

# サンプリングをずらすことによる機械学習の訓練データ水増し手法の検証

## An Investigation of Machine Learning Training Data Padding Method by Shifting Sampling

○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)

Kazuki FUJIMOTO, Saitama University, k.fujimoto.423@ms.saitama-u.ac.jp

Sho SAKAINO, Saitama University/JST PRESTO

Toshiaki TSUJI, Saitama University

A large amount of data is necessary to obtain desired results in machine learning in which a robot substitutes humans' manual labors. In addition, it is an explicit problem that it takes a lot of time and operating costs for it. In this research, we focus on down-sampling and decimation used in system identification and propose a method to reduce teacher data by padding data in machine learning.

**Key Words:** Robot, Manipulation, Machine learning, Imitation learning, Bilateral control, Data padding, System identification

### 1 序論

近年、工場だけでなく人間の生活する実環境においてもロボットによる自動化が期待されている。従来のロボットは設計された軌道を再現するようにプログラムされているため、作業対象物や周囲の環境の変化に対応できない。そのため、人間には単純労働に思われる作業が未だに人間の手作業によって行われる。ロボットがこれらの変化に対応するには、それぞれの状況に合わせてすべての行動を設計しなければならない。しかし、それには無限と通りの組み合わせがあるため多大なコストと時間を必要とする。そこで機械学習による「End-to-End Learning」が注目を浴び研究されている。End-to-End Learning は、得られた入力と出力から、欲しいものを直接学ぶことができる。Levine らは、この概念に基づく強化学習によって複数のオブジェクトを掴むことに成功した [1]。しかし、この学習は 80 万回の操作に 2 か月かかり、その実用性は低い。そこで、近年は模倣学習や LfD (Learning from Demonstration) と呼ばれる人間の操作技能を模倣することによって訓練データを大幅に削減する研究が報告されている [2]。その他にも位置情報を基に軌道生成を行う研究は行われている [3] [4]。しかしながら、ロボットのすべての動作が位置制御と力制御の組み合わせとして記述できることが証明されていることから考えて [5]、力加減の調節を要するタスクには適応できない。また、遠隔操作を用いた peg-in-hole 実験では、作業への反力のフィードバックで作業効率が向上することが報告されている [6]。そのため、動作計画には力情報を用いる必要がある。しかしながら、機械学習では力情報とその情報を使用した力制御の双方を利用したものはほとんどない [7] [8]。そこで、著者らは 4ch 型バイラテラル制御という遠隔操作技術を用いてロボットに位置と力の情報をもつデータを取得することで力加減を考慮したロボットの自律的動作生成を行った [9][10]。さらには、ロボットの動作計画における適した制御系の在り方を検証した [11]。しかしながら、モデルの学習には膨大な訓練データが必要であり、教師データの取得には多くの時間が必要である。いずれにしても教師データの取得には多くの時間と動作コストを有する。そこで、本研究では、システム同定においてノウハウ的に知られるデータの増しの手法を機械学習に応用して、取得した訓練データの数を水増しする手法を 2 つ提案し、比較実験によりその優劣を評価する。さらに、その有効性を検証すべく、従来の学習と更なる比較を行い、評価する。

本論文の構成を以下に示す。2 節では本研究で用いた制御システムについて説明し、3 節でシステム同定を説明し、4 節では本研究におけるネットワークの学習について述べる。5 節で実験を説明し、6 節で本論文の結論を示す。

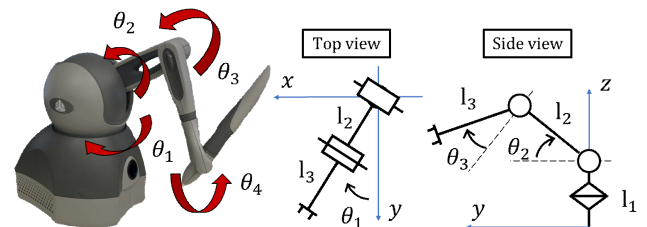


Fig.1 Definition of robot joints and coordinate system.

