解 説

人工知能技術のロボット産業応用

Artificial Intelligence Technology for Industrial Robot Applications

比 戸 将 平* *株式会社 Preferred Networks Shohei Hido* *Preferred Networks, Inc.

1. はじめに

人工知能はもはや単なるブームではなく、その技術はすでに様々な実問題に適用されつつある。応用先には産業用機械や自動車も含まれており、人工知能技術とロボティクスとの融合が進んでいる。その最先端では、近年の人工知能ブームの中核である深層学習によるロボットの認識技術向上、およびロボティクスと関連が深い強化学習と組み合わせた深層強化学習の物理世界における適用が大きな役割を果たしている。

株式会社 Preferred Networks は2014年に設立されて以来,人工知能技術の産業応用に注力してきた.一般に最も知られているのはオープンソースの深層学習フレームワーク Chainer [1] の開発であるが,それ以外にもトヨタ自動車やファナックなどの世界を代表する企業と協業し,製品やサービスへの展開を進めている.その研究開発には実問題を扱うならではの難しさがあり,単純にアルゴリズムを適用するだけではなく,様々な工夫が必要となる.

本記事ではまずロボット産業応用で重要となる人工知能技術の要素について紹介した後、主に精度向上した認識技術に基づくピッキングロボット最適化の事例について述べる。また、Google DeepMind 社が提案し AlphaGo [2] でも利用された深層強化学習について簡単に紹介したあと、それを用いたラジコンカーとドローンの制御に用いたデモについて触れる。

2. 人工知能技術の産業応用

人工知能を用いるロボティクス応用の例として,まず産業用ロボットを取り上げる.これまで数十年にわたるロボットの産業利用は主に定型的な作業の代行であった.作業工程が厳密に決められたタスクには人間よりも確実に,高速に,疲労することなく作業を繰り返すことができるロボットを割り当て,それ以外の不確実なタスクは人間が行うと

いう分業のもとで、安全性や生産効率の向上が図られてきた。状況に合わせて制御を変更する場合も、事前に適切に設計された系におけるセンサ値に基づくフィードバック制御か、既知の条件に当てはめるルール処理が主流であった。それ以外の不確実性・多様性が大きすぎる状況や、未知の現象が発生する場合、あるいは機械には難しいパターン認識を含む作業は依然人間の担当領域であった。

例として、ピッキングロボットを挙げる、ピッキングロ ボットは多関節からなるアームを持ち、さらにその先に様々 な物体をつかむためのハンド機構を持つ. 目的のタスクと しては、物体をつかむ、持ち上げる、支える、などがあり、 そのために複数種類のセンサが搭載されていることを仮定 する. もしピッキング対象が整然と並べられ常に同一の配 置となっていれば、アームとハンドを毎回同じパターンで 動かすようにプログラムすれば良い。しかしそのためには 事前に部品の山を人手によって並べる必要があり、現実的 ではない. そうではなく同一種類のアイテムを乱雑に積ん だ状態からつかみ出すタスクはバラ積みピッキングと言わ れ、ロボットの固定的な動きでは対応することが難しい. 各 アイテムが異なる回転方向で積まれており、一般的な二指 ハンドでつかむには方向を推定したうえで進入角度を決め る必要がある. 真上から見るとアイテム同士が重なり合っ ている(オクルージョン)ため、アイテムの検出と角度の 推定は簡単ではなく、また効率を考えれば可能な限り上に 積まれたアイテムから取り出す方が最適である. バラ積み では無数の状態が存在するため、箱上部から撮影した画像 センサからのルール処理だけで対応することは難しい. こ のように機械が不得意としてきた不確実性の高いタスクこ そが、人工知能技術との組み合わせによってロボット産業 活用が広がっていく対象である. 製造業以外にも, 自動運 転車の例で言えば、車載カメラなどを用いて周辺の車や歩 行者を認識し、衝突を回避しながら運転行動を取ることが 求められる. こちらのタスクも目の前の車道の状況は時々 刻々と変わるため、人間の運転手であれば難なく危険回避 できるようなところでも、機械にとっては自明ではない.

