

自己モデルの投影に基づくロボットによる模倣動作の自律的獲得

○横矢龍之介[†] 尾形哲也[†] 谷淳[‡] 駒谷和範[†] 奥乃博[†]
([†] 京都大学大学院情報学研究科 [‡] 理化学研究所脳科学総合研究センター)

Imitation by Robot Based on Projection of a Self-model

*Ryunosuke YOKOYA[†], Tetsuya OGATA[†], Jun TANI[‡],
Kazunori KOMATANI[†], Hiroshi G. OKUNO[†]

([†]Graduate School of Informatics, Kyoto University [‡]Brain Science Institute, RIKEN)

Abstract— This paper proposes a novel method which enables a humanoid robot acquires an ability for imitation focusing on the developmental process of human infants. In this work, the robot infant predicts a human parent's motions by reusing its own self-model. For the self-model of a robot, we applied a recurrent neural network with parametric bias (RNNPB) model which can learn the robot's body dynamics. For the other-model of a human, conventional hierarchical neural networks were attached to the RNNPB model as “conversion modules”. We focused on the ternary relationship among self-object-other, and infant imitation by a parent in early childhood. Our target task is an imitation of moving an object, and imitation interaction between a parent and an infant is introduced to our method. Experimental results demonstrated that our method enabled the robot to imitate the human's object manipulations which was similar to what the robot has experienced, which proved its adequacy.

Key Words: Imitation, Humanoid Robot, Recurrent Neural Network, Cognitive Developmental Robotics

1. はじめに

本稿の目的は、ロボットによる模倣能力の獲得プロセスをモデル化することである。近年、人間の認知発達の解明に対するひとつのアプローチとして、認知発達ロボティクスが注目されている。これは、人間の認知に関わる仕組みをロボットに構築し、実環境で動作させることで、認知に関する知見を得ようというものである。模倣能力は他者の意図理解の基盤となっており、人間の知能を探索する上で、特に模倣能力の仕組みの解明は極めて重要な課題である。

ロボットによる模倣に関する従来研究の多くにおいて、ロボットは模倣対象である「他者」といった概念は有していない。その多くは、ロボットと模倣対象間の身体構造の対応を設計者が与え、人間が呈示した動作の軌道を正確に再現することを目的としている [3]。しかし、我々人間が他者を模倣する際には、必ずしも他者の身体軌道を再現しているわけではない。文脈に応じて身体軌道を模倣することもあれば、他者の意図を推定しながら、全く異なる動作をもって模倣する場合もある。人間と同様に模倣によって心的状態を共有するためには、他者の内部状態を推定できる枠組が必要である。

本稿では、他者を自律的に発見し、模倣する発達プロセスを提案する。ロボットは他者である人間の動作予測を、自己モデルを再利用することで行う。牧野ら [2] も指摘しているように、他者を含めた外部環境が、個体自身の内部モデルに類似した特性を部分的にでも有しているならば、このモデルを環境に投影して予測に再利用することで、効率的に多様な環境モデルを保持できることが期待される。具体的には、Recurrent Neural Network with Parametric Bias (RNNPB) と呼ばれる、時系列を学習可能な神経回路モデルをロボットの自己

モデルとして利用する。RNNPBに変換モジュールとして階層型 Neural Network を付加したモデルを他者モデルとして利用する。また、幼児の認知発達過程において重要であるとされる、「三項関係」と「親による幼児の模倣」 [1] に着目することで、物体操作をタスクとし、親子の模倣インタラクションを本手法に導入する。

以下、2. 章では自己モデルの投影に基づく模倣手法と、学習モデルである RNNPB、変換モジュールの構成手法について述べる。3. 章では、ハードウェアとソフトウェアの実装について詳細を述べる。4. 章では、模倣実験とその結果を示し、5. 章で本手法と人間の振舞いとの関連について考察する。最後に 6. 章で本稿をまとめる。

2. 自己モデルの投影に基づく模倣手法

2.1 幼児の発達プロセス

本節では、本稿で想定している、模倣能力を獲得するまでの親と幼児のインタラクションについて述べる。

幼児は生まれた直後もしくは生まれる前から、自己探索活動 (バブリング) を通じて自己モデルを獲得する。生まれた直後から、幼児は客観的に観測可能な情報 (手先や音声等) に着目しながら親とのインタラクションを開始する。本稿では、客観的に観測可能な情報として卓上物体を用いる。インタラクション中は、親による幼児の模倣が高い頻度で発生する。幼児は親とのインタラクションを通じて、親が自己モデルと同等のモデルを有していることに気付き、自己モデルを再利用することで他者モデルを構築する。他者モデルを利用することで、幼児による親の模倣が実現される。

