

B2-6 CMAC 学習アルゴリズムを用いた入力デバイスによる操作特性の学習

福井大学 工学部 知能システム工学科

水谷 謙介 (指導教官: 前田 陽一郎 助教授)

1 緒言

近年、ファジィ推論やニューラルネットワークなどのソフトコンピューティング手法を用いて、ロボットに人間のあいまいかつ柔軟な判断を有する制御戦略の表現や獲得を行うことを目標とする研究が盛んに行われている。

本研究では熟練オペレータの操作入力による教師信号を用いて、人間の制御戦略にあたる操作特性をニューラルネットワークの一手法である CMAC 学習アルゴリズム [1] を用いて獲得する手法を提案する。通常 CMAC 学習アルゴリズムは、1 つの制御パラメータを学習するものであるが、人間の操作特性を考えた場合、一般には学習する制御出力は複数となる。そこで本手法では、一般的な自律移動ロボットの人間による操作を想定し、速度と操舵角速度の 2 つの制御出力を獲得するために、2 操作特性を学習することが可能な CMAC 学習アルゴリズムを提案する。

さらに、本手法の有効性を検証するため、試作した RoboCup 中型ロボットリーグ規格のサッカーロボットを用いてジョイスティックによる人間の操作により、ボール追跡とドリブル行動における速度と操舵角速度の操作特性を学習する実験を行った。

2 CMAC 学習アルゴリズム

小脳皮質内の情報処理機構の数学的モデルとして、J.S.Albus により提案された CMAC (Cerebellar Model Arithmetic Computer) がよく知られている。ニューラルネットワークの一手法である CMAC 学習アルゴリズムは、複雑な入力空間マップの要素をいくつかの細胞 (セル) に分散記憶させることにより、効率的に学習できるアルゴリズムである。一般に、CMAC 学習アルゴリズムは以下に示すような特徴を持つ。

- 高い汎化性を持つ。
- 従来のニューラルネットワークに比べ、はるかに少ない記憶容量で足りる。
- 学習速度が比較的速い。
- 入力が多次元の入力マッピングに向いている。

3 2 操作特性を獲得するための CMAC 学習アルゴリズム

自律移動ロボットにおいて、一般に CMAC 学習アルゴリズムを用いて人間の操作特性を学習することを考えた場合、人間は速度や操舵角速度など複数の制御パラメータを同時に操作している場合が多く、通常 1 つの制御パラメータしか学習できない CMAC 学習アルゴリズムを複数パラメータを学習可能なアルゴリズムに改良する必要がある。そこで、本研究では速度と操舵角速度という 2 つの制御パラメータを同時に獲得する CMAC 学習アルゴリズムを提案する。

本手法における学習は以下のような処理手順で行われる。*Step1* ~ 4 を繰り返し、オペレータの操作入力による学習を繰り返す。そして学習終了後、CMAC マップはオペレータの操作特性を盛り込んだ戦略マップとして得られる。

1) *Step0*

センサ入力により得られる値を入力空間と仮定し、学習する制御パラメータはすべて 0 として初期マップの生成を行う。(本ステップは初回のみ)

2) *Step1*

時刻 t においてオペレータにより操作された制御パラメータの値を教師信号 $\hat{v}_t, \hat{\omega}_t$ として入力する。 $(\hat{v}$: 目標速度, $\hat{\omega}$: 目標操舵角速度)

3) *Step2*

以下の 2 式をともに満たしたとき学習を終了する。

$$|\hat{v}_t - v_t| \leq \varepsilon_v \text{ or } \frac{1}{10} \sum_{i=t-9}^t \left| \frac{\hat{v}_i - v_i}{\hat{v}_{i-1} - v_{i-1}} \right| \leq 1 - \delta_v \quad (1)$$

$$|\hat{\omega}_t - \omega_t| \leq \varepsilon_\omega \text{ or } \frac{1}{10} \sum_{i=t-9}^t \left| \frac{\hat{\omega}_i - \omega_i}{\hat{\omega}_{i-1} - \omega_{i-1}} \right| \leq 1 - \delta_\omega \quad (2)$$

v_t, ω_t は、それぞれ時刻 t における入力に対する現在の CMAC マップにおける速度と操舵角速度を示し、 $\varepsilon_v, \varepsilon_\omega, \delta_v, \delta_\omega$ はそれぞれ速度と操舵角速度における許容誤差を示している。通常の CMAC では式 (1), (2) の前半の式が終了判定として使用されるが、本手法では 2 制御パラメータの学習を対象としているため、学習による改善がなされなくなった場合も加えることにより、比較的制限を緩めた終了条件とした。

