

人間参加型のロボット行動学習

概要

生成AIをロボット制御に適用する際、人から与えられた行動データから学習する場合と、人が行動に与える評価から学習する場合の2通りがありますが、一長一短でそのままでは学習効率などの問題があります。本研究では両者を強化学習の枠組みで統一し、どちらの方法も利用できる方式を考えることで学習効率や最終的な制御性能を改善することに成功しました。

特徴

- 人から与えられた行動データから制御則を学習する生成模倣学習と、行動データに対する人の嗜好から評価関数を学習する報酬学習を、一つの基準で統合したアルゴリズムを開発しました。
- 生成模倣学習は制御則だけでなく環境のダイナミクスを同時に学習し計算機内でのシミュレーションが可能となります。従来法と異なり、部分的嗜好データから報酬を学習できます。
- 生成模倣学習や報酬学習単独で用いる場合よりも、少ないデータ数で学習できるようになりました。また報酬学習時に経過時間や生成行動の滑らかさを同時に提示することで学習効率を改善することがわかりました。

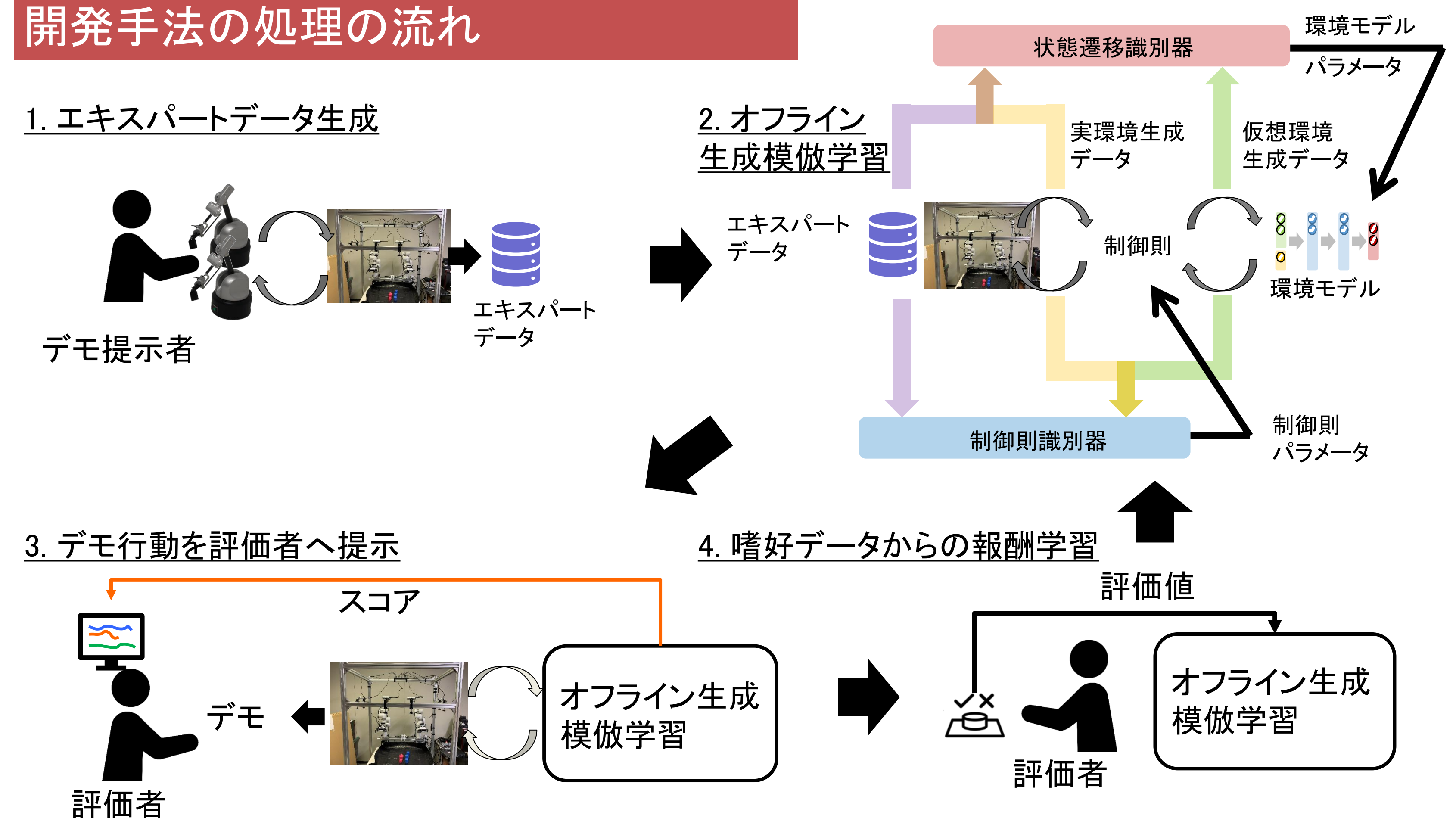
