

バイラテラル制御と機械学習に基づく人間規範型ロボット制御技術の研究

神戸大学 大学院海事科学研究科 海事科学専攻

元井 直樹

1. はじめに

近年、日本では少子高齢化による労働力不足が大きな問題となっている。その問題の解決に向けて、ロボットが人間の行うさまざまな作業を代替することが期待されている。しかし、現状の産業用ロボットではものづくりにおける単純作業には対応できているが、環境変化や複雑な動作の自動化は未だ実現できていない。

人間のように柔軟な作業を実現するには、位置制御に加え力制御を組み合わせた運動制御を行う必要がある。そこで力覚伝送を有する遠隔制御であるバイラテラル制御技術が鍵技術となりうる。バイラテラル制御とは、人間が操作するリーダシステムと、環境に接触するフォロワシステム間で、位置と力情報を双方向に伝達する制御方法である。また、動作環境の変化に対応するために、模倣学習を用いたロボットの自律動作生成手法に関する研究が活発に行われている。模倣学習は機械学習の手法の一つであり、人間により教示されたデータを手本として学習するものである。模倣学習を用いることにより、学習済みの環境に加えて未学習の環境での自律動作も可能となる。

本研究では、バイラテラル制御と機械学習に基づく人間規範型ロボット技術の構築を目的とする。人間動作の一例として塗布作業に着目し、バイラテラル制御による人間の塗布作業動作の抽出、機械学習に基づく動作再現を行い、人間動作に基づくロボット運動制御理論の構築を目指す。

2. モデリング

本章では、本研究で用いる多自由度ロボットであるアールティ社製の CRANE-X7 について説明する。本研究ではバイラテラル制御を実現するために、CRANE-X7 を 2 台用いた。図 1 に 2 台のロボットを並べて配置した様子を示す。本構成において、一方をリーダシステム、もう一方をフォロワシステムとして動作させる。CRANE-X7 は物体を把持することのできる 2 指エンドエフェクタが装備されている。マニピュレータの自由度は 7 自由度、ハンドの自由度は 1 自由度であるため、合計して 8 自由度のロボットである。

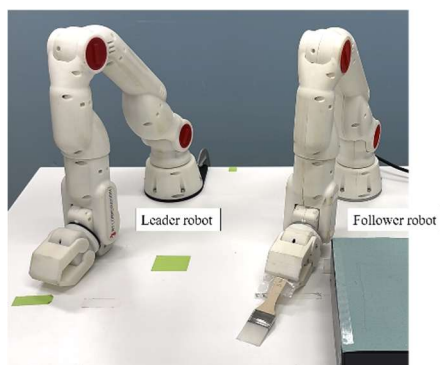


図 1. バイラテラル制御システム

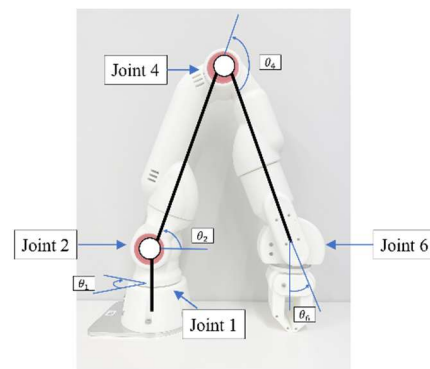


図 2. 座標系

図 2 に本実験で使用したマニピュレータの関節と回転方向を示す。各関節を土台から順に関節 1, 関節 2, ..., 関節 8 と定め, 各関節に対応する角度を $\theta_1, \theta_2, \dots, \theta_8$ と定義する。本実験ではベースリンクから 3 番目, 5 番目, 7 番目の関節およびハンドを固定し, 関節 1, 2, 4, 6 のみ可動する 4 自由度マニピュレータとして使用した。

3. バイラテラル制御

バイラテラル制御とは, 人間が操作するリーダシステムと, 環境に接触するフォロワシステム間で, 位置と力情報を双方向に伝達する制御方法である。図 3 に 4ch バイラテラル制御のブロック線図を示す。バイラテラル制御を実現するためには, 式(1)および式(2)で示す位置の同期と作用・反作用の法則を同時に満たす必要がある。ただし, ここではリーダ・フォロワシステム間での通信遅延はないものとする。

