

バイラテラル制御に基づく模倣学習の 多自由度ロボットアームへの拡張

○七種勇樹（筑波大） 境野翔（筑波大） 辻俊明（埼玉大）

1. 序論

ロボットが社会に普及し、人間が行う家事や肉体労働をロボットが代替することが期待されている。家事や肉体労働を行うロボットには、周囲の環境や操作対象物に合わせた柔軟な動作が求められる。しかし、環境やロボット自身の動特性をあらかじめ数式化し、タスクに合わせた動作を設計することは非常に困難である。

そこで近年、機械学習を用いた動作生成が盛んに研究されている [1]。機械学習を用いた動作生成手法は、センサの情報から環境の特徴量を自動で抽出可能であるため、人間による動作の設計が不要である。なかでも、人間によって取得された教師データをもとに教師あり学習を行う模倣学習が近年注目されている [2]。模倣学習は、人間の柔軟な動作を模倣することが可能であり、他の機械学習手法と比較して効率的な学習が可能である。また、教師データ収集時に人間の力の情報を保存しておくことで、力の情報まで模倣することが可能となっている [3]。力の情報を用いることで位置の変動に対してロバストな動作生成が可能となることが報告されている [4]。

従来、力の情報を用いた模倣学習の一つとしてバイラテラル制御に基づく模倣学習が提案されている [5]。[5] では、4ch バイラテラル制御を用いて教師データを収集し、教師データ収集時と同じ制御系によって自律動作を行うことで従来の模倣学習 [2, 3, 4] では不可能であった人間同等の速度での自律動作が可能となることを示した。さらに、[5] では人間と協調して物体を運搬するタスクを行うことで、より定式化困難な人間との動的なインタラクションを学習可能であることが実証され、有効性が示されている。しかし、バイラテラル制御に基づく模倣学習に関する従来研究 [5, 6] では、3 自由度のマニピュレータを用いた検証を行っており、ロボットの動作する空間が非常に制限されている。ロボットを多自由度化することでロボットの動作空間はより広がり、ロボットの関節間の相互作用を考慮する必要も生じる。したがって、多自由度ロボットによる複雑な動作を対象とした場合における本手法の有効性は示されていない。

そこで本研究では、ホワイトボードの線を消すタスク、人とロボットの協調タスクを 7 自由度のロボットアームを用いて行うことで、バイラテラル制御に基づく模倣学習が多自由度ロボットアームの動作生成に適用可能であるか検証する。

本論文の構成を以下に示す。2 章では実装した制御系について説明し、3 章ではバイラテラル制御に基づく模倣学習を説明する。4 章では本研究で行った実験を述べ、最後に 5 章で本研究のまとめを行う。

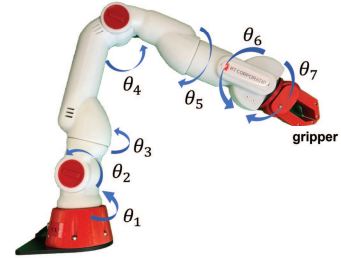


図 1 7 自由度ロボットアームの関節の定義

2. 制御システム

2.1 ロボットアーム

本研究では、図 1 に示す株式会社アールティの CRANE-X7 を 2 台用いた。7 自由度ロボットの各関節に対応する角度 $\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7$ を図 1 に示すように定義する。

2.2 バイラテラル制御

バイラテラル制御とはリーダー、フォロワーと呼ばれる 2 台のロボットを同期させる遠隔操作技術であり、フォロワーの力応答をリーダーにフィードバックすることが可能な技術である [7]。本論文では [5] 同様に、4ch 型バイラテラル制御を用いた模倣学習の検討を行った。4ch 型バイラテラル制御はリーダー、フォロワー双方に位置制御と力制御を有する技術である。4ch 型バイラテラル制御の制御則は、ロボットの角度応答値ベクトル θ^{res} とトルク応答値ベクトル τ^{res} を使用して、

