

## 強化学習を用いたセミトレーラの自動駐車モデルの作成

日大生産工(院) ○蘇 達治 日大生産工 風間 恵介  
日大生産工 丸茂 喜高

### 1. 緒言

近年、トラック業界ではドライバーの高齢化、若い就労者の減少という問題が顕在化している。さらに、2024年4月からトラックドライバーの時間外労働の960時間上限規制と改正改善基準告示が適用され、労働時間が短くなることで輸送能力が不足し、「モノが運べなくなる」可能性が懸念されている<sup>1)</sup>。したがって、セミトレーラに代表される連結車両の重要性がより高まると考えられる。セミトレーラは連結点があるため、駐車時に連結状態で後退する際、トレーラを後方からトラクタで押す形態となる。このことから、トラクタとトレーラは連結点で折れ曲がろうとするため、曲がりたい方向とは逆にハンドルを切る必要がある。そのため、運転操作が極めて難しく、普通の大型車より複雑な操作が要求される<sup>2)</sup>。このような背景から、運転支援システムや自動化が求められている。

セミトレーラの自動後退駐車を実現する方法としては、幾何学的手法<sup>3)</sup>や深層学習手法<sup>4)</sup>などいくつか提案されている。幾何学的手法では、セミトレーラの自己位置と駐車目標の地点までの軌道を幾何学的に生成して、それに追従するように車両を制御することで自動後退駐車を行っている。しかし、演算負荷が高く、複雑な環境や多くの制約がある場合、計算時間が増え、駐車ができなくなるという課題がある。深層学習手法では人間の操作データを取得し、取得したデータを用いて、ニューラルネットワークより自動後退駐車を学習をさせているが、人間の操作データを必要としている点が欠点である。十分なトレーニングデータを収集するためには、データ収集にコストと時間を要する場合がある。

一方、強化学習は学習データが不要で、自動的に最適な行動を見つけられるという利点がある<sup>5)</sup>。強化学習は、モデル自身が試行錯誤しながら価値を最大化するための行動を学習するものである。そこで本研究では、セミトレーラの自動後退駐車に強化学習を適用する。Unity<sup>6)</sup>を用いてシミュレーション環境の作成を行い、強化学習により得られたセミトレーラの自動後退駐車モデルの評価を行う。

### 2. シミュレーション条件

Fig.1に学習に用いるセミトレーラの車両諸元を示す。いすゞ自動車の「ギガ(トラクタ)」<sup>7)</sup>と、日本トレックスの「一般貨物タイプ3軸(トレーラ)」<sup>8)</sup>の車両諸元を参考にして、全長 $l_1=12.5\text{m}$ 、幅 $d=2.5\text{m}$ と設定した。トラクタの長さが $l_2=5.5\text{m}$ 、トレーラの長さが $l_3=10.0\text{m}$ とした。トラクタの質量は5000kg、トレーラも5000kgであり、無積載状態を仮定した。各駐車スペースのサイズは、一般的な大型トラックの駐車スペースのサイズである長さ15.0m、幅4.0mとした。

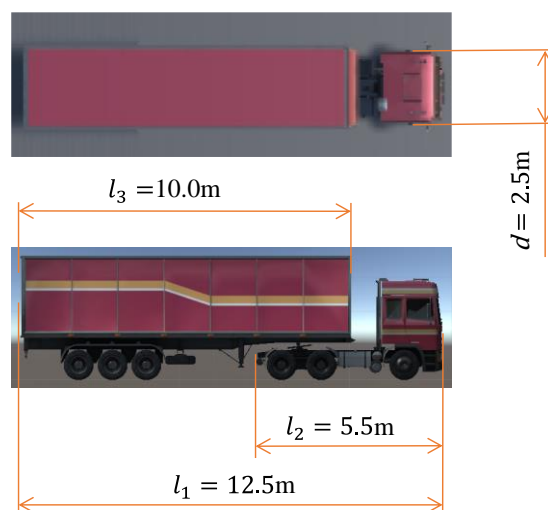


Fig.1 車両諸元

学習能力を確認するために、9つの初期条件を用意した。Fig.2に駐車開始時のセミトレーラの初期位置と初期姿勢を示す。初期位置は9箇所を設定し、駐車スペースの前端から21.5mの位置に①番を設定し、X方向10mごとに3箇所、Y方向6.5mごとに3箇所、合計9箇所用意した。初期姿勢は駐車スペース前端中央と初期位置を結ぶ線分がなす角と設定した。例えば、⑤番の初期位置は駐車スペースの前端からX方向に10.0m、Y方向35.5mのところに、角度 $\varphi$ を19.7 deg傾けた状態とした。設定した各初期位置に対して駐車能力を獲得できるか検証を行った。表1にシミュレーション開始時のセミトレーラの初期位置と初期姿勢を示す。

