

再帰型神経回路モデルを用いた観察視点の獲得による ロボットの模倣学習

Imitative Learning of Robot by Acquiring a Viewpoint Utilizing Recurrent Neural Network

中條亨一 村田真悟 有江浩明 尾形哲也
Ryoichi Nakajo*¹ Shingo Murata*² Hiroaki Arie*¹ Tetsuya Ogata*¹

*¹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

*²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

In imitative learning, acquiring viewpoints of others' actions is a crucial problem. In cognitive developmental robotics, existing models have been requiring coordinate transformations provided by human designers or involved modules to acquire viewpoints. This paper introduces an imitative model that enables a robot to predict viewpoints of the self and others from training data. To allow a robot to acquire representations of behaviors and viewpoints autonomously, we use a recurrent neural network model. Several neural units describing information about actions or positional relationships between the robot and the demonstrator are added to the network model. The training data are labeled with types of behaviors and viewpoints, and are linked to each internal state. We implemented this model in a robot and carried out training of object manipulation. According to the model evaluation, we confirmed the representations of behaviors and viewpoints were self-organized in a space composed of the neural units.

1. はじめに

人間の認知能力に対する研究アプローチの一つとして、認知ロボティクスが挙げられる。これは、人の認知プロセスの原理に関して、仮説とロボットを用いる検証の繰り返しから探求する構成論的アプローチに基づく手法であり、ロボットのもつ身体による実環境との相互作用とその経時的発展から認知に関する知見を得ようと試みる手法である [Asada 01]。

本研究では、認知発達ロボティクスの視点から、人間の認知能力の中でも、模倣能力の獲得に着目する。模倣能力とは、他者の行動を観察し、同じ行動をする能力である。人間は模倣能力を持つために、試行錯誤によらない効率的な行動学習が可能である。特に、他者の行動を模倣することは発達段階の幼児が新たな知識を学習するために重要な能力である。

模倣学習では、他者行為における観察視点の獲得が大きな課題である。人間が模倣を行なう際、正面からだけでなく、様々な向きから他者の行為を観察した場合でもその行為を模倣することができる。しかし、従来の他者行為の観察視点に関する研究の多くは、他者行為の視点を自己の視点に切り替える座標変換を設計者が呈示するものであった [Billard 04, Lopes 05]。また、Ogataらは、他者行為の位置関係を複数の変換モジュールに獲得させることで、自己の行為との対応付けを行なった [Ogata 09]。しかし、人間が他者の行為を模倣するとき、(1)座標変換の規則を事前に設計することは困難である、(2)自他の多様な位置関係を対応づけるモジュールをすべて準備することは困難である、という問題がある。自他の位置関係や多様な動作を、事前情報なしに獲得し推定可能なモデルが求められる。そこで本研究では、ロボットが他者の観察視点を教示データから獲得することを目的とし、再帰型神経回路モデル (Recurrent Neural Network, RNN) を用いる。

2. 提案手法

ロボットが与えられる教示データから他者行為の観察視点を獲得するために、RNNを用いる。RNNの内部状態は学習される時系列ごとに初期値を持っており、この初期値は Back Propagation Through Time により学習される。そのため、時系列の情報は初期値に埋め込むことが可能である。

RNNの内部状態の初期値に動作と観察視点の情報を埋め込むため、自他の動作と観察視点を表現する少数のニューロンを加える。それらのニューロンを加えたモデルを図1に示す。

