

強化学習による連結車両の自動後退駐車能力の獲得

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、若い就労者の減少が深刻な問題となっている。さらに、物流の2024年問題であるドライバーの時間外労働を年間960時間に制限する規制が適用されることで、労働時間が短縮され、輸送能力の確保が困難になることが懸念されている。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。セミトレーラには連結点があるため、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。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²⁾。セミトレーラのような複雑な操作が必要な車両にこそ、自動駐車システムが必要だと考えられる。

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³⁾や深層学習手法⁴⁾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させているが、人間の操作データを必要としている点が課題である。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。

それらの課題に対して、我々はこれまでに、強化学習を適用して、特定の位置からの自動後退駐車を実現した。しかし、1つの位置で学習させて作成したモデルは、別の位置で自動後退駐車できなかった⁵⁾。そこで本研究では、転移学習を用いて、1つの位置で学習して作成したモデルに、さらに複数の条件で学習させてモデルを作成し、様々な位置からの自動後退駐車能力の獲得を実現する。

2. 駐車能力の獲得の基礎検討

基礎検討として、特定の条件下で作成したモデルを用いて、別の条件で駐車可否を検証し、汎化性能を獲得できるかを確認する。学習能力を確認するために、6つの初期条件を用意した。Fig. 1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。駐車スペースの前端から前方21.5 mの位置に条件②を設定し、セミトレーラの重心をその位置とした。前後方向Zの位置は、駐車スペースの長さとしてセミトレーラの後端から重心までの長さの合計に対応している。条件②の左右(X)方向に±10 mに2箇所(条件①・③)、さらに①～③をZ方向に+6.5 m移動させた3箇所(条件④～⑥)の、合計6箇所を用意した。姿勢角は、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、初期条件③は駐車スペースの前端からX方向に+10.0 m、Z方向に+21.5 mの地点であり、姿勢角φは+21.8 degとなる。Table 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。

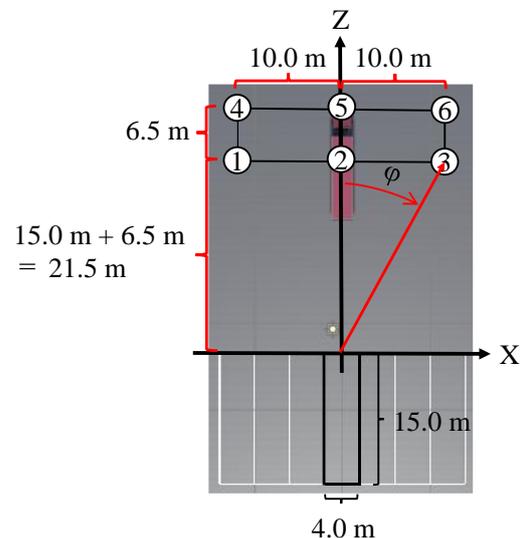


Fig. 1 The initial position and initial attitude of the semi-trailer at the start of parking

本研究では、Unity⁶⁾に用意されているセミトレーラを用いて、シミュレーションを行った。Fig. 2に学習に用いるセミトレーラの車両諸元

Automated Backward Parking of Articulated Vehicles with
Reinforcement Learning

Dazhi SU, Keisuke KAZAMA and Yoshitaka MARUMO

を示す。全長 $l_1=12.5$ m, 幅 $d=2.5$ mである。トラクタ部の長さは $l_2=5.5$ m, トレーラ部の長さは $l_3=10.0$ mであり, 一般的なセミトレーラのサイズとなっている。トラクタおよびトレーラの質量は5,000 kg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15.0 m, 幅4.0 mとした。

Table 1 Initial position and attitude of the semi-trailer

Initial condition	X [m]	Z [m]	φ [deg]
①	-10.0	21.5	-21.8
②	0.0	21.5	0.0
③	10.0	21.5	21.8
④	-10.0	28.0	17.6
⑤	0.0	28.0	0.0
⑥	10.0	28.0	-17.6

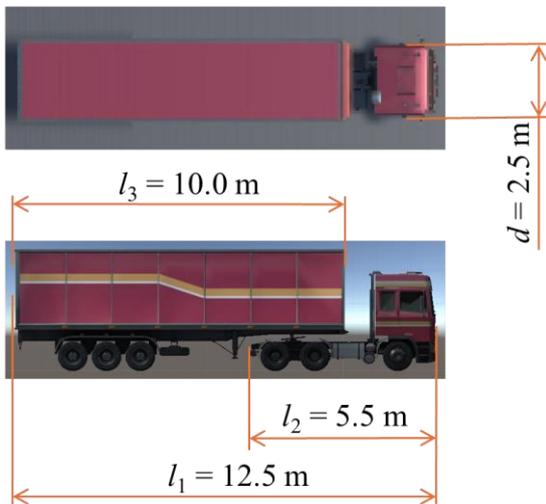


Fig. 2 Semi-trailer specifications

また, 過学習を防ぐために, Fig. 3に示すような初期位置と初期姿勢に乱数を加えた。乱数の範囲は, X方向とZ方向に ± 1.0 m, 姿勢角に ± 10.0 degとした。

