

画像認識と位置・力情報を用いた模倣学習

Imitation Learning Using Image Recognition and Position / Force Information

○学 笹川 歩 (埼玉大) 学 藤本 和樹 (埼玉大)
正 境野 翔 (埼玉大/JST さきがけ) 正 辻 俊明 (埼玉大)

Ayumu SASAGAWA, Saitama University, a.sasagawa.997@ms.saitama-u.ac.jp
Kazuki FUJIMOTO, Saitama University
Sho SAKAINO, Saitama University/JST PRESTO
Toshiaki TSUJI, Saitama University

This paper proposes imitation learning using position, force, and visual information. Imitation learning enables robots to substitute for human works. In conventional research, the performance of robots is improved using position and force information in imitation learning. However, imitation learning using all of position, force, and visual information are not verified. Therefore, this research demonstrates imitation learning with visual information in addition to position and force information. By using visual information, tasks that robots can execute are expanded. In this paper, it is verified that the real time behavior decision by deep learning using position, force, and visual information is effective. As a result, autonomous operation of an eraser at unknown height and position was successful by the proposed method.

Key Words: Robot, Manipulation, Machine learning, Bilateral control, Robot vision

1 序論

近年、先進国では少子高齢化に伴う労働力不足が懸念されている。そこで、人間の肉体労働を代替する方法としてロボットの導入が期待されている。

ロボットが人間の物体操作を模倣する手法として、機械学習を用いた End to End Learning が注目を集めている。この先行研究として、Yang らは遠隔操作を用いてロボットに人間の物体操作を模倣させることにより、少ない試行回数で布の折り畳み動作の模倣に成功した [1]。しかし、位置情報と視覚センサ情報からのみロボットの動作を決定しており、力情報を考慮していないため成功率が 8 割に満たなかった。

人間は視覚情報や位置情報のみだけではなく、物体の剛性や質量などの力情報を考慮した上で物体操作を行う。遠隔操作による物体操作の研究ではあるが、操作者に力情報のフィードバックをすることにより操作者の操作性や成功率が向上することが示された [2]。すなわち、ロボットにおける物体操作の成功率向上にも力情報を考慮することが重要であると考えられる。

しかし、これまでのロボットへの教示は、ダイレクトティーチングや遠隔操作を用いた位置情報のみの教示であり、ロボットへの力情報の教示は困難であった。その理由として、力情報の取得が困難であることが挙げられる。これは、制御入力である操作者の作用力と応答である物体からの反作用力を分離できないためである。また、触覚デバイスを用いて人間の物体操作における力加減を抽出する研究が行われたが、人間の力加減を模倣する手法の確立には至っていない [3][4]。そこで、藤本らは作用反作用力を分離可能な 4ch 型バイラテラル制御による遠隔操作を模倣学習に応用することを提案した。これにより、位置・力情報の取得が可能となり、未知環境での物体操作に成功した [5][6]。しかしながら、藤本らは視覚情報を用いていないため、その適応範囲は限られている。ロボットが動作する際、人間と同様に視覚情報を用いることで、より多様な作業をこなすことができるとともに、成功率の向上が期待できる。Ochi らはバイラテラル制御によりタスクの動作データを取得し、画像認識を含んだ物体操作を試みた [7][8]。しかし、ロボットの動作決定には位置情報のみを用いており、成功率の向上には至っていない。

そこで、本研究では従来の 4ch 型バイラテラル制御を用いた位置・力情報の模倣学習 [5][6] に視覚情報を加え、画像認識を含んだ模倣学習の手法を提案する。視覚情報を含むことで、従来行うことが困難であった力加減と視覚情報を必要とするタスクを遂行す

る。本研究では、4ch 型バイラテラル制御でタスクの動作データを取得するとともに、ロボット上部に設置したカメラから RGB 画像を取得する。RGB 画像とタスクの動作データを複数のニューラルネットワーク (Neural Network: NN) を用いて学習することにより、ロボットの動作を決定する。本論文では、提案手法の有効性を示す実験として、ホワイトボードに書かれた線を消す動作の模倣学習を行う。実験では任意の線の位置に対して字消し動作を行い、視覚情報を用いることにより、線の位置に応じて動作するかを確かめることで、提案手法の汎用性を検証する。

