

箱サイズが変わっても再学習が不要な箱詰め作業向け深層強化学習手法

Deep Reinforcement Learning Method for Robots Facilitating Container-Size-Agnostic Picking and Placing Tasks

深層強化学習による把持・箱詰めの同時計画により、ピッキングロボットを用いた箱詰め作業で高効率な行動計画を実現

労働人口の減少を背景に、物流システムでの自動化や効率化が求められています。ピッキングロボットを用いた箱詰めの自動化では、これまで箱詰め計画の最適化が着目されてきましたが、把持・箱詰め計画を同時に最適化することで、更なる効率向上が期待できます。しかし、把持・箱詰めの同時計画は、取り得る行動の選択肢が多く、現場に応じた箱サイズごとの学習に手間が掛かるという問題がありました。

そこで東芝は、箱サイズが変わっても再学習が不要な深層強化学習手法を開発し、高効率な把持・箱詰めの同時計画を実現しました。開発した手法により、物流現場における自動化・効率化の実現への貢献を目指します。

背景

人手不足対策やコスト削減などのため、物流現場ではロボットを活用した作業の自動化が進んでいます。ピッキングロボットを用いた箱詰めもその対象の一つであり、輸送効率を向上させるため、できるだけ箱の充填率(箱の体積のうち、箱詰めしたアイテムが占める割合)を高めることが求められます。

これまでの箱詰めの自動化では、箱詰め計画の最適化だけに着目しており、把持計画は別の作業対象として扱われてきました。しかし、ロボットがアイテムをどのように把持したかによって、箱詰め可能な位置・姿勢が変わるため、両者を同時に最適化することが必要になります。

ロボット動作の学習という観点では、近年、深層強化学習の研究に注目が集まっています(図1)。強化学習は、行動選択のルール(方策)を試行錯誤により学習する手法であり、強化学習に深層学習を組み合わせた、深層強化学習の研究が盛んに行われています。

しかし、把持・箱詰め計画は、取り得る行動の選択肢が多いため、試行錯誤により学習する強化学習の適用が困難という問題がありました。そこで東芝は、箱サイズが変わっても再学習が不要な深層強化学習手法を開発してこの問題を解決し、把持・箱詰めの同時計画を実現しました。

ピッキングロボットによる箱詰め

箱詰めの大まかな流れは、以下のようになります。まず、アイテムが格納された箱(Bin)の中を認識し、どのアイテムのどこをつかむかを決定(把持計画)した後、アイテム把持のためのロボットアームの軌道計算・制御を行います。次に、アイテムをつかんだ後、アイテムを格納する箱(Tote)の中を認識し、どこに配置するかを決定(箱詰め計画)した後、アイテム配置のためのロボットアームの軌道計算・制御を行います。

箱詰めを自動的に計画する手法として、様々な発見の手法が提案されており、例えば、DBL(Deep Bottom Left)法と呼ばれる、できるだけ深く、手前に、左側から詰め込む手法があります。しかし、ロボットがどのようにアイテムを把持したかによって、箱詰め可能な位置や姿勢だけでなく、ロボットアームの軌道が制限されることもあるため、把持と箱詰めに同時に計画しなければ最適な計画とはなりません。また、把持の仕方が箱詰め計画にどのように影響するかは、ロボットシステムの仕様によって異なるため、個別に調整が必要になります。

