

A2-4 全方向移動ロボットによるCMAC学習を用いた ヒューマンスキル獲得に関する研究

福井大学大学院 工学研究科 知能システム工学専攻 進化ロボット研究室
加藤 進 (指導教員：前田 陽一郎)

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なうことを目標とする研究が盛んに行なわれている。島崎らはファジィ制御、自己組織化マップSOMを用いて車椅子型自動移動ロボットに搭乗者の操作法を学習させる手法を提案している [1]。田伏らはファジィ推論ニューラルネットワークを用いて人間が実行した操作データからテトリスのスキルを自動学習する手法を提案している [2]。本研究室においても、動的な環境においてロボットに自律的行動を学習させるために、動物の調教に用いられる Shaping の概念を取り入れた Shaping 強化学習を基に、調教者が報酬を適宜与えることで人間の望むロボットの制御戦略を獲得する手法を提案してきた [3]。しかしながら調教者である人間は常に学習過程を監視し、適切なタイミングで報酬を与えなければならないという問題点もあった。

そこで本手法では学習における人間の負担を軽減するための学習ゲイン調整手法と、4つの環境情報からオペレータの3つの操作特性を獲得するCMAC学習アルゴリズム [4] を提案する。ここでは提案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全方向移動機構を有する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なった [5]。

2. 全方向移動機構の走行制御

今回使用するサッカーロボットは、駆動部に3輪駆動RollerWheel方式の全方向移動機構を有し、本研究室で提案された全周にわたって物体の正確な距離と方位を計測することができる全方位カメラ3台で構成されたマルチ全方位ビジョンシステム(MOVIS)[6]を搭載している。図1は3輪駆動全方向移動機構の構造をモデル化したものである。車輪の半径を r 、ロボット

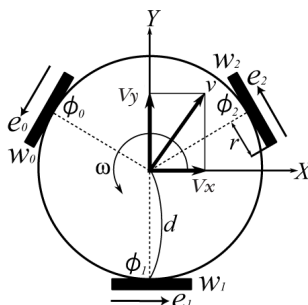


図1: 3輪駆動全方向移動機構のモデル

中心から車輪までの距離を d 、左前の車輪から時計まわり3つの回転角を ϕ_0, ϕ_1, ϕ_2 とし、各車輪の回転によって進む方向の単位ベクトルを $e_0 = (-\frac{1}{2}, -\frac{\sqrt{3}}{2})$, $e_1 = (1, 0)$, $e_2 = (-\frac{1}{2}, \frac{\sqrt{3}}{2})$ とする。

ロボット中心が速度 v で並進運動し、ロボットが ω ($\dot{\phi}$)で回転運動しているとすると、各車輪の回転速度とロボットの移動速度との関係式(1)が得られる。実際の制御ではジョイスティックでロボットを制御するにあたり、ジョイスティックのX軸方向、Y軸方向への傾き、Z軸方向へのスライド操作から得られる値をロボットの目標速度として (V_x, V_y, ω) に対応させた。この値をもとに式(1)により各車輪の回転速度が求められ、無線通信でロボットへ送ることでジョイスティックの信号がロボットの目標速度となる。

$$\begin{bmatrix} \dot{\phi}_0 \\ \dot{\phi}_1 \\ \dot{\phi}_2 \end{bmatrix} = \frac{1}{r} \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & d \\ 1 & 0 & d \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} & d \end{bmatrix} \begin{bmatrix} V_x \\ V_y \\ \omega \end{bmatrix} \quad (1)$$

3. 3操作特性を獲得するためのCMAC学習アルゴリズム

本研究で用いた速度(X方向とY方向)と操舵角速度という3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以下に示す。オペレータの操作入力による学習を繰り返し、学習終了後獲得された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(13×13 セル)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なう。(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 $\hat{v}_{xt}, \hat{v}_{yt}, \hat{\omega}_t$ として入力する。 (\hat{v}_x, \hat{v}_y) :各軸の目標速度, $\hat{\omega}$:目標操舵角速度)