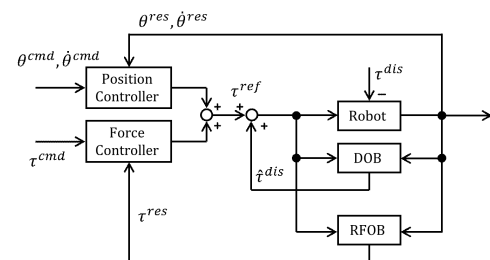
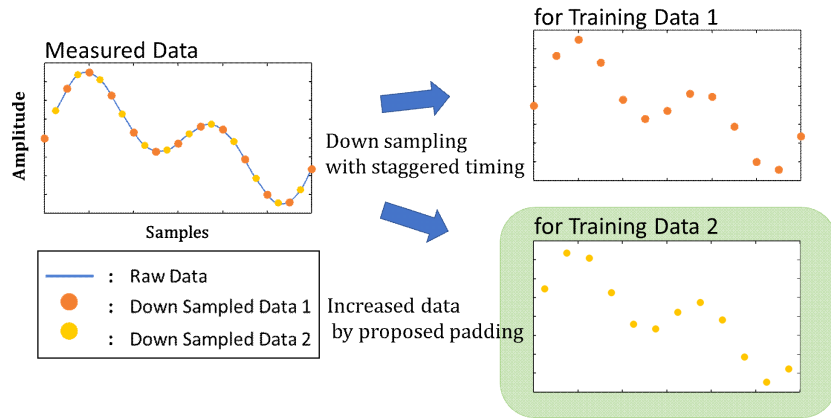


Fig.2 The manipulator's block diagram.

## 2 制御システム

### 2.1 マニピュレータ

本研究で使用するマニピュレータは 3D System 社製の Geomagic Touch であり、図 1 に示す。ここで、ロボットの関節角度を図 1 の右図のように定義する。4 番目の軸は受動関節であるため固定した。Geomagic Touch の関節角度は内蔵されている角度センサによって測定し、角速度はその疑似微分によって計算した。また、図 2 に示される  $\theta$ ,  $\dot{\theta}$ ,  $\ddot{\theta}$ ,  $\tau$  は、それぞれロボットの関節角度、角速度、各加速度、トルクを示し、添え字の  $cmd$  や  $res$ ,  $ref$  はそれぞれ指令値、応答値、参照値であることを表し、 $\tau^{dis}$  は環境接触による反力や重力などによる外乱トルクであることを表す。また、下付き文字の  $m$ ,  $s$  はそれぞれマスタロボット (以下「マスタ」という)、スレーブロボット (以下「スレーブ」という) であることを示す。(マスタ・スレーブについては 2.2 章で述べる。) また、 $\tau_{dis}^{\hat{}}$  は外乱オブザーバ (Disturbance Observer :



**Fig.3** Padding method by shifting sampling

**Table 1** Gains of robot controller

Parameter	value
$K_p$	Position feedback gain
$K_d$	Velocity feedback gain
$K_f$	Force feedback gain
$g$	Cut-off frequency of pseudo derivative [rad/sec]
$g_{DOB}$	Cut-off frequency of DOB [rad/sec]
$g_{RFOB}$	Cut-off frequency of RFOB [rad/sec]

DOB) [12] によって計算した。また、トルク応答値は反力推定オブザーバ (Reaction Force Observer : RFOB) [13] によって計算した。詳細は 2.3 章で述べる。図 2 に示すように制御器は位置制御器と力制御器によって構成されており、位置制御器は比例微分 (PD) 制御、力制御は比例 (P) 制御であり、制御周期は 1 msec である。

## 2.2 4ch 型バイラテラル制御

遠隔操作制御にはいくつかの方式が、本研究では力の伴う物体操作に適した 4ch 型バイラテラル制御を実装する [14]。バイラテラル制御はマスタ・スレーブと呼ばれる 2 台のロボットを用いた遠隔操作技術であり、一般に、マスタを操作者が操り、スレーブが追従して作業を行う。本研究で使用する 4ch 型バイラテラル制御は、マスタとスレーブの双方で角度  $\theta$  の同期と、トルク  $\tau$  の作用反作用の法則の実現を制御目標とし、以下の式によって示される。

