ P_1 に設定するとともに、学習モジュール 10_5 をパターン学習部111から削除する。

【0189】

10

20

30

40

50

なお、図 17 は、2つの学習モジュール 10₁と学習モジュール 10₅を1つの学習モジュール 10₁に統合する例であるが、統合する学習モジュールの数は2つに限定されない。例えば、3つの学習モジュールが互いに閾値 $D_{threshold}$ より小さいと判定された場合には、その3つの学習モジュールを1つの学習モジュールに統合してもよい。この場合、統合後の学習モジュールのモデルパラメータは、統合対象の3つの学習モジュールのモデルパラメータの平均値とすることができる。

【0190】

図 17 に示される学習モジュール 10_iのモデルパラメータ P_i は、式(4)の $p_{i,1}$ 乃至 $p_{i,Q}$ 全体を表し、モデルパラメータ P_1 とモデルパラメータ P_5 の平均値とは、 $p_{1,1}$ と $p_{5,1}$ の平均値、 $p_{1,2}$ と $p_{5,2}$ の平均値、 $p_{1,3}$ と $p_{5,3}$ の平均値、 $p_{1,4}$ と $p_{5,4}$ の平均値、 $\cdot$ 10
 $\cdot \cdot \cdot p_{1,Q}$ と $p_{5,Q}$ の平均値を、それぞれ、統合後の $p_{1,1}$ 、 $p_{1,2}$ 、 $p_{1,3}$ 、 $p_{1,4}$ 、 $\cdot \cdot \cdot$ 、 $p_{1,Q}$ とすることを表している。なお、平均値以外の演算結果を、統合後に存続させる学習モジュールのモデルパラメータに設定してもよい。すなわち、統合後に存続させる学習モジュールのモデルパラメータを、統合対象の複数の学習モジュールのモデルパラメータの平均値以外の演算によって求めることが可能である。

【0191】

以上のように、図 12 の学習装置 101によれば、規模拡張性があり、同時に、汎化特性を有するパターン学習モデルを得ることができるとともに、新しい学習データ(学習サンプル)が供給された場合に、その新しい学習データに対して、モジュール作成部 121が無条件に新たな学習モジュールを作成(追加)するので、学習モジュール追加の要否を 20
 判断する必要がない。また、学習(更新学習)処理後、類似性の高い学習モジュールどうしについては統合するので、学習モジュール数の不要な増加を抑制することができる。

【0192】

なお、学習装置 101に供給された新しい学習データに対応して、学習モジュールを作成する場合に、その作成する学習モジュールのモデルパラメータの初期値は、乱数等によって決定された値を与えるようにしてもよいし、既存の全学習モジュールのモデルパラメータの平均値を与えるようにすることもできる。既存の全学習モジュールのモデルパラメータの平均値を、追加した学習モジュールのモデルパラメータの初期値として与えた場合には、例えば、乱数等によって与えた場合のように、既存の学習モジュールのモデルパラメータとは関係なく初期値が与えられたときと比較して、追加された学習モジュールが、 30
 既存の各学習モジュールが有するパターンの共通性を既に有している状態となるので、高速に学習することができる。

【0193】

上述した一連の処理は、ハードウェアにより行うこともできるし、ソフトウェアにより行うこともできる。一連の処理をソフトウェアによって行う場合には、そのソフトウェアを構成するプログラムが、汎用のコンピュータ等にインストールされる。

【0194】

そこで、図 18 は、上述した一連の処理を実行するプログラムがインストールされるコンピュータの一実施の形態の構成例を示している。

【0195】

プログラムは、コンピュータに内蔵されている記録媒体としてのハードディスク 205やROM 203に予め記録しておくことができる。

【0196】

あるいはまた、プログラムは、フレキシブルディスク、CD-ROM(Compact Disc Read Only Memory)、MO(Magneto Optical)ディスク、DVD(Digital Versatile Disc)、磁気ディスク、半導体メモリなどのリムーバブル記録媒体 211に、一時的あるいは永続的に格納(記録)しておくことができる。このようなリムーバブル記録媒体 211は、いわゆるパッケージソフトウェアとして提供することができる。

