

2025年7月14日

報道機関 各位

国立大学法人東北大学

触覚情報により表裏を判断し ロボットの両手運動を自動生成する AI 技術を開発 — 視覚と触覚情報の統合で人間に近い動きを実現 —

【発表のポイント】

- ロボットがカメラ画像（視覚）に加えて触覚センサーの情報から、環境に応じた両手操作の動作手順を自動生成し運動制御するフレームワークを開発しました。
- 視覚情報のみでは判断が難しい、対象物の質感に基づく操作判断が可能となり、より柔軟で適応性の高いロボット制御が実現します。
- 本成果は、視覚・聴覚・触覚など複数の感覚を統合処理するマルチモーダル AI^(注1)の実現に向けた重要な一歩であり、今後のロボットの高度化・実用化に貢献すると期待されます。

【概要】

近年、人工知能（AI）を支える機械学習を用いて人間の動作パターンを学習し、ロボットが調理や掃除などの日常タスクを自律的に実行する研究が進んでいます。その技術の進歩は日進月歩です。しかし、これらは主に視覚情報に依存しており、素材の質感や表裏の違いなど触覚的な判断を伴う作業には限界がありました。

東北大学大学院工学研究科の Ningquan Gu 大学院生と林部充宏教授は、香港大学および香港の研究機関である Centre for Transformative Garment Production (TransGP) の小菅一弘教授（東北大学名誉教授）との共同研究により、視覚と触覚の情報を統合し、ロボットの両手操作を環境適応的に自動生成するフレームワーク「TactileAloha」を開発しました。これにより、マジックテープ^(注2)やジップタイ^(注3)のような表裏の違いや接着性が重要となる作業においても、ロボットが触って判断し、両手を使って適切に操作することが可能になります。従来の視覚ベースの手法と比較して高いタスク成功率を達成し、実用化への道を大きく前進させました。

本研究成果は、米国電気電子学会が発行するロボット工学の専門誌 IEEE Robotics and Automation Letters に2025年7月2日付けで掲載されました。

【詳細な説明】

研究の背景

近年、機械学習技術を用いて、人間の動作デモンストレーションからロボットが動作パターンを学習し、カメラ画像に基づいて調理や掃除などの日常タスクを実行する研究が進んでいます。米国スタンフォード大学が開発したシステムの ALOHA (A Low-cost Open-source Hardware System for Bimanual Teleoperation) はその代表例で、低コストで汎用的な両腕ロボットの遠隔操作と学習を可能にしました。ハードウェアとソフトウェアがオープンソースとして公開されており、広く研究者に活用されて研究の発展に寄与しています。

しかし、これらの研究はカメラ画像による視覚情報に依存しており、素材の質感や表裏の違いなど触覚的な判断を伴う作業には限界があります。人間が日常的に行う多くの作業が、視覚だけでなく触覚を頼りにした繊細な判断と動作に基づいていることを考えると、ロボットにも同様の感覚が求められています。

今回の取り組み

本研究では、ロボットのグリッパーに触覚センサーを搭載し、カメラ画像と触覚情報の双方を統合的に扱う AI フレームワーク「TactileAloha」を開発しました。視覚と触覚の情報を融合するために視覚—触覚トランスフォーマー^(注4)モデルを採用し、ロボットが環境に応じた両手の動作手順を自律的に生成できるように設計されています。

このフレームワークの有効性を検証するために、マジックテープの固定とジップタイの挿入という二つのタスクを対象としました。マジックテープの操作では、ロボットが触覚情報に基づいて表裏を判別し、正しい面同士を接着するように両手を動かします。ジップタイの操作では、対象物の向きを適切に調整しながら挿入を行う必要があります。これも触覚による質感の判断が重要となります。

実験の結果、提案手法は従来の視覚情報のみに基づく手法と比べて、高いタスク成功率を達成し、触覚情報の導入による明確な利点を示しました。これにより、質感や接触状態に応じて柔軟に動作を変化させる、より人間に近いロボット制御の実現に近づく成果となりました。

今後の展開

今回の成果は、ロボットのためのマルチモーダル AI 開発における重要なステップとなります。人間が日常的に行っている無意識の動作には、視覚や触覚など複数の感覚が連携して働いています。今回開発した技術により、家庭内作業、衣類の取り扱い、リハビリ支援、触覚が必要な産業応用タスクなどさまざまな分野で複数センサー情報を統合した高次判断を伴う作業が可能となり、ロボットの有用性と適応性が飛躍的に向上することが期待されます。