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

$$\hat{\tau}_l^{ext} + \hat{\tau}_f^{ext} = 0 \quad (2)$$

ここで, $\theta = [\theta_1, \theta_2, \theta_4, \theta_6]^T$, $\tau = [\tau_1, \tau_2, \tau_4, \tau_6]^T$ は角度ベクトルおよびトルクベクトルである。また, 下付き添え字 o_l および o_f はリーダシステム, フォロワシステムをそれぞれ示す。上付き添え字 o^{res} および o^{ext} は応答値および外力トルクを表す。 \hat{o} は推定値を示す。ロバストな運動制御を実現するために外乱オブザーバ (DOB) を実装し, 加速度制御系を実現している。また, 反力推定オブザーバ (RFOB) を実装することで, 外力トルクをセンサレスで推定する。

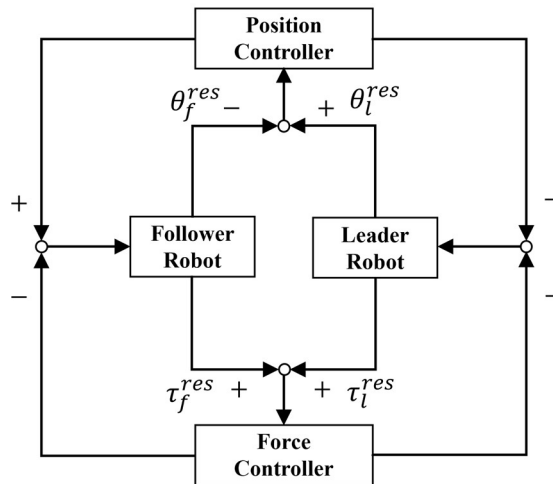


図 3. バイラテラル制御のブロック線図

4. バイラテラル制御に基づく模倣学習

本章では, バイラテラル制御に基づく模倣学習について述べる。

〈4・1〉 模倣学習

模倣学習は, 人間の動作を教師データとして用い, 現在のステップを入力, 次のステップの応答値を出力として教師あり学習を行う手法である。模倣学習はロボット制御や自律移動などの分野で活用されており, 人間の直感的な操作や知識をシステムに取り込むことが可能である。また, 報酬設計が難しい環境でも効率的に学習を行うことができるという特徴がある。

〈4・2〉バイラテラル制御に基づく模倣学習

バイラテラル制御から得られたリーダー・フォロシステム的位置・力情報を用いて模倣学習を行う。ロボットの関節角度やトルクなどの時系列データの応答値を教師データとして収集し、現在のステップから次のステップの応答値を予測するモデルを作成する。そして、自律動作の際には、モデルで予測した応答値を所望の指令値としてロボットに入力する。

本研究では Long Short Term Memory (LSTM) を用いて模倣学習を行った。LSTM は再帰的構造である Recurrent Neural Network (RNN) を拡張し、長期的な依存関係が学習可能である。ロボットによる塗布動作は一連の時系列動作であるため、長期的な依存関係の学習に優れた LSTM が有効であると判断した。

5. 模倣学習に基づく塗布作業の実現

本章では模倣学習に基づく塗布作業の方法について述べる。図 4 にデータ収集から自律動作までの流れを示す。本模倣学習は図 4(a)に示すデータ収集段階、図 4(b)に示すモデル訓練段階、図 4(c)に示す運用段階に分かれている。下付き添え字 o_{f_t} は時刻 t におけるフォロシステムの値を $o_{l_{t+10}}$ は時刻 t から 10 サンプル後のリーダーシステムの値を示す。