原稿受付 2017年2月6日

 $\div - \mathcal{I} - \mathcal{V}$: AI, Deep Learning, Deep Reinforcement Learnig, Chainer

^{*〒 100-0004} 千代田区大手町 1-6-1

^{*}Chiyoda-ku, Tokyo

3. パターン認識技術の進化

複雑な実環境でロボットが目的のタスクを実行するためには三つのプロセスが必要となる。一つめが現在の自分と目的物、環境の状態を知覚する「認識」、その上で目的を達成するための「判断」、そして実際の「行動」である。特に多くのロボットのタスクではまず周辺の物体の検出や分類、姿勢や動きの推定が必要となる。認識を誤るとそのあとの判断と行動をいかに正確にしたとしても求める結果は得られないため、認識精度の向上は極めて重要である。そして、深層学習の登場により特にこの認識の部分が最近5年間で急激に進化している。〈深層学習による画像認識〉にもあるように画像認識がその中心である一方、LIDAR やミリ波レーダなどの深度センサの高度化・低価格化に合わせて、三次元状態を認識する深層学習をはじめとする技術が日々発展を続けており、タスクによっては人間の認識精度を超えつつある。

4. 認識技術ベースの事例

4.1 バラ積みロボット

バラ積みピッキングロボットは、箱の中に乱雑に積まれた対象のアイテムを正しく持ち上げ、コンベアなど後段の処理に流すものである。今回用いたのは FANUC 製品で、アームには3軸の自由度があり、ハンドにはバキュームが付いており、重心付近を吸い付けることで、様々な形のアイテムを持ち上げることできる。センサとしては深度付きカメラが箱上部に固定されており、そこから RGB-D の4チャネル画像情報が得られる。また、ピッキングを施行したときにバキューム内の圧力を測ることにより、アイテムの吸着に成功したかどうかを判断することができる。タスク実行としてはまず深度付き画像から箱の中の状況を把握し、どのポイントをピックするかを選択(判断)し、実際にアームを動かして吸着を試す(行動)、ピッキングポイント

を決めるとアイテム表面における法線方向が推定され、バキュームの吸着角度は自動的に決定されるものとすると、この問題は深度付き画像内で最も吸着可能性が高い点を選ぶという問題に帰着することができる。ここでは教師付き機械学習として扱い、各ポイントの成功/不成功を予測する。ランダムなポイントへのピックを繰り返すことで、その周辺局所画像と吸着の成功不成功から、正解情報データを蓄積できる。対象アイテムとして円柱型の金属を用い、2時間弱かけて1,000回の試行で得られた正解データから CNNを用いて学習したモデルの予測結果を図1(a)に示す。

明るく色づけされた部分が成功確率が高いと予測された 領域であるが、おおよそ円柱側面の中心付近を捕らえられ ていることが分かるものの、一部で円柱同士の境界など、お そらく失敗するであろう予測結果も含まれている. 図1(b) はさらに8時間までデータ収集時間を延長し、合計5.000サ ンプルを用いて学習した改良版のモデルである. 予測結果 を見ると先ほどの境界付近の予測はなくなるとともに、円 形の側面でも一部成功するという予測が得られている. 実 際、このモデルでは約90%という、熟練者が数日から1週 間かけて既存のルールシステムを最適化した場合とほぼ同 等の精度が得られている.また、機械学習ベースの強みと して、対象アイテムによらず同じデータ収集と学習の仕組 みを用いることができるため、形や大きさや材質が異なる 様々なアイテムについて熟練者なしに正しく動作するロボッ トをすぐに稼働させることができるという大きなメリット がある.

4.2 Amazon Picking Challenge

ロボットのピッキング技術のさらなる産業応用を目指す取り組みとして Amazon Picking Challenge (APC) を取り上げる.「アマゾン」のみが主催するコンテストで,第2回となる2016年は16チームが世界各国のロボット関連の研究室やスタートアップが参加しドイツ・ライプツィヒで7月に開かれた. コンテスト専用の棚前に置かれたロボット

