

現場作業の自律動作能力を拡張するAIロボット技術

#ヘルスケア・QoL向上 #AI・生成AI #ロボティクス

執筆者

野口 直昭 (Noguchi Naoaki)

日立製作所 研究開発グループ コネクティブオートメーションイノベーションセンター 所属
現在、ロボットシステム、メカトロニクスなどの研究マネジメントに従事
博士（工学）
日本ロボット学会会員、日本機械学会会員

山本 健次郎 (Yamamoto Kenjiro)

日立製作所 研究開発グループ コネクティブオートメーションイノベーションセンター ロボティクス研究部 所属
現在、ロボットシステムなどの研究開発に従事
日本ロボット学会会員、日本機械学会会員

伊藤 洋 (Ito Hiroshi)

日立製作所 研究開発グループ コネクティブオートメーションイノベーションセンター ロボティクス研究部 所属
現在、ロボット知能化に関する研究開発に従事
博士（工学）
日本ロボット学会会員、日本機械学会会員、計測自動制御学会会員、人工知能学会会員

一粟 秀行 (Ichiwara Hideyuki)

日立製作所 研究開発グループ 脱炭素エネルギーイノベーションセンター 原子力システム研究部 所属
現在、ロボットシステム、福島第一廃止措置向けの燃料デブリ取り出し技術の研究開発に従事
博士（工学）
日本ロボット学会会員、日本機械学会会員、計測自動制御学会会員

ハイライト

日立は、インフラ、交通、製造分野のDXを推進するとともに、フロントラインワーカーのウェルビーイング向上をめざしている。また近年、生成AIとロボティクスの融合により、これまで困難だった現場作業の自動化に注目が集まっている。そこで、深層予測学習モデルを進化させ、視覚や力覚などのマルチモーダル情報とロボットの動作情報を統合的に学習するAI技術を開発した。さらに、人が操作した動作を模倣学習させるロボット教示について、全方位移動機構と双腕マニピュレータから成る独自のロボットを開発し、その有用性を確認した。

本稿では、これらのロボット作業能力の拡張により、危険作業や定常業務の負荷低減に加え、時間や場所の制約を緩和した人の新たな働き方の実現をめざす日立の取り組みについて述べる。

1. はじめに

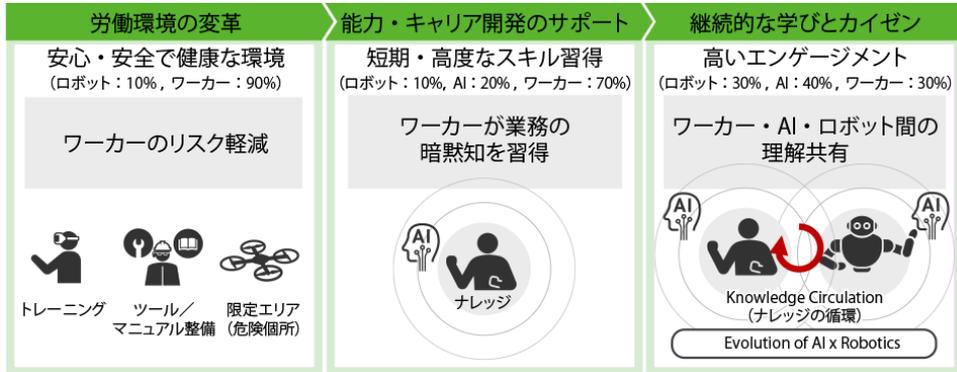
これまで、産業ロボットやサービスロボット、および、これに関する自律、協調、人とのコミュニケーションなどの研究開発が産官学の連携の下で行われてきた。日立は、工場や倉庫などの自動化に加え、都市化、インフラ整備などに関する施工や保守分野で活躍するフィールドロボティクスの開発を進めている¹⁾。一方、これらの事業領域では、現場で働くフロントラインワーカー（以下、「ワーカー」と記す。）の労働力の偏在への対応や、ウェルビーイングの向上も社会課題として顕在化している。現場を支えるワーカーは、状況に応じた認知・判断や作業に関するさまざまな能力・経験を活用して業務を遂行しており、今後はこれらの業務の負担軽減や省人化が求められると同時に、人間が本来やるべきことを追求していく社会が到来すると考えられる。ロボティクスは、この社会を実現するための一つのキーアイテムとして期待されており、昨今では、人間による自然言語指示やシナリオに基づき、生成AI（Artificial Intelligence）を活用して動作を生成するなどの新たな機能も開発されている。

