# しなやかなロボットをうまく制御する ~モデルベース学習

## 濵屋 政志

我々は、柔軟要素を持つしなやかなロボットを使用して、より容易に実現できる組立自動化を目指している。柔らかさによって対象物体との馴染みが生じ、部品の破損やロボットの緊急停止が起きにくくなり、その結果、本来必要であったセンシングや制御計画に対する労力を大幅に削減できると期待される。一方で、柔らかさによってダイナミクスが複雑となるため、学習技術の活用が必要不可欠となるが、生産現場の早い立ち上げのためなるべく少ない試行回数で学習したい。そこで我々は、モデルベース学習に基づく効率的なロボット学習制御手法を提案した。本稿では、我々の最新の研究成果である、組立作業の区分化と柔らかさの活用による学習効率化手法と、柔らかいロボットに応用が期待できる Sim-to-Real モデルベース転移学習手法について解説する。

# Learning Control for Soft Robots: Efficient Model-based Learning Approaches

HAMAYA Masashi

We develop an automatic assembly system using a soft robot. Thanks to the physical softness, the robot can interact with the environment safely, resulting in a significantly easier setup for the robotic system. Meanwhile, as the dynamics model becomes complex due to the softness, we need to use learning-based approaches. However, a large amount of taring data might not be preferable for real robot applications. To deal with this problem, we developed model-based learning approaches, which are efficient and suitable for real robot applications. This article introduces our recent studies using model-based learning approaches: 1) lower-dimensional dynamics models by leveraging subtask segmentation and physical softness and 2) model-based Sim-to-Real transfer learning for multiple environments.

#### 1. まえがき

ロボットによる組立自動化は、産業的価値が極めて高く、これまで数多くの研究開発が行われてきた<sup>1)</sup>。組立の自動化研究が今日まで広く行われてきた理由として作業の難しさが考えられる。一般的に、組立は部品同士の接触を伴い、部品同士の公差も小さい。もしセンシングの不確実性などから制御の誤差が少しでも生じた場合、部品同士に過剰な力が生じ破損などの致命的な失敗につながる可能性がある。

現在の産業用ロボットの多くは位置制御により溶接や塗装を行うために使用されている。このような位置制御に基づくロボットを組立に適用する際には、可能な限り誤差や不確実性を排除するため、入念なセンシング設定と動作計画が必要となる。一方で、力制御を用いた手法も数多くの

Contact: HAMAYA Masashi mailto:masashi.hamaya@sinicx.com

研究開発がされてきており、接触力を制御することで安全な組立作業を実現することができる。しかし、一般的に高精度な力・トルクセンサや、高制御周期のコントローラなど、高性能なハードウェアが必要となる上に、作業に応じて適切なパラメータ設定を行なければならない<sup>2)</sup>。総じて、ロボットの組立自動化は、莫大な労力が必要となるため、依然として工場制手工業に依存している<sup>3)</sup>。

この問題に対し、我々は物理的な柔らかさをロボットに持たせることに着目する。柔軟要素が対象部品になじむことで、誤差が生じても致命的な失敗を引き起こす可能性が軽減される。さらに、力・トルクセンサや、高制御周期のコントローラを導入することなく組立作業が実現でき、結果ロボットの立ち上げにかかる労力を大幅に削減できると期待される。そこで我々は、柔軟要素を持つ手首を開発した4)。この手首は従来の柔軟手首に比べてより軽量で大きな変位量を持つため組立に適しており、様々な市販の汎用

ロボットに搭載することができる。

しかし、物理的な柔らかさと接触を含むダイナミクスは 非線形性を持ちモデル化が極めて難しいことが懸念され る。複雑なダイナミクスに対処するために、機械学習のア プローチを適用することが考慮されるが、訓練データ取得 時における部品の摩耗や、学習環境の初期化などの労力も 無視できないため、なるべく少ない実環境とのインタラク ションで組立作業動作を学習できるようにしたい。

