



多脚ロボットの遠隔操縦システムの開発 - 強化学習による 6 脚ロボットの段差対応歩容の獲得 -

Development of a Remote-Control System for Multi-legged Robot - Acquisition of Stair-climbing Gait for a Hexapod Robot through Reinforcement Learning -

大山溪介¹⁾, 高橋秀智¹⁾

Keisuke OYAMA, Hidetomo TAKAHASHI

1) 東京工科大学 工学研究科 (〒192-0982 東京都八王子市片倉町 1404-1, takahashihdtm@stf.teu.ac.jp)

概要: 現在, ロボットの遠隔操縦で自動化技術の導入が進められている. 本研究では, 多脚ロボットの遠隔操縦におけるオペレーターの環境に応じた歩行パターンを自動生成するシステムの開発を目指す. 仮想空間上に 6 脚ロボットが段差登り動作を学習する環境を作成し, SAC と呼ばれる強化学習のアルゴリズムを導入した. 強化学習により 6 脚ロボットが高さ 0.03 m の段差を乗り越える歩行パターンを獲得することができた. 獲得した動作を実機ロボットに実装して実際の環境下での検証を行って学習の効果を確かめた.

キーワード: 6 脚ロボット, 強化学習, 遠隔操縦

1. 研究背景

近年, IT およびロボット技術の発展により様々な分野でロボットの活用が進められている. 特に災害現場などが立ち入ることの難しい場所では, ロボットを用いることで安全に作業を行うことが期待されている.

多脚式ロボットは, 車両型ロボットなどに比べて不整地での走破性に優れており, その制御方法について様々な研究がなされている. しかし, 多脚式ロボットの歩容獲得では, 足の動かし方が複雑で正解例は歩くときの状況に応じて無数に存在する. また不整地での歩行となると, 不確実性が生じるため制御規則をプログラムすることはさらに困難になる. この問題に対し, 強化学習を用いてロボットの歩容を獲得する研究が行われている. 強化学習は, 報酬を最大化する方策を試行錯誤によって学習することができるため, 複雑で不確実性のある問題を解くことに向いている.

歩容獲得手法として強化学習を導入した研究として大下の研究[3]が挙げられる. 大下は 6 脚の蜘蛛型ロボットが不整地での踏破能力を獲得するために学習戦略フュージョンと呼ばれる手法を導入した. 学習戦略フュージョンは単一のタスクに対して異なる種類の学習戦略を自動的に複数回組み合わせながら適用することで, タスクの遂行能力を高めていく手法である. この研究では, 3 種類の異なる環境で匍匐動作を獲得する実験を行い, それぞれの環境を推定して適切な方策を選択することに成功している.

本研究では, 多脚ロボットの遠隔操縦におけるオペレーターの負担を軽減するため, 環境に応じた歩行パターンを自動生成するシステムの開発を目指す.

2. 対称とするロボットとそのモデル化

2.1 実機ロボット

本研究では, 対象ロボットとして Big Hexapod Robot Kit[5]を使用する. Hexapod は長さ 0.16m, 幅 0.19m, 高さ 0.06m の胴体と 3 つの関節を持つ脚 6 本, カメラを搭載した頭部から構成され, 関節にはサーボモータが搭載されている. 胴体に取り付けられた raspberry pi によってサーボモータを制御し, 歩行動作を実現している.

2.2 目的とする動作

本研究では, ロボットの不整地歩容の一例として, Hexapod が目前の段差を登る歩容の獲得を目指す. Hexapod の前方に 0.03 m の段差を設置しその中央をゴールとして設定した. Hexapod の中心から段差の中央までの距離は 0.3 m とした.