本稿では、先端ロボティクスの社会実装を加速するコア技術について述べる。

2. ロボットの社会実装に向けた課題と取り組み

ロボットの導入ステップ例を図1に示す。ワーカーの危険などのリスクを軽減するために、ロボットが危険作業を代替する領域、生成AIを活用してワーカーの暗黙知を習得して人とロボットが協働する領域、さらに両者が互いに意図や経験を共有することで継続的な学びを得る領域など、いくつかの段階的な導入ステップがある。一方で、さまざまな現場の業務を遂行するワーカーは、安全、品質、環境などの社会要請に対して、五感、コミュニケーション力、思考力、作業力などで対応している。人同士のコミュニケーションをベースにして、思考、五感、作業などの人間が持つ能力を先端技術で拡張することで、前述のロボット導入が加速できると考える。本章では、これらの人間能力の拡張とも関わりが深い、ロボットの自律化や、人とロボットのインタラクション技術について述べる。

図1 | ロボットの導入ステップ例



注: 略語説明 AI (Artificial Intelligence)

ワーカーとロボットの業務分担の割合を例にして、ロボットの導入が業務やワーカーにもたらす変革や、ワーカーとロボットの関係性を示す。

2.1 社会実装に向けた課題と研究ターゲット

ロボットをオープンな環境で社会実装するには、人や周囲とロボットの接触時の安全性、周囲の環境変化に応じたロボット動作の柔軟性、通信途絶や作業失敗時のリカバリー対応など、多岐にわたる検討が必要である。また多くの場合、人の動作をロボットに教示する必要があり、この作業や調整には多くの工数を要する。近年では、ヒューマノイド型のロボットを用いて、さまざまな業務をこなす汎用ロボットの研究も活発化している。

ここでは、ヒューマノイド型ロボット、多関節アーム型ロボット、さらに自由度の少ない移動機械も対象として、人が何気なく無意識に実行している動作のロボットによる自律化、人が現場の状態を見ながら熟慮して実行している動作のロボットへの教示について紹介する。ここに述べる技術は、いずれもリアルタイムでのフィードバックなどを含んだロボットの高度な制御に関するものである。

2.2 複雑な動作を実現するマルチモーダルの活用

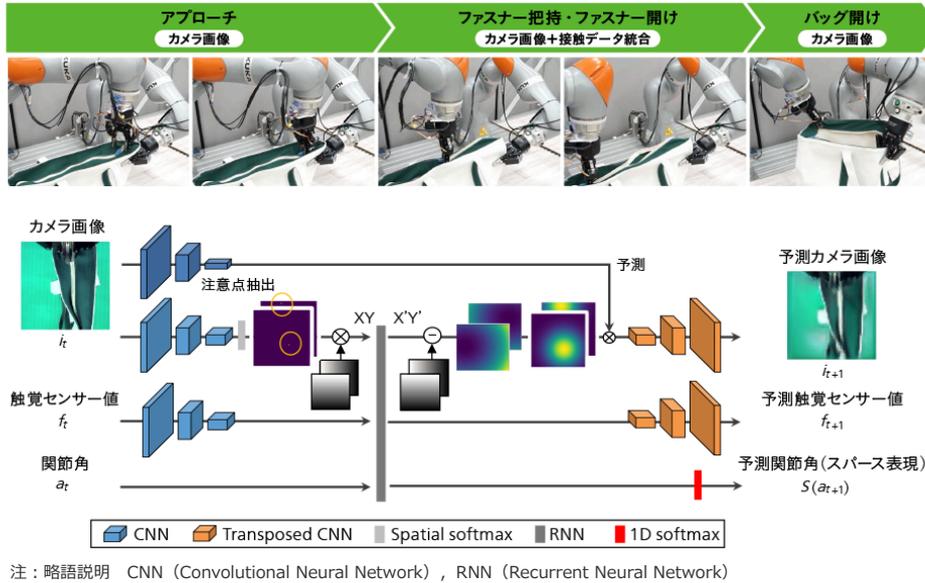
複雑な動作を実現するには、位置や関節情報の指令のみならず、視覚や力覚などのマルチモーダルの情報と動作を基に、近未来の状況を予測し、現実との誤差を最小とるように次の動作を生成することが有効である。特に、マルチモーダルなセンサー情報の重要度を作業の内容や環境に応じてリアルタイムに切り替えることは、後述する学習の効率化の観点でも重要であり、高速なリアルタイム処理を少数の学習データで実行可能であることが求められる。

