

再帰結合神経回路モデルを用いた 予測可能性による段階的対象選択学習

信田 春満[†] 西出 俊[†] 奥乃 博[†] 尾形 哲也[‡]

[†] 京都大学大学院 情報学研究科 [‡] 早稲田大学 基幹理工学部

1. はじめに

自律ロボットの実現において、ロボットが経験から自身の視覚運動系の学習を行う事が必要不可欠である。従来の視覚運動系学習の研究では視野内のどの物体が身体であるかを既知として運動指令と身体や物体の予測モデルの学習を行ってきた [1][2]。我々はこれに対し、視野内の身体領域が未知という条件下で予測可能性に基づいて身体・非身体を識別するモデルを提案した [3]。本稿では P.Y.Oudeyer らの提案する内発的動機づけモデル [4] を学習対象（身体・非身体）選択に応用し、効率的な身体モデル学習について提案する。

2. 関連研究：内発的動機付けモデル

幼児は周囲の環境を学習する際、単純なものに興味を持ち、より複雑なものへと興味の対象を移していく。White は探索的行動はそれ自体が報酬元となると説いた [5]。探索行為のようにそれ自体が報酬元となるものは内的動機付けと呼ばれる。

Oudeyer[4] らの内発的動機付けモデルは感覚運動を学習する通常の予測器と、通常の予測器の予測誤差の遷移を予測するメタ予測器の2つから成る。メタ予測器は次に学習する行動と予測誤差減少の関係を予測する。予測誤差が最も小さくなると期待される行動を選んでいくことで、ロボットの学習対象が単純なものからより複雑なものに遷移していくことを示している。

3. 提案モデル：物体係数付き学習モデル

3.1 システム入出力

本稿で提案する物体係数付き学習モデルでは、ロボットが手先を動かし、その運動指令と視野内の複数の物体の動きの関係について学習を行う。

システムの入出力は、運動指令（ロボットの関節角速度3軸）、視野情報（自身の手先（身体）、外部の物体のそれぞれの x, y 座標）の時系列である。

3.2 神経力学モデル MTRNN

本稿では視覚運動系の予測器として谷らによって提案された Multiple Timescale Recurrent Neural Network (MTRNN) [6] を用い、その学習には本稿で提案する物体係数付き学習を用いる。MTRNN は各ノード群毎に時定数を持つ RNN(Recurrent Neural Network) であり、現在の状態を入力として次状態を出力する予測器として用いられる。各ノード毎の時定数が大きいノードほど内部状態の更新が緩やかになる。本稿で使用する MTRNN の各ノード群は図1のようになっている。本実験では時定数を入出力ノード：2, Cf:5, Cs:70 と設定した。

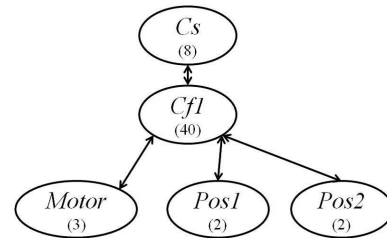


図 1: MTRNN のノード構成。括弧内はノード数を表し、矢印は全結合を表す。Motor:運動指令入出力ノード, Pos1,2:視覚座標入力ノード, Cf,Cs: コンテキストノード

3.3 物体係数付き学習モデル

本稿で提案する物体係数付き学習では、学習時に過去の各物体の予測誤差の推移から、物体毎に学習係数（これを物体係数と呼ぶ）を設定し、最も学習が進みそうな物体を優先的に学習する。学習は以下に行われる。

1. 教師データを基に n_s ループの学習をおこなう。
2. n_s ループ前に比べ各物体の動きを表す各 IO ノードの予測誤差がどのように変化したかを計算する
3. 2 の変化に応じて各 IO ノード毎の学習係数を変更する