図 1 Hexapod の外観

2.3 モデル化

本研究では、Hexapod を再現したシミュレーションモデルを MJCF と呼ばれる記述法で作成し、xml ファイルとしてシミュレーション環境に読み込んでいる。MJCF は、箱・楕円・球・長円形といった形状と、接続方法を定義することでモデルを作成する記述法である。Hexapod のシミュレーションモデルでは胴体を箱形状で定義し、脚の形状は長円形で定義した。脚の接続方法を 1 自由度の回転で定義することで、各関節を再現した。関節の可動範囲は-30 度から 30 度に設定した。図 2 は Hexapod の xml モデルを一部抜粋したものである。3 行目の torso_geom 似て胴体部分を、5 行目の front_left_leg で左前脚を定義している。joint は形状同士の接続方法を定義するもので、6 行目の base_1 は胴体と左前脚の z 軸方向の回転を表している。9 行目の hip_1 と 12 行目の ankle_1 で脚の屈伸動作に関わる関節を定義している。

```

<body name="torso" pos="0 0 0.1">
  <camera name="track" mode="trackcom" pos="0 0 0" xyaxes="1 0 0 0
  1"/>
  <geom name="torso geom" pos="0 0 0" size="0.19 0.16 0.06" type="box"/>
  <joint armature="0" damping="0" limited="false" margin="0.01"
  name="root" pos="0 0 0" type="free"/>
  <body name="front_left_leg" pos="0.08 0.14 0">
    <joint axis="0 0 1" name="base_1" pos="0.0 0.0 0.0" range="-30 30"
    type="hinge"/>
    <geom fromto="0.0 0.0 0.0 0.03 0.03 0.0" name="aux_1_geom"
    size="0.05" type="capsule"/>
    <body name="aux_1" pos="0.0 0.0 0">
      <joint axis="0 1 0" name="hip_1" pos="0.0 0.0 0.0" range="-30 30"
      type="hinge"/>
      <geom fromto="0.08 0.08 0.0 0.17 0.17 0.0" name="left_leg_geom"
      size="0.02" type="capsule"/>
      <body pos="0.17 0 0">
        <joint axis="-1 1 0" name="ankle_1" pos="0.0 0.0 0.0" range="-30
        30" type="hinge"/>
        <geom fromto="0.0 0.0 0.0 0.0 -0.1" name="left_ankle_geom"
        size="0.03" type="capsule"/>
      </body>
    </body>
  </body>
</body>

```

図 2 Hexapod の xml モデル (一部)

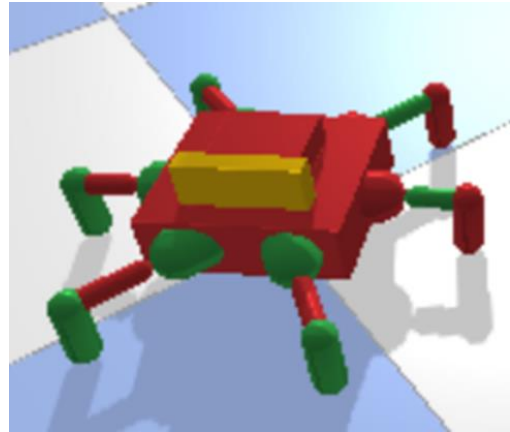


図 3 Hexapod モデルの外観

3. 強化学習による歩行パターンの獲得

強化学習とは、試行錯誤を通じて最適な行動を学習する手法であり、行動を行う主体をエージェントと呼び、エージェントが制御しようとする対象を環境と呼ぶ。本研究では、Hexapod をエージェントとして設定し、その行動は関節を 1 つずつランダムに動かす動作を 1step として設定した。100step を 1episode とし、50000episode の学習を行った。強化学習のアルゴリズムとしては SAC を採用した。

強化学習を行った episode のうち、ゴールに到達した episode のモデルを保存し、そのモデルを用いて 1episode だけのシミュレーションを行って step ごとの関節の角度を CSV ファイルに記録した。

3.1 生存判定と報酬設定

強化学習における生存判定とは、エージェントがより効率的に試行を繰り返すために設定する条件であり、エージェントがどのような状態のとき試行を終了するか決定するものである。本研究では、転倒の判定、移動範囲の条件、接地状態の条件を設けることで、転倒や範囲外への移動などが発生しないようにした。

報酬は、エージェントの行動に対して環境から与えられる評価である。報酬には正の報酬と負の報酬があり、エージェントが良い行動をしたときに正の報酬を与えるように設定する。エージェントはより大きな報酬を得られるように行動を改善することで、最適な方策を学習していく。本研究では、生存報酬、距離報酬、進行方向による報酬、安定動作による報酬を設定することでロボットがより安定した動作で段差の中央に進むことで報酬を獲得するようになった。なお、報酬値が 1000 以上でゴール到達となる。

