

実環境における自律移動ロボットシステム*

3D-10

～その3：強化学習による運搬ロボット制御～

溝口文雄† 平石広典† 高山学† 前野守宏†

東京理科大学 理工学部†

1 はじめに

多くの強化学習の研究は移動ロボットを対象にその有効性を示している [1]. しかし、実際のロボットに適用した例は少なく、つまり、実環境下で何らかのタスクを遂行する自律移動ロボット開発のための力強い手法の一つであることを証明している例は少ない.

そこで実際のロボット“Nomad200”における“運搬タスク”に強化学習を適用する. 運搬タスクとは「ある特定の物体をある指定された場所に移動させること」である. ここで、移動ロボットが指定された場所へ移動するという観点からナビゲーションタスク [2][3] が考えられる. しかし、運搬タスクとナビゲーションでは移動方法が大きく異なる. ある環境の中で物体を確実に運搬するためには、運搬物に敵した運搬動作が求められる. 本稿では、いくつかの形態の運搬物での実験を通して、自律移動ロボットのための力強い制御戦略を実証する.

2 運搬ロボット

本研究で対象とする運搬ロボット (図1) は、半径9インチであり、側面に運搬装置を備えている. 運搬装置はロボット本体から9インチ突き出し



図1: 運搬ロボット (Nomad200)

3 強化学習によるロボット制御

物体を運搬する際に、ロボットの回転とともに運搬物の方向を変化させるために、ステアリングとターレット

の間にある一定の角度差を持たせる (図2左). 実際には目的地点の方向と現時点のターレットの向きの差と同等の角度差を、ターレットから目的地点の方向とは逆の方向にステアリングの角度差とする (図2右).

また、視覚を用いて運搬物上部に取り付けられた点を常に注視し、運搬中に運搬物の動きを監視する.

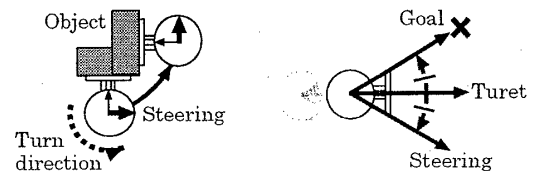


図2: ターレットとステアリングとの角度差

3.1 運搬動作の強化

強化のための学習時間は、状態空間の指数関数オーダーであるために、実際のロボットで強化学習を行なう際に、状態空間及び行動空間をできる限り小さくする必要があるのである [4]. そこで本研究では、直進速度は50インチ/秒に固定し、ステアリングとターレットの角度差 (状態空間) に対して、その時の回転速度 (行動空間) の強化を行なう.

状態空間

22.5度間隔で0度から67.5度までの4つの状態空間を設定する. つまり、ゴールとの方向の差が、0から22.4度の間は角度差を0度の状態、22.5度から44.9度の間は22.5度の状態、45.0度から67.4度の間は45.0度の状態、それ以上は角度差を67.5度の状態とする.

行動空間

0度/秒から22.5度/秒 (最大速度の1/2) を2.5度/秒毎に区切った10個の行動空間を設定する.

強化値

画像中の運搬物の点の動きによって強化値を決定する. 点が右の方向に移動してれば、ロボットに対して物体が右の方向に移動してしまっていることになる. どのような形態の運搬物に対しても、点が画像の一部に固定さ

*The autonomous mobile robot system in the real world.

†The robot control system for transport tasks using reinforcement learning

†Fumio MIZOGUCHI, Hironori HIRAISHI, Gaku TAKAYAMA, Morihiko MAENO

†Faculty of Sci. and Tech. Science University of Tokyo

れるような移動が最も安定した運搬につながる。実際に与える強化値は以下のようにする。

点が移動 (10 ピクセル以上)	—	-100
点が固定 (10 ピクセル以下)	—	10

各々の状態における各々の行動の強化値の初期値は全て0である。最初の選択は0度/秒から22.5度/秒へと順番に選択する。そして、その後の状態に対する行動の選択は、得られた強化値が最大の速度を選択する。

