

ロボットの身体経験に基づく RNNPBを用いた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)

Experience-Based Imitation Using RNNPB

*Ryunosuke YOKOYA[†], Tetsuya OGATA[†], Jun TANI[‡],

Kazunori KOMATANI[†], Hiroshi G. OKUNO[†]

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

Abstract— Robot imitation is a useful and promising alternative to robot programming. Robot imitation involves two crucial issues. The first is how a robot can imitate a human whose physical structure and properties differ greatly from its own. The second is how the robot can generate various motions from finite programmable patterns (generalization). This paper describes a novel approach to robot imitation based on its own physical experiences. Let us consider a target task of moving an object on a table. For imitation, we focused on an active sensing process in which the robot acquires the relation between the object's motion and its own arm motion. For generalization, we applied a recurrent neural network with parametric bias (RNNPB) model to enable recognition/generation of imitation motions. Experimental results demonstrated that our method enabled the robot to imitate not only motion it has experienced but also unknown motion, which proved its capability for generalization.

Key Words: Imitation, Active Sensing, Humanoid Robot, RNNPB

1. はじめに

本研究の最終目標は、人間を模倣するロボットを開発することである。人間の大人は、他人の動作を容易に見真似ることができる。幼児でさえ、表情や身振り手振りを真似ることができる。人間はこのような模倣能力を用いることで、他人から短時間のうちに新たな動作を獲得することができるのである。ロボットによる学習の観点から考えると、人間を模倣するロボットの開発は学習プロセスの著しい高速化につながると言える [1]。実ロボットは耐久性の面で脆弱であるため、学習負荷は重大な問題である。また、人間が手作業であらゆる動作をロボットにプログラムすることは、現実的に困難である。

ハードウェア技術の進歩に伴い、ヒューマノイドロボットは二足歩行や走行、起立動作等、今や様々な動作を生成することが可能となってきた。一部のロボットは、人間の模倣にも挑戦している。中澤らは、モーションキャプチャを用いて舞踊を模倣するロボットを開発している [2]。しかし、ロボットは人間の身体の軌道を再現することを目標としている。そのため、人間の身体構造との違いから、一部のモーターには非常に大きな負荷がかかってしまうと考えられる。また、この研究も含め、従来研究の多くは認識プロセスをパターンクラスタリングとして設計しており、動作生成プロセスは認識プロセスから分離されている。

ロボット模倣の実現のため、本稿では“ミラーニューロン”と“身体バブリング”の2つの要素に着目する。

ミラーニューロンはもともと猿の脳の運動前野で発見されたニューロンであり、ある特定の動作を自身で生

成した際と、同様の動作を他者が生成している様子を観察した際に、共通して発火するものである [3]。これは、脳内で認識と生成が同じシステムとして機能していることを示唆している。本稿では、認識と生成を同じ機構の中で実現できる、Recurrent Neural Network with Parametric Bias (RNNPB) [4] と呼ばれるニューラルネットワークモデルを使用する。

身体バブリングは、幼児が自身の筋肉の動きとそれに伴う身体部位の構成の関係性を獲得する学習過程である [5]。これに対応するものとして、本稿ではアクティブセンシングに基づく学習を導入する。これは、ロボットが自身の動作と対象の挙動との関係性を獲得する学習過程である。

実験タスクは卓上物体の移動であり、提案手法をヒューマノイドロボット Robovie-IIs に実装する。

以下、2. 章ではアクティブセンシングに基づく提案手法と、学習モデルである RNNPB 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. 章で本手法の汎化能力について考察する。最後に 6. 章で本稿をまとめる。

2. アクティブセンシングに基づく模倣手法

2-1 模倣プロセスの概要

ここでは、ロボットがアクティブセンシングに基づき人間を模倣する手法の概要を述べる。

模倣プロセスは学習、観察、動作生成の3つのフェーズで構成される (Fig.1 参照)。プロセスの概要は以下の通りである。

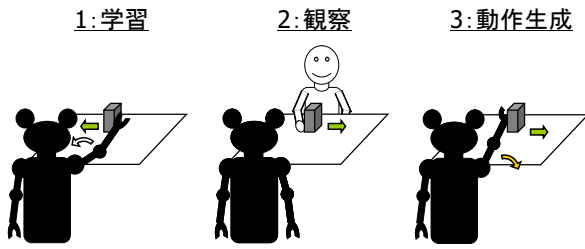


Fig.1 模倣プロセス

1. 学習 (物体認識)
ロボットがプログラムされた動作を生成することで物体を操作し、自身の動作と物体の挙動を結び付ける (アクティブセンシング)。
2. 観察 (動作計画)
ロボットが人間によって操作される物体を観察し、同様の物体操作となるようなアーム動作を計画する。観察する際、人間の動きではなく物体の挙動のみに着目する。
3. 動作生成 (模倣)
ロボットが観察フェーズで計画した動作を実際に生成する。