【0197】

なお、プログラムは、上述したようなリムーバブル記録媒体 211からコンピュータに 50

インストールする他、ダウンロードサイトから、デジタル衛星放送用の人工衛星を介して、コンピュータに無線で転送したり、LAN(Local Area Network)、インターネットといったネットワークを介して、コンピュータに有線で転送し、コンピュータでは、そのようにして転送されてくるプログラムを、通信部 208 で受信し、内蔵するハードディスク 205 にインストールすることができる。

【0198】

コンピュータは、CPU(Central Processing Unit) 202 を内蔵している。CPU 202 には、バス 201 を介して、入出力インタフェース 210 が接続されており、CPU 202 は、入出力インタフェース 210 を介して、ユーザによって、キーボードや、マウス、マイク等で構成される入力部 207 が操作等されることにより指令が入力されると、それにしたがって、ROM(Read Only Memory) 203 に格納されているプログラムを実行する。あるいは、また、CPU 202 は、ハードディスク 205 に格納されているプログラム、衛星若しくはネットワークから転送され、通信部 208 で受信されてハードディスク 205 にインストールされたプログラム、またはドライブ 209 に装着されたリムーバブル記録媒体 211 から読み出されてハードディスク 205 にインストールされたプログラムを、RAM(Random Access Memory) 204 にロードして実行する。これにより、CPU 202 は、上述したフローチャートにしたがった処理、あるいは上述したブロック図の構成により行われる処理を行う。そして、CPU 202 は、その処理結果を、必要に応じて、例えば、入出力インタフェース 210 を介して、LCD(Liquid Crystal Display)やスピーカ等で構成される出力部 206 から出力、あるいは、通信部 208 から送信、さらには、ハードディスク 205 に記録等させる。

【0199】

ここで、本明細書において、コンピュータに各種の処理を行わせるためのプログラムを記述する処理ステップは、必ずしもフローチャートとして記載された順序に沿って時系列に処理する必要はなく、並列的あるいは個別に実行される処理（例えば、並列処理あるいはオブジェクトによる処理）も含むものである。

【0200】

また、プログラムは、1のコンピュータにより処理されるものであっても良いし、複数のコンピュータによって分散処理されるものであっても良い。さらに、プログラムは、遠方のコンピュータに転送されて実行されるものであっても良い。

【0201】

また、本発明の実施の形態は、上述した実施の形態に限定されるものではなく、本発明の要旨を逸脱しない範囲において種々の変更が可能である。

【0202】

即ち、本発明は、ある特定の空間パターン、時系列シーケンス及びパターンに特化した方法ではない。したがって、コンピュータのユーザインタフェースにおけるユーザ入力のパターン、ロボットのセンサ入力及びモータ出力のパターン、音楽データに関するパターン、画像データに関するパターン、言語処理における音素、単語、センテンスなどのパターンの学習及び学習結果に基づくパターンの予測や分類に適用することができる。