2.2 本手法の概要

本節では、2.1 節で述べた発達プロセスのモデル化について、その概要を述べる。

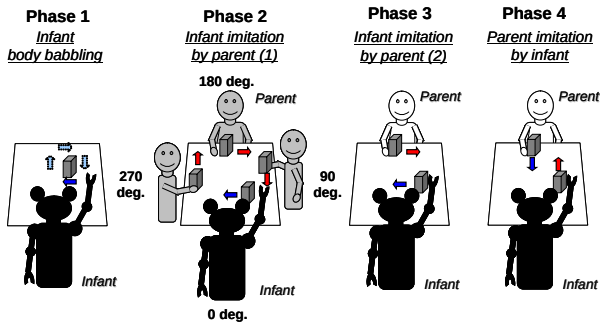


Fig.1 模倣プロセス

本稿では、複数の時系列を単一のモデルで学習可能な RNNPB を自己モデルとして用いる。自己モデルを他者の動作予測に再利用する際、主観的情報 I_s と客観的情報 I_o のデータ変換が必要となる。そこで、RNNPB の入出力部分にデータ変換用の階層型ニューラルネットワークを付加することで、自己モデルの投影を行う (Fig.2 参照)。本稿では、変換用ネットワークを “変換モジュール” と呼ぶ。

本手法は、4 フェーズから構成される (Fig.1 参照)。プロセスの概要は以下の通りである。

1. 幼児の身体バブリング (自己モデルの学習)
ロボットは動作を生成することで、物体操作を行いながらデータを取得し、自己モデルの重みを学習させる。ここで、自己の動作と物体の挙動を結び付ける。
2. 親による幼児の模倣 (1) (変換モジュールの学習)
人間による複数の模倣行為をロボットが観察し、データを取得する。ただし他者の動きではなく、操作対象である物体の挙動のみに着目する。投影モデル中の自己モデルの重みを固定し、観察した物体の挙動を予測できるよう変換モジュール群の重みのみを学習させる。ロボットは、人間による模倣動作の種類を既知であるとする。
3. 親による幼児の模倣 (2) (変換モジュールの選択)
ある位置で人間が行う模倣行為をロボットが観察し、データを取得する。得られたデータを投影モデルに入力することで物体の挙動の予測を行い、最もよく予測を行う変換モジュールを同定する。本フェーズが、模倣すべき他者の発見に相当する。
4. 幼児による親の模倣
人間が提示した動作を、フェーズ 3 で選択された他者モデルを使用して認識し、模倣動作を生成する。ここでは、既に我々が開発している、模倣手法を利用する [6]。

2.3 自己モデルと他者モデル

本節では自己モデルと他者モデル、及びそれらの学習アルゴリズムについて述べる。

2.3.1 RNNPB

RNNPB は谷らによって提案された学習モデルであり [5]、現状態 $S(t)$ から次状態 $S(t+1)$ を予測する予測器である (Fig.2 参照)。このモデルは、従来の Recurrent Neural Network (RNN) に、PB 層と呼ばれるユニットを追加したものである。PB 層は他の入力層とは異な

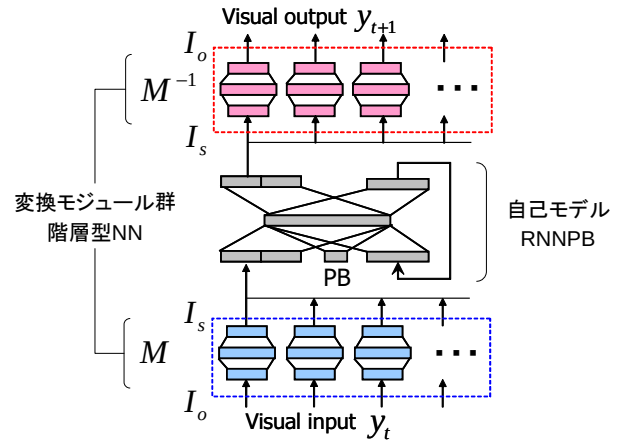


Fig.2 他者モデル

り、各ステップにおいて定数が入力される。RNNPB は時系列を学習することで、各系列の特徴を符号化した PB 値を自己組織的に獲得することができる。

RNNPB は通常の RNN と同様、Back Propagation Through Time と呼ばれる学習アルゴリズムを用いて重みの学習を行う [4]。また重みの学習と同時に、PB 値の更新が行われる。