3) Step2

時刻 t における環境からの入力情報に対する現在のCMACマップにおける速度と操舵角速度 v_{xt}, v_{yt}, ω_t と、オペレータによって与えられる教師信号 $\hat{v}_{xt}, \hat{v}_{yt}, \hat{\omega}_t$ を比較し、以下の式(2)~(4)をもとに学習ゲインを調節する。 $g_{v_x}, g_{v_y}, g_{\omega}$ 、 $\Delta g_{v_x}, \Delta g_{v_y}, \Delta g_{\omega}$ はそれぞれの出力の学習ゲインと修正値である。

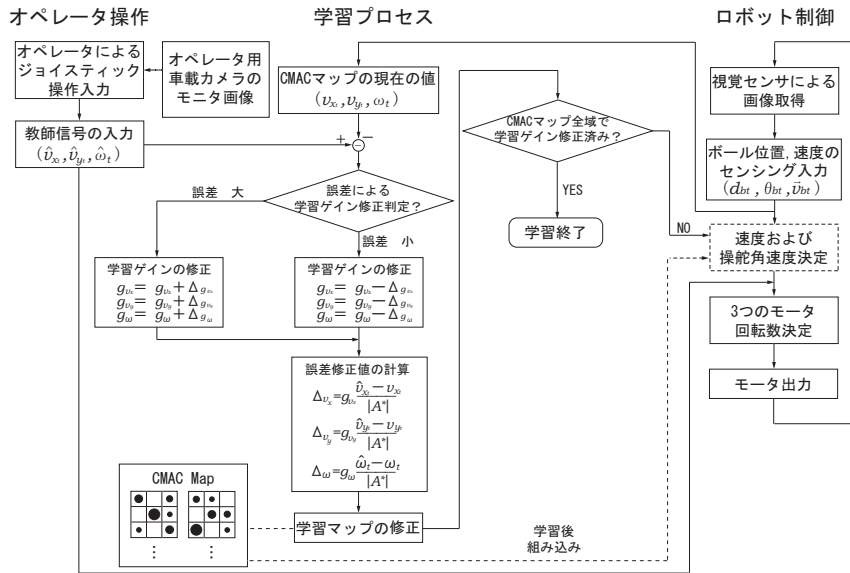


図 2: 学習のアルゴリズムフロー

$$\text{IF } \frac{1}{10} \sum_{i=t-9}^t |\hat{v}_{x_i} - v_{x_i}| \quad |\hat{v}_{x_t} - v_{x_t}| \quad (2)$$

$$\text{THEN } g_{v_x} = g_{v_x} + \Delta g_{v_x}$$

$$\text{ELSE } g_{v_x} = g_{v_x} - \Delta g_{v_x}$$

$$\text{IF } \frac{1}{10} \sum_{i=t-9}^t |\hat{v}_{y_i} - v_{y_i}| \quad |\hat{v}_{y_t} - v_{y_t}| \quad (3)$$

$$\text{THEN } g_{v_y} = g_{v_y} + \Delta g_{v_y}$$

$$\text{ELSE } g_{v_y} = g_{v_y} - \Delta g_{v_y}$$

$$\text{IF } \frac{1}{10} \sum_{i=t-9}^t |\hat{\omega}_i - \omega_i| \quad |\hat{\omega}_t - \omega_t| \quad (4)$$

$$\text{THEN } g_{\omega} = g_{\omega} + \Delta g_{\omega}$$

$$\text{ELSE } g_{\omega} = g_{\omega} - \Delta g_{\omega}$$

4) Step3

速度と操舵角速度のそれぞれの誤差修正値 $\Delta v_x, \Delta v_y, \Delta \omega$ を次式で算出する。 $|A^*|$ は発火したセルの数である。