$$\theta_l^{res} - \theta_f^{res} = 0, \quad (1)$$

$$\tau_l^{res} + \tau_f^{res} = 0, \quad (2)$$

と表される。ここで、下付き文字 l, f はそれぞれリーダーとフォロワーを表し、上付き文字 res は応答値を表す。本研究では 7 自由度ロボットを用いたため、角度応答値ベクトル θ^{res} は $\theta^{res} = [\theta_1^{res}, \theta_2^{res}, \theta_3^{res}, \theta_4^{res}, \theta_5^{res}, \theta_6^{res}, \theta_7^{res}]^T$ 、トルク応答値ベクトル τ^{res} は $\tau^{res} = [\tau_1^{res}, \tau_2^{res}, \tau_3^{res}, \tau_4^{res}, \tau_5^{res}, \tau_6^{res}, \tau_7^{res}]^T$ である。4ch 型バイラテラル制御のブロック線図を図 2 に示す。

2.3 制御器

本研究で用いた制御系を図 3 に示す。図中 $\theta, \dot{\theta}, \tau$ はそれぞれ角度、角速度、トルクを表し、上付き文字 cmd, res, ref, dis はそれぞれ指令値、応答値、参照値、外乱

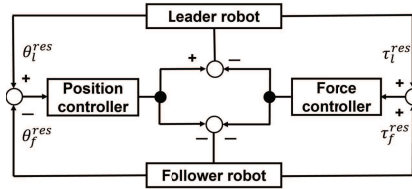


図2 4ch型バイラテラル制御のブロック線図

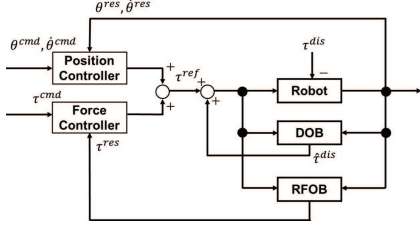


図3 本研究で使用した制御器のブロック線図

値を表す。4ch型バイラテラル制御による式(1),(2)を満たすリーダー、フォロワーそれぞれのトルク参照値ベクトル τ^{ref} は次の式で算出される。

$$\tau_l^{ref} = -C_p(\theta_l^{res} - \theta_f^{res}) - C_f(\tau_f^{res} + \tau_l^{res}), \quad (3)$$

$$\tau_f^{ref} = C_p(\theta_l^{res} - \theta_f^{res}) - C_f(\tau_f^{res} + \tau_l^{res}), \quad (4)$$

ここで、位置制御器のゲインベクトル C_p は、位置制御ゲインベクトル K_p と速度制御ゲインベクトル K_d 及び慣性ベクトル J を用いて $C_p = \frac{J}{2}(K_p + K_d s)$ である。 s はラプラス演算子である。力制御器のゲインベクトル C_f は、力制御ゲインベクトル K_f を用いて $C_f = \frac{1}{2}K_f$ である。位置制御ゲインベクトル K_p 、速度制御ゲインベクトル K_d 、力制御ゲインベクトル K_f 、慣性ベクトル J はそれぞれ各関節を表す下付き文字を利用して、

$$K_p = \text{diag}[K_{p1}, K_{p2}, K_{p3}, K_{p4}, K_{p5}, K_{p6}, K_{p7}], \quad (5)$$

$$K_d = \text{diag}[K_{d1}, K_{d2}, K_{d3}, K_{d4}, K_{d5}, K_{d6}, K_{d7}], \quad (6)$$

$$K_f = \text{diag}[K_{f1}, K_{f2}, K_{f3}, K_{f4}, K_{f5}, K_{f6}, K_{f7}], \quad (7)$$

$$J = \text{diag}[J_1, J_2, J_3, J_4, J_5, J_6, J_7] \quad (8)$$

である。各関節の角度はロボットのエンコーダにより取得し、角速度はその疑似微分によって算出した。外乱トルク τ_{dis} は外乱オブザーバ (Disturbance Observer: DOB)[8] を用いて算出し、トルク応答値 τ_{res} は反力推定オブザーバ (Reaction Force Observer: RFOB)[9] を用いて算出した。制御やRFOBに使用する物理パラメータの同定はYamazakiらの手法[10]を応用した方法で行った。使用した各ゲインとカットオフ周波数、各パラメータは (<https://mclab.iit.tsukuba.ac.jp/parameter.php>) に掲載した。