4. 実機との対応

4.1 ロボットのシステム構造

Hexapod は raspberry pi によって制御を行っている。歩行制御を行うプログラムの重要な箇所について示す。歩行動作の制御は Control クラスの run メソッドによって定義

されている。各関節の角度は `setServoAngle` メソッドで定義されたあと、信号に変換されサーボモータへと送られる。本研究では、`setServoAngle` に学習した角度データを入力することで実機で動作の検証を行うことができるようにした。

4.2 モデルと実機の対応

学習で得られた各関節の角度データを実機のサーボへ適用するために、モデルと実機の対応関係と初期姿勢の関節角度を調査した。対応関係表を表 1 に示す。

表 1 モデルの関節と Hexapod のサーボの対応関係

	モデル上の名前	サーボモータのチャンネル	初期状態の角度 [deg]	回転方向
右前脚	base_1	15	60	-
	hip_1	14	70	-
	ankle_1	13	80	+
左前脚	base_2	16	90	-
	hip_2	17	70	+
	ankle_2	18	80	-
左後脚	base_3	22	70	-
	hip_3	23	80	+
	ankle_3	27	90	-
右後脚	base_4	9	60	-
	hip_4	8	70	-
	ankle_4	31	60	+
左横脚	base_5	19	75	-
	hip_5	20	60	+
	ankle_5	21	60	-
右横足	base_6	12	80	-
	hip_6	11	70	-
	ankle_6	10	50	+

5. 結果と考察

図 4 は 50000 試行の学習を行ったときの AverageRetur の推移を表したグラフである。10000 回を超えてから平均報酬値が上昇し、50000 回近くでは 3000 付近の値で安定している。報酬を多く獲得できる動作を学習できていることが読みとれる。