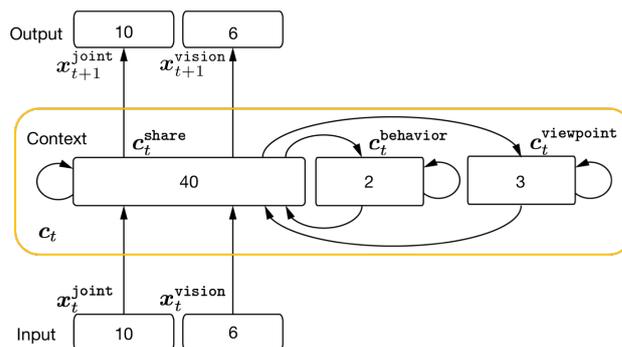


図1: RNNの構造. 数字はニューロン数を示す. Context層を3つに分け、行為と観察視点を表現するニューロンを加えた。

2.1 教示データへのラベリング

図1に示すモデルの内部状態に行為と観察視点の情報を埋め込むため、教示データに事前処理を施す。本稿では、事前に行為と観察視点が学習者(ロボット)に与えられていると仮定し、教示データを行為と観察視点によってラベリングする。ラベリングにより、同じ動作・観察視点の時系列に対して、内部状態の初期値を共有する拘束を加えることができる。

2.2 内部状態の初期値へのバインドを考慮した学習

教示データへのラベリングにより、ある運動時系列にラベル l が与えられているとする。ラベル集合 \mathcal{L} を考え、同じラベルの時系列について、内部状態の初期値を近づけるバインドを加える。ラベル $l \in \mathcal{L}$ である時系列 s について、初期値 $\mathbf{u}_0^{(s,l)}$ のバインドを次式に定める。

$$\mathbf{u}_0^{(s,l)} \rightarrow \rho_{\mathcal{L}} \overline{\mathbf{u}_0^{(l)}} + (1 - \rho_{\mathcal{L}}) \mathbf{u}_0^{(s,l)} \quad (1)$$

ここで、 $\overline{\mathbf{u}_0^{(l)}}$ はラベル $l \in \mathcal{L}$ であるすべての初期値の平均ベクトルであり、 $\rho_{\mathcal{L}}$ はラベル集合 \mathcal{L} のバインド率である。式 (1) の第 1 項は同じ行為・視点のラベルを持つ時系列で初期値を近づける役割を持ち、第 2 項はそれぞれの時系列の予測誤差を減らす役割を担う。本稿では、それぞれのコンテキストニューロンの集合にバインド率を次のように定める。

$$\rho_{share} = 1.0, \quad \rho_{behavior} = \rho_{viewpoint} = 0.7. \quad (2)$$

3. 実験

本実験では、ヒューマノイド NAO を利用し、物体操作運動の時系列を教示データとして与える。物体の操作者にはブロックを、(1) 上下に動かす (*move-up*), (2) 左右に動かす (*move-side*), (3) 左手で触れる (*touch-left*), (4) 右手で触れる (*touch-right*) のいずれかが与えられる。ロボットは同じ初期姿勢から運動を始め、運動の始めと終わりで、腕は初期姿勢を取る。

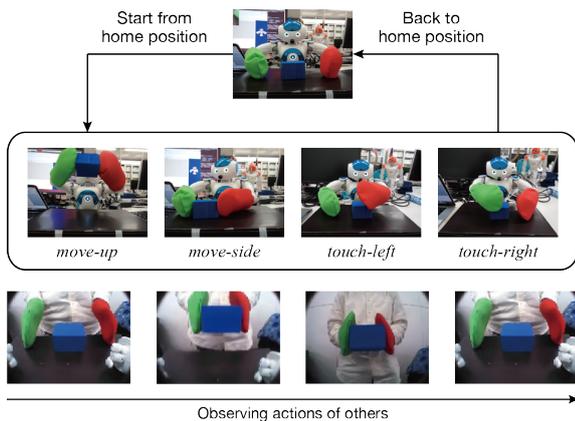


図 2: ロボットによる動作 (上) と他者動作の観察 (下)

3.1 教示データ

教示データはロボットの関節角度 10 次元 (両腕 8 自由度、首 2 自由度) と、ロボットのカメラから撮影する画像の色重心座標 6 次元 (赤, 緑, 青の xy 座標) を $[-0.8, 0.8]$ にスケーリングして記録する。また、他者の位置関係は、ロボット自身の行為を 0° 、正面から観察する他者行為を 180° として、反時計回りに、 $0^\circ, 120^\circ, 150^\circ, 180^\circ, 210^\circ, 240^\circ$ を与える。他者の物体操作時、ロボットの腕の関節角度は初期姿勢に固定され、首の関節角度と画像のセンサ入力を記録する。教示データは、4 動作および 6 視点から 24 種類の時系列を作成する。