3. バイラテラル制御に基づく模倣学習

バイラテラル制御に基づく模倣学習は、教師データの収集、モデルの学習、自律動作と3つのフェーズに分かれているため本章では各フェーズそれぞれの説明を行う。

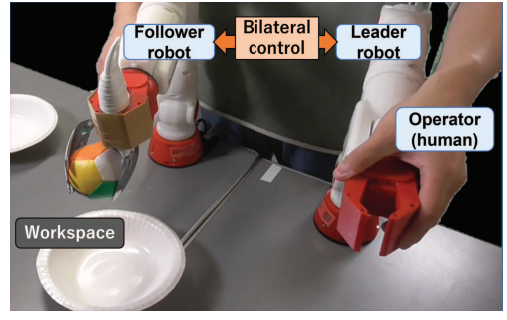


図4 データ収集時

3.1 教師データの収集

訓練データの取得時は、図4に示すように教示者がバイラテラル制御におけるリーダーを操作し、ワークスペースに配置したフォロワーを遠隔操作することでタスクを実行する。タスク実行時のリーダー、フォロワーそれぞれのロボットは500 [Hz]の制御周期で制御されており、角度、角速度、トルク応答値を教師データとして保存した。本研究では、ロボットのグripperは制御しておらず、タスクの実行に使用するスプーンやホワイトボードクリーナーはグripperに直接固定した。また、7自由度のロボットアームは冗長性をもつため、7自由度うち図1における θ_3 の関節は一定の目標値で独立した位置制御を行い、残りの6自由度でバイラテラル制御を行った。

3.2 モデルの学習

取得した教師データをもとに、時刻 t のフォロワーにおける6自由度の角度、角速度、トルク応答値の18次元の情報から、時刻 $t + 20$ [ms] のリーダーの角度、角速度、トルク応答値の18次元の情報を予測するようにニューラルネットワーク (Neural Network: NN) の学習を行った。ニューラルネットワークモデルは時系列情報が学習可能な Long Short-Term Memory (LSTM) と全結合層により構成された。モデルの詳細は4.で述べる。Loss関数には、最小二乗誤差を使用した。入力となる関節角度、角速度、トルクはそれぞれ次元が異なるため前処理として Min-Max Normalization による正規化を行い、全データが最大値1、最小値0となるようにした。

3.3 自律動作

学習済みのモデルを使用してフォロワー1台で自律動作を行った。自律動作時にはフォロワーの応答値がリアルタイムでセンシングされ、NNモデルへと入力された。その後、学習時と同様に入力から20 [ms] 後のリーダーの応答値が予測され、式(4)に従ってフォロワーの指令値としてフォロワーの制御系へ入力した。

4. 実験

4.1 実験1(ホワイトボードの線を消すタスク)

4.1.1 タスクデザイン

本タスクの実験環境を図5に示す。フォロワーはホワイトボードの上で、ホワイトボードに書かれた線を消す前後の往復動作を行う。60秒間ロボットが停止することなくタスクを実行し、動作中にホワイト

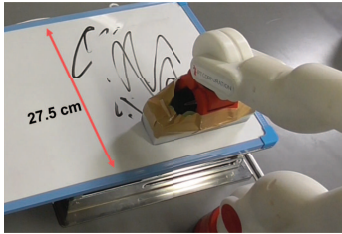


図5 ホワイトボードの線を消すタスクのセットアップ

ボードクリーナをホワイトボードから離さなかった場合をタスク成功とした。これをホワイトボードの高さ 9.5 [cm], 11.0 [cm], 12.5 [cm], 14.0 [cm], 15.5 [cm] の 5 段階の高さで各 5 回ずつ試行した。このうち、ホワイトボードの高さ 11.0 [cm], 14.0 [cm] は未学習の高さである。