日立は、これまでに開発した深層予測学習技術²⁾をさらに進化させ、マルチモーダルの知覚情報を複数の時系列ネットワークモデル(深層予測学習モデル)に入力し、注意すべき知覚情報を学習・選択することで、次の取るべき状態に基づく動作指令を逐次的に生成する方式を開発した。なおこの技術は、さまざまな動作を仮想空間で実行するシミュレータの活用によって、さらに学習効果を高めることができる。

2.3 マルチモーダル情報を統合した深層予測学習

視覚や触覚を利用して、変形する布バッグのファスナーを開ける動作の例を図2に示す。同図下部は、それに用いた深層予測学習のモデルである³⁾。その特徴として、カメラ画像の動的な変化に対しても安定したロバストな動作生成を行うために、視覚注意モデルを組み込んでいる。spatial softmax^{※1)}による特徴マップの強調処理と、soft argmax^{※2)}による最大値を有する点の座標抽出により、注意点を画像中のピクセル座標として得る。また、触覚のセンサーデータが画像と類似することに着目し、触覚処理にもCNN (Convolutional Neural Network) を用いた。CNNが画像の位置シフトにロバストであることから、指先(触覚センサー)の位置ずれに対しても同様のロバスト性が期待できる。さらに、連続的な関節角値を複数のニューロンで確率的に表現するスパース表現を採用することで、学習の効率化を図った。

図2 | 布バッグのファスナー開けタスクの例と深層学習モデルの概要



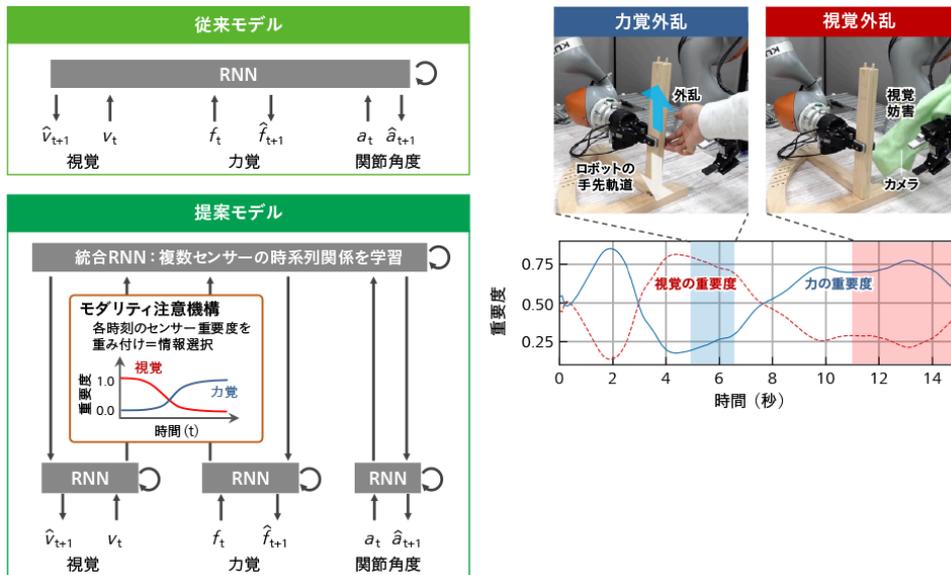
カメラ画像の動的な変化やファスナーを把持した際の触覚センサーの位置ずれに対して、それぞれ視点注意モデルや特徴マップの強調処理などにより学習の効率化を図る。

- ※1) 複数の入力値を0から1の確率値に変換し、複数の出力値の合計が1になるsoftmax関数を、画像などの二次元データに適用可能なように拡張した関数。
- ※2) 最大値のインデックスを返す関数argmax (arguments of the maxima) を微分可能な形式で近似した関数で、ニューラルネットワークの内部演算に利用。