2.3.2 PB 値の計算手法

学習時、RNNPB の重みの更新と同時に、PB 値は以下の式に基づいて更新される。

$$\delta \rho_t = k_{bp} \cdot \sum_0^T \delta_t^{bp}, \quad (1)$$

$$p_t = \text{sigmoid}(\rho_t) \quad (2)$$

ただし、 k_{bp} は学習定数、 T は時系列長、 δ_t^{bp} はステップ t において PB ノードに逆伝播された学習誤差、 ρ_t はステップ t における PB ノードの内部値、 p_t はステップ t における PB 値である。式 (1) により、全ステップ T 分の学習誤差が積算される。式 (2) より、内部値をシグモイド関数を通すことで PB 値が得られる。

2.3.3 Mixture of Experts による変換モジュール群の学習

フェーズ 2 では、変換モジュール群を自己組織的に獲得するために Mixture of Experts による学習を行う。学習アルゴリズムは以下の通りである。ただしフェーズ 2 ではロボットのモータ値は得られないため、モータ値は Closed ループで計算される (Fig.3 参照)。

1. 複数の変換モジュール $M_i (i = 0, 1, \dots, N_M)$ を用意し、それぞれの重みをランダムに初期化する。
2. 学習データ $D_j (j = 0, 1, \dots, N_D)$ から、 D_L をランダムに選択する。
3. データ D_L について、全ての変換モジュールで予測を行い、最も予測誤差の小さい M_{min} を選択する。
4. 変換モジュール M_{min} の重みを、データ D_L を用いて学習させる。
5. 全ての学習データについて、予測誤差が収束するまで 2~4 を繰り返す。

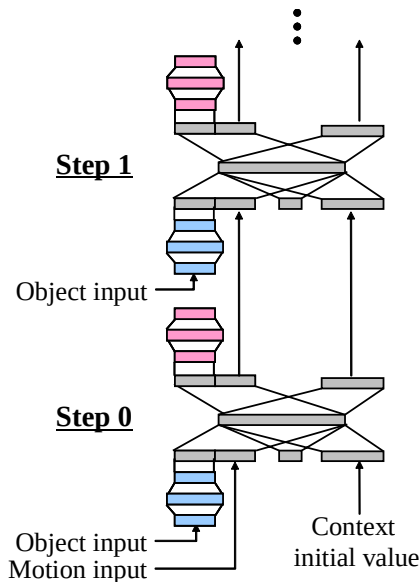


Fig.3 他者モデルによる Forward 計算

3. システムの設計と実装

3.1 テストベッドと対象物体

本稿で用いるテストベッドは、ATR 製のヒューマノイドロボット、Robovie-IIIs である。Robovie-IIIs は首に 3 自由度、各腕に 4 自由度のモータを有しており、頭部には CCD カメラを搭載している。操作対象は、上面を赤青の 2 色に色分けした箱型の物体である。

3.2 実験システム

本実験では首のモータは常に固定しており、物体の初期位置は物体重心がカメラの中心に位置するよう定めた。取得するデータは全て 8 sec 分である。自己モデルとして用いる RNNPB のサイズは入出力層が 7、中間層が 35、コンテキスト層が 10、PB 層が 2 である。他者モデルは 8 組の変換モジュール群を有しており、各変換モジュールのサイズは入出力層が 4、中間層が 4 である。変換モジュールへの入力データは画像情報のみである。RNNPB への入力データとして、以下のセンサデータが収集される。

画像情報 (4 次元)

使用する CCD カメラは左眼のみであり、カメラの解像度 500×400 pixel である。色情報 (RGB) から物体の赤色と青色の領域を抽出し、各領域についてカメラ内での重心座標 xy 2 次元 (0~1 の値) をそれぞれ推定する。

モータ情報 (3 次元)

左腕の 3 自由度 (肩のロールとヨー、肘ピッチ) を使用する。実験において、肩のピッチと首の全モータは常に一定の値に固定されている。

以上のデータの同期をとり、各取得データの最小値及び最大値に基づいて 0.1~0.9 の値に正規化する。データは各操作について 800 msec 毎に 1 ステップ、計 10 ステップ分を取得する。

