

小型ヒューマノイドロボットにおける水泳行動への強化学習の適用

公立はこだて未来大学大学院 ○小林 世弥、公立はこだて未来大学大学院 鈴木 恵二

要 旨

本研究では、ヒューマノイドロボットの活動範囲の拡張を目的として、小型ヒューマノイドロボットに水泳行動を達成させることを目標とする。基礎研究として、小型ヒューマノイドロボットの直線水泳行動へ強化学習を適用させ、水泳行動を獲得することを目指す。

1. 序論

現在、ヒューマノイドロボットの進化は著しい。しかし、産業的には極めて厳しい状態であることも事実である。これは現状でのヒューマノイドロボットの研究対象や使用可能領域のほとんどが室内や平らな屋外に限定されているため、実世界で使用可能なヒューマノイドロボットが広く普及していないことが要因の一つと言える。ヒューマノイドロボットは極限作業や、付加の高い人間の作業を代替することを期待することができる。また、ヒューマノイドロボットはその形状から様々なフィールドでの行動が可能である。例えば、水中で行動していたヒューマノイドロボットが陸に上がったとしても、すぐに行動することが出来る。行動する場所に合わせた行動を行う、すなわち行動の切り替えが可能である。それゆえ、屋外のみならず水辺、水中での行動能力の付与が期待されている。

本研究では、ヒューマノイドロボットの応用領域を広げることが目的に、小型ヒューマノイドロボットによる水泳行動の実現を図る。海洋領域や水中での作業をヒューマノイドロボットで代替することを念頭に、ここでは基礎的研究として、防水対策をした小型ヒューマノイドロボットを水槽内において直線に遊泳させる。遊泳の際、強化学習を用いることで波の影響を受けても直線に行動を行う事の出来るようなロバスト性を持った行動の獲得の検証を行う。

2. 方法

本研究では、小型ヒューマノイドロボットを防水スーツで全身を覆っている。このロボットに対して、外部 PC からサーボを動かすコマンドを送信することにより水泳行動を実現する。

2.1 強化学習

本研究では、ロバスト性を持った水泳行動を行うために、自身の行った水泳の行動を評価し、その評価を反映する必要がある。それゆえ、強化学習の1つである Q-Learning を用いた。学習に必要な行動価値関数である Q 値は以下の式で計算させることが出来る。

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_{t+1} + \gamma \max_a Q(s_{t+1}, a_t) - Q(s_t, a_t)] \quad (1)$$

時間 t でのロボットの行動は a 、状態は s である。報酬 r_{t+1} はロボットがゴールにいる場合は 0、それ以外の領域にいる場合は -1 とする。

本研究でのエピソード数は 30 試行、ステップ数は 100 回とした。

行動選択の手法として、本研究では ϵ -greedy 手法を用いた。 $(\epsilon = 0.1)$ それ以外の値として、学習率 (α) は 0.1、減衰率 (γ) は 0.9 とした。

3. 実験

3.1 ヒューマノイドロボット

本実験では小型ヒューマノイドロボットとして Robovie-M を使用した。Robovie-M の主な仕様は高さ 290(mm)、幅 240(mm)、体重は 1.9(kg)、関節は 21 自由度であり、外部 PC との接続方式は RS232C である。

3.2 防水スーツ

ロボットを水泳を行わせるために、防水スーツ (図 1) を作成した。スーツの材質はネオプレンで、人間のダイブスーツにも使用されている材質である。スーツを製作する際には、スーツ内に水が浸入しないように作成する事、もう一つはロボットの自由度に制限がかからない様にする事を留意した。

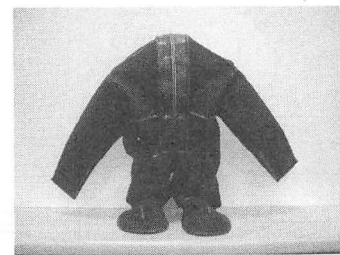


Fig. 1 ヒューマノイドロボットのダイブスーツ

3.3 位置検出と目標領域の決定

ロボットの現在いる位置の検出は水槽上部に USB カメラを設置し、このカメラから取得された画像を処理して検出されている。目標領域はロボットのスタート位置から直線前方 1.0m 離れた点を中心とする一辺 0.1m の正方形とした。

3.4 行動

行動と上半身の手と下半身のバタ足に分けて設定する。下半身のバタ足は角度とバタ足の時間の違いによって4パターン設定した。(表1)

