

# 並列シミュレーション環境を用いた コンセント挿入動作の獲得に関する研究

CHO CHANYEONG\*, 山崎公俊\*

## A Study on Acquiring Plug Insertion Actions Using Parallel Simulation Environments

Cho Chanyeong\* and Kimitoshi Yamazaki\*

This paper proposes a reinforcement learning-based method for robotic plug insertion using parallel simulation environments. Robotic plug insertion is a complex task that requires handling uncertainties such as initial misalignment and environmental noise. This study aims to address these uncertainties by focusing on sensor-free operation and enabling the robot to perform the insertion task through trial-and-error learning, rather than relying on precise control. This allows the robot to adapt to uncertain conditions while efficiently gathering large-scale data, improving the learning speed and enabling robust plug insertion across various initial conditions and noise levels.

**Key Words:** Reinforcement learning, Insertion actions, Isaac Gym (parallel simulation)

### 1. はじめに

近年、ロボット産業は急速な進展を遂げており、日常生活で利用可能なロボットの開発が求められている。特に、何かを挿入する操作は日常生活において必要不可欠なものであり、代表的な例としてコンセントへのプラグ挿入が挙げられる。この動作は非常に精密な制御を要し、センサのノイズやずれ、環境条件によっては、ロボットが正確にプラグを挿入することは極めて難しい。

本研究の目的は、こうした課題を克服し、ロボットが試行錯誤を通じてコンセントへのプラグ挿入動作を獲得することである。

本研究の貢献は以下の3点である。

- ・強化学習と並列シミュレーション環境を組み合わせ、ロボットが効率的にプラグ挿入動作を学習できる手法を提案したこと。
- ・初期位置や角度にノイズが含まれていても、ロボットが適切な操作を行えることを示した点。
- ・提案手法により、高齢者や障害者の日常生活の支援や感電事故の防止など、安全性向上に寄与する可能性を示唆したこと。

### 2. 関連研究

ロボットが何かを挿入してアセンブリをする動作は高度の技術力を要し、長年の課題である[4, 9]。模倣学習を通じて解決しようとした研究もあるが、多様な環境に対応できないことが問題とされている[3, 5]。コンタクトリッチな組み立てタスクに関するシミュレーションから現実世界への転送は、これまでにいくつかの研究で試みられている。特に、[1]は棒の挿入タスクにおいてシミュレーションで学習した動作を現実世界に適

用する手法を提案しており、本研究ではこれを基盤としてコンセント挿入タスクに発展させた。従来の単純な挿入タスクを超え、日常生活に密接したプラグ挿入という複雑な課題に挑戦している。

また、NVIDIAのIsaac Gymは並列シミュレーション環境を提供するツールであり[2]、多くの研究で採用されている[6, 7, 9]。本研究でも、これを採用することで大規模データを短時間で生成し、効率的な学習を実現している。

最後に、[8]ではシミュレーションにノイズを加えることでSim2Realギャップを縮小する手法が提案されており、本研究でも同様にノイズを適用して、現実世界への実装可能性を高める工夫を行っている。

### 3. 問題設定とアプローチ

#### 3.1 問題設定

コンセント挿入タスクにおける現実世界での課題として、まず初期位置や角度におけるセンサの誤差やノイズの影響が挙げられる。これにより、ロボットがプラグをコンセントの穴に正確に挿入することが困難になる。また、センサに依存するシステムはコストが高く、システムの複雑性も増す。挿入動作に必要な高精度な制御は、センサを使用しても完全に解決できない課題である。

これに対して、本研究ではセンサに依存せず、ロボットの動作に注目するアプローチを採用する。これにより、センサのエラーやノイズの問題から自由になり、システムを単純化してコストを削減することができる。ロボットの挿入動作は「コンセントを見つける」「プラグを近づける」「挿入する」という3つのステップに分かれるが、本研究では特に「挿入する動作」に焦点を当てている。

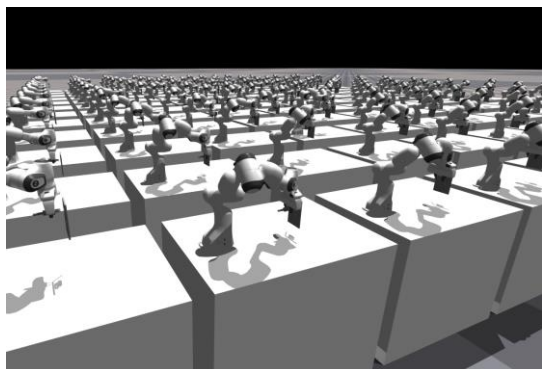


Fig. 1 Robots Learning at Isaac Gym

### 3.2 アプローチ

上記の課題に対して、本研究では強化学習を用いたアプローチを採用する。強化学習は、ロボットが環境と相互作用し、報酬を基に試行錯誤を通じて最適な動作を学習する手法であり、特に精密な操作作業に有効である。NVIDIA の Isaac Gym を使用することで、大量のデータを短期間で収集し、効率的な学習が可能となる。

