

身体バブリングと再帰結合型神経回路モデルによる道具身体化 ～深層学習による画像特徴量抽出～

Tool-body Assimilation Model based on Body Babbling and Recurrent Neural Networks
-Image Feature extraction by Deep Neural Networks-

高橋城志*¹
Kuniyuki Takahashi

尾形哲也*²
Tetsuya Ogata

Hadi Tjandra*³

野田邦昭*⁴
Kuniaki Noda

村田真悟*¹
Shingo Murata

有江浩明*²
Hiroaki Arie

菅野重樹*³
Shigeki Sugano

*¹早稲田大学 理工学術院 創造理工学研究科 総合機械工学専攻

Department of Modern Mechanical Engineering, Graduate School of Creative Science and Engineering, Waseda University

*²早稲田大学 理工学術院 基幹理工学部 表現工学科

Department of Intermedia Art and Science, School of Fundamental Science and Engineering, Waseda University

*³早稲田大学 理工学術院 創造理工学部 総合機械工学科

Department of Modern Mechanical Engineering, School of Creative Science and Engineering, Waseda University

*⁴早稲田大学 理工学術院 基幹理工学研究科 表現工学専攻

Department of Intermedia Art and Science, Graduate School of Fundamental Science and Engineering, Waseda University

We propose a tool-body assimilation model based on body babbling and neuro-dynamical system for robots to use tools. Since almost all of existing studies for tool handling required predetermined motions and tool features, the motion patterns are limited and they cannot use novel tools. To counter these problem, we took the following two approaches: (1) a humanoid robot model to generate random motion based on human body babbling, (2) recurrent neural network and deep neural networks based on human recognition mechanism. Experiments were conducted with multiple tools for manipulating a cylindrical target object. The results show that the tool-body assimilation model is capable of recognition of tools.

1. はじめに

ロボットが人間社会で活躍するためには、道具使用が必要となることが予想される。道具使用能力の獲得のための1つとして、人間の認知発達課程に基づいたモデル構築の方法がある。近年、神経心理学の分野で道具身体化が注目されている[Iriki 04]。道具身体化とは道具を使用しているときに、道具の先まで身体が延長しているように感じる現象のことである。我々は、身体モデルにニューロンを付加することで、身体モデルを変調し、道具を使用させることを考えた[Nishide 2012]。

道具使用の研究の多くは、(1) 事前に動作パターンを与える、(2) 事前に道具や動作の特徴量を与える、という問題があった。これらの問題に対して我々は、(1) ヒューマノイドロボットモデルを用いた身体バブリング、(2) 人間の認知機構を参考にした神経回路モデル、を用いた[Takahashi 2014]。しかし、この研究ではself-organizing map (SOM) を用いて画像特徴量の抽出としているため、複数の道具を扱えないという課題があった。本研究では深層学習による特徴量抽出を行うことで、複数の道具を使用可能な道具身体化モデルの提案と検証を行う。

2. 深層学習

深層学習 (Deep Neural Network: DNN) とは、一般的に5-10層以上のNeural Networkのことである。Hintonら[Hinton 2006]に提案されたプレトレーニングにより学習が可

能となった、砂時計型DNNのオードエンコーダを使用することで、高次元のデータを低次元に圧縮し、特徴量の抽出が可能となる。ロボット分野でのDNNの使用が少ない理由として、DNNの学習には大量の高次元の教師データが必要とされるためである。ロボットでこのような教師データを取得するとなると、長時間の使用に耐えかねてロボットが壊れる可能性がある。これら問題を解決するために、我々はシミュレーション上のロボットに身体バブリングを行わせることで解決する。

DNNの学習方法として、Martens [2010 Martens]により提案されたHessian freeを基にして 2^{nd} -order optimizationを発展させた方法がある。この方法では学習が効果的で効率的である。そのため我々はMartensの提案手法を使用する。

3. 道具身体化モデル

道具身体化モデルを図1に示す。はじめに素手による身体バブリングを行い、そのときの関節角度とカメラ画像を取得する。DNNを用いて、そのときのカメラ画像から画像特徴量の抽出を行う。次に抽出された画像特徴量と関節角度のデータを再帰結合型神経回路モデル (multiple time-scale recurrent neural network: MTRNN) により身体モデルの学習を行う。次に、MTRNNに接続されたparametric bias (PB) ノードで道具の特徴量の学習を行う。これにより、身体モデルを変更することなく、PB値を変更することで道具の使用が可能である。身体モデルと比べて非常に少ないPBノードで道具特徴量を表現することが可能である (本研究ではPBノードが5個、身体モデルが98個である)。そのため、PBノードは身体モデルと比較して、非常に短い時間で学習することが可能である。