4.1.2 教師データと学習モデル

ホワイトボードの高さ 9.5 [cm], 12.5 [cm], 15.5 [cm] の 3 段階の高さでそれぞれ 20 回ずつ計 60 回教師データの取得を行った。取得した教師データに対してデータの水増しを行うことで最終的に教師データの数は $600(3[\text{高さ}] \times 20[\text{試行回数}] \times 10[\text{水増し}])$ サンプルとした。学習モデルには 5 層の LSTM とそれに続く 1 層の全結合層を用い、中間層のユニット数は 200 とした。7 秒に相当する 350 ステップの時系列サンプルを 1 データとし、100 個のデータでミニバッチ学習を行った。

4.1.3 実験結果

実験結果を表 1 に示す。実験には学習回数 1000 回のモデルを使用した。表 1 より、バイラテラル制御に基づく模倣学習を多自由度ロボットに用いることで学習済みの高さおよび未学習の高さの両方で 100% タスクを成功させることが可能であった。未学習の高さにおいてもタスクを成功させる結果が得られたことから、バイラテラル制御に基づく模倣学習は多自由度ロボットにおいても環境との接触を生じるタスクにおいて、位置と力の情報を利用した適切な動作が学習可能であることが示された。

4.2 実験 2(人とロボットの協調タスク)

4.2.1 タスクデザイン

本タスクは、[5] において 3 自由度のマニピュレータで検証された人とロボットの協調タスクを 7 自由度ロボットアームを用いて行った。タスク実行の様子を図 6 に示す。フォロワーは初期位置から右の皿に移動し、皿の上で人と協調動作を行い対象物体を運搬する。運搬後に左の皿に対象物体を置く動作が完了した場合にタスク成功とした。実験は図 7 に示す 8 種類のオブジェクトに対して各 10 回ずつ行った。本タスクでは、人間がどのように動くかはあらかじめ設計されていないため、人間とのインタラクションを考慮した高速で柔軟な動作がロボットに求められる。

4.2.2 教師データと学習モデル

図 7 上部に示す 4 種類のオブジェクトに対してそれぞれ 60 回ずつ計 240 回教師データの取得を行った。さらに 4.1 と同様にデータの水増しを行うことで最終的

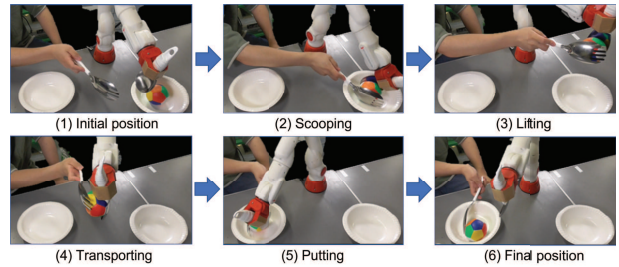


図6 人とロボットの協調タスクのタスク実行の様子

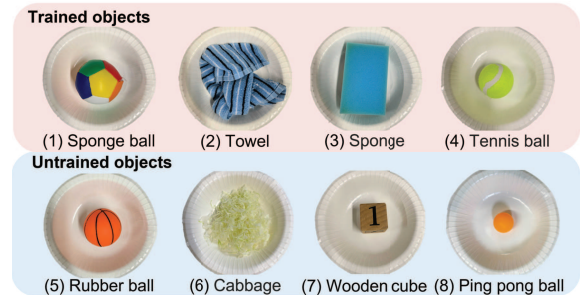


図7 人とロボットの協調タスクにおける学習済みと未学習のオブジェクト

に教師データの数は $2400(4[\text{オブジェクト数}] \times 60[\text{試行回数}] \times 10[\text{水増し}])$ サンプルとした。教師データのタスク完了時間は約 13 秒であった。学習モデルには 8 層の LSTM とそれに続く 1 層の全結合層を用い、中間層のユニット数は 200 とした。7 秒に相当する 350 ステップの時系列サンプルを 1 データとし、100 個のデータでミニバッチ学習を行った。