この問題に対し、我々はモデルベースによる学習手法を利用することで活用することで非線形性を有する複雑なダイナミクスを記述することを試みる。モデルベースの手法は、一般的にモデルフリーの手法と比較してサンプル効率がよいことが知られており、様々なロボット制御に応用されている<sup>5)</sup>。

我々は、このモデルベース学習に基づいて、さらに学習の効率化を試みており、その最新の研究成果について本稿で紹介する。第2節では、モデルベース学習における背景を説明する。第3節では、組立作業を区分化し、各小区分におけるモデルを単純化することで学習を効率化させる手法<sup>6)</sup>を紹介する。第4節では、柔らかいロボットや組立作業にも十分応用が期待できる、シミュレーションでダイナミクスモデルを事前学習し、実環境に転移させることで、実環境でのデータ収集の労力を削減する Sim-to-Real 転移学習手法<sup>7)</sup>について解説する。第5節では、まとめと今後の課題について紹介する。

# 2. モデルベース学習

本節では、モデルベース学習における問題設定を簡単に 説明する。一般的には、対象のロボットや操作される物体 のダイナミクスは、現在の時間の状態と行動から次時間の 状態に遷移するという、マルコフ決定過程(Markov decision process: MDP)によって以下のようにモデル化される。

$$\mathbf{s}_{t+1} = f(\mathbf{s}_t, \mathbf{a}_t) + \boldsymbol{\omega} \tag{1}$$

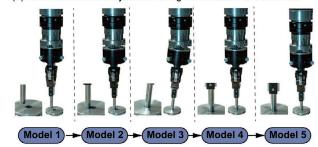
ここで、sは状態(ロボットまたは対象物体の手先位置、関節角、速度など)、aはロボットの行動(ロボットの速度、またはトルク指令など)、 $\omega$ はノイズである。モデルベース学習は、このモデルに基づき、累積期待コスト $\Sigma_{t=0}^{T}$   $\mathbb{E}[c(s_t,a_t)]$  を最小にするロボットの行動を決定することを目的とする。最適行動はモデルを使用して将来の状態を予測しながら求められる。

対象の環境が、極めて単純な場合は事前にモデルを設計することができるが、実世界におけるロボットの環境では、厳密な接触や摩擦などのモデルを解析的に求めることは極めて困難であり、柔軟要素が加わった場合は、その非線形性によりダイナミクスがさらに複雑になるのでより問題が難しくなる。このため、ダイナミクスモデルを機械学習に基づく手法で学習し、ロボット制御に応用するアプ

ローチが存在する<sup>5)8)9)</sup>。これらの手法は、モデルをガウス 過程やニューラルネットワークで学習し、複雑なダイナミ クスを表現しつつ、効率的にロボットの制御器を学習する ことを示した。

しかし、ダイナミクスの自由度が大きくなる、あるいはモデルがより複雑になるなどの場合、モデル学習に必要なデータ数も増加する問題点が生じ、実ロボットのへの応用が困難になる。この問題を解決する手段の一例として、ロボットが作業を遂行可能な範囲で、モデルをできる限り単純に表現するものや、過去に得られたモデルを新しい環境に適用する手段が考えられる。そこで、我々は、組立作業の環境を区分化し、そこに柔らかさによる物体のなじみを考慮することで、ダイナミクスモデルを低次元で表現する手法を開発し、各小区分をモデルベース強化学習で学習する手法を提案した(図 1a)。また、シミュレーション上で大量のデータを収集し、シミュレーション上でモデルを自己教師あり学習で学習し、多様な実世界の環境に追加の訓練なしに適応できる Sim-to-Real 転移学習手法を提案した(図 1b)。次節以降で、これらについて解説する。

