

注視制御を組み込んだ全方位視覚を持つ移動ロボットの 強化学習による行動獲得

Behavior acquisition by reinforcement learning of a mobile robot with an omnidirectional vision and an attention control

加藤 龍憲 (阪大) 正 鈴木 昭二 (阪大) 正 浅田 稔 (阪大)

Tatsunori KATO, Osaka University, 2-1, Yamadaoka, Suita, Osaka

Sho'ji SUZUKI, Osaka University

Minoru ASADA, Osaka University

Abstract: The authors apply Q-learning, one of major method of reinforcement learning, for a mobile robot with an omnidirectional vision. We propose an attention control for the omnidirectional vision by controlling the focal length of the camera in order to reduce learning time. The robot learns a goal keeping behavior of soccer, and the result is shown by simulation.

Keywords: behavior acquisition, reinforcement learning, omnidirectional vision, attention control, keeper

1 はじめに

ロボット自身が経験を通じて合目的な行動を獲得するための手法として強化学習が注目されている¹⁾。強化学習の枠組では、ロボットは観測した環境情報に基づいて自身のとるべき行為を決定し、経験を通じて行動を獲得する。著者らは、視覚を持った単体の移動ロボットに強化学習の代表的な手法であるQ学習を適用し、ボールをゴールにシュートする行動を獲得させた²⁾。しかしながら、ロボットに搭載されたカメラの視野が狭いため環境のごく一部しか見えず、転がるボールを見失いやすくシュートの成功率は高くはなかった。

そこで、著者らはロボットの視覚に注視制御と全方位視覚を導入した上でQ学習によりロボットの行動獲得を実現する。従来、全方位視覚は、全方位画像からのシーンの生成⁴⁾、移動ロボットのナビゲーション⁵⁾⁶⁾などに利用されており、主に三次元のシーンや幾何モデルの再構成が行われてきた。著者らはこのような再構成は行わず全方位視覚の画像情報を直接利用しロボットにQ学習を適用する。また、全方位画像において取り扱われてこなかった注視制御の実現に取り組む。

本報告では、焦点距離の制御による全方位視覚のための注視制御を提案する。提案した手法を移動ロボットに組み込みQ学習によりサッカーにおけるゴール守備行動を獲得させる。

2 ロボットの設定

図1(a)に著者らの移動ロボットを示す。ロボット本体は、PWS(Power Wheeled Steering)方式の移動機構を持ったラジコンを利用しており、これに全方位視覚を搭載している。ロボットの大きさは、幅360[mm]長さ300[mm]高さ350[mm]であり、ロボットの最高速度は4.8[m/s]である。ロボットの行為としては図1(b)に示す8通りの行為と停止行為が予め実装されており、ホストコンピュータから無線を通じて実行される。

全方位視覚は、双曲面鏡とCCDカメラにより構成され、ロボットの周囲にある物体は図2(a)に示すように双曲面鏡を通じてカメラの撮像面上に投影される。図2(b)に全方位視覚により得た画像の例を示す。全方位視覚は、ロボットが回転しても視野の中心が変化しないように、視野の中心がロボットの中心に一致するように取り付けら

れている。

ロボットに与えるタスクは、サッカーにおけるゴール守備行動を扱い、これをQ学習によりロボットに獲得させる(図??)。環境は、幅4575[mm]長さ4110[mm]の水平な平面を仮定する。ゴールの大きさは幅1500[mm]高さ600[mm]とする。環境内にはゴールとボールのみが存在し、他のロボットは敵味方を含めて存在しないものとする。画像上ではゴールとボールは色により識別できるものとし、それぞれ青と赤に色分けされている。また、ボールの大きさは直径200[mm]で床面上を転がる。

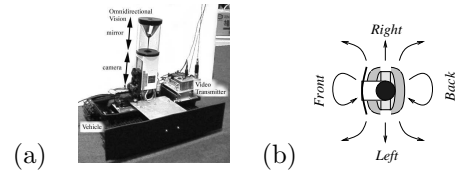


Fig.1 The robot system

3 全方位視覚における注視制御の実現

全方位視覚および観測する物体は水平面上に存在すると仮定し、座標系および各種パラメータを図3のように定義する。この時、実世界内の点 $P(R, \theta, Z)$ は撮像面上の点 $p(r, \theta)$ に投影され、それらの間には以下の関係が成り立つ。

$$\begin{aligned} Z &= R \tan \alpha + c + h \\ \tan \gamma &= \frac{b^2 + c^2}{b^2 - c^2} \tan \alpha + \frac{2bc}{c^2 - b^2} \frac{1}{\cos \alpha} \\ r &= \frac{f}{\tan \gamma} \end{aligned} \quad (1)$$

ここで、 a, b は双曲面の式 $\frac{R^2}{a^2} - \frac{Z^2}{b^2} = -1$ を決めるパラメータで $c = \sqrt{a^2 + b^2}$ の関係がある。ここで用いた全方位視覚では、 $\alpha^2 = 233.3[mm]$ 、 $\beta^2 = 1135.7[mm]$ 、 $h = 250[mm]$ である。

