

強化学習的手法を用いての視覚サーボに基づいた移動ロボットのホーミング行動の実現

Acquisition of Homing Behaviour of a Mobile Robot Based on Reinforcement Learning.

大阪大学 金石貴志, 高橋泰岳, 鈴木昭二, 浅田稔

Takashi Kaneishi, Yasutake Takahashi, Sho'ji Suzuki and Minoru Asada
Osaka University

Abstract Homing behaviour is one of the primitive behaviour of a mobile robot. We apply Parti-Game Algorithm which is a kind of reinforcement learning for autonomous acquisition of robot's behaviour. We evaluate the proposed method by computer simulation.

1 はじめに

自律的に移動するロボットにとって、環境中の任意の位置から特定の位置に戻るタスク(ホーミング)は重要である。ホーミングはバッテリーチャージや学習時の初期位置への移動、サッカーにおけるホームポジションへの移動のタスクを行なう基本行動である。

従来のホーミングは正確な環境の地図に基づき、センサを用いて精密にロボットの現在位置を計測し、地図との対応をとることにより実現されてきた。これらの手法は地図の作成、位置計測にコストがかかり、地図や位置の誤差による影響を受けやすいという問題があった。これに対して、センサから得られる情報を基にしてホーミングを行なう手法がある。特に視覚を用いることにより複雑な環境下でロボットのホーミングを行なうことが期待できる。

視覚に基づいてのホーミングを行なう場合、戻るべき特定の位置は、画像パターンにより示される。ロボットはカメラに写った画

像パターンと目標画像パターンの画像特徴量の差を少なくするように移動し、ホーミングを実現する。ただし、移動機構の制約によりホーミングに必要な移動と画像特徴量の変化を対応づけるのは容易ではない

そこで、本研究では強化学習的手法を用い、ロボットに経験を通して自律的な学習を行なわせる。その際、外界センサとして単眼カメラを用い、ホーミングを行なうというタスクを与える。そのための手法として強化学習的手法のひとつである Parti-Game Algorithm¹⁾を用い、シミュレーションを通して視覚を持ったロボットのホーミング行動獲得の際の問題点について検討する。

2 視覚を持った移動ロボットのホーミングの実現

2.1 ロボットの構成

台車は4輪駆動、4輪操舵の車輪移動型の台車を想定しているが、基本的には1駆動1ステアリング方式と同じである。すなわち、

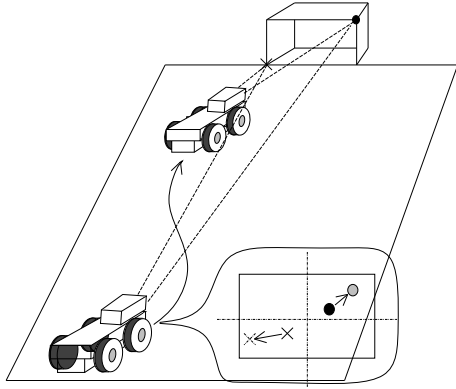


図 1: ホーミング行動

ロボットの移動は前後進，左右旋回，静止という行動ができる．外界センサとして単眼カメラを用い，カメラは床面に水平に取り付ける．ロボットはサッカーゴールの高さ，カメラの高さ，焦点距離，ロボットの速度，回転半径を知っているものとする．

2.2 Parti-Game Algorithm

Parti-Game Algorithm は，ロボットが必要に応じて状態空間を分割しながらゴール状態に達する行動を獲得するアルゴリズムである．ここでの状態空間はセンサで得られた画像特徴量で張られた空間であり，状態は状態空間を分割した一領域である．ロボットは局所コントローラを持っており，これによって隣接する状態への遷移を試みる．ただし，必ずしも目的の状態に移動できる保証はなくてもかまわない．

ロボットは $OUTCOMES(i,j)$ ， $NEIGHS(i)$ というテーブルを持っている． $NEIGHS(i)$ の出力は状態 i に隣接する状態のセット， $OUTCOMES(i,j)$ の出力は状態 i において状態 j に向かって遷移しようとしたときに結果として遷移する可能性のある次状態のセットである．これは，経験によって獲得されるものであり，