4. 模倣実験

4.1 実験タスク

卓上物体の移動をタスクとする。用いる操作は、{ 操作 1: 左から右への平行移動 (左→右), 操作 2: 左回転, 操作 3: 手前から奥への平行移動 (手前→奥), 操作 4: 右回転 } の 4 種類である。

4.2 実験方法

フェーズ 1 操作 1~4 の 4 パターンについて、意図する物体操作となるようプログラムしたモーションをロボットが生成し、データを取得する。ただし、各動作を 4 種類の物体初期位置 (Fig.4 参照) に対して行うため、学習データは 16 パターンとなる。これは、物体の初期位置のみから変換モジュールを同定できることを避けるためである。取得されたデータから自己モデル RNNPB を 20 万回学習させる。

フェーズ 2 人間がロボットの位置から反時計回りに 0° , 90° , 180° , 270° の、4 種類の位置において、ロボットの操作 1~3 (物体初期位置 1) の動作を模倣する。人間による模倣動作をロボットが観察し、8 組の変換モジュール $M_i (i = 0, 1, \dots, 7)$ を、2.3.3 節で述べた手法により延べ 120 万回学習させる。ただし、ここで他者モデルに入力される PB 値は、フェーズ 1 で得られた同一動作に対応する PB 値である。

フェーズ 3 フェーズ 2 と同様に、4 種類の位置で人間が模倣する操作 1~4 をロボットが観察し、各動作について、フェーズ 2 で構成した変換モジュールのうち画像情報の予測誤差が最小となるものを選択する。

フェーズ 4 人間が提示する動作をロボットが観察し、フェーズ 3 で選択された変換モジュールを使用し、PB 値を計算する。得られた PB 値を自己モデル RNNPB に入力することで、模倣動作を生成する。

4.3 実験結果

フェーズ 3 において、各動作について選択されたモジュールを Table 1 に示す。

フェーズ 1 及びフェーズ 4 において得られた PB 値を Fig.5 に示す。ただし、フェーズ 1 において得られた PB 値は、物体初期位置 1 に対する動作のみについて示す。変換モジュール群の学習に用いていない右回転については多少のばらつきが見られるものの、各 * に同色の点がクラスタリングされていることが確認できる。これは、観察した人間の動作が、ロボット自身の同一動作と同等であることが認識されたことを示している。

フェーズ 4 において、観察した人間の全ての動作について、ロボットが人間の視野に立った適切な模倣動作を生成することができた。模倣動作に伴う物体の各色領域のカメラ内重心座標の軌跡と、フェーズ 1 における学習動作に伴う軌跡を Fig.6 に示す。ただし、人間が

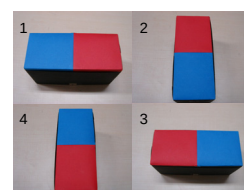


Fig.4 物体の初期位置

Table 1 選択された変換モジュール

操作 \ 位置	0 deg	90 deg	180 deg	270 deg
左→右	M_1	M_2	M_5	M_7
左回転	M_1	M_2	M_0	M_4
手前→奥	M_1	M_2	M_5	M_4
右回転	M_1	M_2	M_5	M_4

