

過去の経験を利用した多様な路面に対応した歩行動作の獲得*

深谷 亮†

芝浦工業大学 情報工学科‡

1. はじめに

現在、社会に進出しているペットロボットなどの歩行ロボットは事前に用意された路面や環境に対応した制御プログラムを使用して移動を行う。しかしこの手法は多数の不整地に対応した制御プログラムを事前に用意する必要があり、これは非常に困難である。

これを解決するにはロボットが自律的な対応をする必要がありそのために試行錯誤を通して環境に適応する強化学習という手法が広く用いられている。しかし、強化学習による動作の獲得には時間がかかる。そこで本研究では、ロボットがある環境で歩行動作を獲得した時のニューラルネットワークの結合加重を利用することで、別の環境での歩行動作獲得にかかる時間を減少させることを目指す。

2. 実験内容

本研究ではニューラルネットワーク(以下NN)と逆誤差伝搬法を使用し、ある環境で学習した知識を傾斜の有無や摩擦係数等の要因が変化した新たな環境で再利用することで、学習時間を短縮することを目指す。

実験では、Russell Smithらが開発したOpen Dynamics Engineを使用し、ある環境パターンでの歩行動作について学習させた後、学習した動作の知識を抽出し、それを新たな環境でも自律的に選択・利用することで、その環境での歩行動作獲得にかかる時間を減少させることを可能にする。

3. ニューラルネットワーク(NN)

NNは人間の脳細胞であるニューロンを模したニューロンモデルを接続して作られるネットワークである。特徴として学習能力、非線形処理、並列処理などがある。

4. 逆誤差伝搬法

前述のNNにおける学習モデルの一つ。入力に対して教師信号と呼ばれる回路網が出力すべき望ましい出力が外部から与えられる。出力と教師信号の誤差を計算し、それが学習に用いられる。また、入出力に0から1への連続値を用いることができる。

5. Open Dynamics Engine(ODE)

フリーのC/C++用動力学シミュレーションライブラリであり、仮想現実環境で動力学シミュレーションを行うことができる。オープンソースでライセンスはGNUのLGPLかBSDとなる。また、3次元グラフィクス(Open GL)と衝突検出機構が組み込まれているのでそれらの知識を必要とせずに使用できる。計算速度は高速だが、精度は低い。

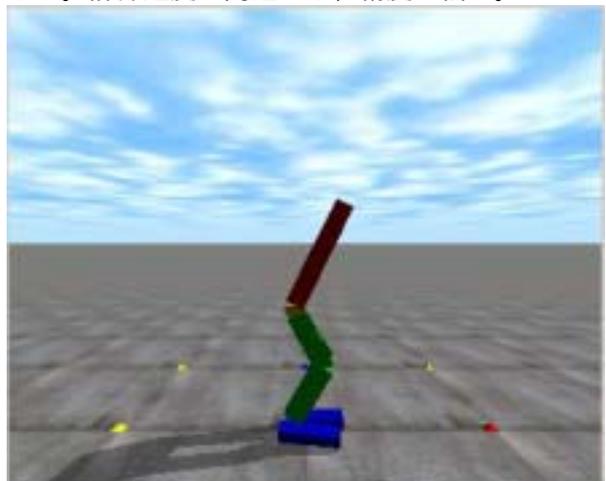


図1 Open Dynamics Engine

6. 有効な知識の抽出と再利用

6.1 知識の抽出

本実験では、中間層ニューロンの要素番号、一定の数値で区切った入力値、閾値以上の結合強度、出力値などを知識として保存する。その際の手順は

- (1) 時刻 t において選択した行動の評価値が閾値以上だった場合、その際に最も出力が大きかった中間層の要素に注目する
- (2) 入力層や出力層の各要素に対する結合強度が閾値以上だったものの結合強度や入力値などを記録する。

となる。

6.2 知識の再利用

前述のように抽出した知識を別の環境で再利用する。しかし、入力値が同じであっても環境や目的が異なれば、同じ行動を取っても同様の評価を得られるとは限らない。そこで本研究では

- (1) 抽出された知識の入力値と現在のNNへの入力値が一致する知識を選択する
- (2) 抽出された知識の出力で行動する

* 「Learning walk movement on various road surface by using the experience」

† 「Ryo Fukaya」

‡ 「Department of Information Science & Engineering, Shibaura Institute of Technology」