一般の視覚システムにおける注視は、注視する対象物を画像中心のように画像内の特定の位置に見続けるように、

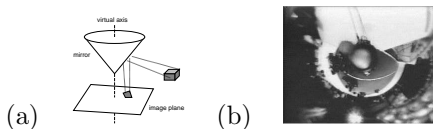


Fig.2 A sample of an omnidirectional vision

カメラのパンやチルトを制御することにより実現される。全方位視覚においては、対象物と画像の対応がとりにくくなるためにカメラを動かして注視を実現することは困難である。そこで、著者らは、焦点距離を変化させることにより注視する対象物を画像中心から一定の距離にある円周上に見続ける方法を提案する(図4)。画像中心からの距離 r を一定とした時、物体までの距離 R とカメラの焦点距離 f の関係は式(1)より求められ、物体までの距離が変化しても焦点距離を制御することにより画像中心からの距離を一定に保つことが可能である。ここで、 u_f をカメラの焦点距離の入力値、 l_{rd} を注視物体の画像上での目標値(注視円半径)、 l_r を現在の画像上の位置とし、注視制御の制御則を $u_f = K(l_{rd} - l_r)$ と定める。例えば、図4に示すように、 l_r が小さい場合には f を大きくすることにより l_{rd} に近付けることができる。

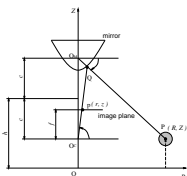


Fig.3 The projection of the hyperbolic mirror

4 学習による行動の獲得

Q学習を用いてロボットに行動を獲得させるためには、ロボットの得られる視覚情報から行動に対して有効な特徴に着目し、状態を定義しなければならない²⁾。ゴール守備行動獲得のための状態を表1の通り定義した。

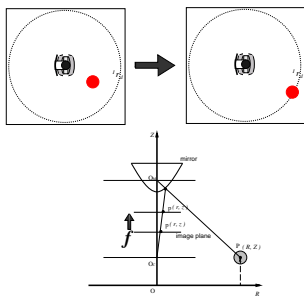


Fig.4 Attention control on the omnidirectional vision

注視制御を行わない場合は、ボールの画像中心からの距離と画像上の方位(図5)、およびボールの移動方向を知るために距離と方位のそれぞれに関する変化を状態とした。ゴールに関しては方位をボール同様とし、距離を2状態(遠い, 近い)と定義した。その結果、注視制御を行わない場合の総状態数は、 $\times 8 \times 3 \times 3 \times 3 \times 8 \times 2 = 3456$ となった。

注視制御を行う場合は、注視制御により画像中心からのボールの距離は一定となるため、状態を定義する必要がなくなる。そのため、ボールに関しては画像上での方位と方位の変化に対してのみ状態を定義すればよく、学習のための状態を削減することができる。ゴールに関して

は注視制御なしの場合と同様に距離と方位が必要である。したがって、総状態数は、 $8 \times 3 \times 2 \times 8 = 320$ となった。

	注視制御なし	注視制御あり
画像上でのボールの方位	8	8
ボールの方位の変化	3	3
画像中心からのボール距離	3	—
ボールの距離の変化	3	—
画像上でのゴールの方位	8	8
画像中心からのゴール距離	2	2
総状態数	3456	320

Table 1 Definition of substates

5 実験結果

シミュレーションにより、ロボットにゴール守備行動を学習させた。ロボットの初期値はゴール前のある空間、ボールの初期値はゴール中心とコーナを半径とする円周上に配置した。ボールはゴール目がけてロボットの最高速度の半分の速度で減衰することなく転がる。一試行は、ボールがゴールに入った場合とボールがフィールド外に出た場合に終了する。報酬はボールがゴールに入った場合には負の報酬を与え、ゴール守備をする上で最悪の行動を選択しないように学習を行わせた。

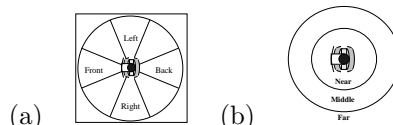


Fig.5 Substates

注視制御を行う場合と行わない場合のシミュレーション結果を図6に示す。図は、10000試行, 20000試行, 30000試行, 4000試行後の学習結果によるシュートの阻止率を示す。これより、注視点制御を行いながら学習した場合の方がより少ない試行で行動が獲得できることが示された。

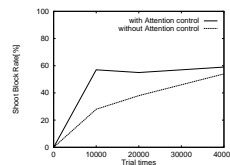


Fig.6 Result

6 まとめ

本稿では、全方位視覚に焦点距離の制御による注視制御を実現し、これを搭載した移動ロボットにサッカーのゴール守備行動を学習させた。その結果、注視制御が状態空間の削減に貢献し学習の時間を短縮できる場合があることが示された。しかしながら、今回の結果は一つのケーススタディであり、獲得すべき行動と学習に効果的な注視制御の関係は明らかではない。今後、この点を明らかにしていくことが課題である。

参考文献

- [1] Connel, J. H., Mahadevan, S.: Robot Learning. Kluwer Academic Publishers (1993)
- [2] 浅田裕・野田影一・依積田健・細田耕: “視覚に基づく強化学習によるロボットの行動獲得”, 日本ロボット学会誌, Vol.13, No.1, pp.68-74, 1995.
- [3] Y. Yagi and S. Kawato. Panoramic scene analysis with conic projection. In Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems 1990 (IROS'90), 1990.
- [4] V. N. Peri and S. Nayar. Generation of perspective and panoramic video from omnidirectional video. In Proc. of 1997 Image Understanding Workshop, pages 243-245, 1997.
- [5] K. Yamazawa, Y. Yagi, and M. Yachida. Obstacle avoidance with omnidirectional image sensor hyperomni vision. In Proc. of IEEE Int. Conf. on Robotics and Automation, pages 1062-1067, 1995.
- [6] H. Ishiguro. Distributed vision systems: A perceptual information infrastructure for robot navigation. In Proc. of IJCAI-97, pages 36-41, 1997.