

論文内容の要旨

博士論文題目

Sim-to-Real Reinforcement Learning with Low-fidelity Simulators and Extensive Domain Randomization for Complex Tasks

(低忠実度シミュレータと広範なドメイン乱択化を用いた複雑タスクの Sim-to-real 強化学習)

氏 名

角川 勇貴

(論文内容の要旨)

近年、シミュレーション上でロボットの制御方策を学習し実機環境にゼロショットで方策転移する Sim-to-Real 強化学習手法が注目を集めている。この手法では実機環境とシミュレーションの間で生じるリアリティギャップに頑健な方策を獲得する必要がある。そのため、実機性能が高い転移方策を獲得するためには、実機環境に忠実なシミュレータを作成することが要求される。しかし複雑タスクにおいては詳細な物理現象の再現が困難であり、忠実度の高いシミュレータの作成は難しい。よって、リアリティギャップを減らすことができないことから、複雑タスクでは Sim-to-Real 強化学習の適用が困難である。この問題に対して本研究では、複雑タスクに対する Sim-to-Real 強化学習のためのフレームワークを提案する。具体的には、シミュレータの各構成要素（ロボットや操作物）のモデルを単純化し、各構成要素の形状や物理パラメータの組み合わせをランダム化して学習するドメイン乱択化を網羅的に適用することでリアリティギャップへ対処する。さらに、網羅的なドメイン乱択化に対して強化学習の方策更新が不安定化する課題に対して、学習を頑健化させるために新たな強化学習アルゴリズムを提案する。提案アルゴリズムはランダム化されたパラメータの範囲をいくつかの小さなサブドメインに分割し、各サブドメインにローカル方策を割り当てる。そして、ローカル方策の学習は、周期的にターゲットサブドメインを隣接サブドメインに遷移させ、単調改善による方策更新スキームで隣接サブドメインの学習された方策を利用しながら行われる。最後に、学習された全てのローカル・ポリシーは、実機転移のためのグローバル方策に蒸留される。

提案するアプローチの評価として、3つの適用事例において検証を行った。1つ目の事例では、複雑タスクとして粉体秤量タスクの Sim-to-Real 強化学習が可能であることを検証した。シミュレーションで獲得した方策を実機転移実験で評価した結果、異なる粉体材料とターゲット質量に対して、計量タスクを達成できることが実証された。2つ目の事例では、複雑タスクとして土砂からの岩石除去タスクの Sim-to-Real 強化学習が可能であることを検証した。シミュレーションで獲得した方策を実機転移実験で評価した結果、複数種類の岩石と土砂環境において頑健に岩石除去を達成できることが実証された。3つ目の事例では、複雑タスクとして柔軟物体のブロック崩しタスクに対して、提案する新たな強化学習アルゴリズムの有効性を検証した。その結果、従来手法と比較して提案手法のみがブロック崩しタスクに適用された広範なドメインランダム化に対して頑健に方策を学習できることを確認した。結論として、広範なドメイン乱択化とそ
ののための強化学習手法によって、複雑タスクに対する Sim-to-Real 強化学習が適用できることが確認できた。

(論文審査結果の要旨)

本論文では、シミュレーション上でロボットの制御方策を学習し実機環境にゼロショットで方策転移する Sim-to-Real 強化学習について検討している。実機環境とシミュレーションの間のリアリティギャップに頑健な方策を獲得するために、忠実度の高いシミュレータが期待される。しかし、非剛体や不定形物を含む複雑タスクに対する詳細な物理現象の再現が困難であり、忠実度の高いシミュレータの作成が難しい課題がある。そこで本研究では、複雑タスクに対する Sim-to-Real 強化学習のためのフレームワークを提案した。提案フレームワークでは、シミュレータの各構成要素（ロボットや操作物）のモデルを単純化し、各構成要素の形状や物理パラメータの組み合わせをランダム化して学習するドメイン乱択化を網羅的に適用することでリアリティギャップへ対処する。さらに、網羅的なドメイン乱択化に対して強化学習の方策更新が不安定化する課題に対して、学習を頑健化する強化学習アルゴリズムを提案した。

提案アプローチの概念検証として、粉体秤量タスク、岩石除去タスク、柔軟物体ブロック崩しタスクの3つの事例について適用実験を実施した。いずれについても、複雑な対象物やロボットとの接触現象として単純化されたモデルを実装しシミュレーション環境を構築した。さらに、各モデルパラメータに対して広範なドメイン乱択化を適用した強化学習により、Sim-to-Real 転移可能な頑健な方策が獲得できることを実験的に確認した。複雑タスクに対して忠実度の低いシミュレータを用いるアプローチの独自性に加えて、挑戦的な3つのタスクに対する実装技術も高い学術的価値が認められる。実ロボットを用いた検証実験結果に裏打ちされた、有用性・実用性も高く評価できる。

公聴会では、提案手法の貢献や位置づけの明確化、複雑タスクにドメインランダム化を適用する知見、忠実度とランダム化の関係性、先行研究との比較などについて審査委員との質疑が行われた。その後の最終審査において、すべての質問への回答と論文修正について説明が行われ、適切な対応がとられたことを審査員全員で確認した。

本論文は、複雑タスクに対する Sim-to-Real 強化学習フレームワークを提案した。シミュレータのモデルを単純化し、各構成要素の形状や物理パラメータの組み合わせをランダム化して学習するドメイン乱択化を網羅的に適用しリアリティギャップへ対処する。シミュレーションおよび実ロボットを用いたタスクなどに対する評価実験を実施していることに鑑み、新規性および有用性の観点から一定の学術的意義があるものと評価できる。よって、本論文は博士（工学）の学位論文として価値あるものと認める。