

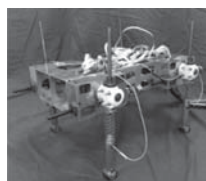
ロボティクス・メカトロニクス 講演会 2018 に参加して

福原裕志

Hiroshi FUKUHARA

機械システム工学専攻修士課程 1年

表 1 Details of the Robot



Length [mm]	550
Width [mm]	414
Leg Length [mm]	250
Weight [kg]	11.8

図 1 Four-legged Robot

1. はじめに

私は 2018 年 6 月 2 日から 5 日にかけて福岡県の北九州国際コンペティションゾーンで行われた「ロボティクス・メカトロニクス講演会 2018」に参加し、4 日に「強化学習を用いた 4 脚走行ロボットの斜面における移動方策の獲得」というテーマでポスター発表を行った。

2. 研究内容

2.1 緒言

近年自律型ロボットの開発が盛んになっているが、現状では、想定外の事態や変化する環境には対応が難しいという問題がある。この問題に対し有効であると考えられるアプローチの一つに、強化学習という手法がある。強化学習は、機械学習手法の一つであり、環境や状況に応じて試行錯誤を行うことで、目的達成のための最適行動を獲得する手法である。これを用いると、環境の変化があっても、それに応じて行動パターンを変更することが可能であるため、多様な環境への適応が可能となる。

そこで本研究では、強化学習を 4 脚走行ロボットの制御に適用することで、斜面環境での高速な移動の実現を試みた。

2.2 4 脚走行ロボット

本研究では、図 1 に示す 4 脚走行ロボットをモデル化し、それをコンピュータ上で再現して学習実験を行った。ロボットの諸元を表 1 に示す。

2.3 歩容

動物は脚を一定の法則に基づいて動かしそれを繰り返すことで歩行や走行などの移動を行う。この動作を歩容と呼ぶ。本研究では、馬や犬などの動物が少し速めの移動を行う際に使用する歩容であるトロット歩容になるように各脚の位相関係に制約を加えたうえで、学習実験を行った。

2.4 強化学習

本研究では学習手法として、代表的な強化学習アルゴリズムである Q 学習を用いる。

Q 学習では、エージェント（ロボット等）は環境の情報を状態 s_t として受け取り、それに応じた行動 a_t を行う。行動の結果に応じて報酬 r_t が付与される。この報酬値の情報を用いて、状態や行動の価値 $Q(s_t, a_t)$ (Q 値) を更新していくことで、最終的に累積報酬を最大化させるような一連の行動パターンが得られるという学習手法である。

2.5 学習実験

本研究では、4 脚走行ロボットに強化学習を適用することで、斜面（上り坂）において高速に移動可能な方策を獲得できるか検討することを目的として実験を行った。

2.5.1 基本学習条件

本学習実験では、シミュレーション環境上でロボットが 10 s 以内に初期位置から前方へ 2 m 進むことを課題とした。シミュレータとしては、動力学計算エンジン「Open Dynamics Engine」を用いた。

Q 学習における状態の判定には、ロボットの進行方向の速度を用いて、12 通りの状態を用意した。

行動としては、脚振り動作における脚振り角度及び周期の組み合わせを 50 通り（角度は 2~20° の 10 通り，周期は 0.2~0.6 s の 5 通り）用意した。状態の取得，行動の選択及び報酬の付与は脚振り 1 周期ごとに行う。報酬条件としては，毎回式 (1) のような速度に応じた報酬を与えるとともに，ゴール時には式 (2)，転倒時には式 (3)，時間切れの際には式 (4) のような報酬を与えた。

$$r = 100 v \quad (1)$$

$$r_{goal} = 10000 - t \quad (2)$$

$$r_{fall} = 100 x - 100 \quad (3)$$

$$r_{time\ out} = 10 x \quad (4)$$

v : 進行方向の機体速度 [m/s], t : 経過時間 [ms]

x : 到達距離 [m]

このような状態，行動，報酬の条件を用いることで，ゴールに到達し，かつなるべく機体速度が大きくなるように，機体速度に応じて適切な脚振り角度と周期を選択するという学習が可能となる。

2.5.2 実験 1 : 脚振り角度・周期を行動とする学習

2.5.1 節に示した条件を用いて，まず平地環境で学習を行った。この際の平均速度とゴール確率の推移は図 2，図 3 のようになった。

ゴール確率は 1 に収束し，平均速度についても学習初期に比べ大きくなった。

次に，1~8° の 8 通りの斜面環境で学習を行っ

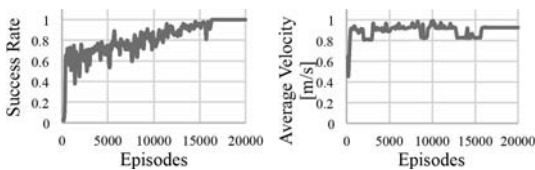


図 2 Success Rate

図 3 Average Velocity

表 2 Average Velocity (in Experiment 1)

Slope Angle [°]	0	1	2	3	4	5
Average Velocity [m/s]	0.93	0.87	0.70	0.45	0.41	0.51

た。結果，1~5° の斜面でゴール可能な方策が得られた。このとき獲得された方策を用いた場合の機体の平均速度を表 2 に示す。

獲得された行動からは，機体速度に応じて脚振り角度を大きくしていくことで，急な斜面においても転倒せず進むことが可能であることが分かった。

2.5.3 実験 2 : 角度・周期・ D_{sw} を行動とする学習

次に，より急な斜面に対応させることを考え，行動を，脚振り角度と周期に加え，周期に占める支持脚時間の割合 D_{sw} についても 3 通り (0.5, 0.7, 0.9) を用意し，合計 150 通りに増やして，その他の条件は実験 1 と同様にして学習を行った。

結果，平地及び 1~7° の斜面においてゴール可能な方策が得られた。 D_{sw} を 0.7 に固定し，脚振り角度及び周期を行動として学習した場合についても 1~7° の斜面においてゴール可能な方策が得られたが，すべての斜面角度における速度の平均としては， D_{sw} を行動として加えた場合の方が 0.075 m/s 速くなった。

3. おわりに

今回は私自身初めての学会参加だった。専門分野外の方も多く，わかりやすく説明することが求められた一方，同分野の方からは内容について様々な意見を頂くことができ，非常に勉強になった。この経験を活かし，今後の研究活動に取り組んでいきたい。