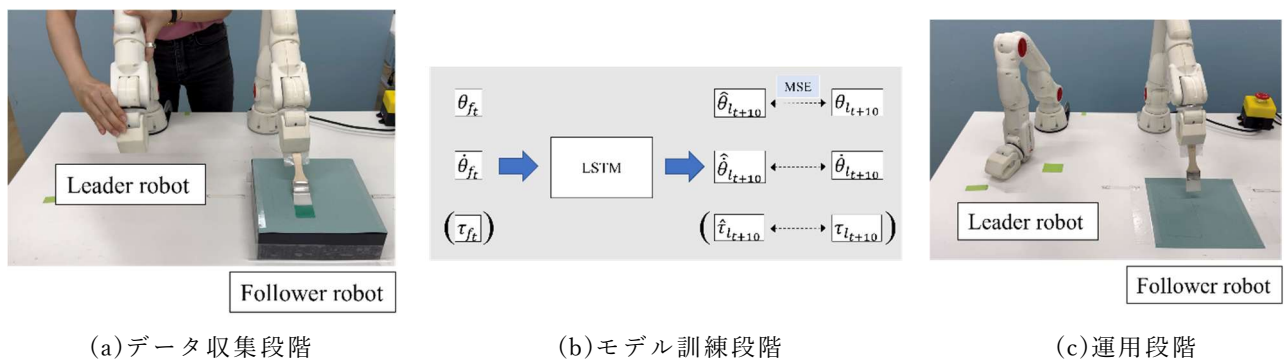


図 4. 模倣学習に基づく塗布作業

〈5・1〉データ収集段階

マニピュレータの手先に刷毛を固定し、テーブルに置かれた紙への塗布動作をバイラテラル制御にて行う。本実験では、台上に設置した紙に対して刷毛で塗布動作を行い、その動作を記録した。教師データとして、バイラテラル制御時のリーダーシステムおよびフォロシステムの関節角度、角速度、トルク応答値を収集した。

〈5・2〉モデル訓練段階

本研究で用いた LSTM モデルは、ロボットアームの関節角度、角速度、トルクを入力とし、次ステップにおける目標状態を出力する構成とした。モデルは LSTM 層と全結合層から構成され、時系列依存関係を捉えることで、過去の動作データから将来の応答を予測できるように設計した。図 4(b) に示すように、時刻 t におけるフォロシステムの状態を入力とし、時刻 $t+$

Δt におけるリーダーシステムの目標状態を出力するように学習を行った。損失関数には、モデルが出力した値と教師データの真値の平均二乗誤差を使用し、Adam 最適化により損失関数を最小化するように学習した。

〈5・3〉 運用段階

図 5 にバイラテラル制御に基づく模倣学習を用いた運動制御のブロック線図を示す。運用段階では、訓練した LSTM モデルを用いて、フォロワシステムの自律動作を行った。本研究ではリーダーシステムを LSTM モデルで置き換え、人間が与えていた操作入力を LSTM モデルが再現し、自律動作に必要な指令を生成する構成とした。

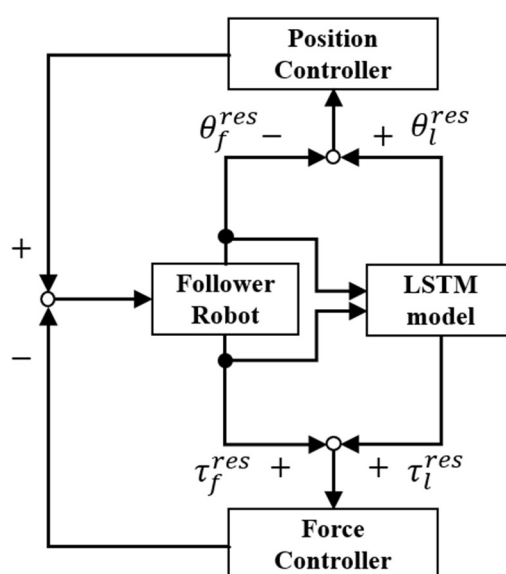
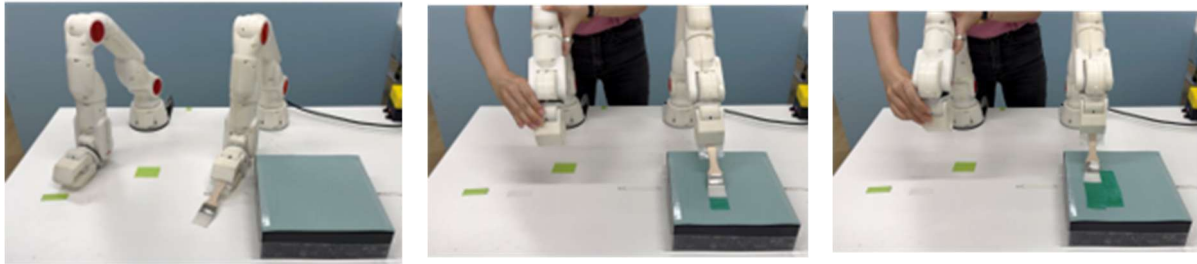