4) *Step3*

速度と操舵角速度のそれぞれの誤差修正値 Δ_v, Δ_ω を次式で算出する。

$$\Delta_v = g \frac{\hat{v}_t - v_t}{|A^*|} \quad (3)$$

$$\Delta_\omega = g \frac{\hat{\omega}_t - \omega_t}{|A^*|} \quad (4)$$

ただし、 g は学習の利得係数、 $|A^*|$ は発火した細胞の数である。

5) *Step4*

誤差修正値 Δ_v, Δ_ω を発火に関係したセルすべてに加え、CMAC マップを修正する。本ステップ終了後、*Step1* に戻る。

4 入力デバイスを用いたオペレータの操作特性の獲得

今回提案する手法のアルゴリズムフローを図 1 に示す。 d_{bt}, θ_{bt} はそれぞれ時刻 t におけるロボットから見

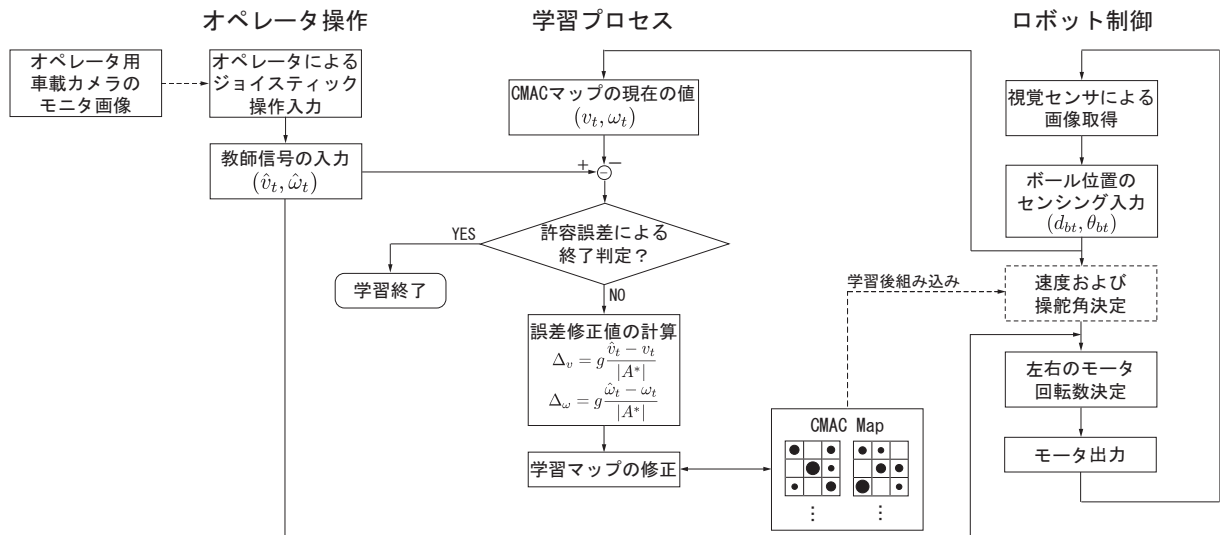


図 1. 学習のアルゴリズムフロー

たボールの距離と方位、 $\hat{v}_t, \hat{\omega}_t$ はオペレータの教師信号による速度と操舵角速度、 v_t, ω_t は学習マップにおける速度と操舵角速度を示している。なお、これらの値はロボットから見た相対座標系で表されており、ロボット正面を 0 とし、右側を正、左側を負とした。

