

### 1. 担当 PM

田中 邦裕（さくらインターネット株式会社 代表取締役社長）

### 2. クリエータ氏名

村松 直哉（筑波大学，ピクシーダストテクノロジーズ株式会社）

### 3. 委託金支払額

2,304,000 円

### 4. テーマ名

機械学習を用いたロボット制御のための汎用システムの開発

### 5. 関連 Web サイト

なし

### 6. テーマ概要

近年、深層機械学習（ディープラーニング）の普及によって、その適用範囲が広がってきた。ロボットの分野においても、フィードバック制御や、動作を事前にプログラミングするだけでなく、ロボットが動作した結果に報酬を与え、強化学習によって動作を決定するというアプローチが広がっている。しかしながら、ロボットの部品故障などのハードウェア条件が変わった際には再学習を行う必要があり、学習を経ないままでは動作を継続させることができず、かといって再学習には時間がかかり、現実的には物理的な修理を必要としている。そのような中で、CPG（Central Pattern Generator）と呼ばれる、脊椎動物固有の反復運動を模した学習パターンを取り込み、再学習にかかる時間を大幅に短縮するというアプローチをとったのが、本プロジェクトである。

### 7. 採択理由

ディープラーニングが普及してきた昨今、ロボットの制御への活用も進んでいるが、全ての制御を大型のコンピュータで計算するのは現実的ではない。だからといって、従来のフィードバック制御では細かな制御を行うのは難しい。その

ような中で、脊椎動物の持つ機能を模倣した制御アルゴリズムをつくり、より短時間でさまざまなロボットを稼働させることを目指す本プロジェクトは、意義深かった。ディープラーニングだけで制御するロボットが脳だけで制御されるものだとすると、今回のプロジェクトは脳と協調して脊椎での自律を実現するものであり、実用現場における自律制御をより発展させることを期待して採択した。

## 8. 開発目標

本プロジェクトにおいては、シミュレータ環境において CPG を実装した仮想のロボットが歩行するまで強化学習を行い、歩行を達成した学習データをもとに、実際のロボットを動かすという提案であった。人間においても、身体部位に支障をきたした際に、頭の中で考えて歩行のイメージを持ったうえで、支障のない身体部位を使用して、歩行を獲得することができるが、実際には脊椎に存在する CPG における歩行の基本パターンが支援することにより、比較的短時間の間に、歩行を獲得することができる。

本プロジェクトにおいても、強化学習において同様の仕組みを込むことにより、比較的短時間で再学習を行うことを目標とした。

## 9. 進捗概要

本プロジェクトでは、ロボット自体の状態変化に対応するための歩容制御システムを開発した。本プロジェクトの特徴は、ロボットが故障しても、変化に対応した歩容ができることであり、成果として、6脚ロボットの複数の脚を故障させても、十分な速さで前に進むことができた。

本プロジェクトにおいて開発を目指したものは、大きく3つある。

- ・ 実機ロボットの学習システム
- ・ シミュレータにおける学習システム
- ・ 故障に対応する歩容制御システム

使用機材の中で主だった主要機材を示す。ただし、4脚ロボットはプロトタイプ作成時に使用し、最終的なシステム動作検証には6脚ロボットを利用している。

- ・ 4脚ロボット：Lynxmotion SQ3U 対称四足歩行ロボット
- ・ 6脚ロボット：PhantomX AX Metal Hexapod Mark III Kit (AX-12A Servos)
- ・ カメラモジュール：Raspberry Pi Camera Module V2
- ・ シングルボードコンピュータ：Raspberry Pi 3 Model B+

本プロジェクトにおいては、シミュレータにおける学習結果をもとに実機ロボットを動作させることには、研究すべき課題も多く、実機をベースに歩容制御アルゴリズムを開発した。最終的な歩容制御システムでは、強化学習を基にした制御アルゴリズムを利用している。強化学習による動作獲得のアプローチでは、ロボットが試行錯誤を繰り返すことによって動作を獲得する。本システムは、シミュレータでの学習をしないこと、脚の可動域を制限することなどにより、学習結果の動きがより予測可能であり、突拍子もない歩行方法を避けることができている。

以上のように、本プロジェクトでは、従来の強化学習手法を、より現実的な現場に合わせたシステムとなった。

## 10. プロジェクト評価

本プロジェクトにおいて当初想定していた内容は、シミュレータ上でロボットの動きを学習させたうえで、その学習結果をもとに実機を稼働させるための汎用システムであった。しかし、実装には紆余曲折があり、プロジェクトの途上においては幾多の困難を乗り越える必要があった。

まず、当初想定していたCPGについて、深層強化学習を行う際にニューラルネットワークへCPGを導入することで、強化学習における探索範囲を狭めることができ、効果的な学習を得られるという成果を認めることが出来たものの、CPGの導入によるロボットの動作精度の向上の効果は限定的であることが分かった。そのため、MMPRL (Map-based Multi-Policy Reinforcement Learning) と呼ばれる別の強化学習手法を導入したが、実機における検証においては想定される性能を見いだせず、実際にロボットをシミュレータ上で深層強化学習を行い、その結果で稼働させることの難しさを垣間見ることとなった。

このプロジェクトにおいては、シミュレータ上でCPGを使った強化学習を行い、その結果をもとに実機を稼働させるという事が未踏性として認識していたが、シミュレーション上と実機とのギャップを埋めることは予想以上に難易度が高いものであった。そのため、シミュレータを使わずに、実機をQ-learningと呼ばれる強化学習手法で稼働させるという手法を選択することになったが、結果としてはシミュレータを使わなくとも想定される動作を獲得することができ、実際にロボットの脚を故意に故障させた状態であっても、歩行を獲得するという成果を達成することができた。成果報告会においても、デモを通じてその成果を説明し、ロボットの脚が壊れた状態でも稼働する事を証明できたことは高い成果である。

## 11. 今後の課題

今回の実装では、シミュレーション上の学習結果をもとに、実機を動かすまでには至らなかった。将来的には、実機での学習ではコストが高いため、よりコス

トを低減化させるためにも、シミュレーション上での学習が欠かせない。今後、シミュレーション上の学習データを、いかに実機に適応するか、加えていかに汎用的に利用できる環境を作れるかが課題である。