

進化計算を用いた仮想ロボットの自律行動獲得

北海道科学大学 ○大江亮介, 北海道科学大学 川上敬

要 旨

物理法則に基づいて運動を行う仮想ロボットの行動獲得を行う。仮想ロボットは直方体を接続した構造を持ち、その運動は物理エンジンを用いて計算する。制御器として複合的人工ニューラルネットワーク (CANN) を用い、粒子群最適化によって学習を行う。CANN の全パラメータを一度に学習した場合、および CANN の選択規則を固定した場合について数値計算実験を行い、結果の比較を行う。

1. 緒 言

ロボットなど動的に状態が変化する対象の制御方法のひとつに、制御対象の数理的記述を必要としないインテリジェント制御がある。人工ニューラルネットワーク (Artificial neural network, ANN) はインテリジェント制御のひとつであり、脳の神経素子ネットワークをモデル化したものである。

ANN の学習方法としては、誤差逆伝播法が用いられることが多い。ただし、誤差逆伝播法は教師あり学習であり、正しい制御信号が未知の場合は適用できない。ANN の教師なし学習方法としては、最適化手法によって ANN のパラメータを学習する EANN (Evolving ANN)¹⁾ などが提案されている。

EANN による学習効率の向上を目的とし、著者らは動的な制御問題で有効な ANN のアンサンブル学習として複合的人工ニューラルネットワーク (Composite ANN, CANN) を提案した。CANN では一つの選択用 ANN によって、複数の学習済み ANN のいずれか 1 つを選択して用いる。円環構造弾性ロボットの行動制御シミュレーション²⁾ や羽ばたき飛翔ロボットの飛翔制御シミュレーション³⁾ により、CANN が単一の ANN よりも効率的であることが示された。

本研究では、CANN の構造が学習に及ぼす影響を明らかにすることを目的とする。そのための基礎的な実験として、物理法則に従って運動を行う仮想ロボットに CANN を適用し、どのような学習結果が得られるかを調査する。この際、CANN の全パラメータを一度に学習した制御器、および CANN における選択規則を固定化した制御器を用い、結果の考察を行う。

2. 複合的人工ニューラルネットワーク

図 1 に CANN の概念図を示す。ここで、CANN における選択用 ANN を高次 ANN、実際に制御信号を出力する学習済み ANN を低次 ANN と定義する。入力ベクトルを \mathbf{x} 、出力ベクトルを \mathbf{y} とすると、CANN の処理は以下の 3 つのステップによって記述される。

Step 1: 高次 ANN の出力計算

入力ベクトル \mathbf{x} を高次 ANN に入力し、出力ベクトル \mathbf{s} を得る。ただし \mathbf{s} の次元は低次 ANN の個数 n に等しい。

Step 2: 低次 ANN の選択

高次 ANN の出力ベクトル \mathbf{s} は各低次 ANN の選択優先度を表す。 \mathbf{s} の中で最大値を取る要素の番号 i を求め、 i 番目の低次 ANN を選択する。

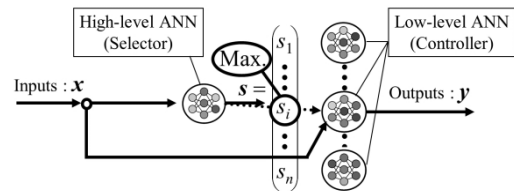


Fig.1 Composite artificial neural network

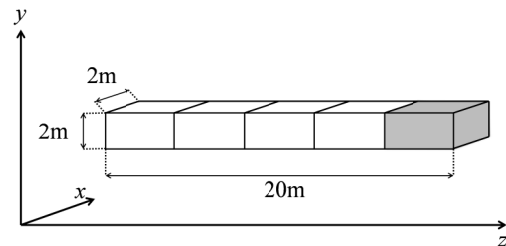


Fig.2 Model of the virtual robot

Step 3: 低次 ANN の出力計算

選択された i 番目の低次 ANN に \mathbf{x} を入力し、得られた出力 \mathbf{y} を CANN の出力とする。

3. 仮想ロボット

図 2 に制御対象である仮想ロボットのモデルを示す。仮想ロボットは幅および高さ 2m、長さ 4m の直方体が 5 個接続された構造である。各直方体は密度 1000 kg/m^3 の剛体であり、他の剛体および床面との接触によって破損や変形はしない。各直方体は隣接する直方体と可動ジョイントによって接続されており、 x 軸向きの回転軸に関して $[-\theta/4, \theta/4]$ rad の範囲で回転を行うことが可能である。また、図 2 の直方体を左から順に直方体 0, 1, 2, 3, 4 と呼ぶと、灰色で塗りつぶされた直方体 4 は頭部の役割を果たす部品であり、直方体 4 のみ重心位置の座標を計測することが可能である。

仮想ロボットの運動は物理エンジンによって計算される。物理エンジンは多体物理を数値積分によって処理するソフトウェアである。本研究では、物理エンジンの数値積分で用いる微小時間を 0.01 s 、重力加速度 g を 9.8 m/s^2 として運動計算を行う。

4. 倒立行動の獲得実験

4.1 実験方法

仮想ロボットが床に接地した状態を初期状態とし、頭部か

ら最も離れた直方体 0 の上方 19m に設置された光源に、頭部の重心を近づける行動を学習させる。

以降の説明では、CANN において高次 ANN の機能を固定化した制御器を領域分割型 ANN (RDANN) と呼ぶことにする。制御器として、一般的な 3 層フィードフォワード ANN および RDANN を用い、それぞれを制御器としたときの比較を行う。RDANN は 5 個の低次 ANN を持ち、頭部の重心の高度に応じて選択的に利用される。高度の範囲はそれぞれ 0m~2m, 2m~3m, 3m~4m, 4m~5m, 5m 以上、である。

