

移動ロボットの行動学習に関する研究

On Behavior Learning of a Mobile Robot

○釜谷 祥弘, 羅 志偉, 渡部 慶二, 遠藤 茂

○Yoshihiro Kamatani, Zhiwei Luo, Keiji Watanabe, Shigeru Endou

山形大学 大学院 理工学研究科

Yamagata University

キーワード: 移動ロボット (Mobile Robot), 行動学習 (Behavior Learning), 予測 (Prediction), 動的環境 (Dynamic Environment)

連絡先: 山形県米沢市城南4-3-16

釜谷 祥弘, Tel 0238-26-3327, E-mail kamatani@ewata.yz.yamagata-u.ac.jp

1 はじめに

本研究室では、移動ロボットによる運動障害物の回避問題や、複雑な摩擦環境における移動ロボットの静止物体に対するリリース行動の学習問題を研究してきた。本研究報告はさらに、移動ロボットによる運動するボールのゴールイン行動学習問題を取り上げる。具体的なボールおよびロボットの動きを図1に示す。

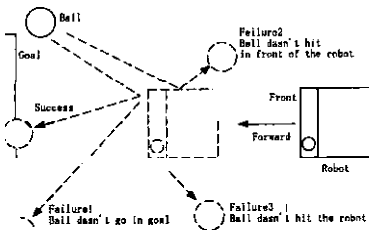


図1: ロボットによるボールのゴールイン行動

Fig.1 A Mobile Robot shoot a ball

ここでは、ボールの運動をCCDカメラで計測し、ボールの運動予測を行う。また、ロボットによるボールのシュート戦略として、ある決められた時刻におけるボールの到達位置を予測する方法や、決められたシュート位置におけるボールの到達時間を予測し、シュート行動を実行するという二種類の方式が考えられるが、本報告では、主に時間予測の戦略を用いる。この問題

において、ボールの重畳、初期状態が未知であるばかりでなく、ボールが転がるグラウンドの摩擦分布は非線形であることや、カメラ観測と画像処理に際して計測されたボールの位置情報に相当のノイズが含まれること、さらには、ロボットによるボールへの作用は一回の弾性衝突しかできないことから、不確かさ、非線形の強い複雑な環境におけるロボットの行動学習問題であることがわかる。この問題のブロック線図表現を図2に示す。

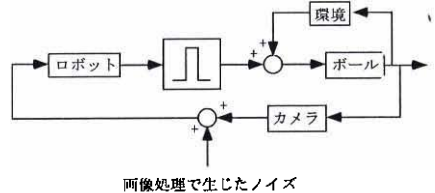


図2: 移動ロボットによるボールのシュート行動のブロック線図

Fig.2 Block diagram of a robot shoot a moving ball

本研究では、まずボールが線形運動方程式に従って運動すると仮定し、一定の時間間隔でボールの位置座標を取り、運動予測を行って目的位置に達する時間を割り出す。これによって移動ロボットの移動速度指令を構成し、ロボットのシュート運動を行わせる。当然、ボールは非線形摩擦の影響で正確な線形運動とはならず、また、たとえロボットの速度指令が正しいと

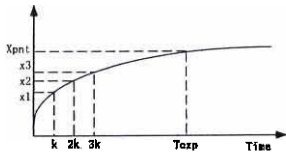


図 6: 初期速度をもつボールの位置運動例
Fig.6 An example of boll's motion

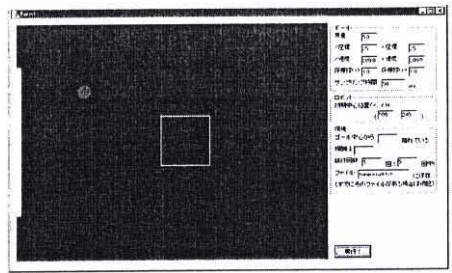


図 7: シミュレーション図
Fig.7 Simulation

となり、ただし、 C_1, C_2, C_3 は常数、 X は時間 T におけるボール座標である。ここで各サンプリング時刻 $t_1 (= k), t_2 (= 2k), t_3 (= 3k)$ それぞれにおけるボール座標 X_1, X_2, X_3 を用いて、 C_1, C_2, C_3 を求め、

$$C_3 = \frac{-1}{k} \log \frac{X_1 - C_1}{C_2} \quad (2)$$

と求められる。これによって、予め決めた目標位置 X_{pnt} へ到着するボールの予測時間 T_{exp} は

$$T_{exp} = \frac{1}{C_3} \log \left(\frac{C_2}{X_{pnt} - C_1} \right) \quad (3)$$