経験がなければ目的の遷移は可能だという仮定をおく．

以上の定義により，最悪のケースを考えた時のゴールまでの最小状態遷移数 $J_{wc}(i)$ は以下のようなになる．

$$J_{WC}(i) = 1 + \min_{j \in NEIGHS(i)} \max_{k \in OUTCOMES(i,j)} J_{WC}(k)$$

ここで， k は状態 i において状態 j に向かって遷移しようとしたときにたどり着いた状態である．

$i = Goal$ の場合は $J_{WC}(i) = 0$ とする． $J_{wc}(i)$ が 0 でなければ $J_{WC} \neq \infty$ となり，このとき，状態 j を目標状態として移動する．ゴールにたどり着けない状態であれば $J_{WC} = \infty$ となるので，状態を分割し，再び $J_{WC}(i)$ を計算する．分割は特徴量ベクトルの最長の軸に対して領域を 2 分割する．

$J_{wc}(i)$ が求まると，次式を満たす状態 j に向かって進む行動をとる．

$$j = \arg \min_{j \in NEIGHS(i)} \max_{k \in OUTCOMES(i,j)} J_{WC}(k)$$

2.3 局所コントローラ

Parti-Game Algorithm を用いるために，ロボットに局所コントローラを与える．局所コントローラは目標画像特徴量と現在の画像特徴量の差を最小とするモータコマンドを出力するコントローラである．これによって目標となる状態への遷移を試みる．

2.4 ホーミングの実現

Parti-Game Algorithm を用いて，ホーミングの実現を考える．特徴点としてサッカーゴールの右上と左下の 2 点をとる．シミュレーション 1 として画像平面上に投影されたサッカーゴールの右上の点の位置 (x_1, y_1) と左下

の点の位置 (x_2, y_2) の4つの特徴量ベクトルをとる．シミュレーション2として画像平面上に投影されたサッカーゴールの右上と左下の midpoint の位置，傾き (x, y, α) の3つの特徴量ベクトルをとる．

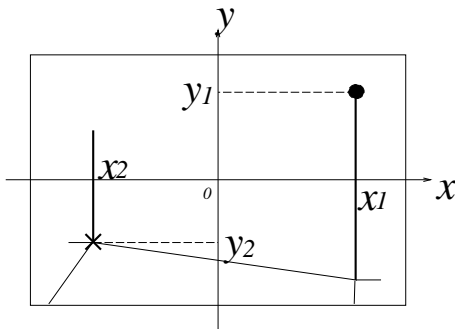


図 2: 状態ベクトルの取り方(状態空間が4次元)

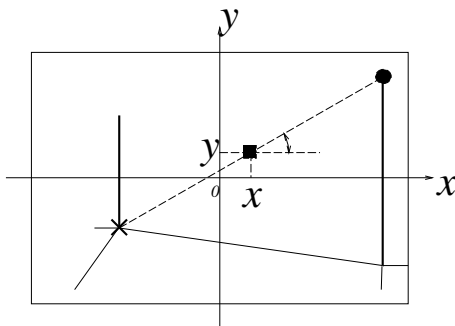


図 3: 状態ベクトルの取り方(状態空間が3次元)

これらの画像平面上の特徴量ベクトルを基にして状態空間を構成し，状態の遷移，分割を行ないながら，ゴール状態への到達を試みる．

3 シミュレーション結果

シミュレーションの結果を図4に示す．(a)はゴール状態への到達回数，(b)は状態数，(c)

は分割によって作られた中間目標状態への到達回数を表す．

4 考察

シミュレーションの結果，状態空間を3次元にした場合(シミュレーション2)は学習が進むにつれゴール状態への到達回数が増加した．状態空間を4次元にした場合(シミュレーション1)は学習が進んでもゴール状態への到達回数が増加しなかった．