本論文の構成を以下に示す。2 章では、バイラテラル制御について説明し、3 章で本提案における実験環境および実験方法を述べる。4 章では実験結果を提示し、5 章で結果に対する考察、6 章で本論文の結論と今後の展望を示す。

2 バイラテラル制御

バイラテラル制御は 2 台のロボットを用いた遠隔操作技術である。スレーブが制御システムを介してマスタを操作し、スレーブが環境からの反力を受けるとマスタに反力がフィードバックされる。

また、本研究で扱う 3 自由度ロボットは Fig. 1 に示すようにモデル化することができる。 τ^{ref} は参照トルク、 τ^{dis} は環境接触による反力や重力による外乱トルク、 θ^{res} は角度応答、 J は慣性モーメントを示す。

2.1 4ch 型バイラテラル制御

バイラテラル制御にはいくつかの方式があるが、本研究では 4ch 型バイラテラル制御を用いる。マスタとスレーブ間で位置を同期し、力情報を双方向にフィードバックするため、物体操作の教示に適した方式である。

4ch 型バイラテラル制御は式 (1), (2) に示すように、マスタの位置 θ_m^{res} とスレーブの位置 θ_s^{res} を同期し、マスタの力 τ_m^{res} とスレーブの力 τ_s^{res} の間に作用反作用力の法則が成立することを制御目標とする [9]。ここで、下付き文字 s, m はそれぞれスレーブとマスタを表す。

$$\theta_s^{res} - \theta_m^{res} = 0 \quad (1)$$

$$\tau_s^{res} + \tau_m^{res} = 0 \quad (2)$$

上式を満たす 4ch 型バイラテラル制御のブロック線図を Fig. 2

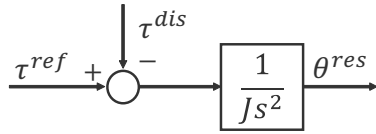


Fig.1 Model of robot

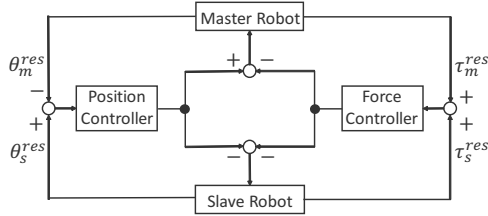


Fig.2 4ch bilateral controller

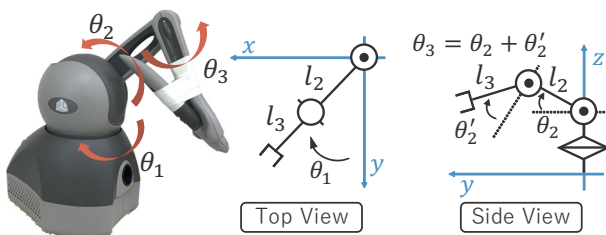


Fig.3 Geomagic Touch

に示す。マスタとスレーブへの入力はそれぞれ次式で表される。

$$\tau_m^{ref} = \frac{1}{2} J C_p (\theta_s^{res} - \theta_m^{res}) - \frac{1}{2} C_f (\tau_s^{res} + \tau_m^{res}) \quad (3)$$

$$\tau_s^{ref} = \frac{1}{2} J C_p (\theta_m^{res} - \theta_s^{res}) - \frac{1}{2} C_f (\tau_m^{res} + \tau_s^{res}) \quad (4)$$

ここで、比例フィードバックゲイン K_p 、微分フィードバックゲイン K_d を用いると、位置制御ゲイン C_p は $C_p = K_p + K_d s$ と表される。また、力フィードバックゲイン K_f を用いると、力制御ゲイン C_f は $C_f = K_f$ となる。

また、予め摩擦等の外乱トルクを Yamazaki らの手法 [10] に基づいて同定し、反力推定オブザーバ (Reaction force observer: RFOB) [11] を用いてトルク応答の算出を行った。

