

## OpenGLを用いた自律移動ロボットの学習シミュレーション The Simulation of Learning of the Autonomous Vehicle using OpenGL

富士原 良平†  
Ryouhei Fujiwara

堂 蘭 浩†  
Hiroshi Dozono

### 1. はじめに

近年、自律移動ロボットに関する研究が盛んに行われている。自律移動ロボットは、外部環境の情報を何らかの方法で取得し、ロボット自身が判断して行動を行うため、強化学習やニューラルネットを用いた機械学習に基づく制御手法を用いることが多い。

一般に機械学習には非常に多くの繰り返し学習が必要となり、自律移動ロボットなどの学習アルゴリズムの開発を実機を用いて行うことは困難であると考えられる。そこで本研究では、空間描画を行うことのできる OpenGL [1] を用いてシミュレーション環境を構築し、その空間内で得られる視覚情報のみで学習を行う方法を提案する。また、学習アルゴリズムは、自己組織化マップ [2] による視覚情報パターンの組織化を行うと同時に、各パターンに対して強化学習 [3] を行うという方法を用いる。

本報告では、視覚画像を元に平面上の目標地点に向けてロボットが自律移動する学習アルゴリズムを、OpenGLを用いたシミュレーションシステムによる実験で検証した結果を発表する。

### 2. OpenGL

3D グラフィック描画を行う OpenGL はウィンドウシステムから独立しており、OS を選ばずに使用できるのが特徴である。OpenGL は、空間内の物体の位置や、視点となるカメラの位置や角度の設定だけで、容易に空間内の視覚画像を得ることができる。そのため、作成したプログラムは簡潔なものとなり効率よく開発を進めることができるのも特徴である。

### 3. 自己組織化マップ

自己組織化マップ(Self-Organizing Map : SOM)は、コホネンによって提案された教師なし学習型ニューラルネットワークであり、多次元からなるデータの解析に使用される。

SOM は、ネットワークが入力層と出力層の 2 層で構成され、入力層で与えられたデータ  $x$  と類似する出力層のデータを探索し、発見した出力データ  $m_{ij}$  (勝利者ニューロン)とその近傍のデータが入力データに近づくようにデータ内容の更新を行う。

$$m_{ij} = m_{ij} + l_s \alpha(d)(x - m_{ij})$$

$$\alpha(d) = \frac{nr - d}{nr}$$

このとき、 $l_s$  は学習係数、 $t$  は繰り返し回数、 $nr$  は近傍の大きさ、 $d$  は勝利者ニューロンと更新ニューロン間の距離を表す。

この結果、得られる出力層のマップでは、データ間で特徴の似たデータは近くに、特徴の異なるデータは遠くに配置されるため、入力データ間の関係性を可視化して表すことができる。

本研究では、入力データとして OpenGL で生成した空間内の視覚画像情報を与え、その画像パターンを各ニューロンの重み値として学習することで、パターンの組織化を行う。

### 4. 強化学習

強化学習は、機械学習の手法の一つであり、学習物体と物体の置かれた環境の相互作用に基づく目標指向型の学習である。その強化学習の代表的手法として Q-Learning がある。これは、環境から現在の状態を観測、学習物体がその状態に基づいて行動し、その行動に対する評価を行うものである。

本研究では、モータで動作する車輪を左右に 2 つ備えたロボットを想定し、視覚画像に基づいた最良のモータの回転方向(正負回転、停止)について学習を行わせる。左右のモータの関係により、前進、後退、回転等の動作を行う。モータの回転方向にはそれぞれ選択確率  $Q$  を与えており、視覚情報からロボットが選択した 2 つのモータの回転方向に対してそれぞれ評価を行うことで、選択確率の値を更新する処理を行う。

・行動に成功した場合

選択された行動パターンについて

$$Q_{ij}^{kl} = Q_{ij}^{kl} + l_r \alpha(d) \sum_{m \neq l}^2 Q_{ij}^{km}$$

選択されなかった行動パターンについて

$$Q_{ij}^{kn} = Q_{ij}^{kn} - l_r \alpha(d) Q_{ij}^{kn}$$

・行動に失敗した場合

選択された行動パターンについて

$$Q_{ij}^{kl} = Q_{ij}^{kl} - l_r \alpha(d) Q_{ij}^{kl}$$

選択されなかったパターンについて

$$Q_{ij}^{kn} = Q_{ij}^{kn} + l_r \alpha(d) \frac{Q_{ij}^{kn} Q_{ij}^{kl}}{\sum_{m \neq l}^2 Q_{ij}^{km}}$$

このとき、 $i, j$  には確率値  $Q$  が格納されている SOM マップ上の座標が、 $k$  にはモータ番号が入る。また、 $l$  は選択されたモータの回転方向であり、 $n$  は選択されたもの以外の回転方向である。そして、 $l_r$  は学習係数、 $\alpha(d)$  は SOM で使用したものと同じ近傍関数を表す。