ここで、3 の IO ノード毎の学習係数の決め方について以下の 2 パターンを試すことにする。

1. 現在予測誤差が最も減っている 物体と運動指令の関係を優先的に学習する
2. 現在予測誤差が最も変化している 物体と運動指令の関係を優先的に学習する

3.4 各ニューロングループの学習

物体係数付き学習は BPTT (Back Propagation Through Time [7]) に物体毎の学習係数を付け加えたものである。

物体毎の学習係数は、学習時に用いられる物体毎の予測誤差の大きさを見かけ上変化させる役割を負い、0.0 ~ 1.0 の値を持つ。物体係数が 1.0 の場合その物体は通常の BPTT と同様の予測誤差を入力され学習を行う。物体係数が小さくなると、学習に用いられるその物体の予測誤差は見かけ上小さくなる。すなわち物体係数の大きさによって学習対象の優先度が決定される。

ここで物体 g の物体係数 β_g は以下の式で計算される。（ただし、 β_g の初期値は全て 1.0 としている。）

$$d_g = \begin{cases} E_g(n - n_s) / E_g(n) & \dots \text{予測誤差の減少比を使う場合} \\ \max(E_g(n - n_s) / E_g(n), E_g(n) / E_g(n - n_s)) & \dots \text{予測誤差の変化比を使う場合} \end{cases} \quad (1)$$

$$\beta_g = \text{softmax}_{g \in IOGroup}(\gamma d_g) \quad (2)$$

Object selection learning by predictability using neuro-dynamical system
Harumitsu Nobuta (Kyoto Univ.), Shun Nishide (Kyoto Univ.), Hiroshi G.
Okuno (Kyoto Univ.), and Tetsuya Ogata (Waseda Univ.)

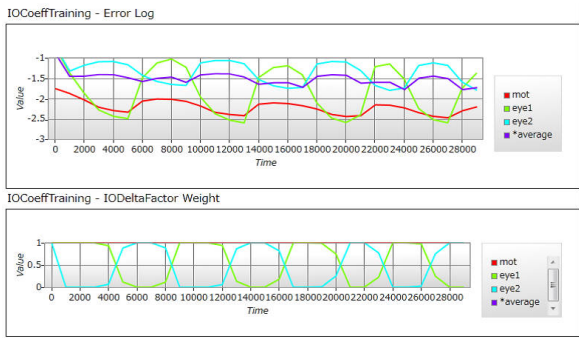


図 2: 物体係数付き学習（減少比によるもの）を行ったもの．上図：予測誤差，下図：IO 係数．mot は運動指令，eye1 は自己身体の位置，eye2 は外部物体の位置を表す．予測誤差は 10 を底とする対数グラフ．

d_g : ニューロングループ g の予測誤差の変化
 n : 更新回数
 n_s : IO 係数更新の間隔
 $E_g(n)$: n ループ学習した時点でのニューロングループ g の平均予測誤差
 $IOGroup$: 物体係数付き学習を行う IO グループ．
 γ : 優先度に偏りを与える係数

4. 実験

視野内に自己の手先と他の移動物体が存在する状態でロボットに手先を動かさせ、物体係数の推移を観察した。

予測誤差の推移と IO 係数の推移を図 2(予測誤差の減少比で物体係数を決めるもの)、図 3(同変化比によるもの)に示す。各図内の上図は予測誤差であり、下図は物体毎の係数である。共に 5000 ループまでは自己身体が優先的に学習されており、5000 ループ時点で身体モデルの学習が進まなくなったため外部物体に学習対象が移る。本実験では外部物体は本質的に運動指令とは無関係なもののため外部物体の学習に失敗し、減少比を用いるモデルではその後自己身体モデルが壊れてしまっている。これに対し変化比を用いるモデルでは、自己モデルが壊れる際に自己モデルの予測誤差が急激に変化するため、自己モデルが再学習され自己モデルが再構築されていることが分かる。

また、通常の BPTT 法の学習による予測誤差の推移を図 4 に示す。最終的な自己身体の予測誤差は通常の BPTT では $10^{-2.4}$ 、変化比を用いる版の提案手法では $10^{-2.7}$ と提案手法では 1/2 倍まで予測誤差が下がっている。また学習の速度も速いことが分かる。予測不可能なものを含む全体の予測誤差は大きくなって、本来学習すべき予測可能な自己身体の予測誤差は従来手法よりも下がる事が確認された。