【図面の簡単な説明】

【0203】

【図1】本発明を適用する学習装置の基本となる学習装置の一実施の形態の構成例を示すブロック図である。

【図2】図1の学習装置の学習処理を説明するフローチャートである。

【図3】パターン学習モデルとしてRNNPBを採用した場合の図1の学習装置の構成例を示すブロック図である。

【図4】パターン学習モデルとしてRNNPBを採用した場合の図1の学習装置の学習処理を説明するフローチャートである。

【図5】シミュレーションの結果を示す図である。

【図6】シミュレーションの結果を示す図である。

10

20

30

40

50

- 【図 7】シミュレーションの結果を示す図である。
 【図 8】シミュレーションの結果を示す図である。
 【図 9】シミュレーションに用いた時系列データを示す図である。
 【図 10】RNNPBのモデルパラメータの共有を模式的に示す図である。
 【図 11】「局所表現」スキーム、「分散表現」スキーム、及び「中間表現」スキームの学習の関係を模式的に示す図である。
 【図 12】本発明を適用した学習装置の一実施の形態の構成例を示すブロック図である。
 【図 13】図 12 の学習装置の追加学習処理を説明するフローチャートである。
 【図 14】図 13 の統合処理を説明するフローチャートである。
 【図 15】パターン学習モデルとしてRNNを採用した場合の図 13 の統合処理を説明する
 フローチャートである。
 【図 16】新たな学習モジュールを追加する処理を概念的に示した図である。
 【図 17】学習モジュールを統合する処理を概念的に示した図である。
 【図 18】本発明を適用したコンピュータの一実施の形態の構成例を示すブロック図であ
 る。

10

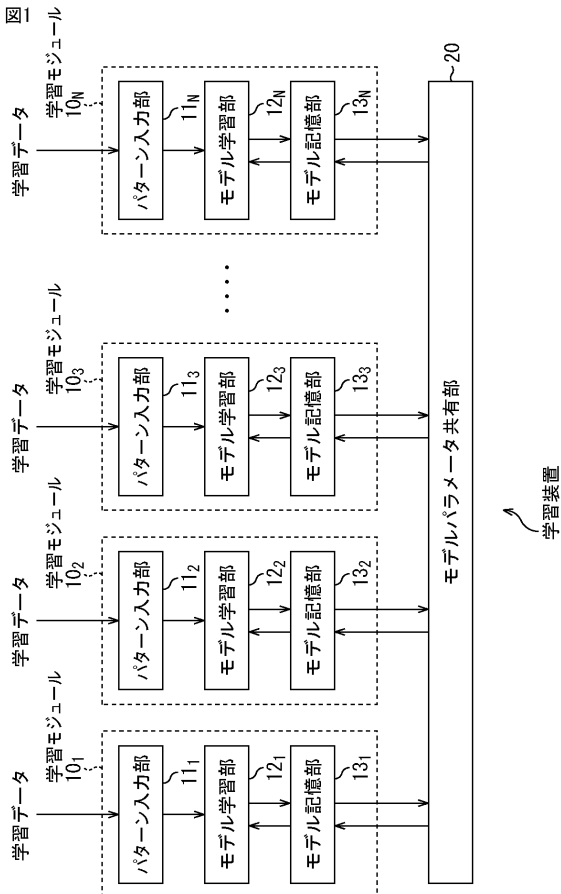
20

【符号の説明】

【0204】

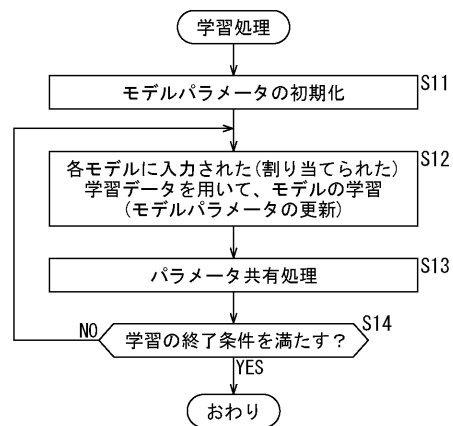
10₁ないし10_N 学習モジュール, 11₁ないし11_N パターン入力部, 12₁
 ないし12_N モデル学習部, 13₁ないし13_N モデル記憶部, 20 モデルパラ
 メータ共有部, 21 ウェイトマトリクス共有部, 101 学習装置, 111 パ
 ターン学習部, 112 学習モジュール管理部, 121 モジュール作成部, 12
 2 類似性評価部, 123 モジュール統合部, 201 バス, 202 CPU,
 203 ROM, 204 RAM, 205 ハードディスク, 206 出力部, 207
 入力部, 208 通信部, 209 ドライブ, 210 入出力インタフェース,
 211 リムーバブル記録媒体