Creating an Automated Semi-Trailer Parking Model  
Using Reinforcement Learning

Dazhi SU, Keisuke KAZAMA and Yoshitaka MARUMO

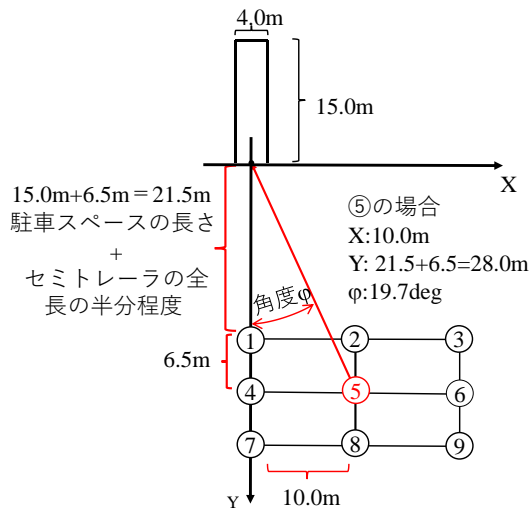


Fig.2 シミュレーション環境

表1 セミトレーラの初期位置・姿勢

	X[m]	Y[m]	$\phi$ [deg]
①	0.0	21.5	0.0
②	10.0	21.5	24.9
③	20.0	21.5	42.9
④	0.0	28.0	0.0
⑤	10.0	28.0	19.7
⑥	20.0	28.0	35.5
⑦	0.0	34.5	0.0
⑧	10.0	34.5	16.2
⑨	20.0	34.5	30.1

後退駐車モデルの作成には、Proximal Policy Optimization (PPO)アルゴリズム<sup>9)</sup>を使用する。式(1)にPPOアルゴリズム学習に使用する報酬関数を示す。ここで、 $\theta$ はポリシーパラメータ、 $E_t$ はタイムステップにわたる経験的期待値、 $r_t$ は過去の学習データと現在の学習データの比、 $A_t$ は時刻 $t$ での行動価値推定量、 $\epsilon$ はハイパーパラメータである。

$$L^{CLIP}(\theta) = E_t[\min(r_t(\theta)A_t, \text{clip}(r_t(\theta)), 1 - e, (1 + e)A_t)] \quad (1)$$

センサで駐車スペースや白線など検出することを想定して、トラクタとトレーラの各四隅と各中心に、合計10個のオブジェクトを設置した。Fig.3に、設置したオブジェクトの場所を示す。トラクタの中心、トレーラの中心および8箇所に設置したオブジェクトは駐車スペースに入ると報酬を与え、静的障害物である壁などに衝突した場合、ペナルティを与える。

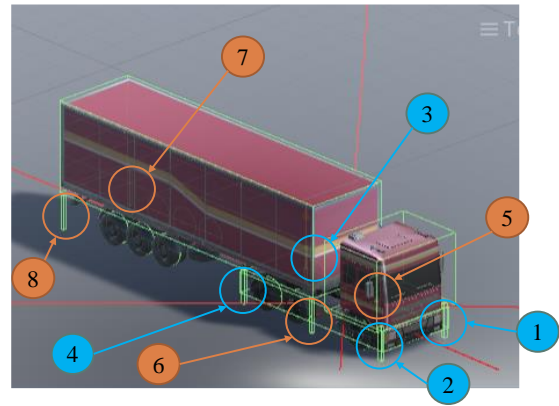


Fig.3 オブジェクト配置図

### 3. シミュレーション

本研究では、エージェントとしてのセミトレーラが、Targetである駐車スペースに向かって自動的に後退駐車するモデルの学習を行った。はじめに、⑤番の位置でシミュレーションを行った。今回のシミュレーションでは総エピソード数を300万回とした。Fig.4に作成したモデルを使った後退駐車の様子を示す。最終的に、自動で後退駐車を行うモデルを作成した。

