

# 自動走行車いすの障害物回避シミュレーション・強化学習環境の実装と評価

坂田 悠馬<sup>†</sup> 長尾 確<sup>†</sup>

名古屋大学 大学院情報学研究科<sup>†</sup>

## 1.はじめに

人と密接な移動手段として、近年パーソナルモビリティに対する需要が高まっている。また、パーソナルモビリティと自動運転を組み合わせることで、より快適な移動が実現可能となる。しかし、人と密接な環境を走行する際には、人に危害が加わらないように自動車よりも正確な障害物回避が必要となる。加えて、障害物回避のテストを何度も実世界で行うことは危険であり現実的ではない。

そこで、本研究ではパーソナルモビリティの走行を評価できるシミュレーション環境の作成を目的とする。本研究ではパーソナルモビリティとして長尾研究室で扱っている自動走行車いす[1]を対象とする。

本論文では、3D シミュレータを使用したシミュレーション環境の作成について述べる。さらに、より正確な障害物回避行動を獲得するため、シミュレーション環境で行った強化学習についても述べる。

## 2.シミュレーション環境の作成

シミュレーション環境の基盤には 3D シミュレータである Gazebo[2]を使用した。Gazebo はロボットのシミュレーションによく使用され、ROS との連携がしやすい。自動走行車いすのプログラムは ROS で作成されているため、実機に移植しやすいという観点からこの 3D シミュレータを使用した。

シミュレーション環境で自動走行車いすを走らせるために、車いすの 3D モデルを作成した。シミュレーション環境の動作を軽量にするためシンプルな形状で構成している。また、大きさや重量に関しては実機のものと同様にしていく。

車いすモデルには環境情報を取得するために 2D の LiDAR が搭載されている。この LiDAR を

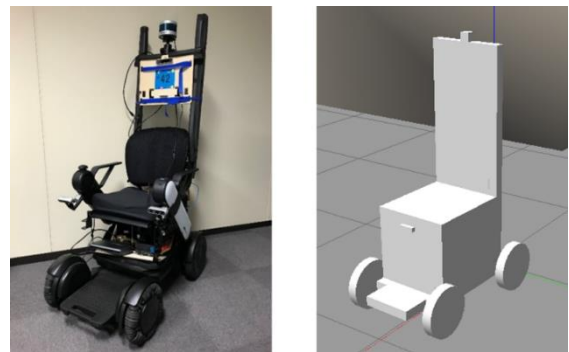


図 1 自動走行車いす(左図)と車いすの 3D モデル(右図)

用いて周囲の環境情報を取得する。センシングの範囲は前方 180 度である。

車いすモデルは後輪駆動で、ROS のトピック通信で速度情報を送信すると Gazebo 上でその速度に合わせて走行する。実機の前輪はオムニホイールとなっているが、シンプルな構成にするため、摩擦力のない円柱タイヤとしている。

図 1 に実機の自動走行車いすと Gazebo での自動走行車いすモデルを示す。

## 3.シミュレーション環境での強化学習

作成したシミュレーション環境で障害物回避の強化学習を行った。

### 3.1.強化学習環境の作成

強化学習では環境の状態を観測し、環境に応じた行動を返す必要がある。

ROS で強化学習を行うためのライブラリとして openai\_ros[3]を使用した。ROS は非同期通信により情報交換を行うが、行動選択や学習は同期的に行われる。openai\_ros は非同期な ROS の情報を統合して強化学習がしやすいプログラム群となっている。

強化学習の観測値として、車いすモデル上部の LiDAR の値と衝突判定を取得できる。Gazebo では当たり判定による衝突を取得することが難しい。そこで、衝突用の仮想 2D LiDAR を車いすモデル下部に用意し、距離が閾値以下になったら衝突と判定できるようにしている。

強化学習のプログラム内で openai\_ros を通じて環境を作成し、環境に対して行動を与えると観

Implementation and Evaluation of Obstacle Avoidance Simulation/Reinforcement Learning Environment for Autonomous Wheelchairs

<sup>†</sup>SAKATA, Yuma (sakata@nagao.nuie.nagoya-u.ac.jp)

<sup>†</sup>NAGAO, Katashi (nagao@nuie.nagoya-u.ac.jp)