# (a) Lower dimensional dynamics using environment and softness



(b) Model-based multiple environment sim-to-real transfer

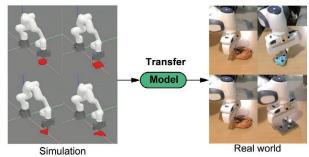


図1 本稿で紹介する研究事例: (a) 環境拘束と物理的柔らかさを活用した低次元ダイナミクス表現<sup>6)</sup> と (b) 多様な環境に即座に適応可能なモデルベース Sim-to-real 転移学習<sup>7)</sup>。

# 3. 環境と柔らかさを活用した低次元ダイナミクス表現による組立学習の効率化

本節では、組立作業の中で代表的なものの1つであるペグ挿入作業に着目し、作業区分化と柔らかさを活用することで、よりシンプルなモデル表現ができ、その結果として学習が効率的化できることを紹介する。

#### 3.1 低次元ダイナミクス表現

ペグ挿入作業では、たとえばペグが地面と接触する、穴 の縁や底に触れるなど、挿入完了に至るまでの環境の物理 的な相互作用から、複数の Manipulation primitives (MP) と いう小課題に区分化できることが示されている10)。この区 分化によって、各区分における状態・行動空間を低次元で 示すことができる。また、柔らかいロボットを活用するこ とで、状態・行動空間をより低次元で示せることが知られ ている。Hangらは、柔軟なグリッパを使用して、薄い物 体を押し付けてスライドさせる制御手法を提案している11)。 柔らかさのおかげで、物体と地面は安定した接触を保障す ることができ、押し付け方向に対する制御を陽に行う必要 がなく、平面上のみを考慮した動作計画が可能となる。ま た、柔らかさによる物体のなじみのおかげで、低次元の制 御入力で、多様な物体の把持が可能になることも報告され ている<sup>12)</sup>。そこで、本研究では、MPと柔らかさを活用す ることで、より状態・行動空間を低次元で表現できるので はないかと考えた。

図 2 に示すように、MP によってペグ挿入作業は 5 種類の小課題に区分化できる。単純のため、矢状面の空間を考慮する。ロボットの状態は、 $y \cdot z$ 方向におけるグリッパの手先の位置と  $\alpha$  方向の回転角及びそれぞれの速度であり、行動はアーム先端の位置もしくは速度指令とする。

- ・n1 Approach: ロボットが、穴表面近くに接近する。この際の状態・行動空間は  $s^{(\text{spproach})} = [y,z]^{\mathsf{T}}$ ,  $a^{(\text{approach})} = [y_d,z_d]$  となる。Approachでは、事前設計しておいた位置制御器を使用する。
- ・ n2 Contact: ロボットが地面と接触する。この際の状態・行動空間は  $\mathbf{s}^{\text{(contact)}} = z$ ,  $\mathbf{a}^{\text{(contact)}} = z_d$  となる。柔軟要素が衝撃を緩和するので、この小課題においても事前設計した位置制御器を使用することができる。
- ・ n3 Fit: ロボットが把持したペグをスライドさせ、穴の縁に当てる。この際の状態行動空間は、 $\mathbf{s}^{(\text{Fit})} = [y, \dot{y}]^{\mathsf{T}}$ 、 $\mathbf{a}^{(\text{Fit})} = \dot{y}_d$ で表現される。柔らかさのおかげで、地面

とペグが安定な接触が保障されているので、z方向の 状態・行動空間を考慮する必要がない。ここから、接 触を伴う作業になりダイナミクスが複雑になるので、 最適な速度指令値をモデルベース強化学習によって学 習する。

- ・ n4 Align: 穴の縁との接触を保ったまま、ペグの姿勢を修正する。状態・行動空間は $s^{(Align)} = [\alpha, \dot{\alpha}]^{\mathsf{T}}$ 、 $a^{(Align)} = \dot{y}_d$ となる。ここでも、柔らかさにより安定した接触を保障できるため、z 方向の状態行動空間は考慮しなくてもよい。さらに、ペグ先端は穴の縁で拘束されており、柔軟要素が 6 自由度の変位が許容できるため、並進運動のみで、ペグの姿勢を修正できる。
- ・n5 Insertion: ペグを穴の底まで挿入する。この際の状態・行動空間は  $s^{(Insertion)} = [z,\dot{z}]^{\mathsf{T}}$ ,  $a^{(Insertion)} = [\dot{y}_d,\dot{z}_d]^{\mathsf{T}}$  となる。入力に y 方向も考慮するのは、ジャミングを回避するためである。