このシミュレーション環境では、様々なノイズやずれがある状況下でロボットが学習を行う設定になっており、リアルタイムでのフィードバックを基に最適な動作を習得することができる。シミュレーションを用いることで、数千台のロボットが同時に学習でき、学習コストの低減にも繋がる。

## 4. 手法の説明

本研究で使用する手法は、NVIDIA の Isaac Gym という並列シミュレーション環境を用いている。Isaac Gym は、GPU を活用した高速シミュレーションツールであり、数千台のロボットを同時に学習させることができる。これにより、大規模なデータを短時間で生成し、強化学習の効率性を高めることが可能となる。強化学習のアルゴリズムとしては、Proximal Policy Optimization (PPO) を使用し、各エピソードで 7-DOF の Franka ロボットアームがプラグ挿入に成功するかどうかを報酬として設定している。

観測に関しては、Franka ロボットの各関節の角度、グリップの位置、プラグピンの中心座標位置、コンセント穴の奥の中心座標位置、そしてそれらの間の距離をロボットが認識している。これらの情報を基に、ロボットはプラグ挿入の動作を行う。なお、Franka ロボットの各関節の速度やトルクの観測は行っておらず、これは実際のロボットで大きなノイズが発生しやすいためである。力学センサを使用するとコストやノイズ、システムの脆弱性に問題が生じることから、観測データは主に位置情報に依存している。

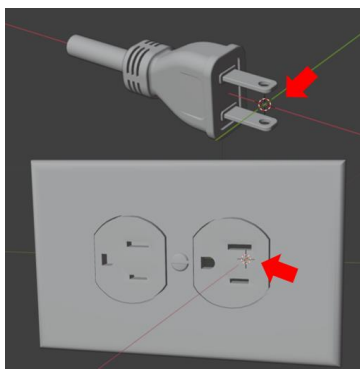


Fig. 2 a Plug and a Socket

行動では、ロボットは各関節の角度を変えながら、プラグを小さな穴に通過させ、コンセントの内部中心座標に向かって動作を行う。この一連の動作は、ロボットが試行錯誤を通じて挿入を試みるものであり、挿入が成功するほど報酬が増加する仕組みである。さらに、プラグとコンセントの間の距離が近づくほど報酬が高くなり、完全に挿入された場合には、最も高い報酬が与えられる。このような報酬設定により、ロボットは精密な制御を行わずとも、効率的にプラグ挿入動作を学習し、試行錯誤を通じてタスクの成功率を向上させることができる。

また、3つのサポートアルゴリズム (SAPU, SBC, SDF) を用いて、ロボットの動作精度をさらに高めている。

• **SAPU (Simulation-Aware Policy Update)** は、シミュレーション環境内での成功距離を基に報酬を調整する手法である。具体的には、ロボットがコンセントの中心に向けてプラグを挿入する際、最終的な挿入距離が成功許容範囲に収まった場合、より高い報酬を与える。このプロセスにより、ロボットは成功した挿入動作に対するフィードバックをリアルタイムで受け取り、試行錯誤を通じて挿入成功率を向上させる役割を果たす。

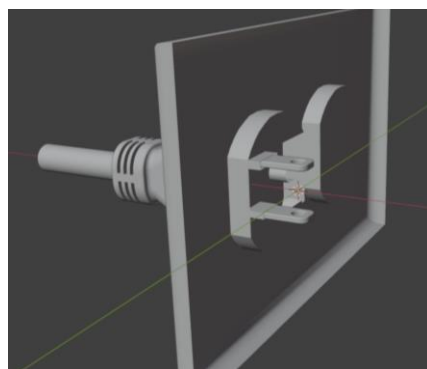


Fig. 3 SAPU Range (Successfully inserted)

• **SBC (Success-Based Curriculum)** は、学習難易度に応じて報酬を変動させる手法である。タスクの難易度が高くなるにつれて、環境に含まれるノイズが増加し、ロボットがより困難な状況で挿入動作を実行することが求められる。このアルゴリズムは、ロボットが困難なタスクに取り組む際、成功時により大きな報酬を与えることで、段階的な学習を促進する。