3 実験

3.1 実験環境

本研究では、3D Systems 社製の 3 自由度ロボット “Geomagic Touch” を 2 台使用した。ロボットの外観と関節角度の定義を Fig. 3 に示す。本論文では、提案手法の有効性を示すタスクとして、ホワイトボードに書かれた任意の線の位置および高さに対する字消し動作の模倣学習を行った。

本実験では Fig. 4 の制御系を構成し、制御パラメータは Table 1 に示す同定したパラメータを用いた。Fig. 4 中の DOB は外乱オブザーバ (Disturbance Observer: DOB) [12] を表し、アクチュエータに働く外乱トルクを推定し、推定外乱トルクをフィードバックすることでロバストな加速度制御系を実現する。

3.2 学習用データセット

実験の様子を Fig. 5 に示す。4ch 型バイラテラル制御をロボットに実装し、マスタからスレーブを遠隔操作することによりホワイトボードに書かれた線を消す動作を行った。4ch 型バイラテラル制御の制御周期は 1 msec とし、スレーブとマスタの 3 つの関節におけるそれぞれの角度応答 $\theta_{s1,2,3}^{res}$, $\theta_{m1,2,3}^{res}$ 、角速度応答 $\dot{\theta}_{s1,2,3}^{res}$, $\dot{\theta}_{m1,2,3}^{res}$ 、トルク応答 $\tau_{s1,2,3}^{res}$, $\tau_{m1,2,3}^{res}$ の 18 状態量を 20 msec 毎に保存した。ここで、各文字の下付き文字 1, 2, 3 はロボットの各関節を表す。以下、3

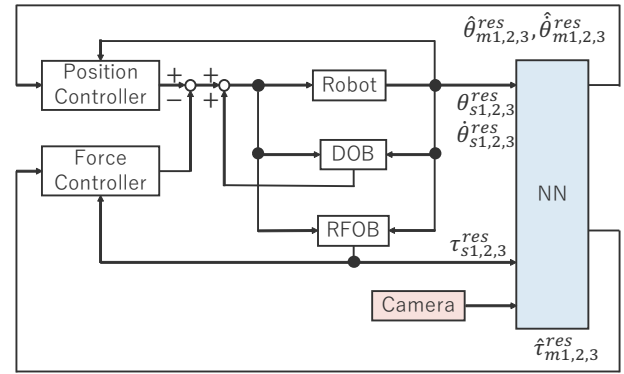


Fig.4 Block diagram

Table 1 Parameter

		Master	Slave
J_{θ_1}	Moment of inertia (θ_1) [mkgm ²]	4.20	4.45
J_{θ_2}	Moment of inertia (θ_2) [mkgm ²]	5.58	4.26
J_{θ_3}	Moment of inertia (θ_3) [mkgm ²]	1.51	1.63
M_2	Gravity compensation coefficient of l_2 [mNm]	135	139
M_3	Gravity compensation coefficient of l_3 [mNm]	98	96
D	Viscous friction coefficient [mkgm ² /s]	12.1	12.7
q	Cut-off frequency [1/s]	40	
K_p	Position feedback gain	121.0	
K_d	Velocity feedback gain	22.0	
K_f	Force feedback gain	1.0	

関節をまとめて表記する場合はこのように簡略化する。さらに、スレーブ上部に設置したカメラから RGB 画像 (120 × 160 × 3chs) を 40 msec 毎に取得した。1 回の字消し動作は 10 秒間行なったため、スレーブとマスタの角度応答、角速度応答、トルク応答はそれぞれ 500 サンプルずつ取得し、RGB 画像は 250 枚取得した。

本実験では、Fig. 6 に示すように 1 つの高さに対して 9 か所の線の位置を訓練位置とし、1 か所につき 4 回ずつ訓練データを取得した。また、ホワイトボード面の高さは Table 2 に示す 3 種類を訓練用の高さとしたため、合計 108 個の訓練データを取得した。

3.3 NN の構成と学習

3.3.1 NN の構成

本研究に用いた NN 全体の構造を Fig. 7 に示し、図中 (a)–(d) の NN の構造を Tables 3–6 に表記する。本研究では、Convolutional Neural Network (CNN)、Long Short-Term Memory (LSTM)、Feed-forward Neural Network (FFNN) の 3 種類の NN を用いることにより学習を行う。