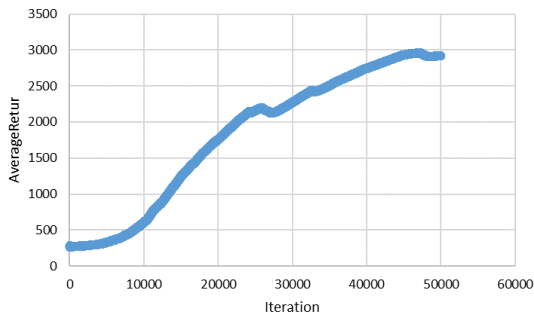


図 4 平均報酬値の推移

5.1 学習した歩行パターンの分析

報酬値が特に大きかった Episode の学習モデルを用いて角度データを取得した。図 5 はゴールに到達した Episode

の再生結果である。50000 回の試行では、ロボットが転倒してしまう動作や、ロボットが範囲外に飛び出してしまう動作も確認された。

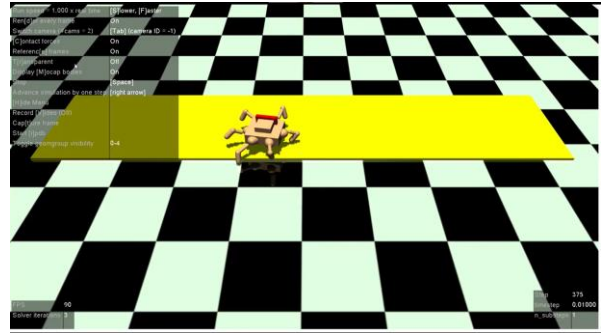


図 5 ゴールに到達した様子

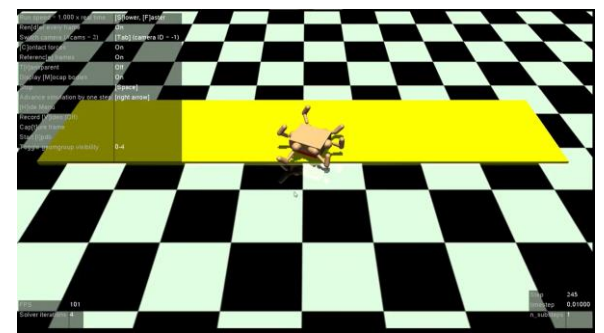


図 6 ロボットが転倒してしまう動作

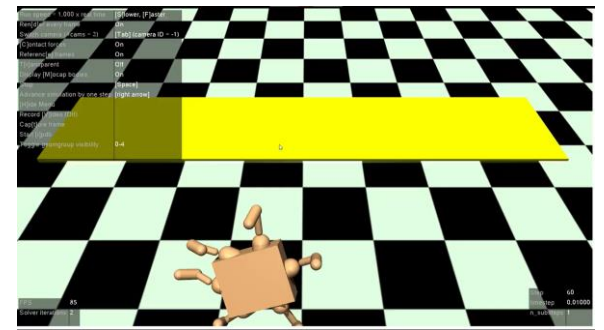


図 7 ロボットが範囲外に飛び出してしまう動作

5.2 実機による検証

ロボットがゴールに到達することができたときの角度データを実機ロボットに適用し実際の環境で検証を行った。実機では胴体を支えることができずゴールに到達することができなかった。

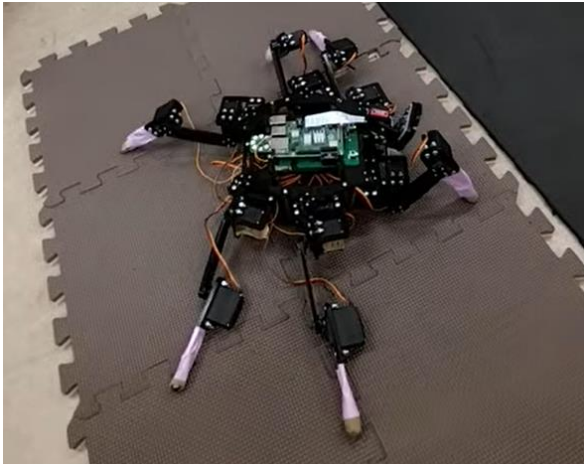


図 8 実機での検証結果

6. 今後の方針

実機で動作できなかった原因としてロボットの初期姿勢と関節の動作範囲の不一致が考えられる。今後はロボットの姿勢と関節の動作範囲を制限したモデルを用いて学習を行うことで、実機で動作可能な歩行パターンを獲得したい。

また、強化学習をより効率的に進める方法として、環境の情報を認識しながら行動するということが考えられる。段差の位置や高さを認識しながら学習を行うことで不要な動作を減らすことができると考えられる。画像データを扱える物理エンジンとして PyBullet[6]を使用して、従来のシミュレーションと同様のモデルを動作させることができたが、学習は行えていない。シミュレーション環境を改善するために、PyBullet を使用して歩容を獲得する環境を構築したうえで、画像データの利用を導入していきたい。

7. むすび

本研究では、仮想空間上に 6 脚ロボットが段差登り動作を学習する環境を作成し、SAC と呼ばれる強化学習のアルゴリズムを導入して、ロボットが段差を乗り越える歩行パターンを獲得し、実機による検証を通じて学習の効果を確かめた。学習した歩行パターンのデータを分析し実機との対応関係を明らかにしたことで、実機を用いて実際の環境で動作を行えるようにすることができた。実機での検証の結果、歩行動作の学習自体は行えていたものの、実際の環境で段差を乗り越えることができない内容であることが分かった。この原因として、学習で使用したモデルと実機の初期状態が一致していないことが考えられる。実機に適応できる歩行パターンを獲得するためには実機をより正確に再現したシミュレーションモデルを用いて学習を行う必要がある。

参考文献

- [1] 武田千逸 強化学習を用いた六足歩行ロボットの段差対応 東京工科大学 工学部機械工学科 2022 年度卒業論文。
- [2] 牧野晴紀 “強化学習を活用した多脚式ロボットの段差対応のための最適歩容の獲得” 東京工科大学工学部機械工学科 2023 年度卒業論文。
- [3] 大下将宗 “強化学習を用いた多様な環境における歩容獲得手法の実機蜘蛛型ロボットにおける検証” 奈良先端科学技術大学 2014 年度修士論文。
- [4] 伊藤多一, 今津義充, 須藤広大, 仁ノ平将人, 川崎悠介, 酒井裕企, 魏崇哲, 現場で使える! Python 深層強化学習入門, 翔泳社, 2019/2019.
- [5] Freenove, <https://www.freenove.com/>
- [6] Bullet Real Time Physics Simulation, <https://pybullet.org/wordpress/>