4.2.3 実験結果

実験結果を表 2 に示す。実験には学習回数 2500 回のモデルを使用した。表 2 より、検証に使用したオブジェクト全体で 89% の高い成功率を示した。タオルやキャベツなどの柔軟物や不定形物に対しても高い成功率を示しており、環境や操作対象物に応じた適切な動作を教師データから学習可能であることが示された。また、運搬中に人間がフォロワーに強い力を加え、押し返すなどの未学習かつ不規則な動作を行った場合にも適切に対応する結果が得られた。こうした人間との動的なインタラクションを考慮した動作をあらかじめ設計することは非常に困難である。特に多自由度ロボットにおいては、ロボット自身の関節間の相互作用も考慮する必要があるためより困難といえる。しかしバイラテラル制御に基づく模倣学習においては、これらの複雑で定式化困難な要素は教師データ収集時に人間が感覚的に補償する。さらにその時の動作データをそのまま保存し学習に利用するため、ニューラルネットワークモデルが人間の操作からあらゆる特徴量を抽出することが可能となる。以上の結果より、バイラテラル制御に基づく模倣学習は多自由度ロボットを対象とした、環境との接触、人間とのインタラクションを生じるタスクにおいてもこれらの要素の定式化不要で動作が学習可能であることから非常に有効であることが明らかとなった。しかし、本研究で学習に必要とされた教師デー

表1 ホワイトボードの線を消すタスクの実験結果

Trained/Untrained [T/U]	T	U	T	U	T
Whitebord height[cm]	9.5	11.0	12.5	14.0	15.5
Success Rate[%]	100(5/5)	100(5/5)	100(5/5)	100(5/5)	100(5/5)

表2 人とロボットの協調タスクの実験結果

Trained/Untrained [T/U]	T	T	T	T	Subtotal	Total
Object	Sponge ball	Towel	Sponge	Tennis ball		
Success Rate [%]	100(10/10)	100(10/10)	90.0(9/10)	100(10/10)	97.5(39/40)	
Trained/Untrained [T/U]	U	U	U	U	Subtotal	Total
Object	Rubber ball	Cabbage	Wooden cube	Ping pong ball		
Success Rate [%]	100(10/10)	80.0(8/10)	40.0(4/10)	100(10/10)	80.0(32/40)	88.9(71/80)

タのサンプル数は240サンプルであり、これを[5]と比較すると、3倍ものサンプル数を要している。動作の設計の簡略化が可能な本手法が今後、より多くの自由度を利用した動作の学習に対して用いられることを想定すると、教師データのサンプル効率も無視できず非常に大きな問題となる。したがってバイラテラル制御に基づく模倣学習においても、より効率的な学習手法を今後検討していく必要があることが明らかとなった。

また、検証結果のうち(7)木のブロックのみ40%と低い成功率となった。今回の検証において、運搬時のロボットアームの手首関節は図6(4)に示すようにワークスペースに対して垂直な状態でタスクを行った。そのため、物体を落下させずにタスクを成功させるためには、物体とスプーンの間に生じる摩擦力和ロボットの発揮力で物体にかかる重力に逆らう必要があった。特に、成功率の低かった木のブロックの表面は非常にめらかで摩擦力が小さく、質量が大きいものであることからロボットには大きい発揮力が求められる。しかし、今回使用したロボットではこの姿勢で運搬可能なほど十分な力が発揮できず、落下させてしまう結果となったと考えられる。この問題を解決するためには、近年盛んに研究されている触覚センサを用いた動作生成[11]を行い、物体の状態に応じて物体が落下しないよう手首姿勢を変化させることでより高い成功率となることが期待できる。