以上により、各区分は状態・行動空間ともに  $1 \sim 2$  次元で表現できることが示された。

#### 3.2 ガウス過程に基づくモデルベース強化学習

区分化された、fit、align、insertionの各区分をモデルベース強化学習で学習する。学習の目的は、Tタイムステップ分の累積期待コストを最小化させるような速度制御器のパラメータを取得することである。

本研究では、サンプル効率の高いモデルベース強化学習 Probabilistic Inference for Learning Control (PILCO) $^{8}$ )を使用する。PILCOはガウス過程回帰を使用し、モデルを学習する。ガウス過程を使用することで、データ不足による不確実性を予測分散として明示的に表すことができる。Tステップ先の状態予測を行うとき、PILCOは解析的モーメントマッチングによる近似推論手法を適用し、過去の状態の不確実性を伝播することができ、その結果効率的な学習を実現している。解析的モーメントマッチングの詳細な説明は松原の解説記事 $^{13}$ )を参照されたい。

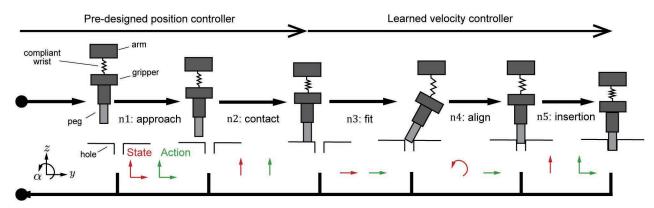


図2 環境と柔らかさを活用した組立作業におけるダイナミクスの低次元表現(文献6)より引用)。

#### 3.3 シミュレーション

提案手法の有効性を確認するため、シミュレーションを 行った。シミュレーションの目的は、提案手法が、低次元 空間表現されていないシステムと比較して、より効率的に 学習できるかを検証することにある。

セットアップ:本研究では、Box2Dという物理エンジンを使用し、2次元平面上のシミュレータを開発した。図 2と同様に、 $\mathbf{y} \cdot \mathbf{z} \cdot \alpha$ 方向の変位が可能である。比較対象は、低次元空間表現を持たないダイナミクスで、状態・行動空間は  $\mathbf{s} = [\mathbf{y}, \mathbf{z}, \boldsymbol{\alpha}, \dot{\mathbf{y}}, \dot{\mathbf{z}}, \dot{\boldsymbol{\alpha}}]^{\mathsf{T}}, \mathbf{a} = [\dot{\mathbf{y}}_{a}, \dot{z}_{a}]^{\mathsf{T}}$ で表現される。

コスト関数は、それぞれの小課題における目標位置・姿勢の誤差である。1学習実験における試行回数は20回とした。

結果:図3にFit、Align、Insertionの各学習試行における、目標位置・姿勢の平均絶対誤差を示す。青線は提案手法、赤線は低次元空間表現を用いない手法である。エラーバーは、30回の学習実験の標準偏差を示す。提案手法を使用した場合、より小さい誤差を示し、3回の試行で収束していることを確認した。この結果から、提案手法を使用することでより効率的に学習できることを示した。

#### 3.3 実機実験

提案手法が実環境でも動作するかを検証するために、実

機実験を行った。柔軟手首を協働ロボット UR5 に搭載した。ロボットのグリッパの位置・姿勢はモーションキャプチャで計測した。実験で使用したペグと穴は、直径10mm、はめあい公差は H7/h7 であった。