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

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

制御目標を満たす位置と力制御器は以下の式によって与えられる。

$$\tau_s^{ref} = \frac{J}{2}(K_p + K_v s)(\theta_m^{res} - \theta_s^{res}) - \frac{1}{2}K_f(\tau_m^{res} + \tau_s^{res}) \quad (3)$$

$$\tau_m^{ref} = \frac{J}{2}(K_p + K_v s)(\theta_s^{res} - \theta_m^{res}) - \frac{1}{2}K_f(\tau_s^{res} + \tau_m^{res}). \quad (4)$$

ここで、 $J$  は慣性モーメントである。本研究では、PD 制御から成る位置制御器と比例制御から成る力制御器のゲインを表 1 の通りに設計した。

## 2.3 制御パラメータ同定

本研究では、DOB のトルク推定値から摩擦及び重力に起因する値を差し引くことでロボットのトルク応答値推定に利用した。本実験で用いたトルク応答値の算出法を式 (5)~(7) に示す。DOB

**Table 2** Identified system parameters

Parameter	Master	Slave
$J_1$	Joint 1's inertia [ m N·m]	4.20
$J_2$	Joint 2's inertia [ m N·m]	5.58
$J_3$	Joint 3's inertia [ m N·m]	1.51
$D$	Friction compensation coefficient [m kg·m <sup>2</sup> /s]	12.1
$g_{c1}$	Gravity compensation coefficient 1 [ m N·m]	135
$g_{c2}$	Gravity compensation coefficient 2 [ m N·m]	98
$g_{c3}$	Gravity compensation coefficient 3 [ m N·m]	123

で推定した外乱  $\tau^{dis}$  から、既知の外乱を引き、ロボットが実際に受けている反力のみを抽出できる。

$$\tau_1^{res} = \tau_1^{dis} - D\dot{\theta}_1 \quad (5)$$

$$\tau_2^{res} = \tau_2^{dis} - g_{c1}\cos\theta_2 - g_{c2}\sin\theta_3 \quad (6)$$

$$\tau_3^{res} = \tau_3^{dis} - g_{c3}\sin\theta_3 \quad (7)$$

同定した物理パラメータはロボットの各関節の慣性モーメント  $J$  と摩擦補償係数  $D$ 、重力補償係数  $g_c$  である。本実験で使用するロボットの物理パラメータは、従来手法を用いて同定を行った [15]。システム同定により算出した各係数は表 2 の通りである。

## 3 システム同定

システム同定は以下のシナリオに沿って進められる [16]。

1. 同定実験の設計
2. 同定実験
3. 入出力データの調整
4. システム同定法の適用

通常、システム同定はシナリオの 3 において、可能な限り長時間の入出力データを収集するため、パラメータは十分量のデータから同定される。そのため、機械学習とは異なりデータの水増しが求められないのである。しかしながら、一部の熟練者の間でノウハウとして用いられる手法がある。それがシナリオ 3 において行われるデシメーションの応用である。デシメーションとはローパスフィルタ (Low-Pass Filter : LPF) とダウンサンプリングから成り、これによりセンサノイズの除去と高周波域への過剰なフィッティングを防止の効果がある。ダウンサンプリングによってデータが間引かれてしまいデータを失うことになるのだが、ダウンサンプリングを行うタイミングを少しずらすことでまた別のデータセットを生成すればデシメーションによって間引かれるデータを有効活用することができる。本研究ではこの手法を機械学習に適用を試みる。