【図 1】

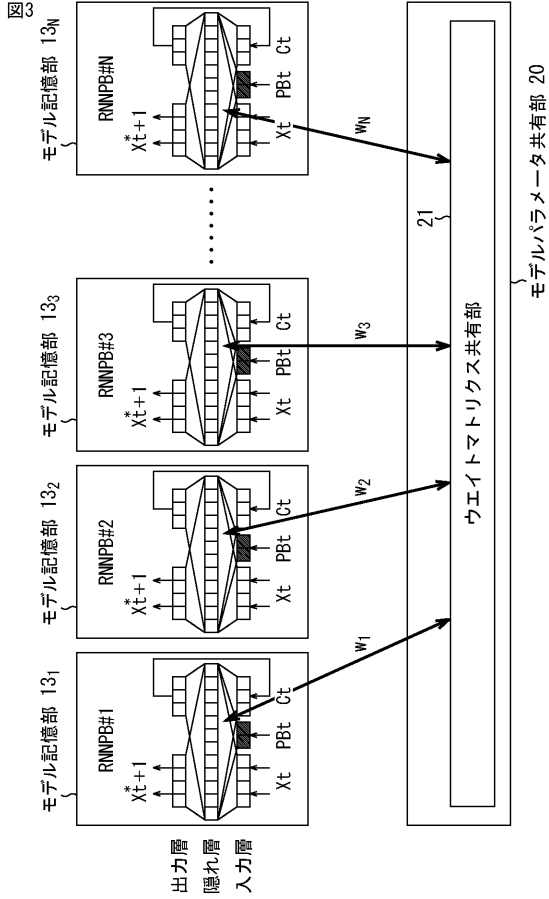


【図 2】

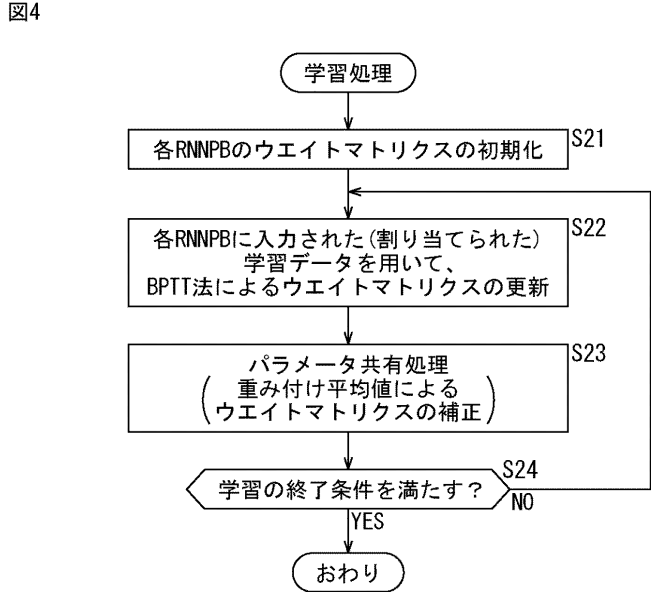
図2



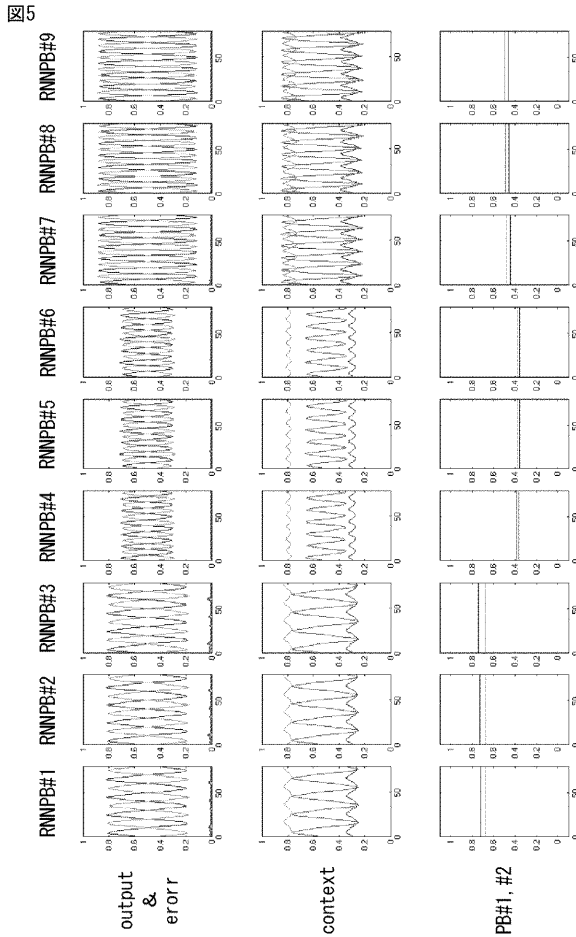
【図 3】



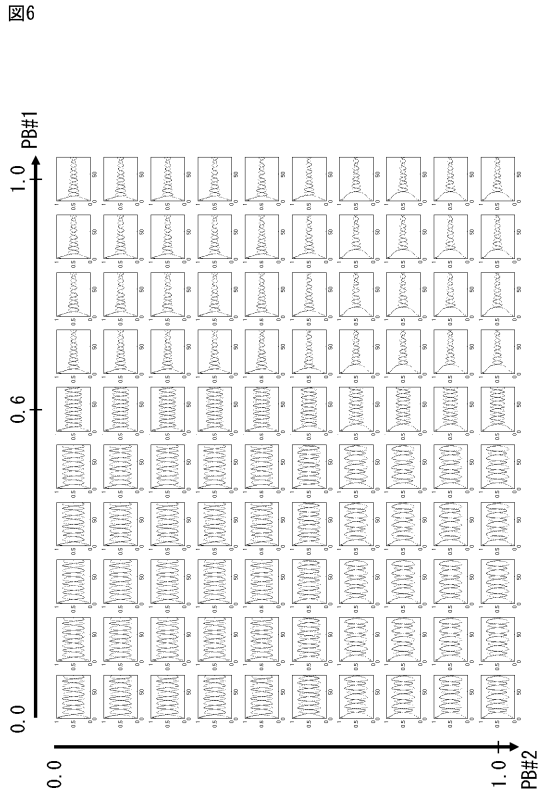
【図 4】



【図 5】

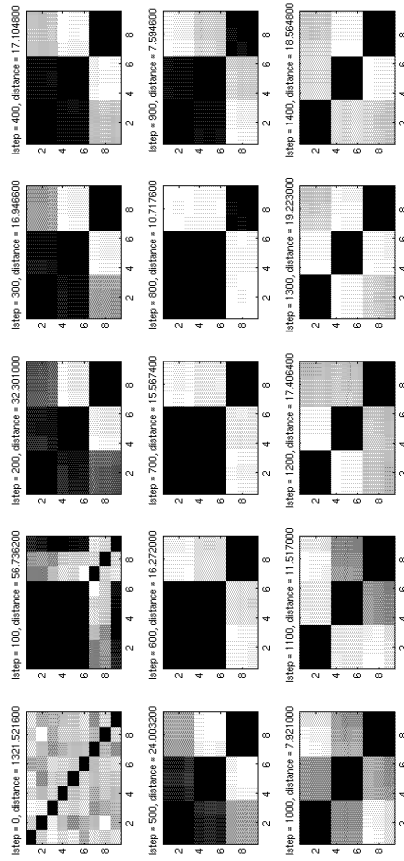


【図 6】



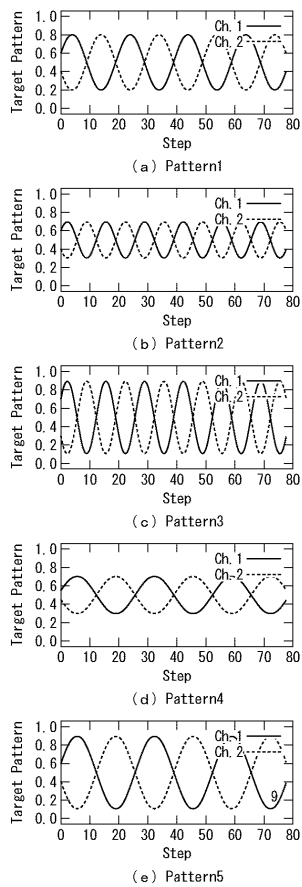