学習後のロボットの動作のスナップショットを図4に示す。シミュレーション同様、ロボットは、Fit、Align、Insertionの小課題を $2\sim3$ 回程度の試行回数で学習でき、学習終了時における成功率はそれぞれ8/10、9/10、10/10であった。以上より、実環境でも少数の試行回数で学習できることを示した。

# 4. 多様な環境に即座に適応する Sim-to-Real 転 移学習

本節では、シミュレータで事前に学習したモデルを多様な実環境に転移させる、Sim-to-Real 学習について解説する。

#### 4.1 多様な環境への Sim-to-Real 転移学習

近年の学習技術の発展によって、ロボットがより効率的に複雑なタスクを学習できるようになってきた。しかし、実環境でのデータ取得にかかる労力は依然として大きいため、シミュレーション学習した事前知識を、実世界に転移させる Sim-to-Real の研究が広く行われている。我々

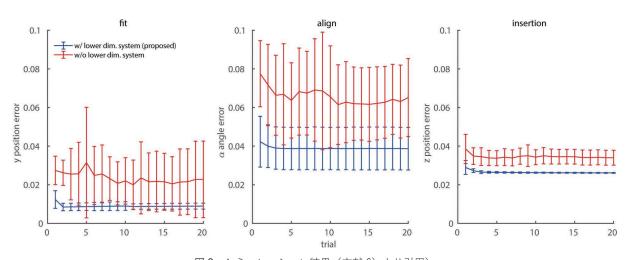


図3 シミュレーション結果 (文献6) より引用)。











図4 学習後のロボットの組立動作(文献6)より引用)。

は、事前学習されたモデルが、多様な環境に対していかに即座に適応できるかについて興味がある。Sim-to-Real の一般的なアプローチとして、Domain Adaptation (DA) $^{14}$ と Domain Randomization (DR) $^{15}$ が挙げられる。しかしながら、DA は未知環境に適応するために追加で訓練する必要があり、DR は環境のパラメータ数が多い時に獲得された制御器もしくはモデルが保守的になることが指摘されている $^{16}$ 。

この問題に対し、我々は、Explicitly and Implicitly Conditioned Network (EXI-Net) というネットワークを提案した $^{14}$ 。図 5 に EXI-Net の概要図を示す。EXI-Net はダイナミクスモデルを、明示的(Explicit)・暗示的(Implicit)パラメータで条件付けされたニューラルネットワークで学習する。明示的パラメータは、質量や摩擦係数など、定量的に表現できるものを表現し、暗示的パラメータは、形状など定量的に示せないものを表現する。

シミュレーションにおける訓練時は、明示的パラメータを与えて、順ダイナミクスモデルのニューラルネットワークのパラメータと暗示的パラメータを学習する。テスト時には、未知環境における明示的・暗示的パラメータを誤差順・逆伝搬を反復的に行いながら推定し、最適な行動を計算する。この反復的計算による推定によって、未知環境にオンランプで適応することができる。

本研究では、ロボットが物体を目標位置まで押す課題に 着目する。この課題は押す物体によって、明示的・暗示的 パラメータが大きく異なるので、多様な環境に適応できる かを検証するのに適している。

#### 4.2 EXI-Net を用いたモデル学習

まず、シミュレーション上でモデルを訓練させる。EXINetによるダイナミクスモデルは  $\mathbf{s}_{t+1} = f(\mathbf{s}_t, \mathbf{a}_t, \mathbf{d}_e, \mathbf{d}_i)$  で与えられる。ここで、 $\mathbf{d}_e$  は明示的パラメータ、 $\mathbf{d}_i$  は暗示的パラメータである。 $\mathbf{d}_e$  は質量・摩擦・重心位置で構成されており、訓練中ランダムに与えられる。一方、同時に複数の異なる形状の物体も与えられ、 $\mathbf{d}_i$  はパラメトリックバイアス<sup>17)</sup> という手法で学習される。また、ランダムに行動を与えて、自己教師あり学習によってモデルを学習する。