ANN の入力は、頭部の姿勢情報、頭部と光源のなす角度情報、各ジョイント間の回転角度および回転速度からなる 13 次元ベクトルである。出力は各ジョイント間の目標回転角度であるが、急激な角度変化による異常な移動行動を防ぐため、0.01 s ごとに $\theta/180 \text{ rad}$ 以上の回転はできない設定とした。中間層のニューロン数は、RDANN の低次 ANN が 2 個、一般的な ANN は 2 個の場合と 10 個の場合の両方を比較対象として用いる。

評価関数 H は式(1)で与える。

$$H = \sum_{s=1}^T h(s), \quad h(s) = \frac{1}{0.01 d^2(s)} \quad (1)$$

ここに、 s は数値積分の時間ステップ、 d は頭部の重心と光源の距離である。また、 T は終了時間ステップであり、 $T = 3000$ とした。

以上の評価関数 H を最大化することで学習を行う。学習には最適化手法のひとつである粒子群最適化(Particle swarm optimization, PSO)⁴⁾を用いる。PSO の粒子数は 20 とし、学習は 500 ステップ実行する。

以上の実験条件で、全パラメータを同時に学習させる CANN (実験条件 a)、RDANN (実験条件 b)、中間層ニューロン数が 2 個の ANN (実験条件 c)、中間層ニューロン数が 10 個の ANN (実験条件 d) それぞれの結果を比較する。

4.2 実験結果

実験条件 a, b, c, d それぞれについて、各 1 回ずつ数値計算実験を行った。図 3 は、それぞれの最良解の評価値の推移である。図より、学習が終了した 500 ステップでは RDANN が最も高い評価を得ており、中間層ニューロン数が 10 個の ANN の評価もこれに近い値となっている。一方、CANN の全パラメータを同時に学習させた場合、および中間層ニューロン数が 2 個の ANN の評価は、200 ステップ付近からほとんど変わっておらず、局所解で学習が停滞したと推測できる。いずれも各 1 回ずつの試行であるため、今後は複数回の試行の平均を用いて結果の評価を行う必要がある。

最良解の行動がどの程度優れているかを調べるため、実験条件 a~d のうち最高の評価を得た RDANN の学習結果について数値データを取得した。図 4 は仮想ロボットの頭部の z 座標および y 座標の時間推移である。仮想ロボットは 10 秒までの間に光源付近まで移動を行い、その後は図 5 に示すように頭部を上方向に持ち上げる動作を行っている。ただし、1 回のシミュレーションが終了する 30 秒より少し前に、横方向に倒れてしまい頭部を持ち上げる行動が取れなくなった。ジョイントの可動軸を増やすなどして、仮想ロボットが横方向のずれを修正できる構造に変更する必要がある。

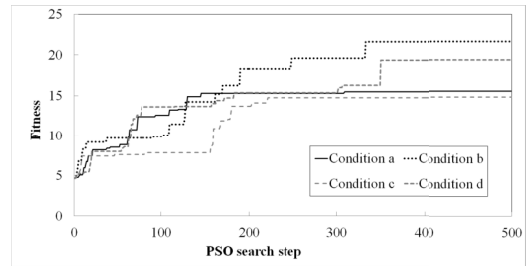


Fig.3 The best fitness for each PSO steps

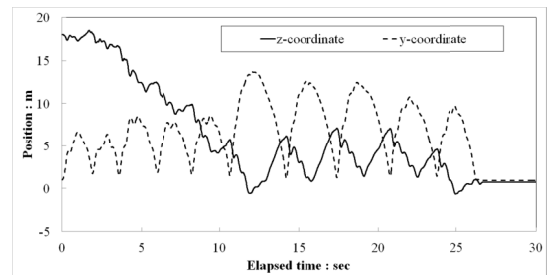


Fig.4 Position of the virtual robot for each time



Fig.5 Snapshot of the virtual robot

5. 結 言

本研究では、仮想ロボットの運動制御に CANN 型の構造を持つ制御器を用い、行動獲得実験を行った。各 1 回ずつの試行であるため統計的な結論は下せないが、実験の結果高次 ANN の機能を固定化させる RDANN が最も高い評価を得た。

最良の学習結果である RDANN を用いた仮想ロボットの一連の行動では、床面に対して水平な移動と、頭部の振り上げ行動が見られた。ただし、理想的な倒立状態は達成できず、シミュレーション終了直前には横方向に倒れて復帰できない状態となってしまった。

今後は、仮想ロボットの構造を再検討して転倒を防止するとともに、中間層ニューロン数などのパラメータを変更した実験を複数回行い、統計的な結論を導き出す必要がある。

参 考 文 献

- 1) X. Yao: Evolving Artificial Neural Networks, Proc. IEEE, 87, 9, (1999) 1423.
- 2) 米陀佳祐, 鈴木育男, 山本雅人, 古川正志: 複合型ビヘイビアによる円環構造弾性ロボットの自律行動獲得, 精密工学会誌, 77, 5 (2011) 476.
- 3) 大江亮介, 鈴木育男, 山本雅人, 古川正志: 複合的人工ニューラルネットワーク—教師なし学習を用いた動的制御における選択型ニューラルネットワークアンサンブル, 精密工学会誌, 79, 6 (2013) 552.
- 4) J. Kennedy and R. Eberhart: Particle Swarm Optimization, IEEE International Conference on Neural Networks, 4 (1995) 1942.