【図 7】

図7



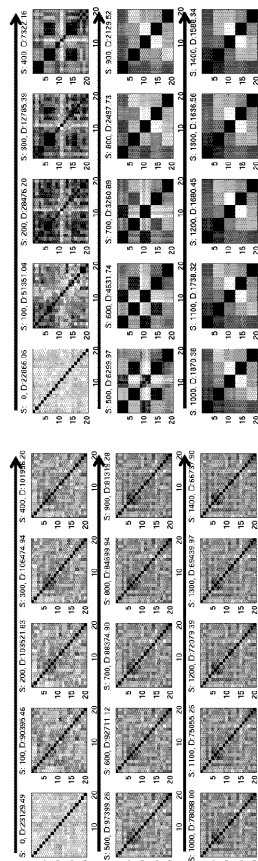
【図 9】

図9



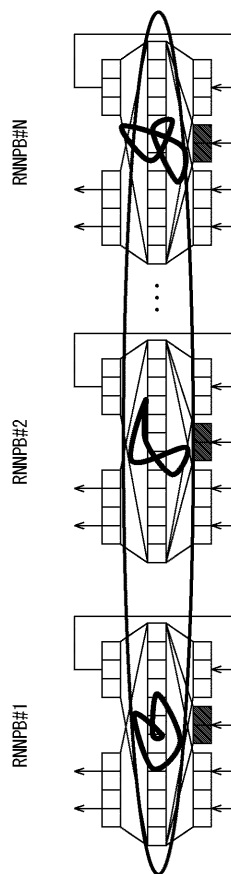
【図 8】

図8



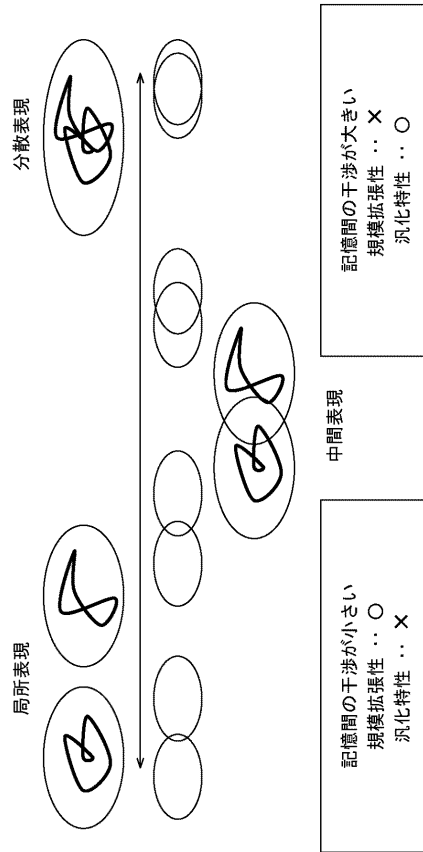
【図 10】

図10



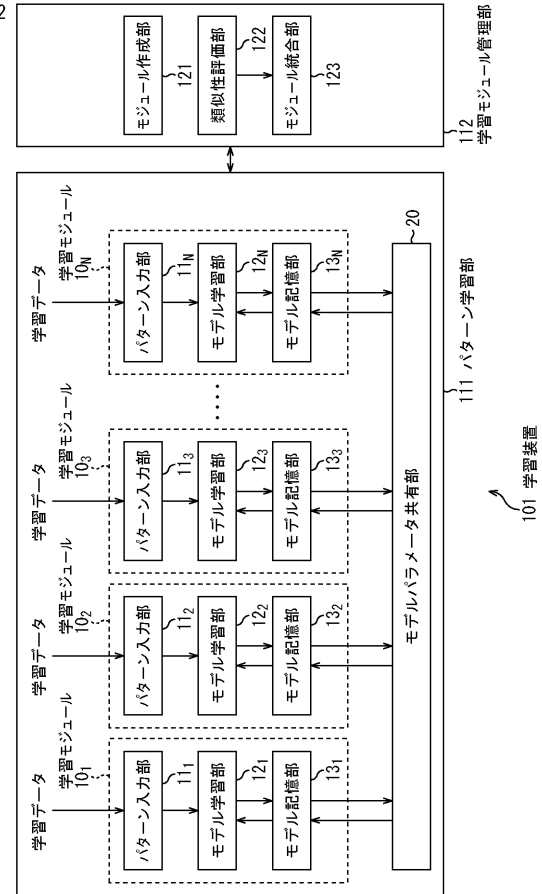
【図 1 1】

図11



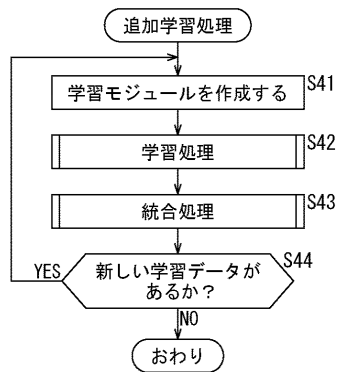
【図 1 2】

図12