次に、テスト時において、EXI-Net が未知物体にどのように適応するかを説明する。前節と同様、EXI-Net も、順モデル予測によって、コスト関数を最小化させる行動を計算する。この手法では、最適な行動は、ニューラルネットワークの誤差順・逆伝搬によって以下のように求められる。

$$\mathbf{g}_{a} = \partial c(\mathbf{s}_{t+1}^{goal}, \mathbf{s}_{t+1}^{pred}) / \partial \mathbf{a}_{t},$$

$$\mathbf{a}_{t} \leftarrow \mathbf{a}_{t} - \varepsilon_{a} \mathbf{g}_{a} / \|\mathbf{g}_{a}\|_{2},$$
(2)

ここで  $\varepsilon_a$  は学習率である。c はコスト関数であり、目標状態と予測状態の 2 乗誤差で定義される。また、未知環境における  $d_e, d_i(d=[d_i^\intercal, d_e^\intercal]^\intercal$ とする)も同様に誤差順・逆伝搬によってオンラインで計算することができる。

$$\mathbf{g}_{d} = \partial c(\mathbf{s}_{t+1}^{obs}, \mathbf{s}_{t+1}^{pred}) / \partial \mathbf{d},$$
  
$$\mathbf{d} \leftarrow \mathbf{d} - \varepsilon_{d} \mathbf{g}_{d} / \|\mathbf{g}_{d}\|_{2},$$
 (3)

ここで、 $\varepsilon_d$  は学習率である。勾配は過去の時間  $\tau$  に獲得された状態・行動・次の状態のペアのデータをランダムに

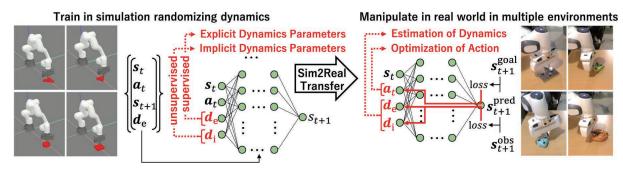


図5 EXI-Netの概要(文献7)より引用)。

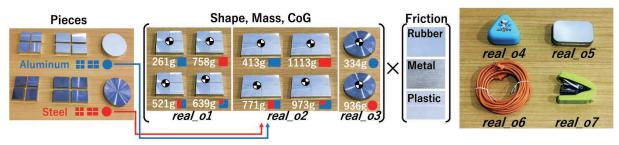


図6 テストに使用した物体(文献7)より引用)。

サンプルして計算する。これにより、未知の環境においても、最適な行動とパラメータ **d**をオンラインで計算することができる。

#### 4.3 実験

提案手法の有効性を検証するために、実機実験を行った。実験の目的は、提案手法が、一般的に使用されている DR の手法と比較して高い成功率を示すかを確認することである。

セットアップ:シミュレータ上では、長方形 2 種類・三角形 3 種類・円の 1 種類、計 6 種類の形状を持つ物体で訓練を行った。実験には Gazebo というシミュレータを使用した。順モデルには、3 層、ニューロン数 100 個の全結合ネットワークを使用した。活性化関数には ReLU を使用した。各物体にそれぞれ、 $60\sim120$  種類の明示的パラメータ  $d_e$ を与えて、ネットワークパラメータと暗示的パラメータ  $d_e$ を学習した。

テスト時における実環境の物体を図6に示す。real\_o1~o3では多様な質量・重心・摩擦係数を設定するために、アルミと鉄のプレートを組み合わせ、底面にはゴムやプラスチックのシートを貼り合わせたものを設計した。また、真のダイナミクスが未知である物体 real\_o4~o7 も検証した。比較対象は、提案手法(EXI)と EXI-Net に真の  $d_e$  が与えられた場合(EXI w/ True  $d_e$ 、real\_o1~3のみ対象)、従来の DR で使用されている、LTSM を使用したネットワーク(LSTM)である。real\_o1~3では、各物体にそれぞれ12条件、12条件、6条件あり、各条件で5試行(計150回)した。また、real\_o4~7では、各物体10試行した。テスト中は、ロボットが物体を15回押す間に指定された目標位置に到達するかどうかを確認した。