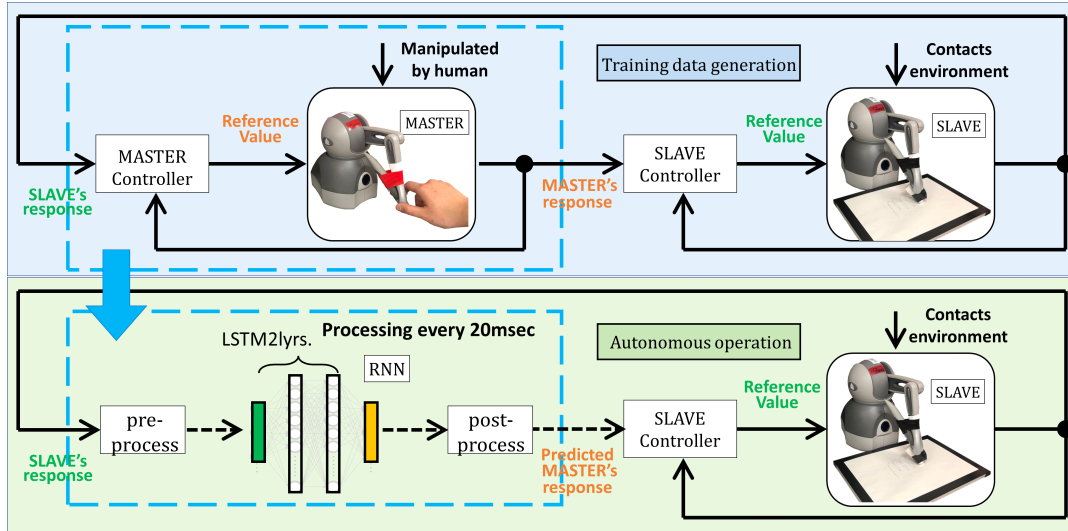


Fig.4 Training data generation and autonomous operation phase

Table 3 Neural Network's structure

Layer	Input	Output	Activation Function
1st layer (LSTM)	9	50	tanh
2nd layer (LSTM)	50	50	tanh
3rd layer (Linear)	50	18	identity mapping
4th layer	9	-	-

## 4 ネットワークの学習

### 4.1 ニューラルネットワークの構造

本研究ではリカレントニューラルネットワーク (Recurrent neural network: RNN) を用いた。RNN はネットワークの出力を別のネットワークの入力として利用するような再帰的構造を持ったニューラルネットワークであり、つまりは過去の情報を基に推測できるということである。本研究で使用するネットワークは LSTM(Long Short-Term Memory) と呼ばれる、長期依存関係を学習できる RNN の一種である。本研究で使用するネットワーク構造を表 3 に示す。モデルの入力にはスレーブ、そして出力にはマスターの各関節の  $\theta_m^{res}$ ,  $\dot{\theta}_m^{res}$ ,  $\tau_m^{res}$  が対応する教師あり学習を行う。ただし、オンライン推論において 20 msec かかることを考慮し、入力値からの 20 msec 先の出力値を出力するよう学習させた。

### 4.2 正規化

本研究では入力間のデータの単位次元が異なるため入力値のスケールに違いが生じる。そこで学習がスムーズに進むよう、各入力値の最大値と最小値を決定し、正規化する前のデータを  $d$ 、正規化したデータを  $d_n$ 、最大値を  $d_{max}$ 、最小値を  $d_{min}$  としたとき、式 (8) にて正規化、式 (9) にて逆正規化を行った。

$$d_n = \frac{d - d_{min}}{d_{max} - d_{min}} \quad (8)$$

$$d = d_n(d_{max} - d_{min}) + d_{min} \quad (9)$$

### 4.3 ミニバッチ

本研究では、学習の効率化を狙ってミニバッチを用いた。学習に利用したミニバッチを以下に示す。なお、本研究では訓練デー

タ数を 15、1 訓練データあたりのサンプル数を 750 個という前提の元説明する。

1. 訓練データ 15 個からランダムに 1 つを選ぶ
2. 1 で選んだ訓練データからランダムに 100 サンプル選ぶ
3. 1, 2 を繰り返し 100 セット用意する

上記の手順を行った後、モデルに入力し、予測マスター出力値と取得済みである訓練データのマスターの値の平均二乗誤差 (Mean Squared Error: MSE) が 0 に近づくよう、モデルの重み等のパラメータ更新を行った。この工程を学習回数 1 回と定義する。また、本モデルの作成に 1000 回の学習を行い、学習には約 6 分程度であった。本研究で使用した PC は CPU : Intel Core i7、RAM : 32 GB、GPU : NVIDIA GTX 1080 Ti である。