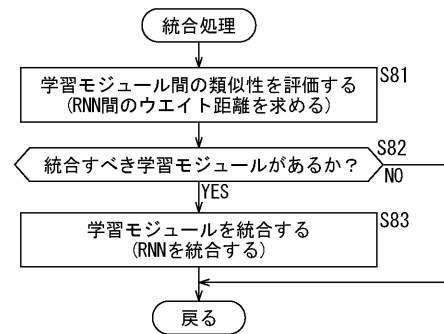
【図 1 3】

図13



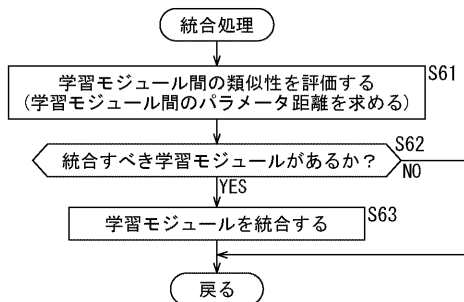
【図 1 5】

図15



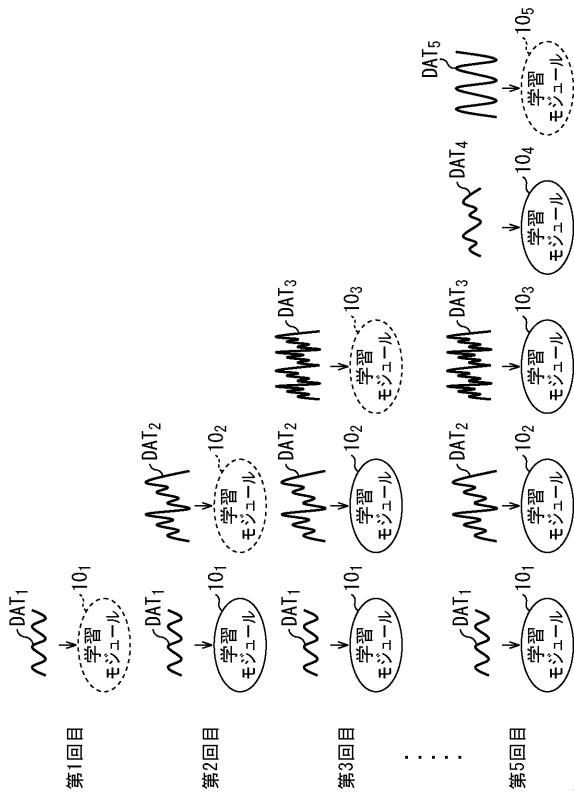
【図 1 4】

図14



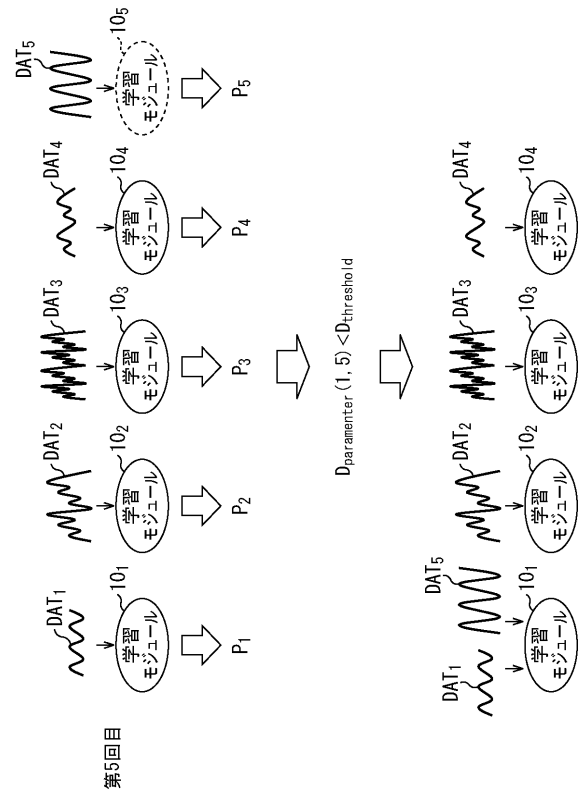
【図 16】

図16



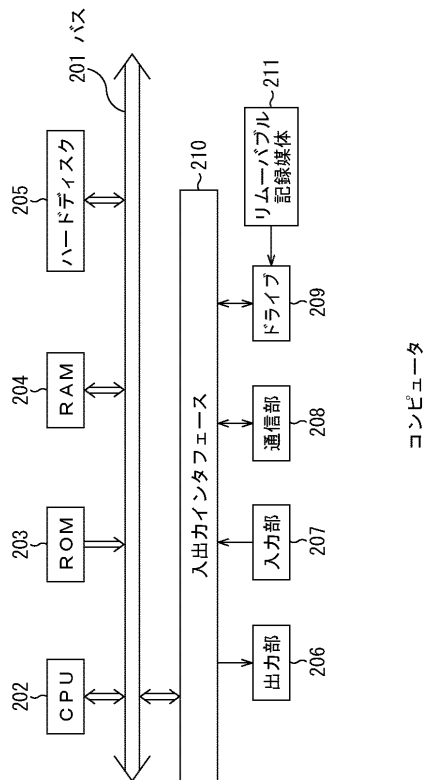
【図 17】

図17



【図 18】