カメラから取得した RGB 画像は CNN に入力され、スレーブの 3 関節における角度応答 $\theta_{s1,2,3}^{res}$ 、角速度応答 $\dot{\theta}_{s1,2,3}^{res}$ 、トルク応答 $\tau_{s1,2,3}^{res}$ の合計 9 状態量を LSTM_{joint} に入力する。出力として FFNN からマスタの 3 関節における予測角度応答 $\hat{\theta}_{m1,2,3}^{res}$ 、予測角速度応答 $\hat{\dot{\theta}}_{m1,2,3}^{res}$ 、予測トルク応答 $\hat{\tau}_{m1,2,3}^{res}$ の 9 状態量を得る。

3.3.2 NN の学習

スレーブの応答 9 状態量と RGB 画像をそれぞれ LSTM_{joint} と CNN のモデルに入力し、FFNN から 20 msec 後のマスタの応答を出力するように学習した。FFNN の出力誤差に応じて (a)–(d) すべてのモデルのパラメータを更新する学習方法を用いた。また、訓練データの RGB 画像は 40 msec 毎に更新し、明るさを変化させた画像を加えて 3 倍に水増しをした。本研究では、NN の入出力データとして角度、角速度、トルクを扱うが、それぞれ次元が異なるデータであるため、学習の前処理として正規化を行った。本研究では、Min-Max Normalization を採用し、全データを最大値 1、最小値 0 となるように正規化をした。なお、正規化に用いる最大値と最小値は取得した訓練データの値から決定した。また、RGB 画像も各画素値が最大値 1、最小値 0 となるように正規化をした。

3.4 タスクの実行

ロボットはリアルタイムでセンサ情報を取得し、学習した NN のモデルを使用することによりタスクを実行する。ロボットの動作実行中に RGB 画像は CNN へ入力され、画像特徴量に圧縮さ



Fig.5 Experimental environment

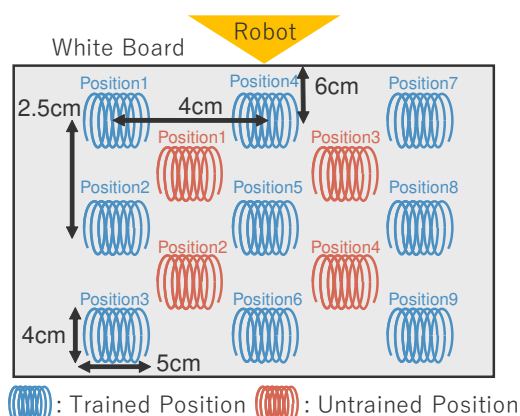


Fig.6 Character position of learning dataset

れる。ロボットのセンサから取得した関節情報は $LSTM_{joint}$ に入力される。FFNN はロボットの次の指令値である関節情報を出力するため、FFNN の出力をロボットへ入力することによりロボットが自律動作を行う。FFNN は正規化後の値を出力するため、実際にロボットへ指令値を入力する前に逆正規化の計算を行う。また、ロボットの制御周期は 1 msec であり、学習時と同様に画像は 40 msec、FFNN の出力は 20 msec 毎に得られる。

4 実験結果

学習済みの CNN, $LSTM_{image}$, $LSTM_{joint}$, FFNN を用いて実機による検証を行った。本実験では、ロボットが任意の初期位置から線の位置へ手先を移動させ、線を消した場合を成功と定義する。なお、成功の判定は、タスク遂行前と後で線の面積を比較し、縮小面積が元面積の 5 割以上だった場合を成功とした。検証は学習済みおよび未学習の高さと線の位置の組み合わせ 6 パターンに対して 1 か所につき 5 回ずつテストを行った。一連の動作を 20 秒間行い、その結果を Table 7 に示し、線の各位置ごとの成功率を Table 8 に示す。ここで、表中の Trained Position および Untrained Position は Fig. 6 に、Trained Height および Untrained Height は Table 2 に示した通りである。

未学習の高さかつ未学習の位置において、20 秒間の字消し動作を行った後のホワイトボードの様子を Fig. 8 に示す。線の位置と高さに応じて適切な位置と力加減で字消しの動作を遂行した。