図の左側はオペレータの操作に関するフロー、中央は学習プロセスのフロー、そして右側はロボット制御のフローを示している。ロボットの状態入力として、視覚センサによりボールの現在位置が計測される。さらに、それをロボット座標系から見た極座標系に変換する。また、オペレータは視覚センサとは別のオペレータ用の車載前方カメラのモニタ画像を見て、操作デバイスにより操作入力を行う。これを教師信号とし、視覚センサからのボールの位置情報を入力として CMAC 学習を行う。学習プロセスでは、現在のボールの状態に関する CMAC マップを、式 (3),(4) において計算した誤差修正値を用いて修正する。これを繰り返し、学習が終了した後、CMAC マップが完成すると、学習プロセスおよびオペレータ操作を切り離し、ロボット制御の速度および操舵角速度の決定部に CMAC マップを組み込めば、ロボットは自律的にオペレータの操作特性に基づいた行動を行う。

5 移動ロボットを用いた操作特性学習実験

本手法の有効性を検証するため、当研究室で試作した RoboCup 中型ロボットリーグ規格のサッカーロボットを用いて、オペレータのボール追跡およびドリブル行動における速度と操舵角速度の操作特性の学習実験を行った。オペレータの操作デバイスには 3 自由度ジョイスティック、オペレータ用カメラにはトランスミッタ内蔵カメラを用いた。また、ロボットの視覚センサには当研究室で開発した全方位カメラを 3 台用いて距離と方位を計測することが可能な MOVIS (Multiple Omnidirectional Vision System)[3] を使用した。MOVIS は全周 360 度にわたり計測が可能であるが、オペレータのモニタ画像は前方 60°しか見えていないため、今回の実験では MOVIS においても前方 60°のみの情報を用いた。

実験の様子を図 2 に、システム構成図を図 3 に示す。オペレータは、UHF で送信されたロボットに搭載された前方カメラのモニタ画像を見て、ロボットと無線 LAN により接続されているパソコンから、ジョイスティックを用いてロボットを操作する。そのオペレータ

のジョイスティック操作入力値を教師信号とし、視覚センサから得られたボールの距離と方位を入力として学習を行う。さらに学習終了後、得られた CMAC マップを搭載してロボットを制御し、ロボットがオペレータの操作特性どおりの行動を行っていることを検証する予定である。

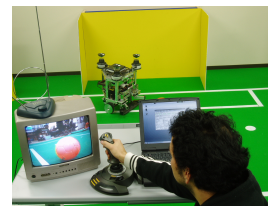


図 2. 学習実験の様子

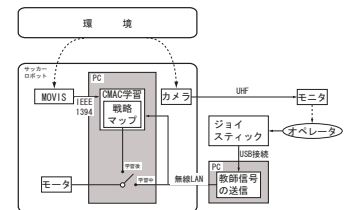


図 3. システム構成図

6 結言

本研究では、CMAC 学習アルゴリズムを用いた入力デバイスによる人間の操作特性を獲得する手法を提案した。試作したサッカーロボットを用いて人間のジョイスティックによる操作特性の学習実験も行ったが、この結果については卒研発表会の際に報告する。

今後の課題としては、今回、CMAC 学習の入力としてボールの位置情報だけを用いて行ったが、今後敵エージェント、障害物、ゴールなどの入力空間が複雑な環境を想定し実験を行っていきたいと考える。さらに、学習する制御パラメータの初期値を 0 ではなく、あらかじめ作成したファジールールなどを初期値とした学習実験を行い、本手法との学習時間の比較などを行っていく予定である。

参考文献

- [1] J.S.Albus, "Brains, Behavior and Robotics", McGraw-Hill (1981)
- [2] 前田陽一郎, 山中猛, "CMAC 学習に基づくオペレータの制御戦略を有するファジィ障害物回避制御," 第 6 回ファジィシステムシンポジウム, pp.531-534 (1990)
- [3] W.Shimizuhira, Y.Maeda, "Self-Localization Method Used Multiple Omnidirectional Vision System", SICE Annual Conference in Fukui, pp.2796-2799 (2003)