4 実験

実験環境は図3に示さる。フローは一般的なタイル張りであり、環境中に障害物は存在しない。スタート地点を座標(0,0)としたときに、ゴール地点を(0,70)に設定した。ロボットの方向はX軸に並行であり、つまりゴール方向との差が90度である。したがってスタート地点からゴール地点までの間に、全ての状態空間を試みることが可能である。

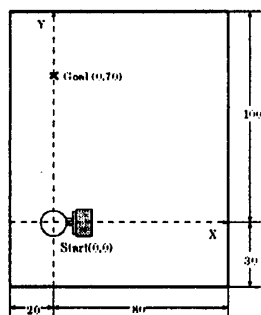


図3: 環境 (単位: インチ)

また、本研究のロボット制御の有効性の証明のために、以下に示す3種類の運搬物に対して実験を行う。

- 物体1: 長方形 (縦12インチ, 横20インチ, 高さ24インチ)
- 物体2: 長方形 (縦14インチ, 横28インチ, 高さ18インチ)
- 物体3: 円柱形 (半径5インチ, 高さ24インチ)

但し、物体1と物体2ではロボットは横の部分に触れて運搬する。

5 結果

結果を図4に示す。横軸に試行回数、縦軸に失敗数を示している。一回の試行回数はロボットがスタート地点を出発してゴール地点に到着するまでを意味し、失敗数とはゴール地点までの間に、マイナスの強化値を受けた回数を示している。

3つの物体共、試行回数が増すに連れて次第に失敗数が減少しており、強化学習の特徴がはっきりと現れている。特に物体1と物体2においては、10回の試行の中で、失敗数が0に到達している。また、物体3の実験の5回目の試行においてフロアに張られたビニールテープ

のために物体が転倒してしまった。そのために、急激に失敗数が増加しており、また、その影響で6回目の試行

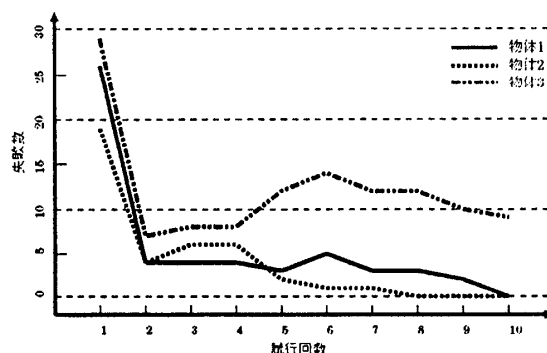


図4: 実験の結果

においても、失敗数が増加している。しかし、その後の強化によって徐々に失敗数が減少しており、最終的には物体1及び物体2同様に失敗数0に到達すると考えられる。本研究のロボット制御によって、運搬物への適応、及び、フロア面への適応が可能であることが示されている。

6 おわりに

運搬ロボットの構築にあたり、運搬動作への強化学習の適用について述べた。また、実環境下で働く運搬ロボットでの実験を通して、強化学習は自律移動ロボットのための力強い制御戦略の一つであることを実証した。本研究の運搬ロボット制御によって、ロボットが運搬可能(重さや大きさに関して)な物体を目的位置まで運搬することが可能となる。今回の実験では環境中に障害物は存在しておらず、また、ロボットは運搬物に触れた状態から始めている。今後の課題として、障害物の存在する環境中で、運搬物を発見し指定された場所へ運搬する一連の運搬タスクを扱う。

参考文献

- [1] Richard Maclin Jude W. Shavlik "Incorporating Advice into Agent that Learn from Reinforcements" *Appears in the Proceedings of the Twelfth National Conference on Artificial Intelligence AAAI-94*)
- [2] 溝口文雄, 高山学, 平石広典, 前野守宏: 実環境における自律移動ロボットシステム~その1: 動的障害物の位置予測による衝突回避~, 第52回情報処理学会全国大会
- [3] 溝口文雄, 前野守宏, 平石広典, 前野守宏: 実環境における自律移動ロボットシステム~その2: レーザースエンサーを用いた位相マップの獲得~, 第52回情報処理学会全国大会
- [4] 浅田稔, 野田彰一, 俵積田健, 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.66-74, 1995