5 考察

実験結果より、カメラから得られる視覚情報とロボットの関節情報を NN が学習し、人間のスキルを抽出したため、未学習の高さと位置においてもタスクを成功した。

本実験における失敗の多くは、望ましい動作幅や力加減で字消し動作をしているものの、ある位置に留まって往復動作をする失

Table 2 Height of whiteboard

Data	Height [mm]
Trained Height 1	20
Untrained Height 1	35
Trained Height 2	50
Untrained Height 2	65
Trained Height 3	80

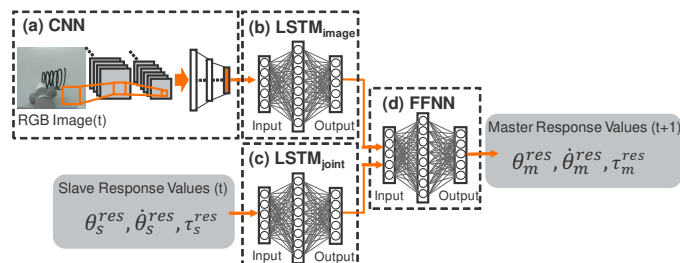


Fig.7 Construction of proposed method

敗であり、大きく場所を間違えて動作する失敗はなかった。ロボットに近い方を下側、遠い方を上側とすると、訓練データは線の下側から上側へ消すように動作を取得した。そのため、検証時に線の中央部から消し始めた場合、再び下側の線を消すことが困難であった。訓練データに上側から下側に向かって消すデータも含むことで改善されることが考えられる。

また、Table 7 より、Case 3 で成功率が低く、特に Table 8 に示すように Trained Position 7 において成功率が低かった。しかし、この位置においても訓練データを 108 個から 216 個に増やした場合、字消し動作に成功した。すなわち訓練データの不足により十分な学習ができなかったと考えられる。その原因として、複数の NN を使っているため NN の規模が大きくなり、学習に大量の訓練データ数が必要になったことが挙げられる。

6 まとめ

6.1 結論

本論文では、ロボットが視覚と位置・力情報から人間の物体操作を模倣する手法を提案した。位置・力情報を用いた模倣学習 [5][6] に視覚情報を加えることにより、視覚情報が必要なタスクへの拡張を行った。提案手法の有効性を示すために、力加減と視覚情報が重要なタスクとして、ホワイトボードに書かれた任意の線の位置に対して字消しの動作を学習し、実機による検証を行った。その結果、NN によってロボットの位置・力と視覚情報の関係性を学習したため、未学習の高さ・位置において字消しの自律動作に成功した。この結果より、NN の学習で汎化性能が向上し、未知環境に対応して物体操作をすることが可能になったと考えられる。

6.2 今後の展望

本論文では学習の訓練データを取得する際に、10 秒間の字消しの動作を合計 108 回試行した。訓練データ取得の労力および時間的制約の観点から、試行回数が多いことは望ましくない。そのため、今後は試行回数の減少を目的に、NN のモデルの構築とデータの水増し方法を検討する。さらに、本論文では 3 自由度ロボットを用いたが、今後は多自由度ロボットにおける模倣学習に拡張し、より幅広いタスクの実現を目指す。

謝辞

本研究は、JST さきがけ、JPMJPR1755 の支援を受けたものである。

参考文献

- [1] P. 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, (2017)

Table 3 Structure of (a) CNN

Layer	Input	Output	Activation Function
1st layer (Convolution)	(120, 160, 3)	(60, 80, 16)	Relu
2nd layer (Convolution)	(60, 80, 16)	(30, 40, 32)	Relu
3rd layer (Convolution)	(30, 40, 32)	(10, 10, 64)	Relu
4th layer (Convolution)	(10, 10, 64)	(2, 2, 128)	Relu
5th layer (fully connected)	512	250	Relu
6th layer (fully connected)	250	30	Relu

Table 4 Structure of (b) LSTM_{image}

Layer	Input	Output	Activation Function
1st layer (LSTM)	30	100	tanh
2nd layer (LSTM)	100	30	tanh