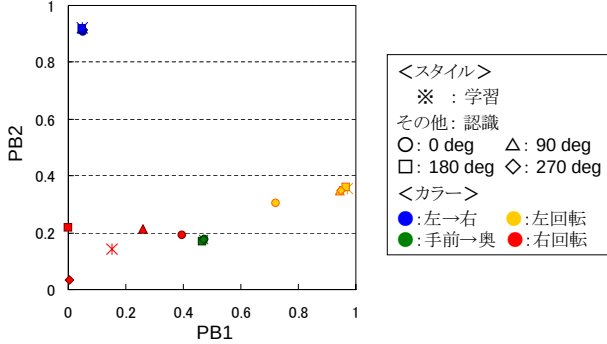


Fig.5 PB 空間

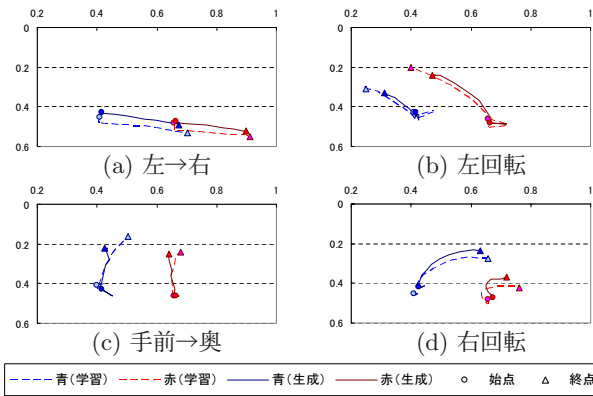


Fig.6 動作生成時の物体軌道

270 deg の位置で呈示した動作に対する模倣結果である。また、実験風景を Fig.7 に示す。

5. 考察

本稿では簡単のため、発達プロセスを4つのフェーズに分割しているが、本来これらはシームレスであるべきである。幼児がフェーズ3にあたる変換モジュールの選択を行った直後に、フェーズ4にあたる親の模倣を実行する場面を考えると、本実験から興味深い現象が説明できる。本実験において、例えばフェーズ3で「手前→奥」の人間による模倣を観察し、選択されたモジュールを用いると、ロボットはその場で人間に呈示される「右回転」を模倣できることになる。これは、幼児が一人遊び等の自己経験がありさえすれば、親子のインタラクション中に現れていない動作を模倣できることを意味する。人間の幼児はある時期から新規動作を模倣できるようになるが、本プロセスは、そのような現象を一部説明できるのではないかと考えられる。

また、本手法において、PB 値が他者の動作認識と自己の動作生成を仲介しているが、これは近年注目されているミラーニューロンの機能に対応する。自己モデル

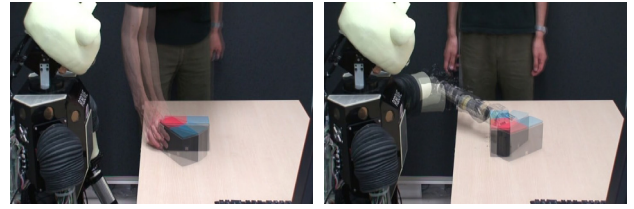


Fig.7 幼児による親の模倣 (右回転)

を他者の動作予測に用いることが、ミラーニューロン発現の起源となっていることが示唆される。

6. おわりに

本稿では、模倣能力獲得プロセスのモデル化を行い、人型ロボット Robovie-IIs をテストベッドとしてその妥当性を検証した。人間の幼児期における「三項関係」と「親による幼児の模倣」に着目した卓上物体の模倣タスクを設定し、自己モデルの投影に基づく模倣手法を提案した。ロボットの自己モデルとして時系列を学習可能な RNNPB を用い、他者モデルとして RNNPB に階層型 NN を付加したモデルを用いた。本手法は4フェーズからなる。まずロボットは身体バブリングにより、RNNPB を用いて動作と物体の挙動を自己組織的に結び付けた。続いて親と幼児の模倣インタラクション時に、構成済の自己モデルを再利用することで、主観的情報と客観的情報の変換を行う変換モジュール群を自己組織的に獲得した。最後に、ロボットが構成済の他者モデルを用いて人間の動作を認識し、模倣動作が生成できることを確認した。また、本手法と人間の認知発達との関連も確認された。

今後は、他者の身体情報を考慮した模倣手法を開発する予定である。

謝辞 本研究は科研費若手研究 A (No. 17680017) 栢森情報科学振興財団設立 10 周年記念特別研究助成、理研の支援を受けた。

参考文献

- [1] S. Jones. Infants learn to imitate by being imitated. In *Proc. IEEE/RSJ Int. Conf. Development and Learning*, 2006.
- [2] T. Makino and K. Aihara. Self-observation principle for estimating the other's internal state. In *Mathematical Engineering Technical Reports METR 2003-36*, the University of Tokyo, 2003.
- [3] A. Nakazawa, S. Nakaoka, K. Ikeuchi, and K. Yokoi. Imitating human dance motions through motion structure analysis. In *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, pp. 2544–2593, 2002.
- [4] D. Rumelhart, G. Hinton, and R. Williams. Learning internal representation by error propagation. In *MIT Press, Cambridge, MA*, pp. 318–362, 1986.
- [5] J. Tani and M. Ito. Self-organization of behavioral primitives as multiple attractor dynamics: A robot experiment. In *IEEE Trans. on SMC Part A*, Vol.33, No.4, pp. 481–488, 2003.
- [6] R. Yokoya, T. Ogata, J. Tani, K. Komatani, and H. G. Okuno. Experience based imitation using rnnpb. In *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 2006.