$$\Delta v_x = g_{v_x} \frac{\hat{v}_{x_t} - v_{x_t}}{|A^*|} \quad (5)$$

$$\Delta v_y = g_{v_y} \frac{\hat{v}_{y_t} - v_{y_t}}{|A^*|} \quad (6)$$

$$\Delta \omega = g_{\omega} \frac{\hat{\omega}_t - \omega_t}{|A^*|} \quad (7)$$

5) Step4

誤差修正値 $\Delta v_x, \Delta v_y, \Delta \omega$ 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 2 に示す。 $d_{bt}, \theta_{bt}, v_{bt}$ はそれぞれ時刻 t におけるロボットから見たボールの距離と方位、ボールとロボットの相対速度である。 $\hat{v}_{x_t}, \hat{v}_{y_t}, \hat{\omega}_t$ はオペレータの教師信号によ

るロボットの目標速度と目標操舵角速度、 $v_{x_t}, v_{y_t}, \omega_t$ は学習マップにおけるロボットの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、左側を負としている。

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。そしてボールの位置情報からボールの速度情報としてボールの速さと向きを得る。また、オペレータは直接ロボットや目標物を見ず車載 CMOS カメラのモニタ画像のみを見ながら操作指令を与える。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 学習を行なう。

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 (5),(6),(7) 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、オペレータの設定した時間 t_{end} まで学習を行い、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実現できる。

5. 全方向移動ロボットを用いた操作特性学習実験

全方向移動機構を有する 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 CMAC 学習実験を行なった。以下に実験方法と結果、考察を示す。

5.1 実験方法

CMAC 学習アルゴリズムで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を学習させた。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を用いる。また、ロボットの視覚センサには MOVIS を使用し、ボールの相対方位と

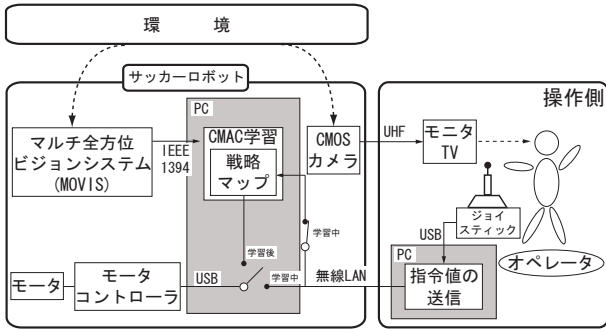


図 3: 実験システム構成図

相対距離を得る。MOVIS は全周 360 度にわたり物体の距離と方位を計測可能であるが、今回の実験ではオペレータが使用する車載カメラ映像がロボットの前方 60°しか見えないため、MOVIS も前方 60°のみの情報を用いた。

実験時のシステム構成図を図 3 に示す。オペレータは車載カメラ映像を見ながら、ロボットに搭載してある制御用パソコンと無線 LAN によって接続されている操作用パソコンからジョイスティックを用いてロボットを遠隔操作する。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なう。今回の実験では「8 の字回り」のボール追跡行動を学習させた。RoboCup 中型ロボットリーグを想定したミニチュアフィールド (6 m × 4 m) 内で実験協力者にボールを転がしてもらい、オペレータがロボットを操作して追跡した。

5.2 実験結果

図 4~6 に gain=0.5 で固定した時と提案した学習ゲイン調整手法による可変学習ゲインの時を比較したグラフを示す。gain=0.5 というのは様々な固定ゲインで事前実験を行い最も良い性能を示したゲインである。これらの図は実験によって得られた誤差を対数近似したグラフである。どのグラフも学習ゲイン調整手法を用いた方が誤差が減少している。