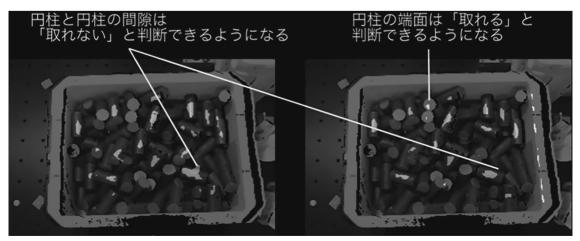


図1 (a) 1,000 サンプル学習後の予測, (b) 5,000 サンプル学習後の予測

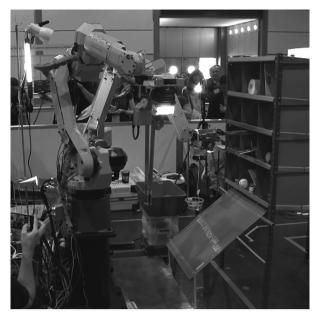


図2 APC 本番の様子

で競うタスクは2部門, Stowing と Picking に分かれてお り、それぞれ箱に入れられた様々なアイテムを棚に収納す ることと、逆に棚から取り出して箱に入れることを目的と していた. 図2は本番の様子である. アイテムはペットボ トルやボール,本,ダンベルなど様々な形状・材質/重さ のものが混ざっており、タスクに成功した場合の獲得ポイ ントにも傾斜が付けられている. 15分のチャレンジ時間の 間は人間側の操作は一切許されず、物体の認識から、ピッ ク対象の選択、それを棚に置く/取り出す作業を全自動で 繰り返す必要がある. ピック失敗によるアイテムの落下や アームの棚への衝突は減点があり、各チームは制限時間で より多くのポイントが得られるよう、本番と同じ棚とアイ テムを用いて約3ヶ月間で開発したシステムを持ち込んだ. ロボットの種類などは重さや接地面積以外はかなり自由度 があり、ロボット研究でよく用いられる RethinkRobotics 社の Baxter から、FANUC や YASKAWA などの産業用 ロボットを持ち込んだチームがあった. PFN の結成間も ないロボットチームは16チーム中、Stowing部門で4位、 Picking 部門で1位と同点の2位と、健闘した、上位チーム の傾向としては、2015年とは異なり、深層学習ベースのビ ジョンシステムを持ち込んで精度の高い物体認識を行うこ とと、産業用ロボットの精度の高い動きを組み合わせるこ とが決め手となったことがある. 今回の結果を見ても、棚 における一般物体の取扱いは依然人間に分があるタスクで あり、ロボットによる倉庫内物流の自動化に向けては技術 的進歩がまだまだ必要ではあるが、その最前線の認識技術 としては深層学習がデファクトとなりつつある.



図 3 APC 参加メンバー

5. 深層強化学習の登場

これまでの事例では、主に認識と判断において、教師付 き学習の設定での深層学習の応用について述べてきた. 一 方, 最後の行動の部分においては, 目的の動作を決定する 部分は依然ルールベース等で決められていた. 例えばバラ 積みロボットであれば、ピッキングを行うポイントを指定 すればあとはアームをどう動かし、バキュームハンドを当 てるかは FANUC のロボット任せであった. しかしながら、 最適な行動のパターンすらも学習によって獲得することが 考えられる. その場合は教師あり学習に変わって、強化学 習の枠組みが用いられることになり、特に深層学習と組み 合わせた深層強化学習に注目が集まっている.強化学習は 外部環境を観測し、アクションを行い、その結果の良し悪 しに応じて報酬を受け取るエージェントの視点から定式化 される. エージェントの目標は、累積報酬を最大化するよ うなアクションの決め方, すなわち最適ポリシーを獲得す ることである.強化学習はアクションによって刻々と変わ る状況(環境)に対する連続的意思決定であり、マルコフ決 定プロセスの一種と見なすことができる. このアクション の部分にロボットのタスクにおける行動を位置づけること で、環境の観測からアクションを決めるプロセス全体を強 化学習の中で扱うことが可能となる。強化学習の利点とし ては, 適切な報酬関数さえ設計すれば逐次的なアクション 選択の最適化が可能であるため、状況の変化に対応しやす いことが挙げられる.強化学習の問題を解くアプローチと して、状態が与えられたときのアクションの条件付き価値 関数を現すQ 関数を推定するものがあるが、これまで問題 ごとに設計して推定されていた Q 関数を, ニューラルネッ トワークを用いてブラックボックスとして扱うことで、特 に環境の観測が画像情報として与えられる場合に、深層学 習における画像認識の精度向上をそのまま利用でき、様々 なタスクを強化学習として表現して効率的に解くことが可 能となった. この DQN (Deep Q-Network) を提案したの は DeepMind という英国のスタートアップであり、オリジ