また、 γ を小さくした場合には物体係数が物体間で極端に偏りにくくなる。そのため自己モデルが壊れる事が少なくなる代わりに、予測不可能な物体の係数も大きくなるため予測可能な物体の最終的な予測誤差は大きくなる事が観察された。

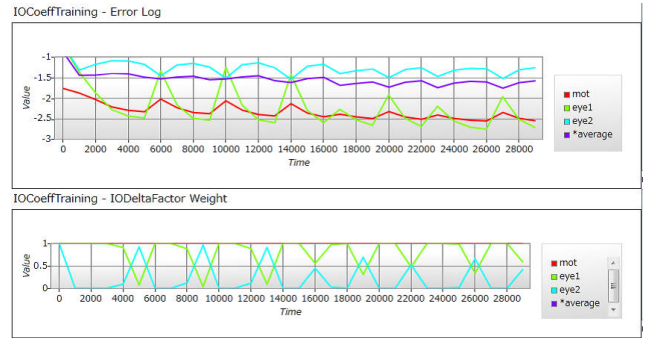


図 3: 物体係数付き学習（変化比によるもの）を行ったもの．図の見方は前図と同様．

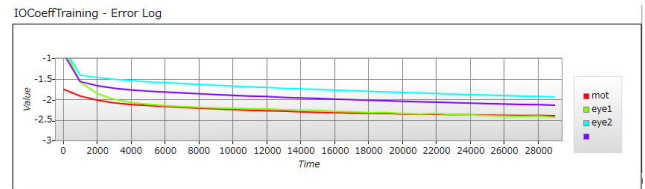


図 4: 通常の BPTT 法による予測誤差の推移．

5. おわりに

本稿では、予測誤差の推移から学習対象の優先度を決定するモデルを提案した。提案手法によりロボットが優先的に自己身体の予測モデルの学習を行う事によって従来の学習モデルより効率のよい学習が行われた。また自己身体の予測モデルの学習が進まなくなった場合には自動的に外界の物体に学習対象を移行する現象が見られた。

謝辞

本研究は、さきがけ領域研究「情報環境と人」、科研費新学術領域研究「構成論的発達科学」(24119003)、科研費若手研究活動スタート支援(23800033)、栢森情報科学振興財団の助成を受けた。

参考文献

- [1] C. Nabeshima, M. Lungarella, and Y. Kuniyoshi, "Timing-based model of body schema adaptation and its role in perception and tool use: A robot case study" Proceedings of the 4th IEEE International Conference on Development and Learning (ICDL-05), 2005.
- [2] M. Ito, K. Noda, Y. Hoshino, and J. Tani, "Dynamic and interactive generation of object handling behaviors by a small humanoid robot using a dynamic neural network model", *Neural Networks*, vol. 19, no. 3, pp. 323-337, 2006.
- [3] Harumitsu Nobuta, Shun Nishide, Hiroshi G. Okuno, Tetsuya Ogata "Identification of self-body based on dynamic predictability using neuro-dynamical system", Proceedings of 2011 IEEE/SICE International Symposium on System Integration (SII2011), pp.256-261, Dec. 20-22, 2011, Kyoto.
- [4] Oudeyer, P.Y. and Kaplan, F. and Hafner, V.V. "Intrinsic motivation systems for autonomous mental development" *Evolutionary Computation*, IEEE Transactions on vol.11 No.2 pp.265-286, 2007.
- [5] R. White, "Motivation reconsidered: The concept of competence," *Psychol. Rev.*, vol. 66, pp. 297-333, 1959.
- [6] Y. Yamashita and J. Tani "Emergence of Functional Hierarchy in a Multiple Timescale Neural Network Model: a Humanoid Robot Experiment", *PLoS Comput. Biol.*, vol.4, no.11, 2008.
- [7] P.J. Werbos "Backpropagation through time: What it does and how to do it", Proceedings of the IEEE - Vol.78, No.10, pp.1550-1560, 2002.