Table 5 Structure of (c) LSTM_{joint}

Layer	Input	Output	Activation Function
1st layer (LSTM)	9	50	tanh
2nd layer (LSTM)	50	9	tanh

Table 6 Structure of (d) FFNN

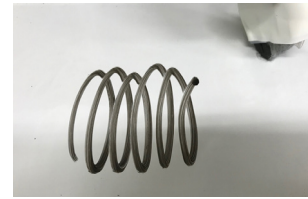
Layer	Input	Output	Activation Function
1st layer (fully connected)	9	200	ReLU
2nd layer (fully connected)	200	9	Identity Mapping

Table 7 Machine experiment result

	Position	Height	Success Rate
Case 1	Trained Position(9 Positions)	Trained Height 2	93.3% (42/45)
Case 2	Trained Position(9 Positions)	Untrained Height 1	84.4% (38/45)
Case 3	Trained Position(9 Positions)	Untrained Height 2	66.7% (30/45)
Case 4	Untrained Position(4 Positions)	Trained Height 2	100.0% (20/20)
Case 5	Untrained Position(4 Positions)	Untrained Height 1	95.0% (19/20)
Case 6	Untrained Position(4 Positions)	Untrained Height 2	95.0% (19/20)
	Total		86.2% (168/195)

Table 8 Success rate of each position

Position	Success Rate
Trained Position 1	86.7% (13/15)
Trained Position 2	93.3% (14/15)
Trained Position 3	86.7% (13/15)
Trained Position 4	73.3% (11/15)
Trained Position 5	86.7% (13/15)
Trained Position 6	86.7% (13/15)
Trained Position 7	60.0% (9/15)
Trained Position 8	86.7% (13/15)
Trained Position 9	73.3% (11/15)
Untrained Position 1	93.3% (14/15)
Untrained Position 2	93.3% (14/15)
Untrained Position 3	100.0% (15/15)
Untrained Position 4	100.0% (15/15)



[1] Before task execution

[2] After task execution

Fig.8 Result of task execution

- [2] C. Pacchierotti, L. Meli, F. Chinello, M. Malvezzi, and D. Praticchizzo: "Cutaneous haptic feedback to ensure the stability of robotic teleoperation systems," *The International Journal of Robotics Research*, (2015)
- [3] L. Roza, P. Jiménez, and C. Torras: "A robot learning from demonstration framework to perform force-based manipulation tasks," *Intelligent Service Robotics*, Vol.6, Issue.1, pp.33–51, (2013)
- [4] P. Kormushev, S. Calinon, and Darwin G. Caldwell: "Imitation Learning of Positional and Force Skills Demonstrated via Kinesthetic Teaching and Haptic Input," *Advanced Robotics*, Vol.25, (2011)
- [5] 藤本 和樹, 足立 強, 境野 翔, 辻 俊明: "人間の物体操作技能に基づくニューラルネットワークを用いた位置・力制御," ロボティクス・メカトロニクス講演会 2018, 2A2-E18, (2018)
- [6] T. Adachi, K. Fujimoto, S. Sakaino, and T. Tsuji: "Imitation Learning for Object Manipulation Based on Position/Force Information Using Bilateral Control," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, (2018)
- [7] H. Ochi, W. Wan, Y. Yang, N. Yamanobe, J. Pan, and K. Harada: "Deep Learning Scooping Motion using Bilateral Teleoperations," *arXiv*, 1810.10414v1, (2018)
- [8] 越智 仁絵, 万 偉偉, 山野辺 夏樹, 原田 研介: "多視点の視覚センサフィードバックを用いた深層学習による物体認識と掬い動作の生成," 第 36 回 日本ロボット学会 学術講演会, 2E2-03, (2018)
- [9] S. Sakaino, T. Sato, and K. Ohnishi: "Multi-DOF Micro-Macro Bilateral Controller Using Oblique Coordinate Control," *IEEE*

Transactions on Industrial Informatics, Vol. 7, No. 3, pp. 446–454, (2011)

- [10] T. Yamazaki, S. Sakaino, and T. Tsuji: "Estimation and Kinetic Modeling of Human Arm using Wearable Robot Arm," *IEEE Transactions on Industry Applications*, (2016)
- [11] 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)
- [12] 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)