図 5. 模倣学習を用いた運動制御のブロック線図

6. 実験

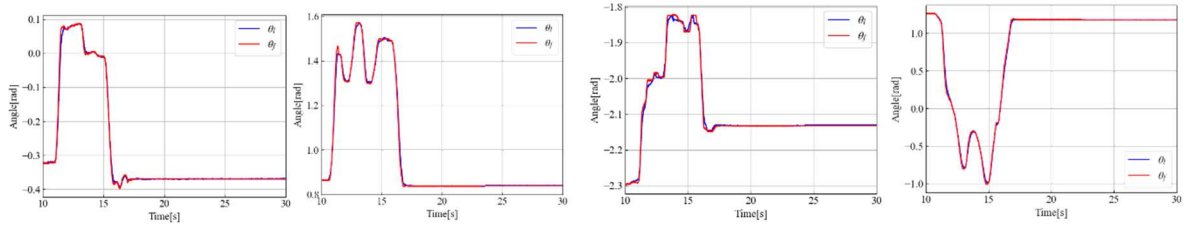
本章では人間動作を規範とした塗布作業における実験結果を示す。図 6 にデータ収集段階におけるバイラテラル制御の実験結果をスナップショットとして示す。データ収集段階では、図 6(a)に示すように、高さ 6cm の台の上に 21cm×21cm の紙を設置し、塗布動作を行った。紙の右半分に約 7cm×11cm の長方形領域を設定し、図 6(b), (c)に示すように、その範囲を刷毛で奥から手前へ片道 1 回の塗布動作を、計 2 回繰り返した。バイラテラル制御による塗布動作をそれぞれ 30 回行い、22 回分の動作データを教師データ、残りの 8 回分の動作データを検証用データとして取得した。制御周期とデータ保存周期は、ともに 1ms とした。学習モデルの推論時間と指令値の送信時間を考慮し、取得したデータを 10ms ごとのデータにリサンプリングした。

データ収集段階で取得したリーダーおよびフォロワシステムの位置応答値を図 7, 力応答値を図 8 に示す。図 7-8(a)は関節 1 の応答, (b)は関節 2 の応答, (c)は関節 4 の応答, (d)は関節 6 の応答を示す。本研究ではこの実験結果を用いて、モデルを訓練した。



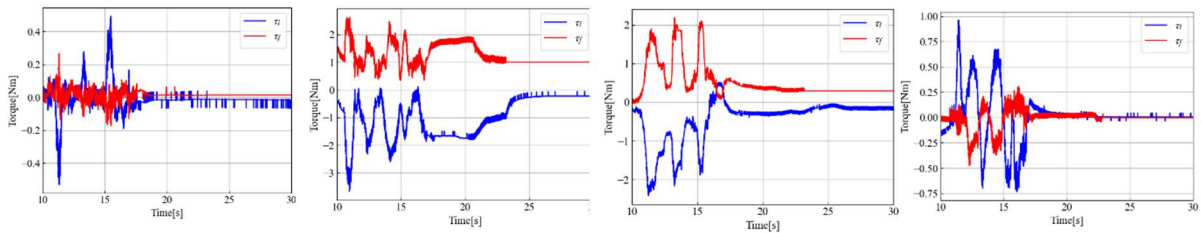
(a) 初期状態 (b) 塗布動作 (1回目) (c) 塗布動作 (2回目)

図 6. スナップショット (データ収集段階)



(a) Joint 1 (b) Joint 2 (c) Joint 4 (d) Joint 6

図 7. バイラテラル制御の位置応答値



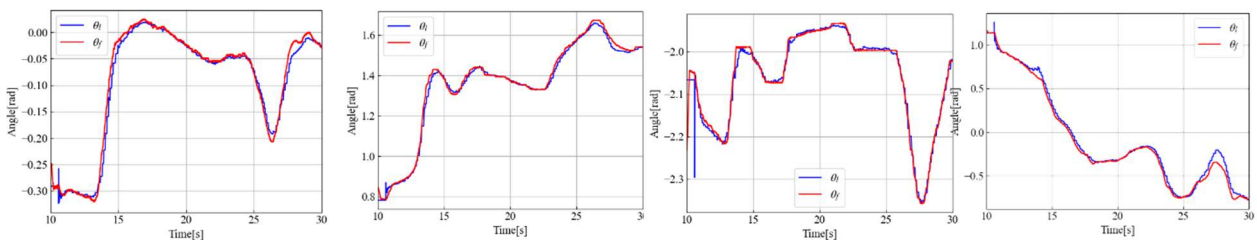
(a) Joint 1 (b) Joint 2 (c) Joint 4 (d) Joint 6