ここで、呈示された物体操作からいかに適切なアーム動作を決定するかが問題となる。ロボットの動作と物体の挙動を結び付ける手法が必要である。ロボットは、有限の学習パターンから動作を生成しなければならない。

RNNPB を用いることで物体の挙動とロボットの動作を結び付けることができる。また、RNNPB は動作プリミティブが自己組織的に獲得される点や、汎化性能が高いという利点も有している。そこで本研究では、RNNPB を用いてロボットの動作と物体挙動を関連付ける PB 値を獲得し、その PB 値によって観察時の物体挙動と生成動作の対応付けを行う。

2.2 学習モデル

本節では、本手法で用いる学習モデルである RNNPB とその学習アルゴリズムについて述べる。

2.2.1 RNNPB

RNNPB は谷らによって提案された学習モデルであり、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。

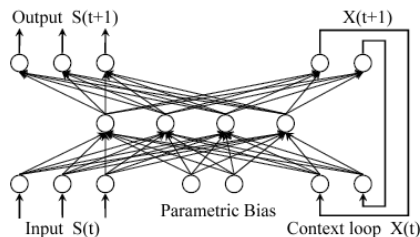


Fig.2 RNNPB

RNNPB は通常の RNN と同様、Back Propagation Through Time (BPTT) 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 [6]。また重みの学習と同時に、PB 値の学習が行われる。

2.2.2 PB 値の計算手法

学習時、一般的に PB 値は以下のように計算される。

$$\delta\rho_t = k_{bp} \cdot \sum_{t-l/2}^{t+l/2} \delta_t^{bp} + k_{nb}(\rho_{t+1} - 2\rho_t + \rho_{t-1}), \quad (1)$$

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

ただし、 k_{bp} と k_{nb} 、 l は定数、 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) の第一項は、出力層から PB 層に伝播された学習誤差 δ_t^{bp} を前後 l ステップ分積算している。これにより、誤差の局所的な影響を緩和し、出力系列が構造的に変化した時にのみ PB 値が変化する。式 (1) の第二項は PB 層の内部値を前後ステップで平均化し、PB 値が急激に変化することを防いでいる。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。

本稿では各物体操作について一意の PB 値を獲得する必要があるため、PB 値の変更は以下の式に基づいて行われる。

$$\delta\rho_t = k_{bp} \cdot \sum_0^T \delta_t^{bp} \quad (3)$$

ただし、 T は時系列長である。式 (3) により、学習誤差は固定ステップ分ではなく、全ステップ T 分が積算される。

2.3 観察及び動作生成時の計算手法

学習フェーズで BPTT によって RNNPB が構成され、PB 値が学習された後、RNNPB は観察と動作生成フェーズにおいて使用される。本節では、観察と動作生成フェーズにおいて RNNPB がどのように使用されるかについて説明する。

2.3.1 観察時の物体操作認識手法

観察時の動作計画における、PB 値の計算手法を述べる。学習時に BPTT によって求められた RNNPB の重みを固定し、式 (2)、(3) に基づいて PB 値のみを学習する。学習時とは異なり観察時は腕のモーターは動かさないため、それらの入力データは存在しない。そこで、初期ステップ (Step0) において腕のモーター値は初期値 (初期姿勢におけるモーター指令値) のみを動作入力層に与え、次ステップ (Step1) 以降はクロズド・ループで forward 計算を行う (Fig.3 参照)。ステップ $t-1$ における出力がステップ t における入力となる。つまり、モーター入力層の BPTT における挙動はコンテキスト層と同様である。

2.3.2 動作生成手法

動作生成時にロボットへ転送されるモーター指令値の計算手法を述べる (Fig.4 参照)。モーター指令値は、RNNPB の forward 計算によって得られる。RNNPB の重みは固定したまま、リアルタイムで各データを入力する。また、各ステップにおいて入力される PB 値は、観察時に物体操作として認識された PB 値である。各

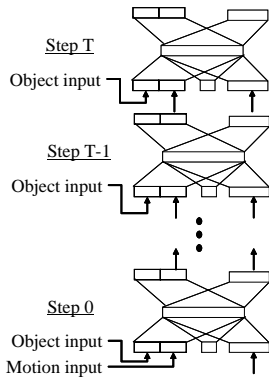


Fig.3 PB 値認識時の forward 計算

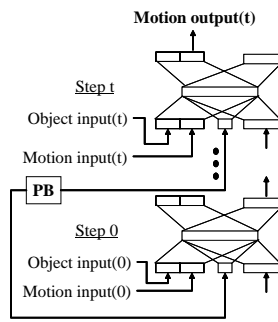


Fig.4 動作生成時の forward 計算

ステップにおけるモーター層の出力値 (次ステップの予測値) が、モーター指令値としてロボットへ転送される。

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-1Is である。Robovie は首に 3 自由度、各腕に 4 自由度のモーターを有しており、頭部には CCD カメラを搭載している。これらから、実験に必要なデータを収集する。

