

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

— 故障に強いロボット制御 —

1. 背景

現在、ロボットは様々な現場で利用し、今後も大きな活用が期待されている。工場内ではロボットアームによる部品組立などが行われている。一般家庭向けにも、お掃除ロボットが販売されている。しかし、これらのロボットは、環境の変化に対応することができない。ロボットアームのような、移動を想定しないロボットであれば、環境が大きく変化するような事態には陥り難い。広範囲の移動を想定したロボットの場合であれば、路面状況など外部環境が場所によって異なる。

これに対応するために作られたのが、脚を使った移動ロボットである。人間やその他多くの動物のように脚を使った移動ならば、多少の段差も乗り越えられ、砂利道、芝生、山岳地形など多様な路面状況に対応した移動が可能である。

脚を持つロボットの欠点として、可動部が多いため機械的に壊れやすいことが挙げられる。各関節は車輪のように、一方向に継続的な動きが続かないため、アクチュエータなどに非常に負荷がかかる。つまり、壊れやすいロボットである。

現在利用されているロボットの多くのプログラムは、ロボットの状態が正常である前提で書かれている。また、ロボットの形状も完全に決められているため、似た形状のロボットであっても、異なった制御プログラムを用意する必要がある。

異なる形状のロボットを似た制御プログラムにより操作可能で、2つの利点がある。1つは、ロボットの開発が簡単になることだ。統一されたプログラムにより、ロボットのマイナーチェンジや似た構成のロボットの開発が行いやすくなる。2つ目は、ロボットの故障に対しても対応できることである。6足歩行ロボットの場合、脚が一本壊れたならば、5足歩行ロボットを制御していると見ることができる。故障したその場で、制御対象のロボットが変わったものとし、再調節を自動的に行うことで、行動可能な状態にできる。

2. 目的

ロボット形状の変化を自動的に吸収するシステムを開発することにより、ロボットの動作プログラム開発が容易になるだけでなく、故障に対応したロボットの開発も可能になる。

本プロジェクトでは、ロボットの形状変化に素早く対応する制御システムの開発を目的とした。実際のロボットを用いて、形状変化に対してより妥当性の高い歩容(歩行パターン)の獲得ができるアルゴリズムの開発を行った。具体的には、ロボットの状態を変化させた場合の、前進動作パターンの獲得を行うシステムの構築である。また、本プロジェクトでは、サーボモータを利用したロボットを対象とした。

プロジェクト開始以前より、形状変化や未知のロボット形状に対する動作方法の獲得には、強化学習アルゴリズムなどを始め、様々な手法が提案されている。これらの手法に対して、それぞれの特徴を明らかにした。そこから、本プロジェクトの目的である素早い動作獲得に最適なアルゴリズム構築を行った。

3. 開発の内容

本プロジェクトの開発成果は、大きく 2 つある。「実機ロボットの学習システム」と「故障に対応する歩容制御システム」である。

3.1. 実機ロボットの学習システム

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。そのため、ロボットの試みた動作に対するフィードバックを行う必要がある。

本プロジェクトでは、ロボットに対して情報タグを設置し、そのタグを外部カメラで読み取ることで、実験環境内での絶対位置座標をロボットに渡している。ロボットおよびカメラには、それぞれシングルボードコンピュータを搭載しており、図 1 の構成では、Bluetooth 通信を利用して、上部カメラで取得したロボットの位置座標を、ロボット側に渡すことで、強化学習を行うことができる。

また、最終的なシステム検証に用いた 6 脚ロボットを利用する場合には、ロボットのアクチュエータ制御にはマイコンボードを利用しているため、XBee を利用した ZigBee 通信により位置情報を送信している。