深層強化学習を用いた把持・箱詰め同時計画

今回、図2に示すように、把持・箱詰め計画を、Tote

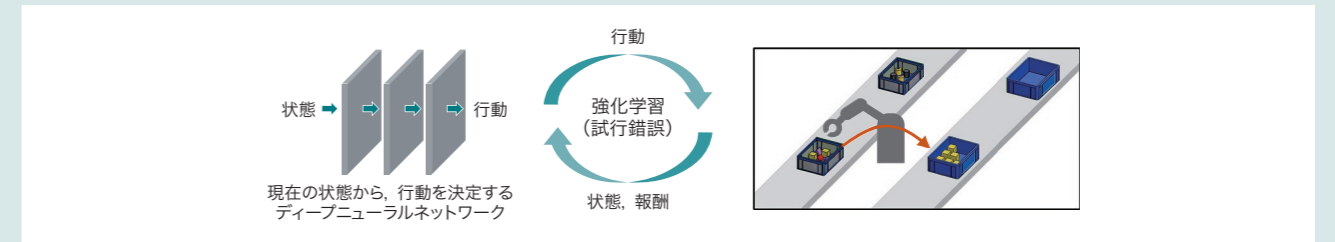


図1. 深層強化学習によるロボット動作の学習

試行錯誤を通じて、ロボットの動作を自律的に学習できます。

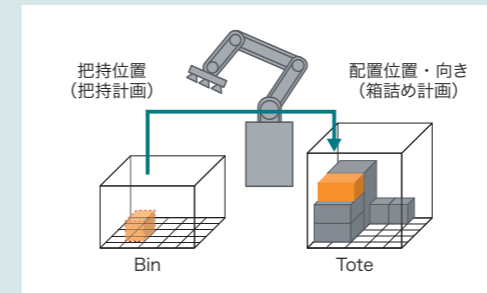


図2. 把持・箱詰めの同時計画問題

把持位置及び配置位置・向きを繰り返し行い、Toteの充填率を最大化することを目的とします。

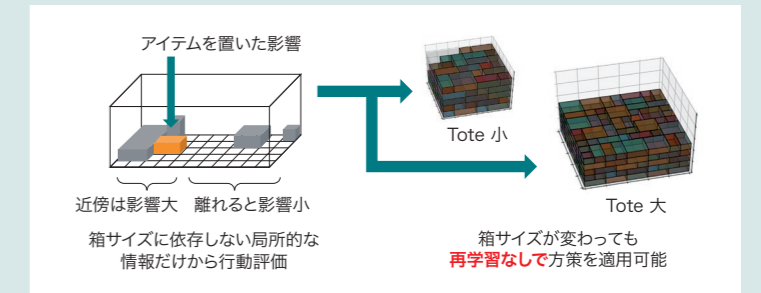


図3. 箱サイズが変わっても再学習不要な深層強化学習

局所的な情報だけに基づいて行動の善しあしを判断することで、小さな箱サイズで学習させた把持・箱詰め計画を、より大きな箱サイズへ再学習なしで適用できるようになりました。

の充填率を最大化する問題として新たに定義し、Bin及びToteを上から撮影したデプス(奥行き情報)画像から把持位置及び配置位置・向きを決定し、充填率を最大化する強化学習アルゴリズムを開発しました⁽¹⁾。

強化学習を用いた手法としては、箱詰め計画だけを学習する手法⁽²⁾が提案されていますが、箱サイズごとに方策の学習が必要になるという問題がありました。また、強化学習は、試行錯誤により学習するため、把持・箱詰め計画のように行動の選択肢が多い場合に問題があります。例えば、縦横30 cmの箱を1 cm幅でグリッド化した場合に相当する、箱サイズ30×30グリッドといった現実的な条件でも、計算量が膨大になるという問題があるため学習が困難でした。

開発した手法では、アイテムを把持・配置する行動を評価する際、箱サイズに依存しないように、局所的な情報だけに基づいて行動の善しあしを判断するようにしました。図3左側に示すように、ある位置にアイテムを置く行動を評価する場合、その位置に近い場所の箱詰め状態は、評価に大きな影響を及ぼしますが、そこから離れた場所の状態は、余り影響を及ぼさないと考えられます。このような近似計算を、局所的な情報を処理する畳み込み層だけから構成されるディープニューラルネットワークで実現しました。その結果、小さな箱サイズで学習させた把持・箱詰め計画を、より大きな箱サイズへ再学習なしで適用できるようになりました。

簡易シミュレーターを用いた定量評価により、開発した手法で得られる充填率は、把持と箱詰めに独立に計画するDBL法よりも高いことを確認しました。また、従来の強化学習を用いた手法では、計算量の問題で学習が困難であった箱サイズに対しても、把持・箱詰め計画が実現できるようになりました。

今後の展望

箱サイズが変わっても再学習が不要な深層強化学習手法の開発により、ピッキングロボットを用いた箱詰めの自動化に向けて、前進したと考えられます。今後は、物流現場における自動化・効率化の実現に貢献することを目指し、開発した手法を適用したロボットシステム実機での検証を進めていきます。

文献

- (1) Tanaka, T. et al. "Simultaneous Planning for Item Picking and Placing by Deep Reinforcement Learning". 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2020). Las Vegas, NV, 2020-10, IEEE, 2020, p.9705-9711.
- (2) Kundu, O. et al. "Deep-Pack: A Vision-Based 2D Online Bin Packing Algorithm with Deep Reinforcement Learning". 2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN 2019). New Delhi, India, 2019-10, IEEE, 2019, p.417-423.

田中 達也

研究開発センター 知能化システム研究所 メディアAIラボラトリー