

# 階段などの段差を含む経路を走行可能な4脚車輪車いす

赤見坂 篤記<sup>†</sup> 長尾 確<sup>††</sup>

名古屋大学 情報学部<sup>†</sup> 名古屋大学 大学院情報学研究科<sup>††</sup>

## 1. はじめに

近年、バリアフリー化が推進されつつある中、様々な理由からバリアフリー化が困難な建築物が存在する。そこで、4足歩行ロボットと車いすを合体させた4脚車輪車いすにより、車いす自体をそのような建築物に対応させるアプローチを採用し、その実現に向けて実験を行う。最終的には、実機の製作も視野に入れて、4脚車輪車いすの設計とそのシミュレータ環境の作成を行い、深層強化学習技術を用いて4脚車輪車いすに適応した歩行を含む移動を可能にする仕組みを目指す。

## 2. ロボットとシミュレータの仕組み

### 2.1. 4脚車輪車いす

本研究で使用する4脚車輪車いすをCADソフトウェアのFusion360を用い、模型サイズで設計した(図1)。設計ではBoston Dynamicsの4足歩行ロボットであるSpot[1]の構造を参考に、前脚にはブレーキ機構とオムニホイールを、後脚には動力機構付きのタイヤを、上には椅子を設置した。



図1 4脚車輪車いす

## 3. シミュレータの仕組み

シミュレータの仕組みについて説明する。シミュレータは4脚車輪車いすで想定される2種類

の移動方法である歩行移動と車輪移動を学習するために使用する。図2にシミュレータの仕組みを示す。物理エンジンにはUnityを使う。Unityに4脚車輪車いすのモデルを読み込む必要があるため、Fusion360からURDF (Unified Robot Description Format)と呼ばれるロボットを記述したxml形式のファイルを生成し、Unityに読み込んだ。Unityはロボットを扱うための様々な機能をUnity Robotics Hub[2]として提供しており、URDFをUnityに読み込む操作は、URDF Importerと呼ばれるパッケージで実現できる。Unity内の4脚車輪車いすはROSネットワークを介してPythonプログラムで制御される。強化学習実験において想定される環境の微調整やアルゴリズムの変更を容易にするために、OpenAI Gym[3]のインタフェース形式でカプセル化を行った。

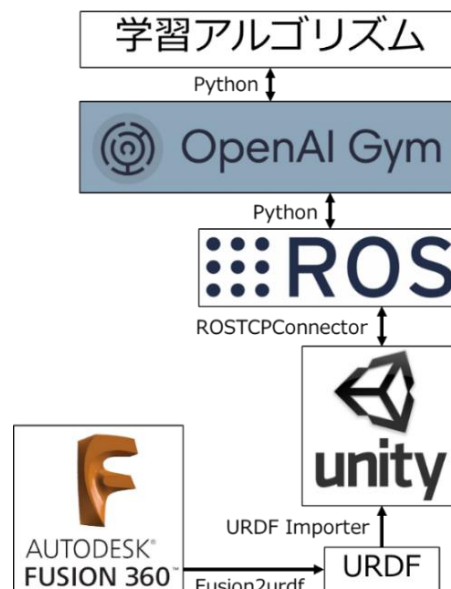


図2 シミュレータの構成

## 4. 強化学習の仕組みと結果

### 4.1. 実験計画

4脚車輪車いすの移動手法には歩行移動と車輪移動があるが、まずは歩行移動に重点を置き実験を進める。まず、現在研究されている歩行移動の実現方法には以下のものがある。出力を脚の制御信号としたとき、強化学習によるエンドツーエンドの手法、高次制御器と低次制御器に

Four-Legged Wheelchair that Enables Smooth Movement on Stairs and Steps

<sup>†</sup>AKAMISAKA, Atsuki (akamisaka@nagao.nuie.nagoya-u.ac.jp)

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

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

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

分ける階層的な手法[4]、軌道生成器を用いる手法[5]、模倣学習を用いる手法[6]などである。近年は、モデルベース強化学習とモデル予測制御(Model Predictive Control: MPC)を用いた歩行移動学習が有効であると考えられている。モデルベース強化学習に対を成す手法としてモデルフリー強化学習があるが、モデルベース強化学習とモデルフリー強化学習を融合させた手法もある。

4 脚車輪車いすは利用者の体格や移動環境などにより変化する環境のダイナミクスに柔軟に対応する能力が必要である。既存研究では、シミュレータと実環境間のダイナミクスの違いに注目して、学習済みモデルを実機に適用する際に sim-to-real[7]と呼ばれる技術を用いる研究が行われている。一方で、本研究のように搭乗者によってダイナミクスが変化する場合は、問題がより複雑になる。これらの問題を考慮したうえで、本研究ではモデルベース強化学習、モデルフリー強化学習、モデル予測制御を融合させた歩行移動学習を目指す。歩行移動学習の実験は段階的に行い、最初にモデルフリー強化学習単体による手法、次にモデルベース強化学習とモデル予測制御による手法、最後にモデルベース強化学習とモデルフリー強化学習およびモデル予測制御を用いた手法を適用する。ここでは、モデルフリー強化学習単体による歩行獲得実験の結果を示す。