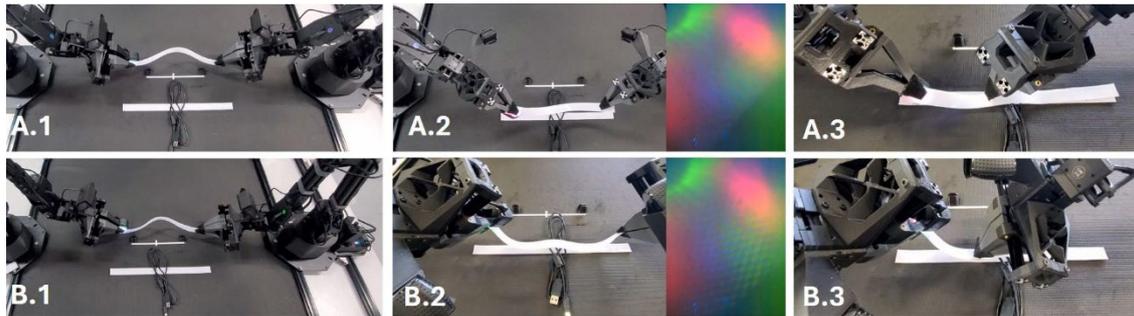


図 1. カメラ情報を基にアームはマジックテープの両端を掴む (A.1、B.1)。触覚情報によりテープの向きを感知し、フック面がループ面と合うように姿勢と角度を調整して位置合わせを行う (A.2、B.2)。マジックテープを固定し、右アームが押し付けて確実に結合させる (A.3、B.3)。与えられたテープの面に適応して、適切な接触面となるよう異なるテープ操作運動が自動生成される。

動画 : <https://www.youtube.com/watch?v=2ohHM6nmqJg>

【謝辞】

本研究は中国香港特別行政区政府創新科技署イノベーション・テクノロジー基金 (the Innovation and Technology Commission of the HKSAR Government under the InnoHK initiative) と東北大学機械科学技術国際共同大学院プログラム (GP-Mech) の助成により実施されました。また、本論文は『東北大学 2025 年度オープンアクセス推進のための APC 支援事業』の支援を受けました。

【用語説明】

- 注1. マルチモーダル AI : マルチモーダル AI は、テキスト・画像・音声・動画など複数の異なるデータ形式を同時に処理・理解する人工知能技術です。従来の単一データ処理と異なり、人間のように複数の感覚情報を統合して総合的な判断を行います。
- 注2. マジックテープ : マジックテープは、表面に微細なフック状突起がある面とループ状繊維がある面を貼り合わせて接着する面ファスナーです。1948 年にスイスで発明され、衣類・靴・バッグなどに広く使用されています。簡単に着脱でき、繰り返し使用可能な便利な接合材料です。
- 注3. ジップタイ : ジップタイ (結束バンド) は、プラスチック製の帯状ストラップで、一方向にのみ締まる仕組みを持つ固定具です。電線の束ね、配管固定、梱包作業などに広く使用され、安価で簡単に使える便利な工業用品として、建設・電気・物流分野で重宝されています。
- 注4. トランスフォーマー : トランスフォーマーは 2017 年に登場したディープラーニングモデルで、アテンション機構により文脈を効率的に理解します。ChatGPT や BERT など現代の大規模言語モデルの基盤技術として、自

然言語処理分野を革新し、翻訳・文章生成・質問応答などで高い性能を発揮します。

【論文情報】

タイトル : TactileAloha: Learning Bimanual Manipulation with Tactile Sensing

著者 : Ningquan Gu, Kazuhiro Kosuge, Mitsuhiro Hayashibe*

*責任著者 : 東北大学 大学院工学研究科 教授 林部充宏

掲載誌 : IEEE Robotics and Automation Letters

DOI : 10.1109/LRA.2025.3585396

URL : <https://ieeexplore.ieee.org/document/11063285>

【問い合わせ先】

(研究に関すること)

東北大学大学院工学研究科 ロボティクス専攻 教授 林部 充宏

TEL: 022-795-6970 Email: hayashibe@tohoku.ac.jp

(報道に関すること)

東北大学大学院工学研究科 情報広報室 担当 沼澤 みどり

TEL: 022-795-5898 Email: eng-pr@grp.tohoku.ac.jp