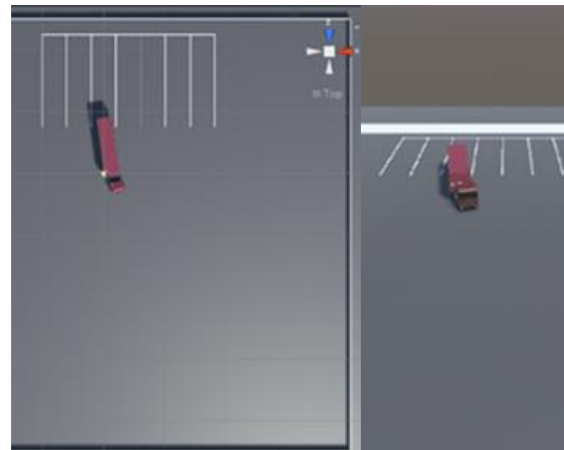


Fig.4 後退駐車の様子

Fig.5 に、エピソード数（学習一回当たりのステップ数）に応じた、平均累積エピソード報酬の推移を示す。横軸は学習のエピソード数であり、縦軸は累積した報酬をそのときのエピソード数で割った値である。一般的に、累積報酬は、継続して小さな振幅で増加することが期待される。Fig.5 から、エピソード数の増加とともに累積報酬は増加していることが確認できる。40万回を超えてから目標の駐車領域を認識し、徐々に得られる報酬が増加していった。200万回程度から報酬値が収束し始め、学習が完了したと考えられる。

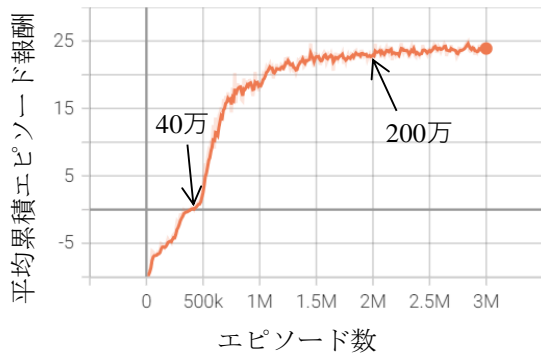


Fig.5 ⑤番における平均累積エピソード報酬の推移

Fig.6 に、エピソード数に応じた、価値の平均損失の推移を示す。横軸は学習のエピソード数であり、縦軸は価値の平均損失である。一般的に、学習が進んでいる間は損失が増加し、報酬が安定していくと減少する傾向にある。しかし Fig.6 では、価値損失は減少しているものの、収束しているとは言えない。ハイパーパラメータや報酬関数を見直すことで、駐車能力のさらなる向上が期待できる。

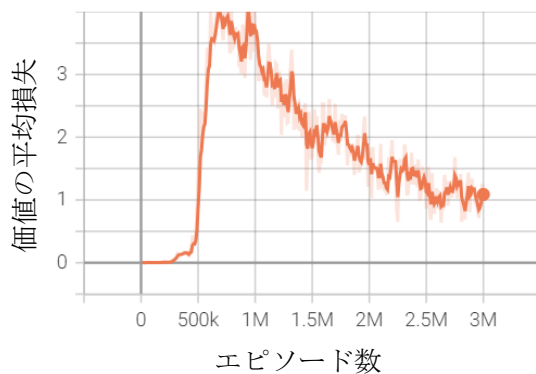


Fig.6 ⑤番における価値の平均損失の推移

続いて、残りの8箇所、セミトレーラの自動後退駐車ができるかどうかを確認した。その結果、①、②、④、⑥～⑧番で駐車できたが、③、⑨番で駐車ができなかった。

Fig.7 に⑥番で学習した結果の平均累積エピソード報酬の推移を、Fig.8 に価値の平均損失の推移を示す。Fig.7, 8 には、報酬が145万回程度から増加し、価値損失は250万回程度から減少する傾向があるが、両方とも学習が完了するまで収束していない。このため、セミトレーラの自動後退駐車が失敗することもある。学習が完全に終了していないためと考えられる。

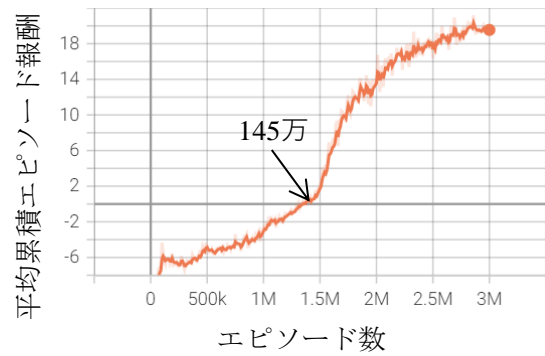


Fig.7 ⑥番における平均累積エピソード報酬の推移



Fig.8 ⑥番における価値の平均損失の推移

Fig.9 と Fig.10 は、③番で学習させた結果の平均累積エピソード報酬の推移と価値の平均損失の推移を示す。駐車スペースに車両が入らず、平均累積エピソード報酬がうまく得られなかったことがわかる。学習が進まなかったため、価値の平均損失も増加していないことがわかる。