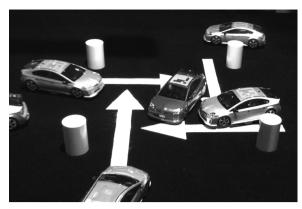


図 4 CES 2016 における深層強化学習によるデモ

ナルの論文[3] においては Atari のビデオゲームを用いて 人間以上のスコアを達成したことを紹介していた. その後. このスタートアップが Google に買収され、2016年囲碁で 人間のトップ選手を破った AlphaGo を作り出したことは 記憶に新しいが、AlphaGoの中でも深層強化学習が用いら れていた[2]. またゲームだけではなくロボットへの応用も 盛んとなりつつあり、物体のピッキングやドアノブの操作 など、人間の手による操作に近いタスクをロボットに獲得 させるような論文が発表されている[4]. ここで強化学習の もう一つの強みとしては、シミュレータとの相性が良いこ とが挙げられる.一般に、ロボット実機を用いた実験はコ ストが高くまた時間が掛かるため、長時間にわたって行う ことが難しい. しかしながら, もし 3DCG と物理演算エ ンジンを用いたシミュレータが用意できれば、まずシミュ レータで強化学習モデルを構築し、実機に適用するという 組み合わせが考えられる.

6. 深層強化学習の事例

6.1 自動運転シミュレータ

深層強化学習を用いた制御の事例として、自動運転を取り上げる。図4はラスベガスで2016年1月に行われたCESのトヨタ自動車ブースでの我々の展示である。正方形のコースの中のミニチュアの車数台が所狭しと走り回っている。銀色の車はすべて自動制御であり、ポイントはその運転ロジックがゼロから強化学習で得られたことにある。学習はまずシミュレータを用いて行われる。シミュレータの中で各エージェント(車)はおよそ300種類のセンサ情報を持つ。例えば360°周囲に発生させた二次元LIDARによる壁あるいは他車への距離情報等である。それらの観測に元に、各時刻において車は運転アクションを決定する。アクションの種類としては直進、左折、右折、ブレーキ、後進がある。強化学習の報酬としては、前方に進むことができれば速度に応じてプラスの報酬、もし壁や他車に衝突した場合は大きなマイナスの報酬が与えられる。各センサの意味のようなも

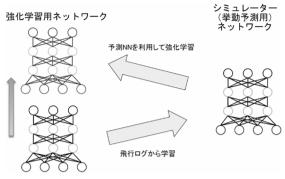


図5 スパイラル学習のイメージ

のは与えられず、すべてが等価の数値情報として与えられ ているため、車はどの変数が前方向でどの変数が後ろ方向 なのかすら最初は与えられていない。状況の評価には実際 に DQN ベースの手法を用いている。学習の最初のフェー ズでは、運転ロジックをまったく獲得していないため、初期 位置からの微小な前進後進や方向転換を繰り返しているだ けである. その中から、やがて前進による報酬獲得を優先 するようになる. ただしまだその動きはぎこちなく、壁に ぶつかったあとの復帰が難しい状況である. さらに学習を 続けると最終的にはスムーズにコースを走るポリシーが得 られる. また. 他車が眼前に現れたときにも. どちらにハ ンドルを切れば、あるいは後進すれば衝突を避けられるか も、まるで目が付いていて、人間のドライバーが判断して いるような動きを見せるようになる. このシミュレータに よるモデルをミニチュア車の制御に用い、二次元バーコー ドを用いて位置推定と疑似センサ情報の生成を行いながら, 実機に起因するキャリブレーションを行うことで、上記の CES でのデモが行えるようになった。さらに人間がリモー トコントロールする赤い車を投入し、たとえその車が危険 な動きをとったとしても, 周囲の自動運転車が対応して回 避行動をとることで、衝突がかなり防げることを示した. も ちろん実際の自動運転車がこの強化学習ベースの手法のみ で得られるわけではないが、自動運転でも多様な状況にお ける判断と行動の最適化は課題であり、ルールや教師あり 学習と組み合わせた精度と安全の向上に強化学習を用いる 可能性が示された.

6.2 ドローン制御のためのスパイラル学習

次に、深層強化学習をドローンに用いた例を述べる。ドローンは物流や外観検査などの産業応用に向けて開発が盛んに行われているが、その安定飛行の制御は容易ではない。例えば、機体の重量バランスの変化や、プロペラの損傷、風向きの変化などがあるととたんに空中に静止してることすら難しくなる。そこで、PFNではドローンの安定的な飛行方法の動的な獲得を目指して、深層強化学習による制御を試み、2016年10月のCEATEC JAPANでデモを展示し