5. まとめ

本研究では、バイラテラル制御に基づく模倣学習を多自由度ロボットアームに拡張した。検証タスクとして、ホワイトボードの線を消すタスクと人とロボットの協調による運搬タスクを行った。実験結果より、バイラテラル制御に基づく模倣学習を多自由度ロボットアームに適用することで、力の情報を利用した柔軟な動作が模倣可能であることが示された。また、人とロボットの協調タスクにおいて人間が未知の動作をした場合にも適切に動作を生成し、タスクを成功させることが可能であることも確認され、環境や人とのインタラクションを考慮した動作が可能となることが示された。以上の結果からバイラテラル制御に基づく模倣学習は、多自由度ロボットの動作生成においても高速で柔軟な人間の動作を模倣可能であること結論付けた。しかし、多自由度化することで学習に必要な教師データのサンプル数は大幅に増加することが明らかとなり、これを改善することを今後の課題とした。

謝辞 本研究はJSTのA-STEP, JPMJTR20RGの支援を受けたものである。

参考文献

- [1] S. Levine, P. Pastor, A. Krizhevsky, J. Ibarz and D. Quillen: "Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection," *The International Journal of Robotics Research*, vol. 37, no. 4-5, pp. 421-436, 2017.
- [2] T. Zhang, Z. McCarthy, O. Jow, D. Lee, K. Goldberg and P. Abbeel: "Deep imitation Learning for complex Manipulation Tasks from Virtual Reality Teleoperation," arXiv:1710.04615v1, 2017.
- [3] P. Kormushev, S. Calinon and D. G. Caldwell: "Imitation Learning of Positional and Force Skills Demonstrated via Kinesthetic Teaching and Haptic Input," *Advanced Robotics*, vol. 25, no. 5, pp. 581-603, 2011.
- [4] A. X. Lee, H. Lu, A. Gupta, S. Levine and P. Abbeel: "Learning Force-Based Manipulation of Deformable Objects from Multiple Demonstrations," *2015 IEEE International Conference on Robotics and Automation*, pp. 177-184, 2015.
- [5] A. Sasagawa, K. Fujimoto, S. Sakaino and T. Tsuji: "Imitation Learning Based on Bilateral Control for Human-Robot Cooperation," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 6169-6176, 2020.
- [6] A. Sasagawa, S. Sakaino and T. Tsuji: "Motion Generation Using Bilateral Control-Based Imitation Learning With Autoregressive Learning," *IEEE Access*, vol. 9, pp. 20508-20520, 2021.
- [7] S. Sakaino, T. Sato and K. Onishi: "Multi-DOF Micro-Macro Bilateral Controller Using Oblique Coordinate Control," *IEEE Transactions on Industrial Informatics*, vol. 7, no. 3, pp. 446-454, 2011.
- [8] K. Ohnishi, M. Shibata and T. Murakami: "Motion control for advanced mechatronics," *IEEE/ASME Transactions on Mechatronics*, vol. 1, no. 1, pp. 56-67, 1996.
- [9] T. Murakami, F. Yu and K. Ohnishi: "Torque sensorless control in multidegree-of-freedom manipulator," *IEEE Transactions on Industrial Informatics*, vol. 40, no. 2, pp. 259-265, 1993.
- [10] T. Yamazaki, S. Sakaino and T. Tsuji: "Estimation and Kinetic Modeling of Human Arm using Wearable Robot Arm," *IEEE Transactions on Industry Applications*, vol. 199, no. 3, pp. 57-67, 2017.
- [11] S. Funabashi, S. Ogasa, T. Isobe, T. Ogata, A. Schmitz, T. P. Tomo and S. Sugano: "Variable In-Hand Manipulations for Tactile-Driven Robot Hand via CNN-LSTM," *proc. of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 9472-9479, 2020.