3.2 運動時系列の学習結果

本研究では学習モデルに Stochastic Continuous Time Recurrent Neural Network (S-CTRNN) を用いた [Murata 13]。ニューロン数を図 1 のように定め、24 種の時系列を 300,000 回学習させたところ、得られたパラメータによる教示データ

の再現を確認した。また、コンテキスト *behavior*, *viewpoint* の初期値に主成分分析 (PCA) を行なうと、4 動作と 6 視点それぞれの内部状態の初期値空間中に自己組織化している。コンテキスト *viewpoint* の初期値に PCA を行なった結果を図 3 に示す。図 3 より、神経回路モデルの内部状態に動作と視点

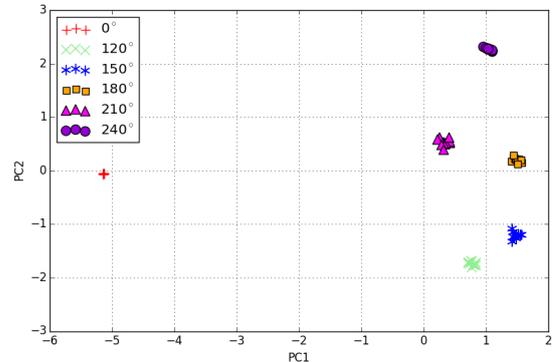


図 3: コンテキスト *viewpoint* の初期値 (PC1-PC2)

表現するニューロンを与え、教示データのラベリングをして学習させることで、ロボットが自他の区別と他者の観察視点を内部状態の初期値空間中に自己組織的に獲得できたとと言える。

4. まとめ

本研究では、模倣学習において、設計者による座標変換や他行為の対応付けを行なうモジュールの獲得によらず、ロボットが教示データから自他の観察視点を獲得することを目的とした。そのため、RNN の内部状態に運動と観察視点を表現する少数のニューロンを追加し、教示データを運動・観察視点ごとにラベリングすることで、それぞれの内部状態の初期値に拘束を加える手法を用いた。学習により獲得された内部状態のうち、観察視点に関する内部状態の初期値を解析したところ、自他と他者行為の観察視pointsの区別が自己組織的に構造化していることが確認された。

謝辞

本研究は、さきかけ領域研究「情報環境と人」及び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 B(26870649) の助成を受けた。

参考文献

- [Asada 01] Asada, M., MacDorman, K. F., Ishiguro, H., and Kuniyoshi, Y.: Cognitive developmental robotics as a new paradigm for the design of humanoid robots: *Robotics and Autonomous Systems*, Vol. 37, pp.185-193 (2001)
- [Ogata 09] Ogata, T., Yokoya, R., Tani, J., Komatani, K., and Okuno, H.: Prediction and Imitation of Other's Motions by Reusing Own Forward-Inverse Model in Robots: *Proceedings of IEEE-RAS International Conference on Robots and Automation*, pp.4144-4149 (2009)
- [Billard 04] Billard, A., Epars, Y., Calinon, S., Schaal, and S. Cheng, G.: Discovering optimal imitation strategies: *Robotics and Autonomous Systems*, Vol. 47, No. 2-3, pp.69-77 (2004)
- [Lopes 05] Lopes, M. and Santos-Victor, J.: Visual Learning by Imitation With Motor Representations: *IEEE Trans. on Systems, Man, and Cybernetics*, Vol. 35, No. 3, pp.438-449 (2005)
- [Murata 13] Murata, S., Namikawa, J., Arie, H., Sugano, S., and Tani, J.: Learning to Reproduce Fluctuating Time Series by Inferring Their Time-Dependent Stochastic Properties: Application in Robot Learning via Tutoring: *IEEE Transactions on Autonomous Mental Development*, Vol. 5, No. 4, pp.298-310 (2013)