#### 4.2. モデルフリー強化学習による歩行学習実験

強化学習モデルとして Soft Actor-Critic(以降 SAC)[8]と呼ばれる手法を用いる。シミュレータ内の 4 脚車輪車いすから SAC へ渡される状態情報は 6 軸 IMU とサーボモータの角度・速度とする。SAC から出力される行動情報は次の時刻におけるサーボモータの目標角度である。即時報酬は式(1)に示す関数により求める。

$$r_t = \alpha(x_t - x_{t-1}) - \beta|\theta_t| \quad (1)$$

時刻 $t$ における 4 脚車輪車いすの位置を $x_t$ 、いすの傾きを $\theta_t$ として、1 時刻で前進した距離に係数 $\alpha$ を乗算したものが即時報酬において正に働き、傾きに係数 $\beta$ を乗算したものが即時報酬において負に働くように設計した。係数の前進距離と傾きのうち重視すべき情報の比率を調節するために使用し、何度か行った事前実験を行い $\alpha = 100$ 、 $\beta = 1$ が最適であると判断した。転倒判定は角度 $\theta_t$ と閾値を比較し、閾値より大きい場合は転倒と判断し環境をリセットする。学習ステップ数は 200 万回とし、1000 ステップごとに 1 エピソードで得られる累積報酬和を計算する。その時の累積報酬和の変化を見て、モデルフリー強化学習

による歩行移動学習の精度を確認する。

#### 5. 実験結果

実験による累積報酬和のグラフを図 3 に示す。学習開始直後は、累積報酬和が急激に上昇し、50 万ステップで上昇が落ち着いている。65 万ステップ辺りから再上昇し、100 万ステップで上昇がほぼ安定した。以降は、目立った報酬の上昇は確認できなかった。

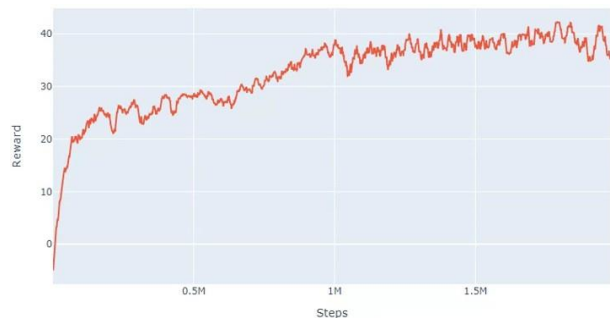


図 3 累積報酬和のグラフ

#### 6. まとめと今後の課題

式(1)で定めた即時報酬は、前進移動が継続する限り正の報酬が与えられるため、転倒しなければ累積報酬和の値は増え続けることが予測される。しかし、実験結果から累積報酬和が増え続ける挙動を示さなかったため、モデルフリー強化学習である SAC を用いて、転倒せずに歩行させることは困難であると考えられる。今後はモデルベース強化学習とモデル予測制御に基づく手法を試してみる予定である。

#### 参考文献

- [1] Boston Dynamics, Spot, <https://www.bostondynamics.com/products/spot>
- [2] ロボティクスのシミュレーション, <https://unity.com/ja/solutions/automotive-transportation-manufacturing/robotics>.
- [3] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang and Wojciech Zaremba. OpenAI Gym. CoRR, Vol. abs/1606.01540, 2016.
- [4] Marko Bjelonic, Ruben Grandia, Oliver Harley, Cla Galliard, Samuel Zimmermann, and Marco Hutter. Whole-Body MPC and Online Gait Sequence Generation for Wheeled-Legged Robots. CoRR, Vol. abs/2010.06322, 2020.
- [5] Atil Iscen, Ken Caluwaerts, Jie Tan, Tingnan Zhang, Erwin Coumans, Vikas Sindhwani, and Vincent Vanhoucke. Policies Modulating Trajectory Generators. CoRR, Vol. abs/1910.02812, 2019.
- [6] Xue Bin Peng, Erwin Coumans, Tingnan Zhang, Tsang-Wei Edward Lee, Jie Tan, and Sergey Levine. Learning Agile Robotic Locomotion Skills by Imitating Animals. In Robotics: Science and Systems, 2020.
- [7] Jie Tan, Tingnan Zhang, Erwin Coumans, Atil Iscen, Yunfei Bai, Danijar Hafner, Steven Bohez, and Vincent Vanhoucke. Sim-to-Real: Learning Agile Locomotion for Quadruped Robots. CoRR, Vol. abs/1804.10332, 2018.
- [8] Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, and Sergey Levine. Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor. CoRR, Vol. abs/1801.01290, 2018.