†佐賀大学大学院工学系研究科

## 5. 実験条件

今回、OpenGL で描かれた三次元の仮想空間内の平面上で学習を行うために、以下のような環境の仮想空間を作成した。

- ・移動可能領域は、縦100横100の平面
  - ・目標物は、座標  $(x, y)$  を中心とした一辺の長さ3の立方体であり、この周囲(座標から  $x, y$  ともに-5~+5の範囲)にロボット(カメラ)が進入することを目的とする。
  - ・移動可能領域の周囲は色の異なる直方体で障害物を作成し、ロボット(カメラ)が障害物に進入するような移動を選択した際には、移動を行わず現在の座標に留まる。
- 作成した仮想空間を図1, 図2に示す。

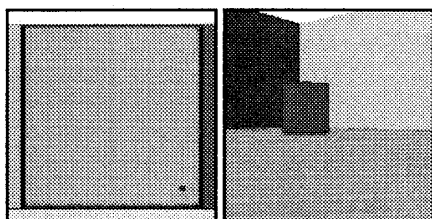


図1：作成した仮想空間の上空から見た図(左)  
図2：作成した仮想空間内で得られる視覚画像例(右)

## 6. SOM 型強化学習

まず OpenGL で作成した仮想空間内に配置したカメラから現在の座標での視覚画像(29×29画素)を取得する。この視覚画像のRGB色データを入力データとし、SOMマップ上から最も入力に近似したデータを探索する。最近似データを発見した後、マップ上の同位置に同じく格納されている自律移動ロボットの行動パターンを用いてロボット(カメラ)の座標を移動させる。行動後はSOMのアルゴリズムによってマップを更新する。SOMパラメータは、

- ・MAPサイズ30×30
- ・初期近傍の大きさ15
- ・初期学習係数0.8
- ・繰り返し回数10000回

としている。また、マップの更新と同時に、選択した行動に対して評価を行い、今後この行動を選択する確率を上下させる。強化学習に関する条件として、

- ・初期学習係数0.5
- ・評価は、目標物に距離が近づくと報酬を与え、遠ざかると罰を与える。

これら一連の処理を繰り返して学習を進めていく。学習の終了である繰り返し回数10000回時のSOMマップの一部を図3に示す。

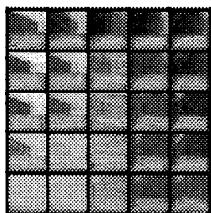


図3：繰り返し回数10000回終了時のSOMマップ(一部分)

ここで、学習が正しく行われたかを確認するため、ランダムに10箇所からスタートして目標を達成できるか検証を行った。スタートから150ステップ以内に目標物周辺の領域内まで到達できなければ、次のスタート地点へと移動させる。繰り返し回数が10回(学習初期)、10000回(学習末期)での平面上のロボットの動作軌跡を図4, 図5に示す。

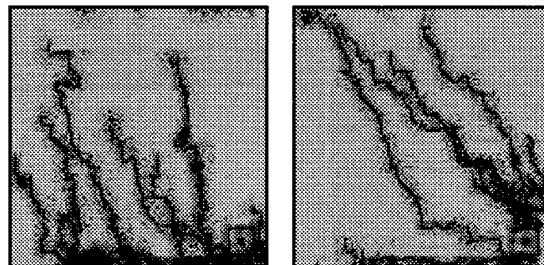


図4：繰り返し回数10回時の動作軌跡(左)  
図5：繰り返し回数10000回時の動作軌跡(右)

また、繰り返し回数による学習結果の違いについて比較を行った。繰り返し回数1000回、5000回、10000回の場合を考え、10箇所からのスタートで目標領域へ何回到達したか回数を調査する。それぞれの場合で得られた目標達成・未達成回数を表1に示す。

表1：学習回数と目標達成回数の比較

学習回数[回]	1000	5000	10000
目標達成回数[回]	5	4	9
目標未達成回数[回]	5	6	1

## 7. まとめ

本研究では、OpenGLを用いて仮想空間を作成し、その空間内で得られる視覚画像を用いて、SOM型強化学習による自律移動シミュレーションを行った。

実験によって得られた学習初期、末期の移動軌跡を比較すると、視覚画像を基に学習を行うことで、目標達成のためのより良い行動選択が行われるようになったことが分かる。しかし、学習回数が10000回も必要となる点は改善の必要がある。例えば、SOMや強化学習で使用した学習係数などは現状が最適な値であるとは限らない。また、今は移動可能な範囲内に障害物が無いが、障害物が多いほうが目印となるものが増えて、学習の効率が上がる可能性もある。

今後の課題は、学習回数の減少に向けたSOMや強化学習の手法の見直し、空間内に障害物がある場合の学習などがある。

## 参考文献

- [1] 三浦憲二郎：“OpenGL3D グラフィックス入門” 朝倉書店、1995
- [2] T. コホネン 著：“自己組織化マップ 改訂版” シュプリンガー・フェアラーク東京、2005
- [3] 電気学会 GA・ニューロを用いた学習法とその応用調査専門委員会 編：“学習とそのアルゴリズム” 森北出版、2002