

2024 年度年次報告書
CRONOS 川原領域
2024 年度採択研究開発代表者

小林 泰介

情報・システム研究機構 国立情報学研究所
助教

データの非規格化を許容するインクルーシブロボット基盤モデル

研究開発体制：

栗田 修平 （情報・システム研究機構 国立情報学研究所 助教）
河原塚 健人 （東京大学 大学院情報理工学系研究科 講師（特任助教））
板寺 駿輝 （産業技術総合研究所 インダストリアル CPS 研究センター 研究員）

1. 研究進捗状況と成果

本研究の目指す「非規格化データを取り扱い可能なインクルーシブロボット基盤モデル」の開発に向けて、2024年度は a) 先行研究の調査・検証および実タスクへの応用, b) 画像から手先動作へのモダリティ変換技術の開発, c) 世界モデルの未来予測性能を向上させる学習理論の導出, d) 多種インターフェイスを切り替え可能な遠隔操作フレームワークの開発, などを中心に研究を進めた. a)に関しては, 現状のロボット基盤モデルの主流と言える Vision-Language-Action (VLA) Model を中心に調査を進め, オープンソース化されている VLA モデルの実装や動作確認を進めている. また, 既存の VLM や LLM モデルを組み合わせることでナビゲーション等の言語指示を要する実タスクへの応用を試みた. b)に関しては, 一人称動画データセットの Ego4D から, カメラ装着者による物体操作を 6DoF 形式でとりだすことで, テキストに対応した物体操作のデータセットを約 3 万軌跡数作成した. c)に関しては, 世界モデルの学習に潜む高次元観測の低次元潜在空間への圧縮・再構成, 潜在空間上でのダイナミクス, 潜在空間の正則化, という多目的性を公平に考慮可能な新たな最適化問題をツァリス統計に基づいて導出し, 学習の安定化や予測性能の改善をシミュレーションで確認した. d)に関しては, ロボット基盤モデルの学習に必要な大量の教示データを多種インターフェイスから収集可能にした上で, モデルの学習からロボットへのロールアウトまでを一貫して行うための人・ロボット協調フレームワーク, OpenHRC を整備・発展させた.

以上