状態数及び中間目標状態への到達回数を比べるといずれの場合も状態数は同様に増加する．しかし，中間目標状態への到達回数はシミュレーション2は増加するのに対しシミュレーション1では変化が見られなかった．本来ならば状態数の増加に伴い中間目標状態への到達回数も増加するはずであるが，シミュレーション1では望ましい結果が得られていない．その原因としては4次元における特徴量ベクトルがロボットの行動に適しておらず，状態空間の分割がうまくいかないためだと考えられる．その結果，ロボットはゴール状態へは到達できず後退を繰り返していた．それに対し，3次元では特徴量ベクトル (x, y, α) がロボットのフィールド上の座標に対応している．すなわち， y がゴールへの距離， x がロボットの向いている方向， α がゴールとの角度に対応している．そのため，分割された状態空間をたどることによってゴール状態へたどりつくことができている．

5 おわりに

本報告では強化学習の一手法である Parti-Game Algorithm を車輪を持った移動ロボットに適用することによって，自律的ホーミング行動を獲得させた．今後の課題として，

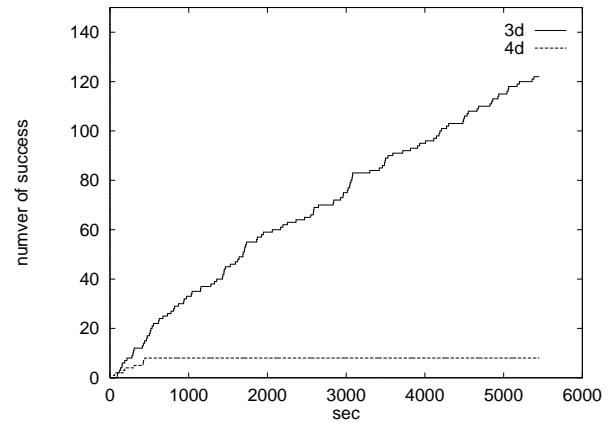
- 特徴量ベクトルの設定

- 状態空間の分割の方法
- 実機における検証

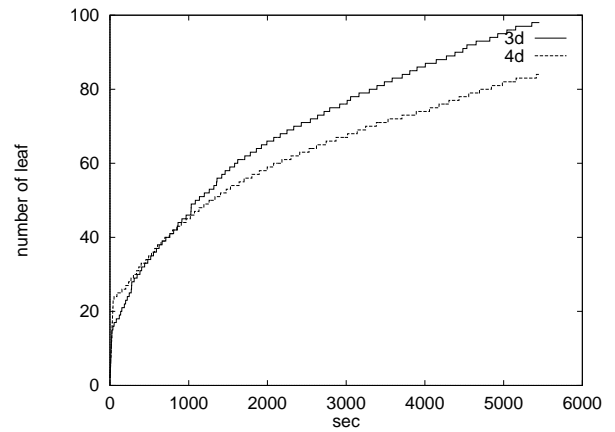
があげられる .

参考文献

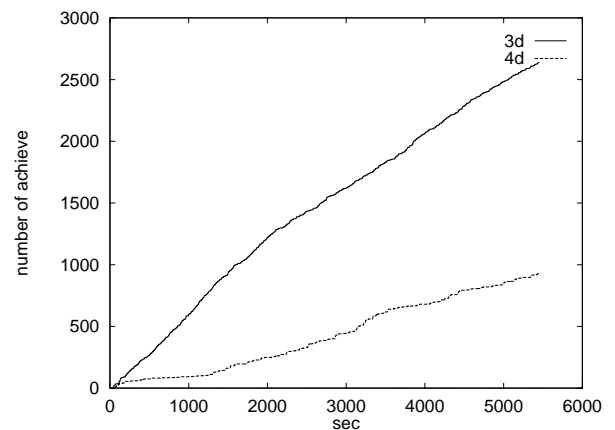
1) A.W.Moore and C.G.Atkeson: Parti-Game Algorithm for Variable Resolution Reinforcement Learning in Multidimensional State-spaces, Machine Learning, Vol21, pp.199-233 (1995)



(a) ゴール状態到達回数



(b) 状態数



(c) 中間目標状態到達回数

図 4: シミュレーション結果