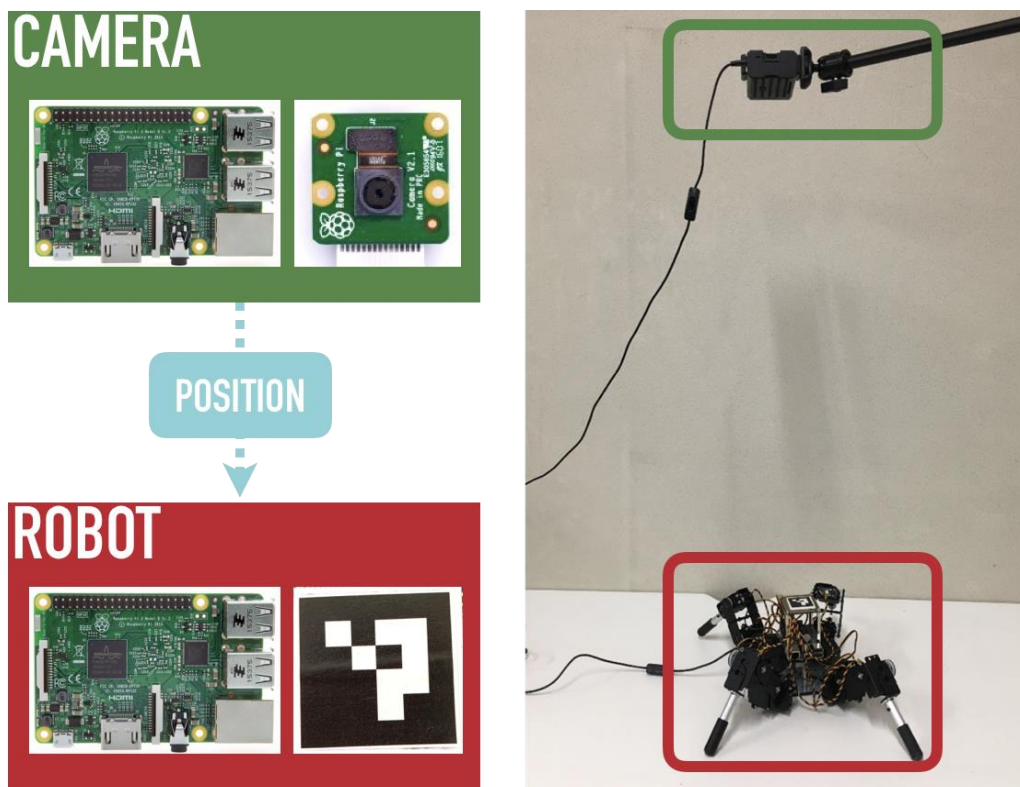


図 1. 実機ロボットの学習システム

3.2. 故障に対応する歩容制御システム

最終的な歩容制御システムでは、基本的な歩行パターンを基に、ロボットの状態が変化した場合の歩容を学習する。本手法は、3 段階に分かれて、歩容方法の学習を行う。

1 段階目は、事前に指定された歩行パターンを利用して、前進を試みる。今回、この時の動作パターンは、図 2 のように全部で 3 つ用意しており、同時に動かす脚の本数が異なる。



図 2. 基本歩行パターン

2 段階目では、脚の初期位置の学習を行う。この時、各関節にかかる負荷が最も少なくなるように、立脚時の初期位置を最適化する。

3 段階目では、各脚の動作順序を学習する。1 段階目で最も性能が高かった歩行パターンを基にして、各脚の動作順序を強化学習により最適化する。この時、各脚は 2 段階目で求めた初期位置を中心に動作する。

以上に、本プロジェクトの歩容制御システムは、3 段階の学習アルゴリズムにより、動作獲得を行う。

3.3. システム評価

実際の 6 足歩行ロボットを用いて、各脚の故障の組み合わせで、動作獲得実験を行った。それぞれが壊れた場合の全パターンを試して、本システムの有効性を確認した。

図 3 は、右中脚と左後脚が壊れた場合の実験結果である。図の下側から上側に前進する動作を獲得した結果である。この 2 本が壊れたまま歩行すると、右後脚に重量負荷が集中してしまい、右側に逸れていってしまう(図 3: 故障時)。これを本システムで改善することにより、正常時と同じように前進することが可能になっている(図 3: 改善結果)。

4. 従来の技術(または機能)との相違

従来のサーボモータを利用したホビーロボットでは、決められた動作を実行するだけであったため、ロボットの状態に対応したフィードバックを行うことが出来ていなかった。そのため、ロボットの故障など、状態が大きく変化する事態に対処できずに、部分的な故障が致命的なロボットの機能停止につながっていた。

また、研究レベルでは、自動的な動作獲得ができるものの、大規模なシミュレーション環境が必要であり、動作獲得の学習までに時間がかかることが多かった。

その上で、本プロジェクトの開発成果の特徴として大きく 3 つある。

1 つ目は、安価な部品で構成されているということである。位置情報取得用のカメラなどの機材も 1 万円以内で揃えられる。本プロジェクトで使用したロボット自体は高価であるが、システムそのものはロボットを規定していないため、ロボット自体の値段はシステムの構成コストに考えられない。