## 5 実験

### 5.1 訓練データの取得

本研究では、提案する水増し手法の有効性を検証すべく、ロボットの先端に固定した消しゴムを用いて字消し動作を行った。図 4 の上図のように操縦者がマスターを操作し、それに追従してスレーブが字を消す。訓練データは紙面高さを 10 mm、40 mm、70 mm の三段階で用意し、各高さで 5 回ずつ、合計 15 個取得した。なお、各データは 5sec の動作を 1 msec ごとに保存した。訓練データの取得には 5 分程度要した。

### 5.2 実機検証

本研究では、システム同定におけるデータの水増しに発想を受け、ニューラルネットワークにおける教師データの増し手法を提案する。提案する 2 つの水増し手法を従来手法と比較することでその有効性を検証した。表 4 に検証した 4 つのモデルの教師データの構成を示す。モデル 1 は従来手法 [9] であり、紙面の各高さにおいて 5 回ずつの計 15 個の訓練データを基に学習を行った。モデル 2 は紙面の各高さにおいて 1 回ずつの試行を行い、それぞれを等間隔にタイミングをずらしたダウンサンプリングを用いて 5 倍に複製した。モデル 3 は 2 と同様の元データに 3 で用いられる通りのタイミングをずらしたデシメーションを行い、5 倍に複製した。このとき使用した LPF のカットオフ周波数  $g_{LPF}$  はサンプリング定理に従い、以下の式によって求めた。

$$g_{LPF} = \pi / st \quad (10)$$

ただし、 $st$  はサンプリングタイムであり、本研究では 20 msec であることから  $g_{LPF}$  は 157 rad/sec とした。また、訓練データの単なる複製ではタスクを学習できないことを証明するために、モデル 4 は対照実験として紙面の各高さにおいて 1 回ずつの試行



**Table 4** Training data of each model

Model	Paper height	Trial number	Padding	Total
1	3	5	0	15
2	3	1	5	15
3	3	1	5	15
4	3	1	5(copy)	15

**Table 5** Executable range of each model

Model	Minimum height [mm]	Maximum height [mm]
1	10	85
2	10	85
3	10	70
4	45	60

を行って作成した計 3 個の訓練データを 5 倍に単純複製して用意した 15 個の訓練データとした。

表 5 に各モデルにおけるタスクの実行可能な範囲を示す。モデル 4 の結果から単なる教師データの複製では学習しきれなかった高さを提案したモデル 2, 3 は達成することができた。また、モデル 2 においては訓練データが 5 倍多い従来手法と同程度の適応範囲を発揮した。同定にて用いられるモデル 3 の水増し手法がモデル 2 に劣る原因として LPF による遅れが考えられる。図 5 に同一入力 ( $\theta_{s,1}^{res}$ ) に対するモデル 2 とモデル 3 の出力を示す。同図右よりわずかではあるがモデル 3 の出力が遅れていることがわかる。この差によりモデル 3 では消しゴムの振幅に遅れが出てしまい、実行可能範囲に差が生じた。

## 6 結論

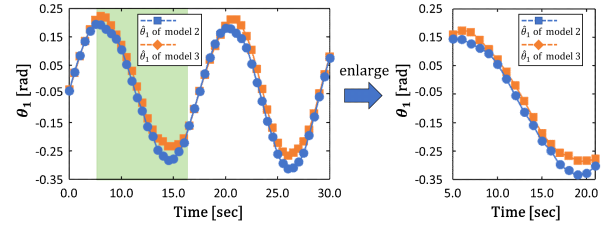
本論文では、学習における訓練データを水増しするための手法としてシステム同定においてノウハウ的に知られるタイミングをずらしたダウンサンプリングやデシメーションによる水増しを機械学習に応用した。従来、紙面の高さ 3 段階を各 5 回ずつ行った計 15 個の訓練データからモデルを作成していたが本提案手法を用いることで 5 分の 1 の訓練データ数で同性能のモデルの作成に成功した。また、ダウンサンプリングとデシメーションを用いた 2 つのモデルの適応範囲の比較からダウンサンプリングによる水増しが望ましいという結果が得られた。デシメーションを用いた水増しはオフラインで行われるシステム同定において有効であるが、オンラインで指令値を生成する機械学習には LPF による遅れが問題であると考えられる。本論文の結果から、可能な限り高い周波数でデータを計測することが教師データの増しには重要であると言える。今後は本手法の性能についてより詳しく検討するとともに他タスクに適応させたい。