後退駐車モデルの作成には, Proximal Policy Optimization (PPO)アルゴリズム⁷⁾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。

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

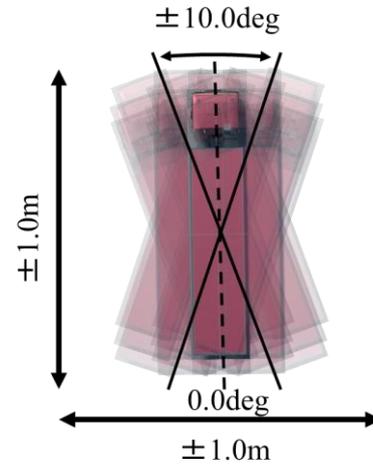


Fig. 3 Range of random numbers for initial position and initial attitude

ここで, θ 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, e はハイパーパラメータである⁸⁾。

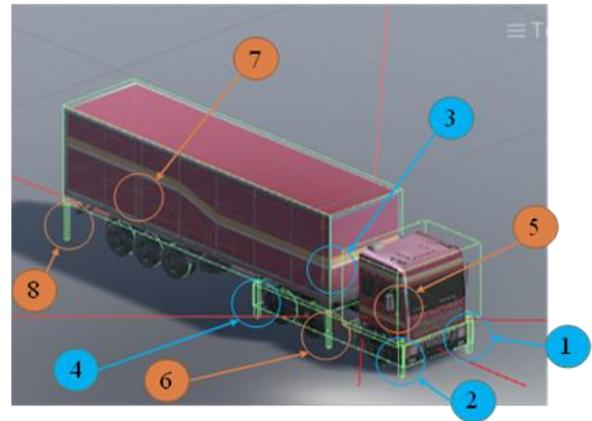


Fig. 4 Layout of sensors

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. 4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。

はじめに, 初期条件②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

ード数を1000万回とした。最終的に、自動で後退駐車を行うモデルを作成した。

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、継続して小さな振れ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。20万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。800万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。

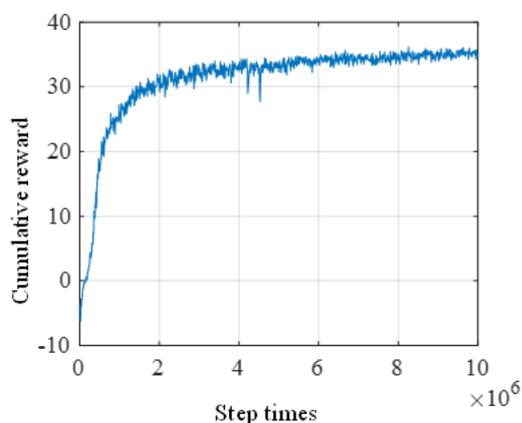


Fig. 5 Transition of average cumulative reward based on the number of episodes

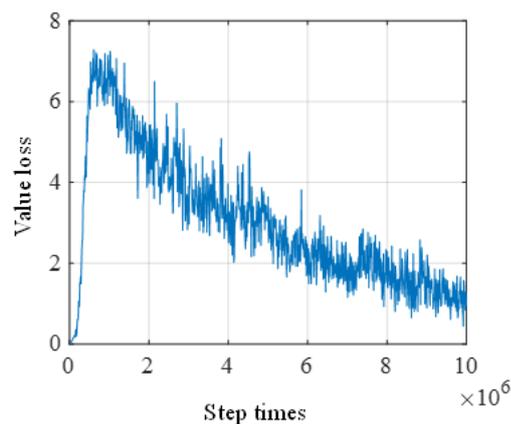


Fig. 6 Transition of average loss of value based on the number of episodes

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。横軸はエピソード数であり、縦軸は価値の平均損失である。平均損失は、損失した価値をそのときのエピソード数で割っ

た値である。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。

次に、条件②をベースに学習して作成したモデルを用いて、条件①から⑥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成功の割合を検証した。条件②の成功割合は100%、⑤は91%であった。しかし、条件①、③、④、⑥の成功割合は0%であった。目標姿勢角が0 degの場合は駐車成功したが、0 degではない場合は駐車ができず、汎化性能を獲得することはできなかった。これは初期状態が一種類しかなく、乱数を加えても過学習の問題が解決できていないと考えられる。