2.4 マルチモーダル統合学習モデルの改良

ロボットは、視野内に遮蔽物が存在する、計画時には存在しなかった力が加わる、といった想定外の事態においても安定的に動作する必要がある。このため、視覚が遮断された場合に力覚情報に注意を向ける機能を実現するため、深層予測学習モデルの改良を行った。従来、モダリティの統合手法としては、入力の複数のモダリティに対し、一つのモジュール内での統合が用いられてきた。この場合、作業上重要でないセンサー情報もロボットの動作決定・実行に必要なとされるため、これらの情報がノイズとなり、ロボット動作のロバスト性が下がるという課題があった。これに対し、生物の脳での感覚の分散・統合処理を参考に、感覚・運動を処理する単一モダリティのRNN (Recurrent Neural Network) を統合・連合する統合RNNを設け、各モダリティの重要度を切り替え可能な構成とした。これをモダリティ注意機構と定義する。この有効性を検証するため、作業を行う手先の位置と力加減の調整が求められる椅子の組み立て作業を行った⁴⁾、⁵⁾ (図3参照)。これにより、部品位置合わせ時は視覚情報、組み立て時は力覚情報の重要度が高いことを確認した。なお、ロボットの作業中の重要度を可視化することで、どのセンサー情報が動作に寄与しているかを解釈することも可能である。

図3 | モダリティ注意機構のコンセプトと椅子組み立てタスクの例



人がロボットに所望の動作を複数回教示するだけで、センサー情報の重要度を学習し、作業の内容や環境に応じて特定のセンサー情報に着目しつつ作業を実行することができる。

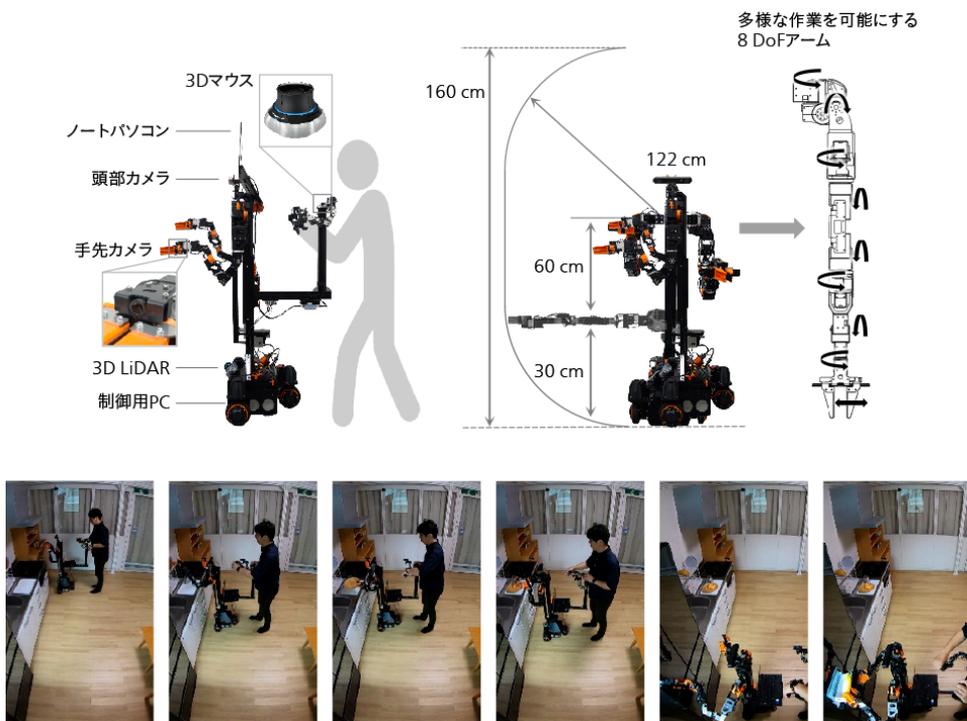
2.5 リアリティギャップに対応する模倣学習

次に、シミュレーションを用いてロボットの動作学習を行う場合に、実世界でのロボットの動作で生じ得る課題について述べる。シミュレーションと実世界には、接触を伴う場合などさまざまな物理量の差異、すなわちリアリティギャップがあり、これを考慮しないと、目標とする動作を完遂できないケースが発生する。

この解決策の一つとして、人がロボットにタスクを教示し、その際に得られるセンサーデータを用いて動作学習する方法がある。従来の教示や学習で獲得した汎化性能は、データの質に大きく依存し、複雑なタスクになるほどデータの質の重要度が高まる。このため、高自由度ながら安価に、かつ直感的な動作教示が可能な独自のロボットを開発した⁶⁾。リーダー・フォロワーの機構を備え、緻密で繊細な作業を少数の学習データで実行可能である。