と計算される。

3 シミュレーション考察

実環境での実験を行う前に、ここで、まずシミュレーションによる学習を行った。シミュレーション画面を図 7 に、初期配置を図 8 に示す。また、計算に使用した各パラメータを表 1 に示す。

表 1: パラメータ
Table.1 Parameter

パラメータ	値
m (ボールの質量)	5.0
f_r (床摩擦)	1.0
f_n (床摩擦ノイズ)	$0.0 \leq f_n \leq 2.0$
Δt (サンプリング時間)	50[ms]

シミュレーションを行う画面の大きさは 640×480 画素で、左上端の座標を原点 $(0,0)$ 、右下端の座標を $(640,480)$ とする。また、原点から横方向を x 軸、縦方向を y 軸とする。ボールの大きさは直径 30 画素、ロボットは一辺 100 画素の正方形で示し、ゴールは $(0,90)$

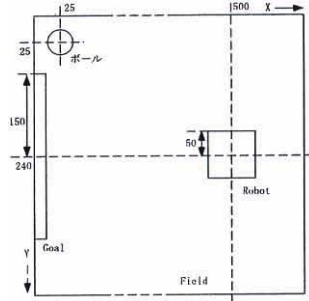


図 8: シミュレーションにおける初期配置
Fig.8 Initial position in simulation

と $(10,390)$ を対角の頂点とする長方形である。このとき、ロボットはゴールの方を向いているように配置する。ボールの初期配置は中心座標が $(25,25)$ になるように配置し、ロボットの初期中心座標は $(500,240)$ に配置する。また、ボールの初期速度 (x, y) は $(200, 200)$ とした。

3.1 ボール運動の予測

2 章のボール予測方式を適用し、 $k=200$ [ms] にしたときの結果を図 9 に示す。この結果よりボール運動が完全に予測できていることが分かる。

3.2 ロボットの行動学習方式

比較するために、本シミュレーションは以下に示す 2 通りの方法で学習を行う。

1. 数回のボール運動を予測し、1 0 回単位でゴール

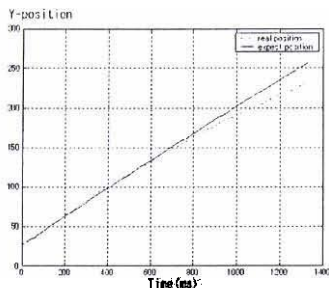


図 9: ボールの実運動と予測
Fig.9 Time response of the boll

が $y = 195$ に到達する予測時間の平均を求め、その時間を元にロボットの移動速度を計算する。

2. ボールが $y = 195$ に到達する時間を予測し、その時間からロボットの速度を決定する。ただし、床摩擦やカメラノイズなどの影響を考え、求められた速度について 0.5 から 1.5 まで 0.1 間隔でランダムにある倍率を掛け、その倍率でロボットの運動を実行させ、実行結果から得られた報酬（次節で説明する）をもとに、ルーレット戦略を行い、次の速度倍率を決定する。

3.3 報酬

報酬は以下のように与えた。

1. ロボットがボールをシュートした場合、報酬は

$$\frac{y - 90}{150} \times 1.5 \quad (y \leq 240)$$

$$\frac{390 - y}{150} \times 1.5 \quad (y \geq 240)$$

に基づいて与える（図 1 で Success、Failure1 の状態）。

2. ボールがロボットと衝突したとき、ボールの中心 y 座標 y_b が $y \leq 175$ を満たすなら、報酬は -1 とする（図 1 で Failure2 の状態）。
3. ボールがロボットと衝突せずに中心 y 座標 y_b が $y_b \geq 305$ を満たすなら、報酬は -2 とする（図 1 で Failure3 の状態）。

これは、ボールがゴール中央 ($y = 240$) に近いほど高い報酬を得ることが出来、ゴールに入れば正の報酬をもらい、最高報酬は 1.5 である事を示す。また、報酬が -1 の時にボールがどの位置でロボットと当たったの