## 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] S. Levine, P. Pastor, A. Krizhevsky, and D. Quillen, Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection, *arXiv: 1511.06732*, 2015.
- [2] P. Yang, K. Sasaki, K. Suzuki, K. Kase, S. Sugano, and T. Ogata, Repeatable folding task by humanoid robot worker using deep learning, in *proceedings of IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 397-403, 2017.
- [3] J. I. Lipton, A. J. Fay, and D. Rus, Baxter's homunculus : virtual reality spaces for teleoperation in manufacturing, in *proceedings of IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 179-186, 2017.
- [4] T. Zhang, Z. McCarthy, O. Jow, D. Lee, X. Chen, K. Y. Goldberg, and P. Abbeel, Deep imitation learning for complex manipulation tasks from virtual reality teleoperation, *arXiv: 1603.02199*, 2016.
- [5] S. Sakaino, T. Sato, and K. Ohnishi, A novel motion equation for general task description and analysis of mobile-hapto, in *proceedings of IEEE Transactions on Industrial Electronics*, vol. 60, no. 7, pp. 2673-2680, 2013.

**Fig.5** Comparison between model 1 and model 2 about  $\theta_1^s$ 's data

- [6] C. Pacchierotti, L. Meli, F. Chinello, M. Malvezzi, and D. Prattichizzo, Cutaneous haptic feedback to ensure the stability of robotic teleoperation systems, *International Journal of Robotics Research*, 2015.
- [7] P. Kormushev, S. Calinon, and D. G. Caldwell, Imitation learning of positional and force skills demonstrated via kinesthetic and haptic input, *Advanced Robotics*, vol. 25, no. 5, pp. 581-603, 2012.
- [8] L. Roza, P. Jiménez, and C. Torras, A robot learning from demonstration framework to perform force-based manipulation tasks, *Intelligent Service Robotics*, vol. 6, no. 1, pp. 33-51, 2013.
- [9] 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.
- [10] T. Adachi, K. Fujimoto, S. Sakaino, and T. Tsuji, Imitation learning for object manipulation based on position/force information using bilateral control, in *proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018.
- [11] 藤本和樹, 足立強, 境野翔, 辻俊明, “人間の物体操作の模倣学習に適したロボットの制御座標の評価”, 第 36 回日本ロボット学会学術講演会, 2E2-02, 2018.
- [12] K. Ohnishi, M. Shibata, and T. Murakami, Motion control for advanced mechatronics, in *proceedings of IEEE/ASME Transactions on Mechatronics*, vol. 1, no. 1, pp. 56-67, 1996.
- [13] T. Murakami, F. Yu, and K. Ohnishi, Torque sensorless control in multidegree-of-freedom manipulator, in *proceedings of IEEE Transactions on Industrial Electronics*, vol. 40, no. 2, pp. 259-265, 1993.
- [14] S. Sakaino, T. Sato, and K. Ohnishi, Multi-dof micro-macro bilateral controller using oblique coordinate control, in *proceedings of IEEE Transactions on Industrial Informatics*, vol. 7, no. 3, pp. 446-454, 2011.
- [15] T. Yamazaki, S. Sakaino, and T. Tsuji, Estimation and kinetic modeling of human arm using wearable robot arm, in *proceedings of Electrical Engineering in Japan*, vol. 199, no. 3, pp. 57-67, 2017.
- [16] 足立修一, “システム同定の基礎”, 東京電機大学出版局, pp. 222-226, 2009.