図 8. バイラテラル制御の力応答値

運用段階では、訓練した LSTM モデルを用いて以下の 2 条件で実験を実施した。

- ・ Case1 : データ収集段階と同じく高さ 6cm に紙を設置した場合
- ・ Case2 : 作業台を外して高さ 0cm に紙を設置した場合

運用段階での実験結果を示す。図 9-10(a)は関節 1 の応答、(b)は関節 2 の応答、(c)は関節 4 の応答、(d)は関節 6 の応答を示す。図 9 に示す Case1 では、塗布動作を行っているが、トルク推定値の推定精度が不十分であったことから、15 秒から 17 秒においてロボットと台との接触が発生し、指令値に対する応答値の追従性が低下した。しかし、人間動作と同様な塗布動作の一部は実現を確認した。Case2 では、作業台を外すことで学習時と紙の高さを変更した。



(a) Joint 1 (b) Joint 2 (c) Joint 4 (d) Joint 6

図 9. 運用段階 (Case1) における位置応答

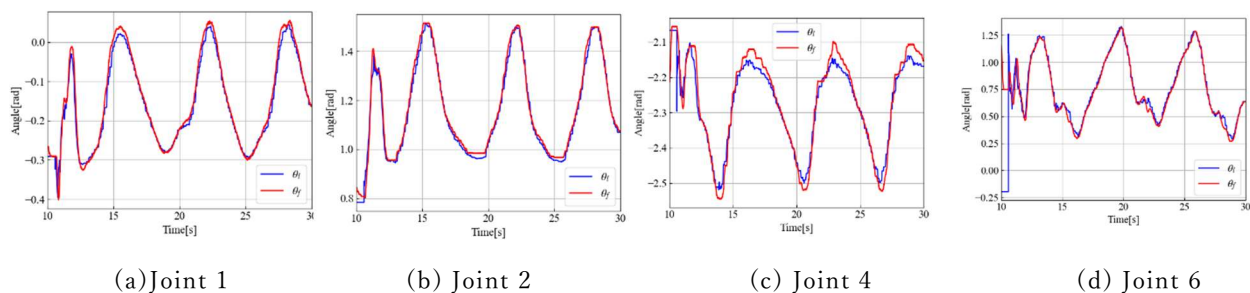


図 10. 運用段階 (Case2) における位置応答

図 10(b)に示すように 19~20 秒および 25~26 秒付で関節 2 の角度が低下しており、刷毛が机に接触していたことが確認できた。そのため塗布動作の一部は確認できたが、十分な達成制度の実現には至らなかった。

7. まとめ

本研究では、バイラテラル制御に基づく模倣学習による人間動作を規範としたロボットの運動制御技術の構築を目指した。人間動作の一例として塗布作業に着目し、バイラテラル制御による人間の塗布作業動作の抽出、抽出データに基づく機械学習の訓練、および機械学習に基づく動作再現を行った。実験結果より、人間と同様の塗布動作の一部の再現は確認できたが、動作環境の差異に対して、人間と同程度の環境適応までは認められなかった。本原因としては、機械学習における推定精度が十分でなく、位置情報の再現精度が低下したと考えられる。そのため、今後の課題としては、機械学習におけるモデル構造の改良による推定精度の向上があげられる。また、摩擦補償を新たに導入することで、精度の高い力覚推定を実現することも必要である。

謝辞

本研究を遂行するにあたり、公益財団法人天野工業技術研究所から多大なご支援を頂きました。ここに記して謝意を示します。

参考文献

- 1) K. Ohnishi, S. Katsura, and T. Shimono, "Motion Control for Real-World Haptics," IEEE Transactions on Industrial Electronics Magazine, vol.4, no. 2, pp. 16–19, 2010.
- 2) P.C. Yang, K. Sasaki, K. Suzuki, K. Kase, S. Sugano and T. Ogata, "Repeatable Folding Task by Humanoid Robot Worker Using Deep Learning", IEEE Robotics and Automation Letters, Vol. 2, No. 2, pp. 397-403, 2017.
- 3) 大西公平, "外乱オブザーバによるロボスタ・モーションコントロール", 日本ロボット学会誌, vol. 11, no. 4, pp. 486-493, 1993.
- 4) T. Murakami, F. Yu, and K Ohnishi, "Torque sensorless control in multidegree-of-freedom manipulator", IEEE Transactions on Industrial Electronics, vol. 40, no. 2, pp. 259-265, 1993.
- 5) 杉山 凜, 元井 直樹, "バイラテラル制御に基づく模倣学習による塗布動作の検討", 電気学会産業計測制御研究会, IIC-25-0021, 2025.