かを示すため、「正面までの距離」という概念を導入した。これは図 10 中の距離①から距離②を引いた値で、0 に近づくほど正面に近いと言うことを示している。

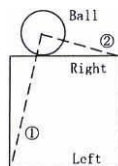


図 10: 正面までの距離
Fig.10 Distance from forward surface

4 シミュレーション結果

4.1 学習方式 1 の結果

予測時間の平均のみを用いた場合の結果を図 11～図 12 に示す。ここでは 10 回連続して $y = 195$ になる時間を予測し、それぞれの予測時間を平均してロボットの移動速度を決定する。また、ここではカメラノイズは含んでいない。

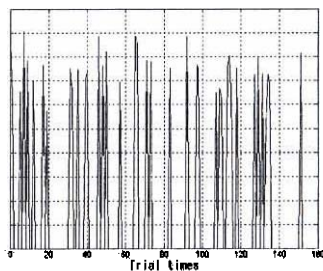


図 11: 試行回数-シュート成功率との関係
Fig.11 Trial times- shoot result relation

この結果、回数を増やしてもボールがゴールインされる成功率が変わらず、また、ボールとロボットの正面距離を示した「正面までの距離」に関しても学習した形跡はみられなかった。

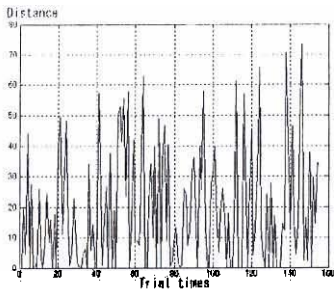


図 12: 試行回数-正面までの距離との関係
Fig.12 Trial times-distance relation

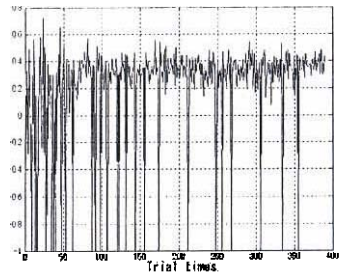


図 14: 試行回数-報酬との関係
Fig.14 Relation of trial times and reward relation

4.2 学習方式 2 の結果 (カメラノイズなしの場合)

次に、カメラノイズがない場合での報酬を基にしたルーレット戦略の学習結果を図 13～図 15 に示す。

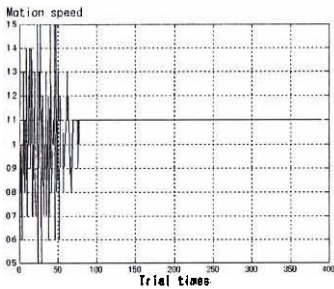


図 13: 試行回数-ロボット速度倍率との関係
Fig.13 Relation of trial times and robot's motion speed

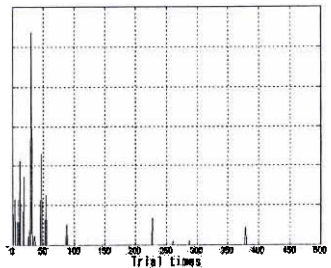


図 15: 試行回数-正面までの距離との関係
Fig.15 Relation of trial times and distance relation

この結果から、500 回の学習を行っても報酬が常に正の値を取るということはなかったが、しかし、倍率が 0.9 倍に収束し、また、正面までの距離が最初に比べて大幅に小さくなったという結果を得ることができた。

この結果より徐々ではあるが確実に学習している様子が分かる。150 回以降でもまれに報酬が -1 となる時があるが、図 15 よりほとんどロボットの右端でボールを蹴るような学習を行っていることがわかる。また、倍率も 1.1 倍に収束している。

4.3 学習方式 2 の結果 (カメラノイズがある場合)

カメラノイズを含んだ場合のルーレット戦略の結果を図 16～図 18 に示す。これは 4.2 の状態にカメラノイズ cn ($-2 < cn < 2$) を加えた状態である。

5 実験考察

実環境での画像処理結果を図 19、図 20 に示し、実環境内で $y = 195$ に一番近くに到達する時間を図 20 中の「REAL ATTACK TIME」で示す。ここでは予測サンプリング時間を $100[\text{ms}]$ とし、2 章の手順で $y = 195$ に到達する時間、およびその時の x 座標を予測した。予測結果を図 20 中の「expect_pos」、実環境で $y = 195$ に一番近い場所を観測したときの座標を「real_pos」、ボールが $y = 195$ に到達すると予測した時間を「EXPECT ATTACK TIME」で示す。また、予測と観測それぞれの時間と座標関係を図 19 に示す。