図 7~12 に学習後の CMAC マップをボールの速度 0~50cm/s でボール方位-30~-150°の時 [ロボット正面左側] と 30~150°の時 [ロボット正面右側] について示す。distance がボールの距離、angle がボールの方位 (°)、Z 軸の V_x, V_y が速度 (m/s)、 $\dot{\theta}$ が操舵角速度 (rad/s) である。速度 V_x の軸は負が左移動、正が右移動の速度を表している。速度 V_y の軸は負が前進方向、正が後退方向の速度を表している。操舵角速度は正の値のとき右旋回、負の値のとき左旋回を表している。ボール方位-30~-150°はボールがロボットの前を左に横切る向きに動いている事を表し、ボール方位 30~150°はその逆を示す。

グラフを見てみると、ボール方位-30~-150°の図 8、9 はボールが左側にある時出力の絶対値が大きくなっており、ボール方位 30~150°の図 11、12 はボールが右側にある時大きくなっている。人間がボールの動きを見て予測的にロボットを操作したことを学習できていることがわかった。

今回は学習環境を簡単にするためにボールがマシンから一定距離以上離れないよう、角度的にもあまり見失

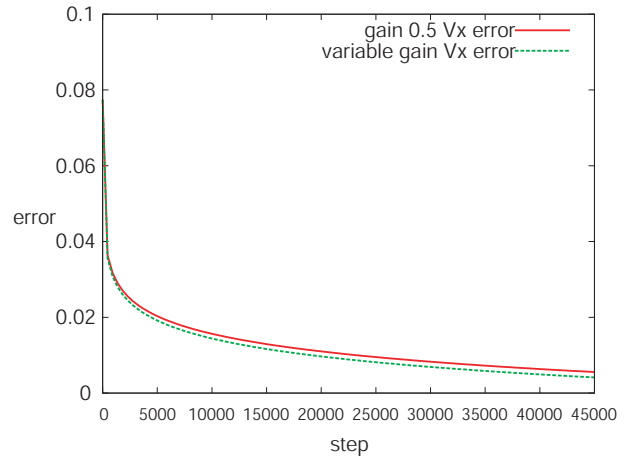


図 4: 固定学習ゲインと学習ゲイン調整手法を用いた実験結果 (V_x についての比較)

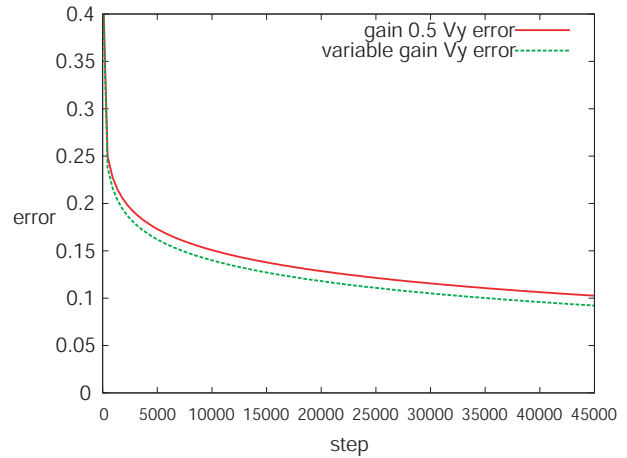


図 5: 固定学習ゲインと学習ゲイン調整手法を用いた実験結果 (V_y についての比較)

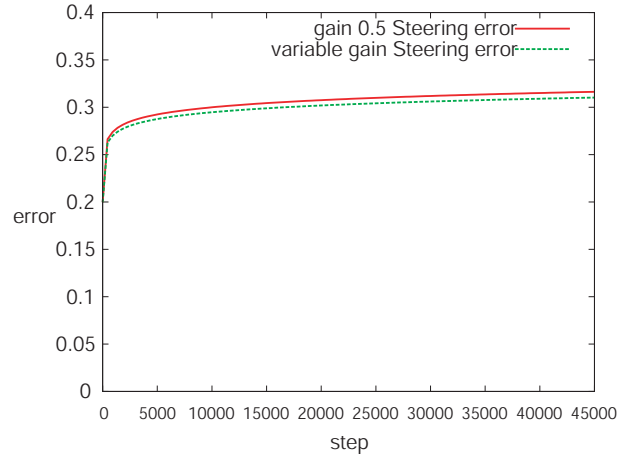


図 6: 固定学習ゲインと学習ゲイン調整手法を用いた実験結果 ($\dot{\theta}$ についての比較)