図18



【手続補正書】

【提出日】平成21年7月10日(2009.7.10)

【手続補正 1】

【補正対象書類名】明細書

【補正対象項目名】0 0 7 0

【補正方法】変更

【補正の内容】

【0 0 7 0】

したがって、式(2)の右辺のサメーション $\sum \beta_{ij}(w_j - w_i)$ は、係数 β_{ij} を重みとした、RNNPB# i のウエイトマトリクス w_i に対するRNNPB#1ないしRNNPB# N のウエイトマトリクス w_1 ないし w_N それぞれの偏差(差分)の重み付け平均値を表し、 α_i は、その重み付け平均値 $\sum \beta_{ij}(w_j - w_i)$ を、ウエイトマトリクス w_i に影響させる度合いを表す係数である。

【手続補正 2】

【補正対象書類名】明細書

【補正対象項目名】0 1 0 0

【補正方法】変更

【補正の内容】

【0 1 0 0】

また、RNNPB#1には、パターンP#1の時系列データにノイズN#1を重畳して得られる時系列データを、RNNPB#2には、パターンP#1の時系列データにノイズN#2を重畳して得られる時系列データを、RNNPB#3には、パターンP#1の時系列データにノイズN#3を重畳して得られる時系列データを、それぞれ、学習データとして与えた。