<sup>†</sup>Graduate School of Informatics, Nagoya University

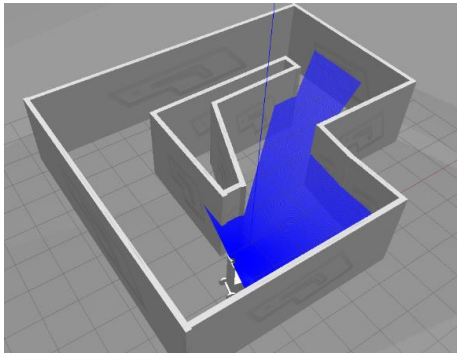


図2 サーキット環境

測値と行動に対する報酬を得ることができる。得た結果はランダムメモリに保存しておき、学習時に使用する。

### 3.2. シンプルなタスクの強化学習

まず、シミュレーション環境で強化学習が可能であることを確認するため、周回できるサーキット環境を作成した。

図2に作成したサーキット環境を示す。

車いすモデルがこの環境で壁にぶつからず走行することを目標に強化学習を行う。行動として前進と回転を与える。報酬には行動に対応する値と、壁に衝突したときのマイナス報酬を与えている。強化学習のアルゴリズムとして、DQN[4]とDDPG[5]を用いた。DQNは行動を離散的に扱い、DDPGは連続的に扱うことができる。

1000 エピソードで学習させた結果、DQNもDDPGもサーキット環境をぶつからず周回する行動を学習することができた。DQNに比べDDPGは連続値を扱えることから滑らかに移動することができた。したがってDDPGが障害物のスムーズな回避に適しているといえる。

### 3.3. 障害物回避タスクの強化学習

次にシンプルな空間で障害物回避を行うための強化学習を行った。環境には様々な形状の静止障害物を複数設置する。そして、一連の行動を始める最初にスタート位置とゴール位置を決め、障害物を避けつつスタートからゴールまで向かうことを目標とする。この学習により、自動走行に存在する局所的な目的地に対して効率的に行動できるかどうかを確認する。

図3にこのタスクで使用した障害物環境を示す。

このタスクではシンプルなタスクでの観測値に加え、現在地から目的地への向きを取得する。これにより車いすモデルがどこに向かうべきかを学習させる。また、報酬は新たにゴール到達報酬を設ける。この値は前進の報酬よりも大きく設定している。学習アルゴリズムにはDDPGを使用した。

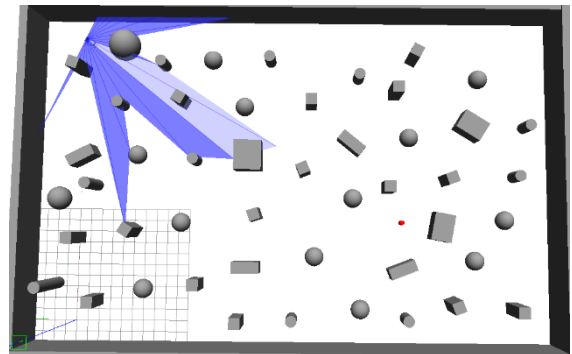


図3 障害物環境

約 3000 エピソードで学習を行った結果、最短とは限らないが、ゴールに対して向かい、障害物を回避する行動を学習することができた。

### 4. 複雑な環境への対応

Gazebo はシンプルな環境を作成しやすいが、移動障害物や複雑な地形を作成することが難しい。また、描画に弱く現実環境に近いとは言えない。

そこで、3DシミュレータをGazeboからゲームエンジンであるUnityに移行した。Unityは物理エンジンを搭載していることに加え、環境の作成が自由で複雑なものにも対応可能である。

UnityでROSを使用するため、ROS#[6]を使用した。これにより観測値や行動などROSを用いた情報のやりとりがUnityでも可能になる。また、車いすモデルの構造をそのままUnity内のオブジェクトに変換する機能もあり、Gazeboからの移行が容易である。

Unity上で同様の障害物環境を作成し、ROSの通信を用いて強化学習を行い、Gazeboと同様な学習結果を得ることができた。

### 5. まとめと今後の課題

本研究では、自動走行車いすのシミュレーション環境を作成し、そこで障害物回避の強化学習を行った。

今後の課題として、自己位置推定や経路生成などの自動走行の仕組みを採り入れて障害物回避の強化学習を行う必要がある。また、現在、名古屋大学東山キャンパスの3Dモデルを作成中であり、この環境で学習を行い、実環境と実機で強化学習モデルを評価する予定である。

#### 参考文献

- [1] Y. Mori and K. Nagao, "Automatic Generation of Multidestination Routes for Autonomous Wheelchairs," *Journal of Robotics and Mechatronics*, Vol. 32, No. 6, pp. 1121-1136, 2020.
- [2] Gazebo Simulator, <http://gazebosim.org/>
- [3] ROS.org, openai\_ros, [http://wiki.ros.org/openai\\_ros](http://wiki.ros.org/openai_ros)
- [4] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, et al., *Playing Atari with Deep Reinforcement Learning*, NIPS, 2013
- [5] T. Lillicrap, J. Hunt, A. Pritzel, N. Heess, et al., *Continuous control with deep reinforcement learning*, ICLR, 2016
- [6] ROS#, <https://github.com/siemens/ros-sharp>