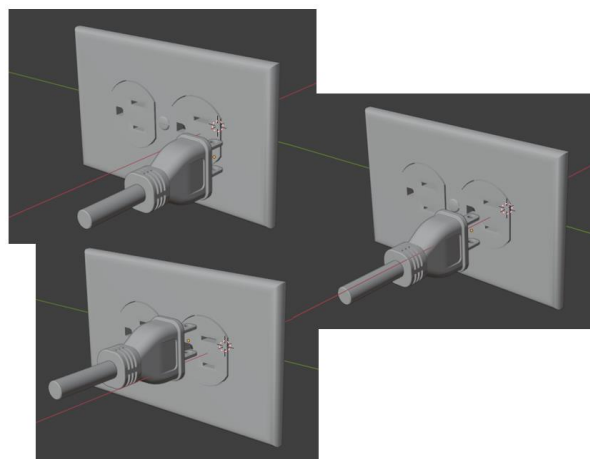


Fig. 4 SBC Noises

• **SDF (Signed Distance Field)** は、プラグとコンセントの間の距離に基づいて報酬を計算する手法であり、特に挿入動作がより遠い位置から成功した場合、より高い報酬を与える。この

手法により、ロボットは困難な状況下での成功を評価され、長距離の操作を通じてタスクの精度を向上させることが可能となる。

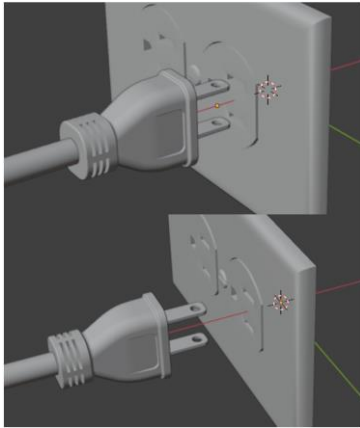


Fig. 5 SDF Initial Distance

## 5. 実験

### 5.1 実験

実験では、ロボットがノイズやずれがある環境でプラグ挿入動作を学習できるかを評価するため、初期位置と角度にランダムなノイズを加えた3つの異なる設定で強化学習を行った。使用したのは7自由度（7-DOF）のFranka ロボットアーム256台で、各エピソードでFranka ロボットアームがプラグ挿入に成功するかどうかを報酬として与える。シミュレーション環境では、ロボットはノイズとずれがある様々な初期位置と角度からプラグ挿入を試みる。

2ピン110Vコンセントと3ピン110Vコンセントの挿入はそれぞれ別々に学習を行い、その後、両方の形状に対応できるように混合条件での学習も試みた。混合条件下では、ランダムに2ピンと3ピンのコンセントが切り替わり、ロボットがどちらの形状にも対応する動作を学習できるかを確認した。

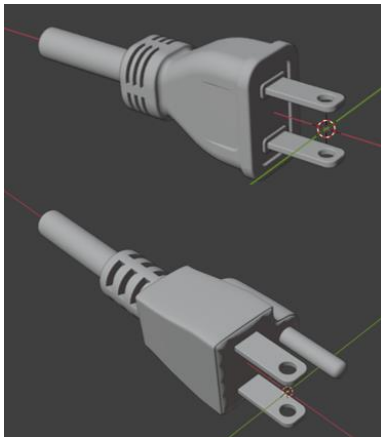


Fig. 6 2-pin and 3-pin 110V plug

これらの設定において、ロボットは各エピソードで成功有無に基づいて報酬を受け取り、挿入に失敗した場合にはペナルティが課された。特に、SBC (Success-Based Curriculum) アルゴリズムを用いることで、段階的に難易度を高めながらロボットが学習を進め、最終的に困難なタスクに適応できるようにした。

学習の成果を評価するため、学習が終わったロボットをIsaac Gym シミュレーション内で4台生成し、動作を再生して

プラグが正しく挿入されるかどうかを肉眼で確認した。

各エピソードの長さは約11秒であり、成功基準として、挿入が始まった後、エピソードが終了する際に完全挿入が予測される場合を「挿入開始」とし、プラグが完全に挿入された場合を「完全挿入」として記録した。