3. 転移学習による駐車能力獲得の改善

続いて、過学習の問題が解決できていないため、初期条件を変更して、条件②で学習したモデルに対して転移学習を行った。転移学習とは、学習済みのモデルに対して、新しい条件を付与して再学習を行うことである。セミトレーラを条件①から⑥でランダムに出現させ、その位置・姿勢に乱数を加えて、2000万回の学習を行った。Fig. 7に学習した結果の平均累積エピソード報酬の推移を示す。報酬が途中何度か低下したが、1900万回ごろから収束している。

Fig. 8に価値の平均損失の推移を示す。Fig. 6と比べると、学習済みモデルから再学習を行っているために、価値損失の立ち上がりは小さく収まったが、収束の勾配も小さくなったことがわかる。

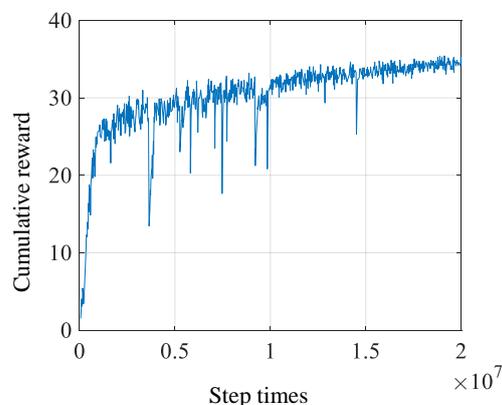


Fig. 7 Transition of average cumulative reward based on the number of episodes

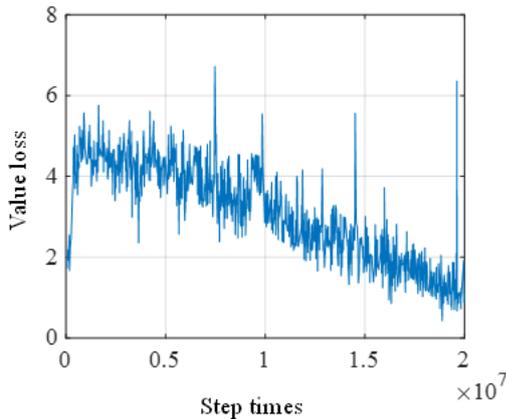


Fig. 8 Transition of average loss of value based on the number of episodes

最後に、作成したモデルを用いて各条件でシミュレーションを100回ずつ行い、成功の割合を確認した結果、各条件の成功の割合は100%であり、1つの作成した自動後退駐車モデルだけで、特定の6条件における自動後退駐車能力を獲得することができた。

4. まとめ

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。直進状態で後退する条件で、初期位置と初期姿勢に乱数を加えて学習を行い、作成したモデルを用いて、6条件下でシミュレーションを100回ずつ行った。その結果、目標姿勢角が0 degの場合は駐車が成功したが、0 degではない場合は駐車ができず、汎化性能を獲得することはできなかった。そこで、初期条件を変更して転移学習を行った結果、作成したモデルを用いて各条件でシミュレーションを100回ずつ行い、成功の割合は全ての条件下で100%であり、1つの作成した自動後退駐車モデルだけで、特定の6条件における自動後退駐車能力を獲得することができた。

今後の課題として、報酬設定やプログラムを改善して、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ことが考えられる。また、シミュレーションで作成したモデルは、実際のセミトレーラの運転にどのように関連するのかを定量的に評価するため、人間が用意した後退駐車モデルと強化学習で得たモデル同士を、ドライバモデル同定を行うことで比較するこ

とも検討している。ドライバモデル同定によって得られる前方注視距離や反応時間などのドライバのパラメータから、実際の運転に還元できるデータがあるかどうかを確認し、セミトレーラの後退駐車時の運転操作について考察する。

参考文献

- 1) 中野剛志, “政策動向のご紹介～物流の2024年問題～”, 経済産業省, (2023), pp.1-5.
- 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”, *精密工学会誌*, Vol.90, No.3 (2024), pp. 298-305.
- 5) 蘇達治, 風間恵介, 丸茂喜高, “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”, 日本大学生産工学部第56回学術講演会講演概要 (2023), pp. 134-137.
- 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) J. Schulman, F. Wolski, P. Dhariwal, A. Radford and O. Klimov, “Proximal policy optimization algorithms”, arXiv preprint arXiv:1707.06347 (2017).
- 8) 布留川英一, 佐藤英一, “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), pp. 25-34.