図 6 CEATEC JAPAN 2016 における自律飛行ドローンデモ

た. ドローンには、仮想の強化学習タスクとして、空中に 現れるバーチャルなフラグに触れて得られる報酬を最大化 する課題を与えた. しかしながら. 突如目の前にバーチャ ルなフェンスが出現することもあり、フェンスに接触する と大きなマイナス報酬となる。そこで、フェンスを避けな がらフラグを回収し続ける自律制御を目指し、ゼロの状態 からの飛行学習を行った. まず初期状態では浮上すること すら困難であるため、シミュレータでの浮上方法の獲得か らスタートすることになる. ただしドローン独自の課題と して、わずかな制御の差異によって墜落してしまう不安定 さがあった. そのため、車の例のようにシミュレータで学 習を終えたモデルを実機に適用してもまともに動かないと いう問題があった. そこで PFN では図5に示したスパイ ラル学習と呼ぶ仕組みを導入し、シミュレータと実機の間 で交互に学習を繰り返すことで、シミュレータで学習した 行動予測ニューラルネットワークの誤差に起因する不安定 さを回避しつつ、徐々に性能を向上させるアプローチをとっ た. その結果、CEATEC本番ではほぼ障害もなく飛び続け てタスクをこなすドローンを展示することができた. 図6 は実際の飛行時の軌跡を頭上から撮影したものである. 最 初左上に存在するドローンが、左下のフラグ(プラスの報 酬)を取りに行くにあたって、間の赤いバー(マイナスの 報酬)を避けるために右側に回り込む経路を選択している 様子が分かる. このように様々なロボットの操作方法をシ ミュレータと深層強化学習によって獲得する技術は、Deep Mind 等もピッキングタスクで取り組んでおり [5], 深層学 習そのものの手法の改善との相乗効果により産業応用の可 能性が大きく広がっていくものと考えられる.

7. ま と め

本稿では、深層学習を中心とする最新の人工知能技術が、 ロボット関連の産業用途でどのようにして応用されるかに ついて紹介した、深層学習が威力を発揮する画像を中心と する認識の部分ではすでに積極的な導入が進んでおり、商 用利用が始まっている. また. ロボットの制御まで含めて 人工知能技術を適用する動きも, ここ数年の深層強化学習 の進歩と成功に従って広がっており、すでに様々なタスク において実機を用いたデモが出揃い今後実際の製品に応用 されることは間違いない. まずはバラ積みロボットなど人 工知能技術の役割が明確で比較的導入が容易な対象から始 まり、今後はさらに複雑なプロセス、また異なる種類のロ ボットが同じ製造ライン上で協調するなどメタレベルでの 学習と最適化においても人工知能技術が重要な役割を果た すと考えられる.しかしながら、訓練データの収集コスト、 実機の制御の難しさ、シミュレータと実機の差など、解決 の難しい課題は残っており、機械学習をはじめとする人工 知能技術だけですべてが魔法のように解けるわけではない. そこでは、機械学習をはじめとする人工知能技術の利点を 最大限活かすようなロボットの設計と、そのロボットのメカ ニズムやダイナミクスに合わせた人工知能技術の適用方法 の双方を進めることが産業応用を進める両輪として必要で ある. そのためには両方の分野でアカデミアとインダスト リとの協業を深めることで、理論的な取り組みを進め、実 機でのテストを行い、その結果をフィードバックしてアル ゴリズムの改良を行うというループを回すことが重要にな るだろう.

参考文献

- [1] S. Tokui, K. Oono, S. Hido and J. Clayton: "Chainer: a Next-Generation Open Source Framework for Deep Learning," NIPS Workshop on Machine Learning Systems, http://chftlncr.org/, 2015
- [2] D. Silver, A. Huang, C.J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel and D. Hassabis: "Mastering the Game of Go with Deep Neural Networks and Tree Search," Nature, no.529, pp.484-489, 2016.
- [3] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra and M. Riedmiller: "Playing atari with deep reinforcement learning," NIPS Deep Learning Workshop, 2013.
- [4] S. Gu, E. Holly, T. Lillicrap and S. Levine: "Deep Reinforcement Learning for Robotic Manipulation with Asynchronous Off-Policy Updates," Proceedings of ICRA, 2017.
- [5] A.A. Rusu, M. Vecerik, T. Rothorl, N. Heess, R. Pascanu and R. Hadsell: Sim-to-Real Robot Learning from Pixels with Progressive Nets, http://arxiv.org/abs/1610.04266, 2016.



年1月より現職.

Ī

比戸将平(Shohei Hido)

米国 Preferred Networks America, Inc., Chief Research Officer. 2006 年京都大学大学院情報学研究科修士課程修了. 同年日本アイ・ビー・エム東京基礎研究所入所. データマイニングと機械学習の研究開発に従事. 2012 年株式会社 Preferred Infrastructure 入社. 2015

JRSJ Vol. 35 No. 3 —18— Apr., 2017