操作対象は箱型の物体である。物体の回転の検出を容易にするために、物体の上面は赤青の 2 色に分けられている。

3-2 実験システム

実験システムの概観は Fig.5 の通りである。物体の重心がカメラの中央にくるよう首モーターを制御することで、カメラは物体を追従する。リアルタイムに処理する必要があるため、モーター駆動のモジュールはロボット内部の PC を、特徴抽出やモーター指令値の計算は外部 PC を用いて行う。学習に用いる RNNPB のサイズは入出力層が 9、中間層が 15、コンテキスト層が 10、PB 層が 2 である。

RNNPB への入力データとして、以下のセンサー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・青色を抽出し、色付きの各面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。

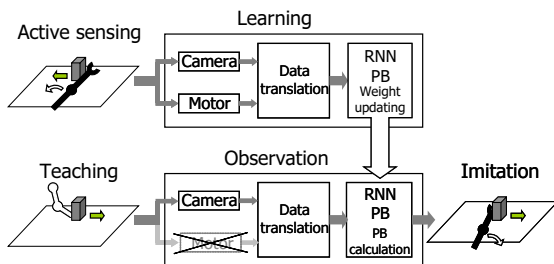


Fig.5 システム概観

モーター情報 (5 次元)

首 (ピッチとヨーの 2 自由度) と 左腕 (肩のロールとヨー、肘ピッチの 3 自由度) を使用する。ただし、実験において肩のピッチと首のロール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 400ms 毎に 1 ステップ、計 20 ステップ分を取得する。

4. 模倣動作生成実験

ロボットにとって学習済の物体操作のみでなく、未知の物体操作の模倣実験を行った。

4-1 実験タスク

実験タスクは箱型物体の移動である。操作としては、左から右への平行移動 (左 右) と右回転を扱い、それぞれについて移動距離の度合を 3 段階設定している。このように各操作について複数のパターンを用意することで、モーター値のゆるやかな変化を学習できることが期待される。Table 1 の学習操作 1~6 を学習時に用いる。また、Table 1 の観察操作 1~3 を観察時に用いる。なお、Table 1 の観察操作 3「右回転させながら左から右へ移動させる」操作が、ロボットにとって未知の物体操作となっている。

4-2 実験方法

まず、Table 1 の学習操作 1~6 の 6 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。各モーションパターンについて 1 データずつ、全 6 データを取得する。取得データから、RNNPB の重みを 20 万回学習させ構成する。学習に要した時間は約 10 分である。

次に、人間が行う Table 1 の観察操作 1~3 をロボットが観察する。各観察操作について 1 データずつ、全 3 パターンを取得する。取得データから、学習時に構成済の RNNPB を用いて 5000 回学習することで PB 値を認識する。PB 値の計算に要した時間は 15 秒である。

最後に、観察時に得られた PB 値を RNNPB へ入力することで動作生成を行う。

Table 1 物体操作

	移動方向	移動距離
学習操作 1	左 右	小
2	左 右	中
3	左 右	大
4	右回転	小
5	右回転	中
6	右回転	大
観察操作 1	左 右	大
2	右回転	大
3*	右回転+左 右	大+大