われないように追跡させた。そのため 3m より先のボールを追うような学習を行なわなかったため、CMAC マップにおいて distance が 200cm 以上のあたりはあまり活性化されなかったものと思われる。図 6 のみグラフが右肩上がりで誤差が増えているが、オペレータが見るカメラ情報からはボールとの距離をつかみにくいため旋回タイミングが掴みにくかったからである。

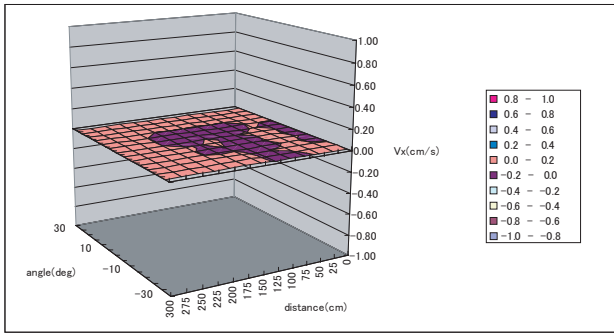


図 7: 学習後の CMAC マップ (45000 試行:Vx)
[ボール速度 0 ~ 50cm/s、ボール方位-30 ~ -150 °]

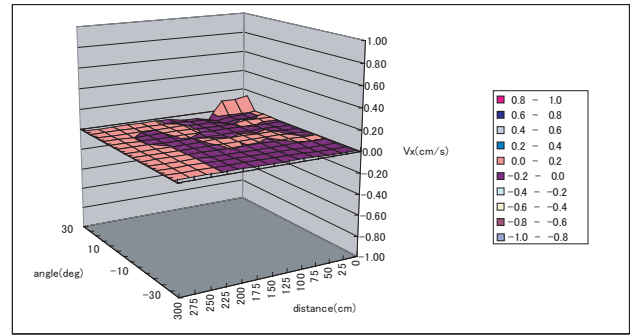


図 10: 学習後の CMAC マップ (45000 試行:Vx)
[ボール速度 0 ~ 50cm/s、ボール方位 30 ~ 150 °]

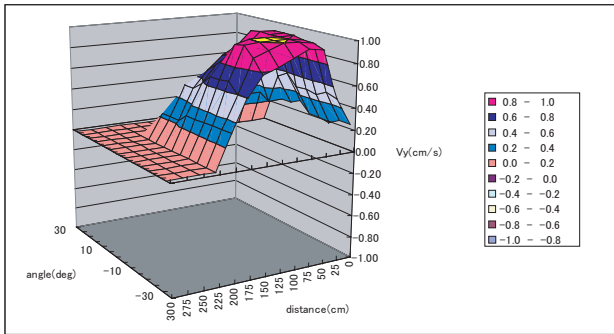


図 8: 学習後の CMAC マップ (45000 試行:Vy)
[ボール速度 0 ~ 50cm/s、ボール方位-30 ~ -150 °]

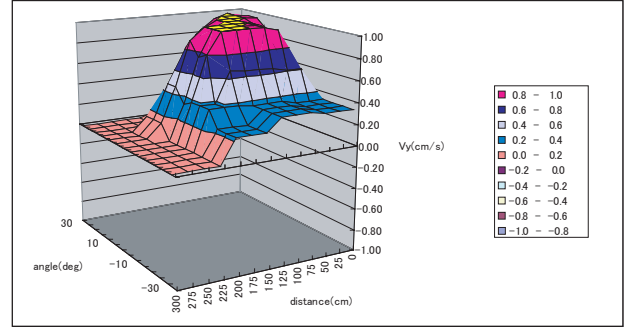


図 11: 学習後の CMAC マップ (45000 試行:Vy)
[ボール速度 0 ~ 50cm/s、ボール方位 30 ~ 150 °]