連絡先: 高橋城志, 早稲田大学 理工学術院
創造理工学研究科 総合機械工学専攻,
takahashi@sugano.mech.waseda.ac.jp

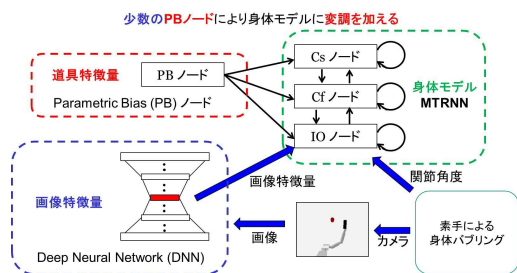


図 1: 道具身体化モデルの概要

3.1 身体バブリングによる自己身体モデルの学習

ロボットはヒューマノイドモデルで身体バブリングを行う。身体バブリングとは人間の幼児に見られるランダムな動作で、多くの動作が必要となる。そのため、シミュレータを用いる。取得された関節角度と DNN で抽出された画像特徴量の関係性から身体モデルの学習を行う。身体モデルの学習には谷らにより提案された MTRNN を用いる [Tani 2008]。MTRNN は複数の時系列データの学習が可能であり、現在の入力から、次の状態を予測することができる。異なる時定数を持っており、発火速度の遅い Cs はデータの時系列が学習され、 Cs の初期値は学習したそれぞれの動作を表現している。

3.2 道具特徴量の学習

道具を把持した状態で、身体バブリングを行い、MTRNN に付加された PB ノードで学習を行う。PB ノードは時定数が無限大であるため、時系列の変化を受けない。学習中は身体モデルの MTRNN の重みや初期値を固定しておき、PB ノードの重みと値の学習のみ学習を行う。これは、道具を把持したときと把持していないときの画像の違いを学習することになる。学習された PB 空間には道具の動的特徴量が表現される。

4. シミュレーションによる評価実験

4.1 実験設定

評価試験として、最も単純である対象物体を素手と道具で操るタスクを行った。道具として、I 字型、T 字型、L 字型の棒状物体を用いた。L 字型道具は未知道具として、学習を行わない。このタスクは道具身体化の研究ではよく用いられる。

道具身体化モデルを評価するために、ロボットシミュレータ OpenHRP3 上でヒューマノイドモデルを作成した。モデルの寸法と自由度配置は ACTROID を基に作成した。関節角度の範囲は人間を参考にしている。モデルの有効性を検証するために、2 次元平面の運動が可能な右腕部 3 自由度を使用する。

身体バブリングは、ロボットは物体の初期値位置に対して、右側と左側をそれぞれ初期位置として、2 つの姿勢を経由する動作をそれぞれ 75 回行う。動作中に関節角度と画像データの取得を行う。データは 6 秒間に 30step 分のデータを取得する。画像データとして、 32×24 ピクセルのグレースケール画像を用いた。この画像データから DNN により抽出される 15 次元の画像特徴量と関節角度 3 次元を MTRNN の入力とした。

4.2 実験結果

主成分分析を用いて PB 値を解析した結果を図 2 に示す。図 2 の学習後では PB 値が道具毎にクラスタリングされている。一部クラスタが重なっている要因は、異なる道具でも、共通な部位を使用するためである。図 2 の認識後では、道具毎にクラスタリングされている。さらに未学習な L 字型の道具も I 字

型と T 字型と重なる分布となっている。以上のことから、ロボットは素手とそれぞれの道具毎に認識できていると言える。

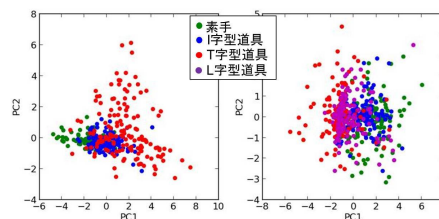


図 2: PB の主成分分析 (左図:学習後, 右図:認識後)

5. まとめと今後の展望

本研究では神経回路モデルを用いて道具身体化モデルによる道具使用について述べた。従来のモデルでは、扱うことが困難な複数の道具のために、画像特徴量抽出に DNN を用いた。道具の特徴量の学習をした PB により、素手とそれぞれの道具において、特徴量のクラスタが自己組織化された。今後の展望として、動作生成の検証を行う。

謝辞

本研究成果の一部は、文部科学省科研費基盤研究 (S) (No. 25220005)、文部科学省科研費新学術領域研究「構成論的発達科学」(No. 24119003)、文部科学省「卓越した大学院拠点形成支援補助金」、JST さきがけ領域「情報環境と人」の助成を受けたものです。本研究成果の一部は、早稲田大学理工研プロジェクト研究「自然と共生する知能情報機械系に関する基盤研究」の一環として行われたものです。ここに謝意を表します。

参考文献

- [Iriki 04] A. Maravita, A. Iriki.: Tools for the body (schema). Trends in Cognitive Sciences, vol. 8, no. 2, pp. 79-86 (2004)
- [Nishide 2012] S. Nishide, et al.: Tool-body assimilation of humanoid robot using neurodynamical system. IEEE Transactions on Autonomous Mental Development, vol. 4, no. 2, pp. 139-149 (2012)
- [Takahashi 2014] Hadi Tjandra, 高橋城志, et al.: 神経力学モデルと身体バブリングに基づく道具身体化と動作生成. 情報処理学会 第 76 回全国大会, 1S-4, 東京 (2014)
- [Hinton 2006] G. E. Hinton, R. R. Salakhutdinov.: Reducing the dimensionality of data with neural networks. Science, vol. 313, no. 5786, pp. 504-7, Jul (2006)
- [2010 Martens] J. Martens.: Deep learning via Hessian-free optimization. In Proceedings of the 27th International Conference on Machine Learning (2010)
- [Tani 2008] Y. Yamashita, J. Tani.: Emergence of Functional Hierarchy in a Multiple Timescales Recurrent Neural Network Model: A Humanoid Robot Experiment. PLoS Computational Biology, Vol. 4, No. 11 (2008)