### 5.2 実験結果

実験結果として、2ピンおよび3ピンのコンセント挿入における各エピソード100回を観察した成功率を比較した。まず、**2ピン110Vコンセント**においては、ロボットが84%の確率でプラグ挿入を開始し、76%の確率で完全な挿入に成功した。また、**3ピン110Vコンセント**では、89%の成功率で挿入を開始し、82%の確率で完全に挿入できた。これに対して、混合条件で学習させた場合、3ピンコンセントへの挿入成功率は56%、2ピンコンセントでは52%と、単独での学習に比べて成功率が低下する結果が得られた。

これらの結果から、個別に学習した場合、ロボットは高い精度でプラグ挿入動作を習得できることが確認されたが、複数の形状を同時に学習させた場合には精度が低下する傾向が見られた。

Table. 1 Results of Individual Learning

	挿入開始	完全挿入
2-pin 110V plug	84/100	76/100
総合	84%	76%
3-pin 110V plug	89/100	82/100
総合	89%	82%

Table. 2 Results of Combined Learning

	挿入開始	完全挿入
2-pin 110V plug	52/100	43/100
総合	52%	43%
3-pin 110V plug	56/100	48/100
総合	56%	48%

## 6. 考察

実験結果から、2ピンと3ピンのコンセント挿入動作において、各形状を個別に学習させた方が高い成功率を達成できることが分かった。一方で、混合条件で学習させた際には、ロボットが異なる形状のコンセントに対応するのが難しく、精度の高い挿入動作を学習することが困難であることが確認された。この理由として、2ピンと3ピンで挿入動作自体が異なることが挙げられる。2ピンの場合は平行移動で挿入するが、3ピンの場合はプラグを傾けて下のピンを先に挿入し、その後残りの2つのピンを挿入する動作が必要となるためである。

この結果は、ロボットが異なる形状のコンセントを扱う際には、事前に物体の形状を認識し、それに応じた動作計画を立ててから挿入動作を試みる必要があることを示している。人間が異なる形状の物体を確認し、適切な挿入方法を選択するプロセスと同様の処理がロボットにも求められる。

現実世界での実装に向けては、ロボットがコンセントの種類を事前に識別し、各形状に応じた動作計画を立てる必要がある。今後は、シミュレーション環境での学習を進めると同時に、現実世界での実装に向けた検証を行い、ロボットがより多様な状況に対応できるよう技術を発展させていく必要がある。

## 7. まとめ

本研究では、並列レーション環境を利用して、ロボットがノイズやずれのある環境下でもプラグ挿入動作を効果的に習得できることを示した。また、強化学習アルゴリズムを組み合わせることで、正確な制御に頼らずとも、ロボットが効率的に挿入動作を学習できることが確認された。

今後の展望としては、より多様なコンセント形状やノイズ条件に対応する手法の開発が課題となる。さらに、現実世界での実装に向けては、ロボットが「コンセントを見つける」「プラグを近づける」「挿入する」という3段階の動作をすべて遂行できるように技術を発展させていきたい。また、これらの機能を搭載したロボットシステムを実際に構築することを目指している。

## 参考文献

- [1] Tang, Bingjie, et al. "Industreal: Transferring contact-rich assembly tasks from simulation to reality." *arXiv preprint arXiv:2305.17110* (2023).
- [2] Makoviychuk, Viktor, et al. "Isaac gym: High performance gpu-based physics simulation for robot learning." *arXiv preprint arXiv:2108.10470* (2021).
- [3] Stepputtis, Simon, et al. "A system for imitation learning of contact-rich bimanual manipulation policies." *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022.
- [4] Kenneth Kimble, Justin Albrecht, Megan Zimmerman, and Joe Falco. Performance measures to benchmark the grasping, manipulation, and assembly of deformable objects typical to manufacturing applications. *Frontiers in Robotics and AI*, 2022.
- [5] Wu, Philipp, et al. "Gello: A general, low-cost, and intuitive teleoperation framework for robot manipulators." *arXiv preprint arXiv:2309.13037* (2023).
- [6] Inoue, Shintaro, et al. "Body Design and Gait Generation of Chair-Type Asymmetrical Tripedal Low-rigidity Robot." *2024 IEEE 7th International Conference on Soft Robotics (RoboSoft)*. IEEE, 2024.
- [7] Huang, Binghao, et al. "Dynamic handover: Throw and catch with bimanual hands." *arXiv preprint arXiv:2309.05655* (2023).
- [8] Chebotar, Yevgen, et al. "Closing the sim-to-real loop: Adapting simulation randomization with real world experience." *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019.
- [9] Chen, Yuanpei, et al. "Sequential dexterity: Chaining dexterous policies for long-horizon manipulation." *arXiv preprint arXiv:2309.00987* (2023).