(3) 行動の結果、得られた評価値が閾値以上であれば、NNの結合強度を抽出された知識の結合強度で上書きするという手順を取る。

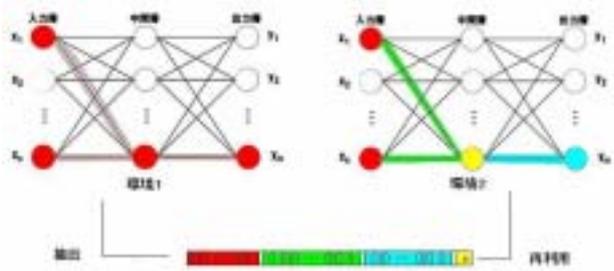


図 2 知識の抽出と再利用

7. 実験

7.1 実験の概要

本研究ではある環境で学習した知識を利用して、類似性のある環境での動作獲得にかかる時間を減少させることを目的としている。そこで、手法の有効性を確認するために、以下の3つの手法を比較する。

- (1) 環境が変化した際に、NNの結合加重をリセットする(独立に学習)
- (2) 環境が変化した際に、NNをそのまま使用する(連続に学習)
- (3) 環境が変化した際に、NNの結合加重をリセットするが、別の環境で有効であった知識を導入する(提案手法)

上記の手法を3つの異なる環境 A、環境 B、環境 C で学習させる。なお、各環境の設定は

環境 A: 低重力

環境 B: 障害物設置

環境 C: 低重力、障害物設置

となっている。

NNへの入力はい各関節の角度・環境の種類・ロボットの位置が与えられ、出力された値から目標姿勢を選択し、その姿勢をとるように行動する。

実験は各環境において行動を 10000 ステップ繰り返し、それまでの間に目的地にたどり着くことを目標とする。これを環境 A から環境 C までを 1 セットとし、30 試行繰り返す。その3つの環境すべてで目的地に到達したことを成功とし、その場合の成功率をそれぞれ S_1, S_2, S_3 とする。また、成功した場合の平均ステップを E_1, E_2, E_3 とする。これらの結果を比較して、各手法の精度や学習速度などを検証する。

7.2 実験の結果

S_1	S_2	S_3
53%	47%	67%
E_1	E_2	E_3
6176.70	5761.16	4318.58

図 3 実験結果

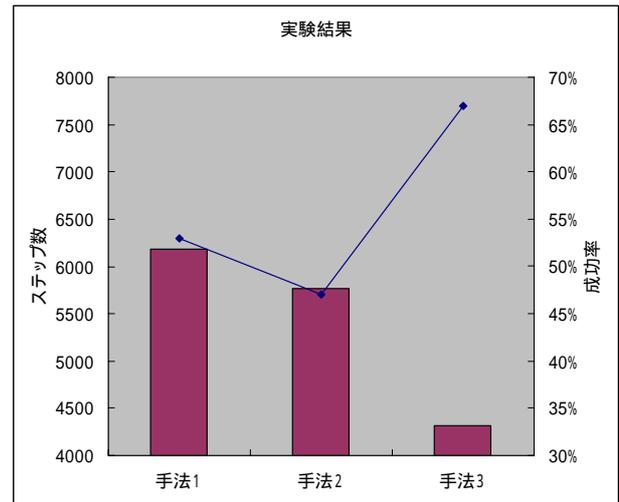


図 4 平均ステップ数、及び成功率

実験の結果である成功率と平均ステップを図. 3 に示す。図から提案手法は他の2つの手法に対して精度・速度共に向上していることが伺える。

この結果から、ある環境で有効であった学習の結果を保存しておき、類似した環境で一部継承することにより、その環境での学習の精度や速度を向上させられることがわかった。また、複雑なタスクに適用させる際、段階的に学習させることが有効であることが分かった。

8. まとめ

本論文では、ある環境で学習した知識を別の類似した環境に一部継承する手法を提案した。実験により提案手法は単独での学習や連続の学習に対して精度・速度が向上することが確認された。

今後は段差などのもっと複雑な環境や、腕などの新たな部位や関節を追加した、より複雑なタスクへの適用について検討していきたい。

参考文献

- [1] 総合研究論文: 井谷久博 古橋武
自律移動ロボットによる教示情報の理解
- [2] 総合研究論文: 近藤敏之 伊藤紀彦 伊藤宏司
拘束ルール抽出機構を用いた自律移動ロボットの段階的行動学習