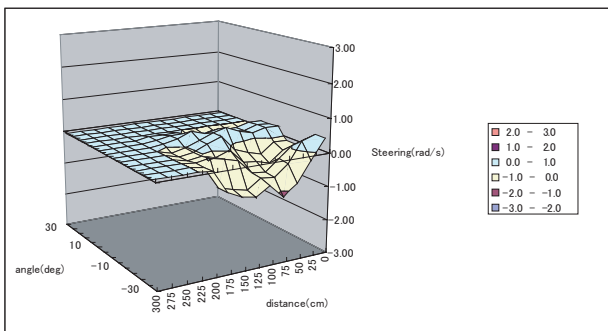


図 9: 学習後の CMAC マップ (45000 試行:)
[ボール速度 0 ~ 50cm/s、ボール方位-30 ~ -150 °]

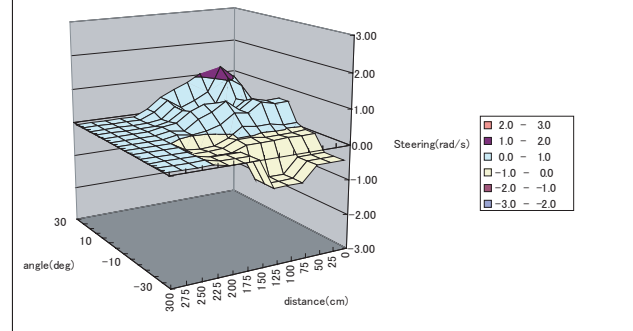


図 12: 学習後の CMAC マップ (45000 試行:)
[ボール速度 0 ~ 50cm/s、ボール方位 30 ~ 150 °]

6. 結言

本研究では全方向移動機構を有するロボットを用い、学習ゲインの自動調整手法を組み込み込んだ。3つの操作特性を獲得する CMAC 学習アルゴリズムを提案し、検証実験を行なった。その結果、オペレータの操作をロボットの動きに制約されることなく伝えることができ、操作特性を獲得したデータを取得することができた。学習ゲインを変えることで誤差の値が変わることが分かり、可変学習ゲインにすることで誤差一定のままでは時間をかけなければ到達できない状態まで誤差を減らすことができ学習時間を短縮できた。

今後はロボットの動きと目標物の動きを別の視点から検証し、よりオペレータの操作特性を反映させることができる実験を行なっていく予定である。

参考文献

[1] 島崎尚史, “自己組織化マップを用いた車椅子型自動移動ロボット DREAM-4 の走行制御に関する研

究,” 第 8 回自己組織化マップ研究会 2007 講演論文集, pp.31-34 (2007)

[2] 田伏未来, 萩原将文, “ファジィ推論ニューラルネットワークを用いたテトリスのスキル獲得のための自動学習,” 日本ファジィ学会誌, Vol.11, No.6, pp.1089-1097 (1999)

[3] 長谷川大樹, 前田陽一郎, “ファジィ状態分割型 Shaping 強化学習を用いたサッカーロボットの戦略獲得,” 第 25 回ファジィシステムシンポジウム, CD-ROM, 2A2-03 (2009)

[4] J.S.Albus, “Brains, Behavior and Robotics,” McGraw-Hill (1981)

[5] 加藤進, 前田陽一郎, “全方向移動ロボットの人間による操作特性の CMAC 学習,” 第 25 回ファジィシステムシンポジウム, CD-ROM, 2A2-02 (2009)

[6] 清水平涉, 前田陽一郎, “マルチ全方位ビジョンを用いた自律移動ロボットのファジィナビゲーション,” 第 19 回ファジィシステムシンポジウム講演論文集, Vol.19, pp.577-578 (2003)