*: 未知操作

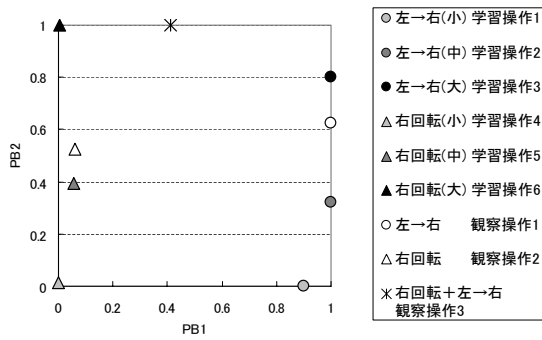


Fig.6 実験において形成された PB 空間

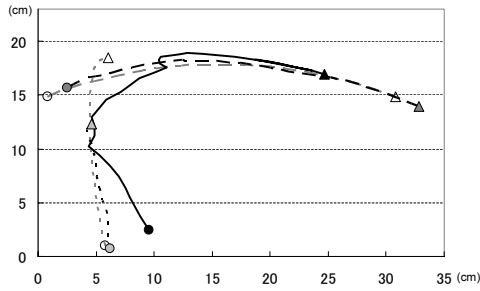


Fig.7 ロボットの手先軌道

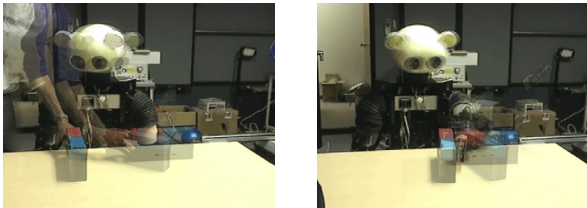


Fig.8 観察と動作生成 (右回転+左 右)

4.3 実験結果

学習及び観察の結果構成された 2 次元 PB 空間を Fig.6 に示す. Fig.6 より, 学習の結果得られた PB 値は物体操作及び移動距離ごとに組織化されていることが確認できる. また, 観察の結果認識された PB 値のうち, 学習済の操作については学習データの同操作に対する PB 値近辺にプロットされており, 未知の操作については中間的な位置にプロットされている.

学習時及び動作生成時におけるロボットの手先の軌道を Fig.7 に示す. ただし Fig.7 は, 手先軌道を机に対して真上から見たものである. 図の下側がロボット本体の位置となる. ロボットが観察操作を観察する様子と, その際得られた PB 値を用いて動作を生成している様子の一例を Fig.8 に示す. Fig.7 より, 学習済である「左 右」と「右回転」については, ほぼ正確に学習時の物体操作が再現されていることが確認できる. また, 未知の物体操作である「右回転+左 右」については, 学習済操作の組み合わせとして再現されている.

5. 考察

学習により, ロボットが動作プリミティブを暗黙的に獲得できたと考えられる. 具体的には, 「左 右」の操作に対しては腕を左から右へ動かす (肩ロール減少) プリミティブが, 「右回転」の操作に対しては腕を伸展させる (肩ロール減少, 肘ピッチ減少) プリミティブが獲得された. また, 未知の物体操作である「右回転+左 右」については, それらの組み合わせとして汎化された動作が生成された.

6. おわりに

本研究では, 物体操作の第一段階として卓上物体の移動をタスクとし, アクティブセンシングに伴う物体の挙動に着目した模倣動作生成手法を提案した. 提案手法は学習・観察・動作生成の 3 ステップからなる. ロボットの学習負荷の軽減のため, 汎化能力を有する RNNPB を学習モデルとして使用した. 具体的には, RNNPB により動作と物体の挙動を自己組織的に結び付け, 観察された物体の挙動から動作の連想を行い, 模倣動作の生成を行った. ヒューマノイドロボット Robovie-2s をテストベッドとし, 動作に伴って観測される画像情報及びモーター情報を取得し, 学習するシステムを構築した. 実装システムにより, 箱型物体を使用した実験を行った. その結果, 学習済の物体操作のみでなく, 未知の操作についても既知の操作の組み合わせとして模倣が可能であることを確認した.

今後は, より多くの動作学習に適用することで本手法の有効性を確認し, より多様な動作パターンを生成する手法を開発する予定である.

謝辞 本研究の一部は, 日本学術振興会科学研究費補助金, 21 世紀 COE プログラム, 栢森研究助成, 理研の支援を受けた.

参考文献

- [1] S. Schaal, "Is Imitation Learning the Route to Humanoid Robots?," Trends in Cognitive Sciences, vol. 3, no. 6, pp. 233-242, 1999.
- [2] A. Nakazawa, S. Nakaoka, K. Ikeuchi, and K. Yokoi, "Imitating Human Dance Motions through Motion Structure Analysis," Proc. IROS 2002, pp. 2539-2544, 2002.
- [3] G. Rizzolatti and L. Craighero, "The mirror-neuron system," Annu. Rev. Neurosci., 27:169-192, 2004.
- [4] J. Tani and M. Ito, "Self-Organization of Behavioral Primitives as Multiple Attractor Motions: A Robot Experiment," IEEE Trans. on SMC Part A, vol. 33, no. 4, pp. 481-488, 2003.
- [5] R. Rao, A. Shon, and A. Meltzoff, "A Bayesian Model of Imitation in Infants and Robots," Imitation and Social Learning in Robots, Humans, and Animals, Cambridge University Press, 2005.
- [6] D. Rumelhart, G. Hinton, and R. Williams, "Learning internal representation by error propagation," in D.E. Rumelhart and J.L. McClelland, editors, Parallel Distributed Processing (Cambridge, MA: MIT Press), 1986.