Table 1 バタ足の組み合わせ

	角度 θ	時間 $t(s)$
Kick1	$\pi/9$	1.0
Kick2	$\pi/9$	2.0
Kick3	$\pi/18$	1.0
Kick4	$\pi/18$	2.0

表1の角度とは上半身を中心としたバタ足の振り幅であり、時間はバタ足を1試行する時間を示す。

上半身の動きは表2の3パターンが用意された。

Table 2 上半身の組み合わせ

	右腕	左腕
Arm1	0	0
Arm2	0	$\pi/2$
Arm3	$\pi/2$	0

表2の腕の角度はロボットの体の側面に腕がある状態を基準として、その基準から腕を上げた角度である。すなわち $\pi/2$ だけ腕を上げている行動とは、体の前面に対して腕を伸ばしている行動である。

ヒューマノイドロボットの行動数は、上半身と下半身の行動を組み合わせた12パターンである。

3.5 状態

ヒューマノイドロボットは水面を泳ぐ。ロボットの状態は、位置と姿勢で表した。これは人間が水泳を行う場合、自分が向かおうとしている目標の位置と自分が今いる位置との関係を考え、自分の姿勢を考慮して泳ぐ事からこのように表した。

位置は直線上、直線より右の位置、直線より左の位置である。図2は水槽上部から見た図である。ここでの直線上とは、スタート地点から目標領域を延長した直線を意味する。目標領域に到達した場合、ロボットは停止し、次の試行となる。

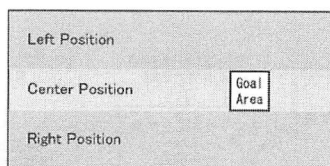


Fig. 2 ロボットの水槽での位置判定

ロボットの姿勢はロボットが目標領域に向かう方向を中心として、左向き、右向きとする。

4. 結果

結果は次のグラフの通りである。

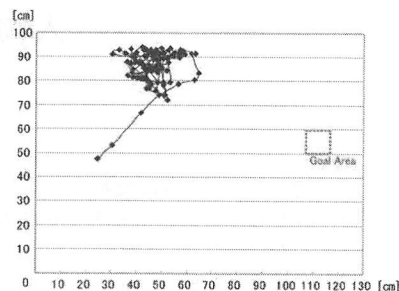


Fig. 3 ロボットの水泳行動 (エピソード0)

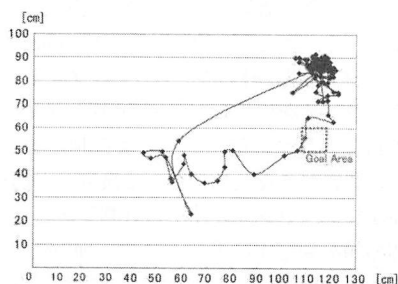


Fig. 4 ロボットの水泳行動 (エピソード30)

図3は水泳上部から見たロボットの水泳行動を辿ったものである。この試行では、ゴールに向かうことすら出来なかった。図4は30試行目の水泳行動を辿ったものである。個の試行では、ゴール地点に近づくことが出来たが、ゴールを通り、止まることが出来なかった。

5. 考察

本研究では、ロボバスト性を持った水泳行動を獲得するために小型のヒューマノイドロボットに強化学習の1つであるQ-Learningを適用した水泳行動を検証した。水泳行動自体は学習が行われた事は確認できたが、本研究の目標である水泳行動のロボバスト性の獲得は見られなかった。

これらの事から、提案手法の再考を行い、よりロボバスト性を持った水泳行動の獲得を目指していく。同時に現段階では2次元である行動を3次元にすることも目指していく。

参考文献

- 1) 三上貞芳 田野浩明 嘉数侑昇:強化学習による多足歩行ロボットの適応的歩様獲得に関する研究, 日本機械学会論文集 60 巻 580 号 (1994.)
- 2) 三上貞芳 皆川雅章:強化学習 (2000.)
- 3) Daisuke Iijima Wenwei Yu Hiroshi Yokoi and Yukinori Kakazu:Distributed Robotic Learning -Adaptive Behavior Acquisition for Distributed Autonomous Swimming Robot in Real-World (1999.)
- 4) N.Kato, J.Ayers and H.Morikawa:Bio-mechanisms of Swimming and Flying(2004.)
- 5) 小林世弥, 鈴木恵二:小型ヒューマノイドによる水泳行動の学習, ロボティクス・メカトロニクス講演会 2005 講演論文集, 1P2-S-066(2005.)