今後の展開

- 本研究において、デモ提示者と評価者が同一である場合の行動学習法が確立されました。今後は提示者や評価者が複数になった場合に、提示者間・評価者間でデータを共有できる方式に拡張する計画です。

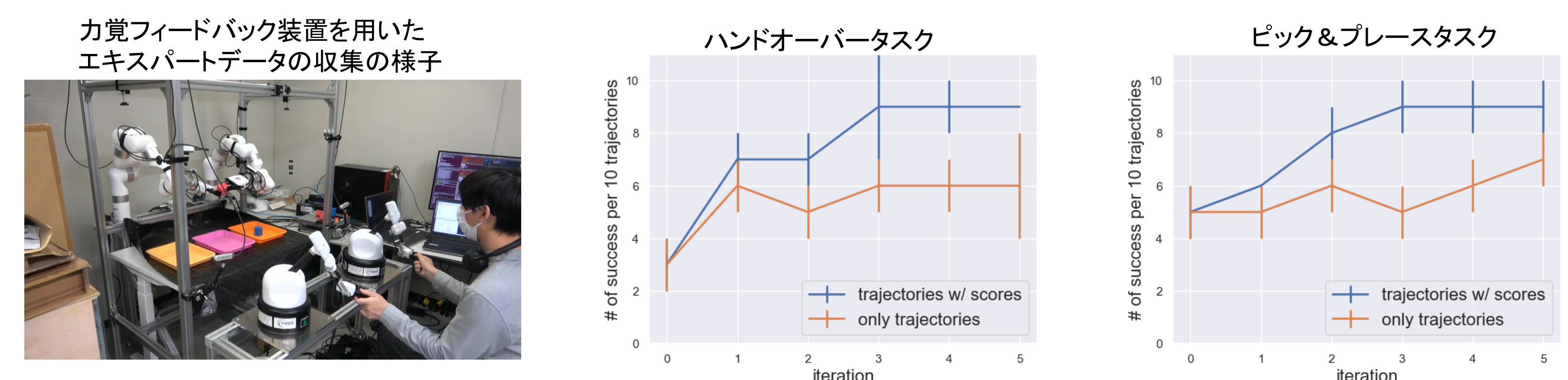
テーマ「万博、そしてその先へ ～科学技術が描く未来～」との関連

- 生成AIの発展は目覚ましいですが、ロボット制御に適用するには圧倒的にデータ量が少ない問題があります。現実的なデータ数で学習するロボットを開発することで、未来社会に貢献できるよう研究を続けております。

開発手法の処理の流れ



実験結果



連絡先: 脳情報研究所 担当 内部英治 E-Mail: uchibe@atr.jp

本研究の成果の一部は国立研究開発法人新エネルギー・産業技術総合開発機構(NEDO)の委託業務(JPNP20006)の結果得られたものです。また本研究の成果は JSPS 科研費 JP21H03527 の助成および JST 未来社会創造事業JPMJMI21B1 の支援を受けたものです。