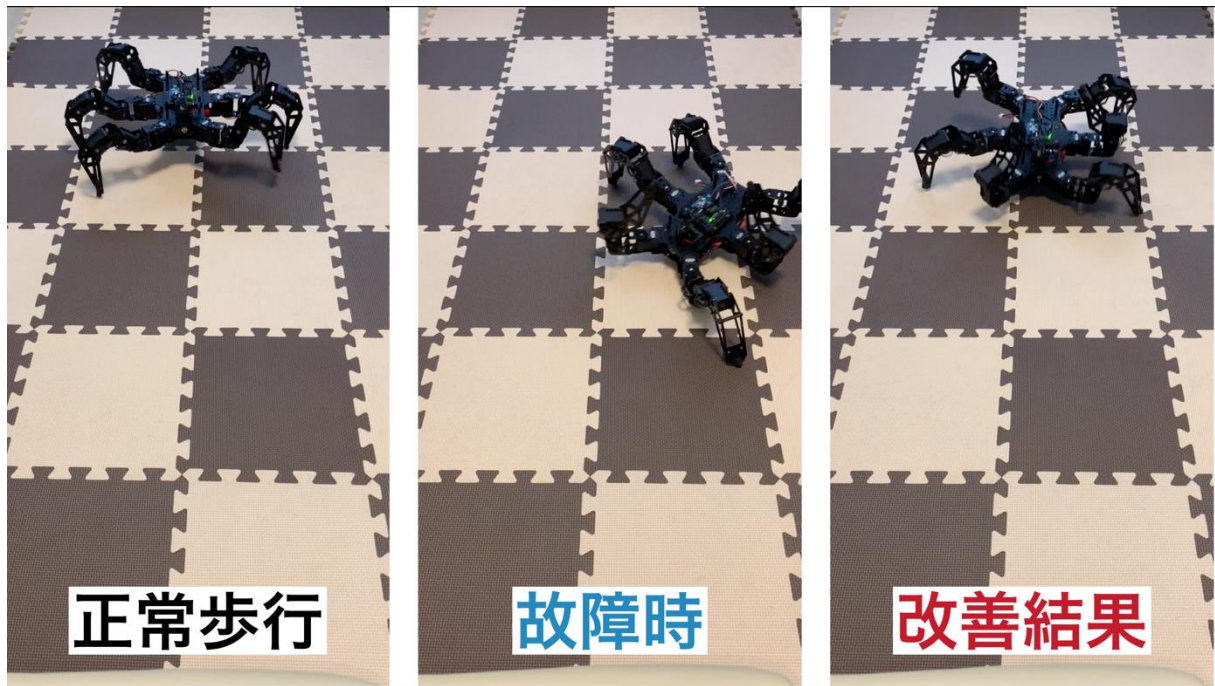


図 3. 右中脚と左後脚が壊れた場合の実験結果

強化学習自体の計算コストも少なく、ロボットに搭載されたマイコンで行うことができる。シミュレーションを使った深層強化学習などの研究でよく見られる高価な計算資源は必要がない。例えば、MMPRL の論文で使用されている計算機器は、数百万円かかる。少ない計算コストで学習できることは、社会普及を考えた時には、大きなアドバンテージとなる。

2 つ目は、短時間の学習が可能だということである。一般に深層強化学習のような方法だと、学習時間は短くても数時間はかかる。学習時間が長いものでは、数週間かかることもある。これらの手法により、たとえ十分な精度が出たとしても、試行錯誤を繰り返す開発現場では、とても使いにくいシステムである。

本システムでは 10～20 分で学習を行うことができる。これは、手法的な一般性を下げることとで実現しているが、短時間で学習できることにより、開発現場で利用しやすいシステムにできた。またホビー向けとしても、要求される計算資源が極めて少ないため、利用が簡単である。

3 つ目は、予測可能な学習を行うということである。強化学習では、与える必要があるのは報酬関数だけである。その一方で、予期しない行動をとることが多々ある。これは、人間が見つめることができなかったより素晴らしい手法の発見につながることもあるが、そのほとんどは好ましくない結果が多い。本プロジェクトにおいても、シミュレーションでの学習時に、シミュレータの計算誤差を利用して、脚をほぼ動かさずに細かい振動で前進したり、ひっくり返ることで速く動こうとしたりした。

これを解決するには、予期しない動きを制限するような形で、報酬関数を設計すれば良い。しかし、今度は各制限の重み付けがハイパーパラメータとなる。慣れていない場合、報酬関数のパラメータ調節は、通常のロボット制御のそれよりも難しい。これでは、簡単にロボット制御を行うというプロジェクトの主旨とは違ってしまい、本末転倒である。

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になっている。

5. 期待される効果

多脚ロボットは、ホビーや産業用途を始め、様々な領域での活用が盛んに研究されている。本システムにより、これらロボットの故障時の動作の獲得方法が簡単にできる。また、故障時だけでなく、4 足歩行や 6 足歩行ロボットなどの多脚ロボットでの動作最適化を行い、動作チューニングの負荷を軽減されることが期待される。

本プロジェクトの成果は、安価な部品で構成されており、高額な計算資源を必要としない。そのため、研究分野だけでなく、実際の組み込み機器での動作や現場での利用も期待される。

6. 普及(または活用)の見通し

本プロジェクトの成果は、当初の現実的な時間での行動獲得という目標を達成するのには十分なものであった。しかし、当初の目論見よりもわずかに汎用性という点において劣っている。本システムの動作は、歩行以外の動作による移動ロボットを対象としていない。

本プロジェクトでは、ロボットの形状が異なった場合でも動作可能になるというものであった。しかし、実際にどの程度の幅で対応できるかについては、正確に調べる必要がある。そのため、これらの成果を十分に整理した上で、ソフトウェアを使いやすい形にし、オープンソースで発表する予定である。これにより、多くの人に使ってもらい、対応できる範囲を調べ、本システムの改善を目指す。

7. クリエータ名(所属)

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)