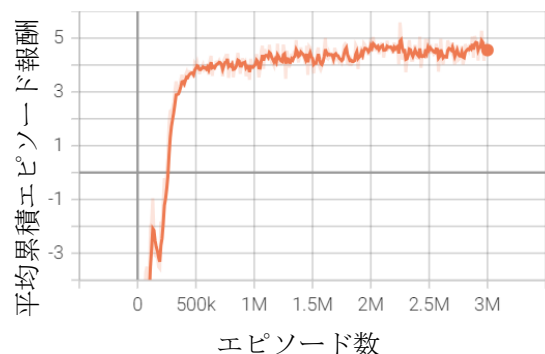


Fig.9 ③番における平均累積エピソード報酬の推移

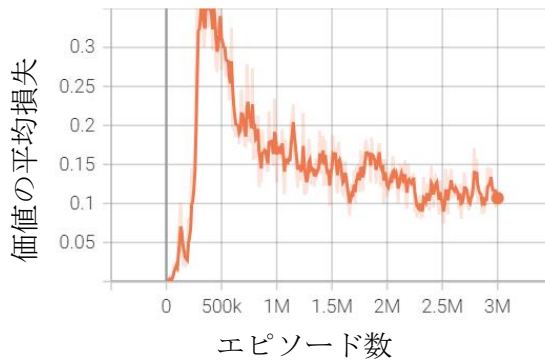


Fig.10 ③番における価値の平均損失の推移

表2 各初期位置での成功割合

各初期位置での成功割合		
①	②	③
98%	94%	0%
④	⑤	⑥
99%	96%	66%
⑦	⑧	⑨
97%	95%	0%

最後に、各初期位置で学習させて、作成したモデルでシミュレーションを100回ずつ行い、セミトレーラの自動後退駐車の実現割合を検証した。その結果を表2に示す。①、②、④、⑤、⑦、⑧番の成功割合はほぼ95%以上だが、⑥番は66%、③、⑨番は0%であった。報酬設定を改善し、任意の初期位置から駐車可能なモデルの作成を試みる。

#### 4. 結語

本研究では、強化学習を用いて、セミトレーラの自動後退駐車を実現することを目的とした。今回のシミュレーションでは、トレーラとトラクタの各中心と四隅にセンサを設置し、そのセンサが駐車スペース内かどうかに応じて報酬を設定した。その結果、特定な位置でセミトレーラの自動後退駐車が実現できた。

しかし、任意の位置で自動後退駐車が失敗したこともあった。また、作成した自動後退駐車モデルは、別の位置で自動後退駐車できないこ

とがあった。今後の課題として、報酬設定やプログラムを修正して、セミトレーラの自動後退駐車モデルの性能を向上させる。

#### 参考文献

- 1) 経済産業省 中野 剛志, 政策動向のご紹介 ~ 物流の2024年問題~, (2023) [https://www.meti.go.jp/shingikai/mono\\_info\\_service/physical\\_internet/hokkaido/pdf/001\\_09\\_00.pdf](https://www.meti.go.jp/shingikai/mono_info_service/physical_internet/hokkaido/pdf/001_09_00.pdf), (参照 2023-10-09)
- 2) GP企画センター, 「トラクター&トレーラの構造」, グランプリ出版, (2010) pp.9-22.
- 3) Y. Hamaguchi and P. Raksincharoensak, “Automated Steering Control System for Semi-Trailer Vehicle in Reverse Parking Maneuver Considering Motion Planning by Simulation of Feedback Control System”, *Journal of Robotics and Mechatronics*, Vol. 32, No. 3 (2020), pp. 561-570.
- 4) 伊尾拓真, 吉田章浩, 風間恵介, 小方博之, “深層学習による連結車両の自動後退駐車能力の獲得”, In *IEICE Conferences Archives, The Institute of Electronics, Information and Communication Engineers*, (2022).
- 5) 布留川英一, 佐藤英一 編「Unity ではじめる機械学習・強化学習 Unity ML-Agents 実践ゲームプログラミング v1.1 対応版」, 株式会社ボーンデジタル, (2020).
- 6) A. Juliani, V. P. Berges, E. Teng, A. Cohen, J. Harper, C. Elion, C. Goy, Y. Gao, H. Henry, M. Mattar and D. Lange, “Unity: A general platform for intelligent agents”, *arXiv preprint arXiv:1809.02627*, (2018).
- 7) いすゞ自動車株式会社, トラクタ展開, <https://www.isuzu.co.jp/product/giga/lineup/tractor/>, (参照 2023-10-09).
- 8) 日本トレクス株式会社, ウイングセミトレーラシリーズ, (2023) <http://www.trex.co.jp/media/20230919-115225-391.pdf>, (参照 2023-10-09).
- 9) J. Schulman, F. Wolski, P. Dhariwal, A. Radford and O. Klimov, “Proximal policy optimization algorithms”, *arXiv preprint arXiv:1707.06347*, (2017).