実験結果:実験結果を表1に示す。各物体における、平均の成功率・目標到達のためのステップ数、真の $d_e$ との誤差を示す。この表より、すべての物体で提案手法が最も高い成功率を示し、少ないステップ数で目標に到達していることがわかる。興味深いことに、真の $d_e$ が与えられた場合と比較しても、提案手法の方が高い性能を示している。これは、シミュレーションと実環境とのモデル化誤差が残っているためであり、提案手法はオンラインパラメータ推定によってこの誤差を減少させていることが原因であると考えられる。

表 1 テスト環境における結果 (成功率/到達ステップ数/真の  $d_e$  との誤差)

Object	EXI w/ True <b>d</b> e	LSTM	EXI
real_o1	0.89/11.3/0.0	0.72/12.8	1.0/8.3/0.52
real_o2	0.95/13.2/0.0	0.69/13.1	0.98/6.2/0.33
real_o3	0.98/10.8/0.0	0.92/13.4	1.0/7.9/0.38
real_o4	_	0.5/13.8	1.0/7.0
real_o5	_	0.6/14.5	0.9/9.5
real_o6	_	0.5/13.3	1.0/7.2
real_o7	_	0.5/13.6	1.0/6.8

以上の結果より、提案手法が従来手法と比較してより高い成功率で物体を操作することができた。この手法は、硬いロボットを使用しているが、現在我々は柔軟手首を持つロボットのシミュレータを開発しており<sup>18)</sup>、柔らかいロボットへの応用や、様々な部品の組立への応用も期待できる。

# 5. むすび

本稿では、組立自動化に向けて柔らかいロボットをより容易に制御するために、モデルベース学習を活用した研究事例について紹介した。モデルベース学習の手法は、モデルフリーの手法と比較してサンプル効率が高く、実ロボットの応用に適している。我々はさらに学習を効率化させるために、作業を区分化し柔らかさを活用することで、よりシンプルなダイナミクスにすることで学習を効率化させることがきた。また、柔らかいロボットにも十分応用ができる、多様な環境に適応できる Sim-to-Real 転移学習に関する手法も紹介した。我々は、EXI-Netという、シミュレータ上でモデルを事前学習し、未知環境のダイナミクスパラメータをオンラインで推定できる手法を提案し、従来の一般的な Domain Randomization の手法より高い成功率を示した。

本稿を読み、モデルベース学習に基づく柔らかいロボット制御や、Sim-to-Real転移学習に興味を持っていただければ幸いである。

#### 参考文献

- Li, R.; Qiao, H. A survey of methods and strategies for high-precision robotic grasping and assembly tasks some new trends. IEEE/ ASME Transactions on Mechatronics. 2019, Vol.24, No.6, p.2718–2732.
- Beltran-Hernandez, C. C.; Petit, D.; Ramirez-Alpizar, I. G.; Nishi, T.; Kikuchi, S.; Matsubara, T.; Harada, K. Learning force control for contact-rich manipulation tasks with rigid position-controlled robots. IEEE Robotics and Automation Letters. 2020, Vol.5, No.4, p.5709-5716.
- 3) 井尻善久, フェリクス フォンドリガルスキ. 産業用ロボッ