全方位移動機構の上に、双腕マニピュレータを付けた双腕型移動マニピュレータを図4に示す。本ロボットは、頭部、胴体部、移動部、遠隔操作部から構成される。胴体部は、8 DoF (Degree of Freedom) のアームと1 DoFのグリッパを備えた双腕マニピュレータと昇降機構から成り、移動部には独立3輪ステアリングを用いた。教示の要となる遠隔操作装置は、リーダーアームと、3D (Three Dimensions) マウス、ノートPCで構成され、ロボットのカメラで取得した視覚情報がPCのディスプレイに表示される。人は無意識のうちにロボットの視野外の情報を使って動作教示する傾向があり、その結果、教示データの質が低下する。そのため、本ロボットのように、二人羽織型でロボット視点の画像を見ながら動作教示することで、模倣学習に適切な高品質データの収集が可能となる。この教示手法により、ドアの開閉、台所の物品の取り出しなどのさまざまな動作が短時間で教示可能であり、ロボットによって自律動作が実行できることを確認した。

図4 | 双腕型移動マニピュレータとその動作教示の様子



注：略語説明 3D (Three Dimensions)

人はロボット視点の画像を見ながら、ドアの開閉、台所の物品の取り出しなどの動作を短時間で教示することができる。これにより、ロボットの自律動作が実行できる。

3. おわりに

ロボティクス研究では、ここで紹介した技術以外にも、ヒューマノイドとそれを対象とした汎用基盤モデル、VR (Virtual Reality) を活用した遠隔操作、自然言語とAIを活用した人とロボットの対話、不規則に移動する物体の操作⁷⁾ など、複雑で多様なタスクの実現をめざすさまざまな試みがなされている。本稿では、特に視覚や力覚などのマルチモーダルな情報と動作に注目し、少ないデータで高い学習効率を実現するAI学習基盤と、人がロボットに教示するにあたって模倣学習に適した高品質なデータを収集できるロボットを紹介した。今後、これらのロボットを用いてさまざまな現場でデータを収集することで、エコシステムを形成し、社会実装を加速していく。

謝辞

本稿で述べた深層学習を用いたロボットの動作学習・動作生成技術の開発においては、早稲田大学の尾形 哲也教授にご協力を頂いた。深く感謝の意を表す次第である。

参考文献など

- 1) 山本健次郎, 外: 現場作業のデジタルソリューションを革新するフィールドロボティクス, 日立評論, 103, 2, 254~259 (2021.3)
- 2) H. Ito et al.: Efficient multitask learning with an embodied predictive model for door opening and entry with whole-body control, Science Robotics, Vol 7, Issue 65 (2022.4)
- 3) H. Ichiwara et al.: Contact-Rich Manipulation of a Flexible Object based on Deep Predictive Learning using Vision and Tactility, 2022 International Conference on Robotics and Automation (ICRA) , Philadelphia, PA, USA, pp. 5375-5381 (2022)
- 4) H. Ichiwara et al.: Multimodal Time Series Learning of Robots Based on Distributed and Integrated Modalities: Verification with a Simulator and Actual Robots, 2023 IEEE International Conference on Robotics and Automation (ICRA) (2023.6)
- 5) H. Ichiwara et al.: Modality Attention for Prediction-Based Robot Motion Generation: Improving Interpretability and Robustness of Using Multi-Modality, IEEE Robotics and Automation Letters, Vol. 8, No. 12, pp. 8271-8278 (2023.12)
- 6) 伊藤洋, 外: 直感的な全方向移動とマニピュレーションを実現するロボットシステム 人生活支援を実現するマルチタスク型ロボットの開発, 日本機械学会 ロボティクス・メカトロニクス講演会 (2024.5)
- 7) K. Yamamoto et al.: Real-time Motion Generation and Data Augmentation for Grasping Moving Objects with Dynamic Speed and Position Changes, Proceeding of IEEE/SICE International Symposium on System Integration (SII 2024) (2024.1)

日立評論

日立評論は、イノベーションを通じて社会課題に応える日立グループの取り組みを紹介する技術情報メディアです。

日立評論Webサイトでは、日立の技術者・研究者自身の執筆による論文や、対談やインタビューなどの企画記事、バックナンバーを掲載しています。ぜひご覧ください。

日立評論(日本語) Webサイト

<https://www.hitachihyoron.com/jp/>



Hitachi Review(英語) Webサイト

<https://www.hitachihyoron.com/rev/>



 日立評論メールマガジン

Webサイトにてメールマガジンに登録いただきますと、
記事の公開をはじめ日立評論に関する最新情報をお届けします。