- トの進化によるものづくりの近未来. 日本ロボット学会誌. 2019, Vol.37, No.8, p.675-678.
- 4) Tanaka, K.; Von Drigalski, F.; Hamaya, M.; Lee, R.; Nakashima, C.; Shibata, Y.; Ijiri, Y. "A compact, cable-driven, activatable soft wrist with six degrees of freedom for assembly tasks". IEEE/RSJ International Conference on Intelligent Robots and Systems. 2020, p.8752-8757.
- 5) Ibarz, J.; Tan, J.; Finn, C.; Kalakrishnan, M.; Pastor, P.; Levine, S. How to train your robot with deep reinforcement learning: lessons we have learned. The International Journal of Robotics Research. 2021, Vol.40, No.4-5, p.698-721.
- 6) Hamaya, M.; Lee, R.; Tanaka, K.; von Drigalski, F.; Nakashima, C.; Shibata, Y.; Ijiri, Y. "Learning robotic assembly tasks with lower dimensional systems by leveraging physical softness and environmental constraints", IEEE International Conference on Robotics and Automation. 2020, p.7747–7753.
- 7) Murooka, T.; Hamaya, M.; von Drigalski, F.; Tanaka, K.; Ijiri, Y. "EXI-Net: explicitly/implicitly conditioned network for multiple environment sim-to-real transfer". Conference on Robot Learning, 2020, https://corlconf.github.io/corl2020/paper\_268/, (参照 2021–06–30).
- Deisenroth, M. P.; Fox, D.; Rasmussen, C. E. Gaussian processes for data-efficient learning in robotics and control. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2015, Vol.37, No.2, p.408-423.
- Lenz, I.; Knepper, R. A.; Saxena, A. DeepMPC: Learning deep latent features for model predictive control". Robotics: Science and Systems. 2015.
- Johannsmeier, L.; Gerchow, M.; Haddadin, S. "A framework for robot manipulation: Skill formalism, meta learning and adaptive control". IEEE International Conference on Robotics and Automation. 2019, p.5844-5850.
- Hang, K.; Morgan, A. S.; Dollar, A. M. Pre-grasp sliding manipulation of thin objects using soft, compliant, or underactuated hands. IEEE Robotics and Automation Letters. 2019, Vol.4, No.2, p.662–669.
- Deimel, R.; Brock, O. A novel type of compliant and underactuated robotic hand for dexterous grasping. The International Journal of Robotics Research. 2016, Vol.35, No.1-3, p.161-185.
- 13) 松原崇充. ガウス過程に基づくロボットの運動制御・学習―解析的モーメントマッチングによる近似推論. システム/制御/情報. 2016, Vol.60, No.12, p.515-520.
- 14) Patel, V. M.; Gopalan, R.; Li, R.; Chellappa, R. Visual domain adaptation: A survey of recent advances. Signal Processing Magazine. 2015, Vol.32, No.3, p.53-69.
- 15) Peng, X. B.;rychowicz, M.; Zaremba, W.; Abbeel, P. "Sim-to-real transfer of robotic control with dynamics randomization". IEEE International Conference on Robotics and Automation, 2018, p.3803-3810.
- 16) Nachum, O.; Ahn, M.; Ponte, H.; Gu, S.; Kumar, V. Multi agent manipulation via locomotion using hierarchical sim2real, arXiv preprint arXiv:1908.05224, 2019, (参照 2021-06-30).
- 17) Ogata, T.; Ohba, H.; Tani, J.; Komatani, K.; Okuno, H. G. "Extract-

- ing multi-modal dynamics of objects using RNNpb". IEEE/RSJ International Conference on Intelligent Robots and Systems. 2005, p.966–971.
- 18) Hamaya, M.; Tanaka, K.; Shibata, Y.; von Drigalski, F.; Nakashima, C.; Ijiri, Y. Robotic learning from advisory and adversarial interactions using a soft wrist. IEEE Robotics and Automation Letters. 2021, Vol.6, No.2, p.3878-3885.

### 執筆者紹介



**濵屋 政志 HAMAYA Masashi** オムロン サイニックエックス株式会社 リサーチアドミニストレイティブディビジョン 専門:ロボティクス、機械学習、強化学習 所属学会:日本ロボット学会、IEEE 博士 (工学)

本文に掲載の商品の名称は、各社が商標としている場合があります。