

視覚と行動の end-to-end 学習により経路追従行動をオンラインで 模倣する手法の提案

経路への復帰行動の解析と復帰行動を強化する教師データ収集法の検討

A proposal for an online imitation method of path-tracking behavior by end-to-end
learning of vision and action

-Analysis of return-to-pathway behavior and a method of collecting training data to enhance
return-to-pathway behavior-

学 今井悠月 (千葉工大)

清岡優祐 (千葉工大)

学 春山健太 (千葉工大)

正 上田隆一 (千葉工大)

正 林原靖男 (千葉工大)

Yuzuki IMAI, Chiba Institute of Technology, s20c1015as@s.chibakoudai.jp

Yusuke KIYOOKA, Kenta HARUYAMA,

Ryuichi UEDA and Yasuo HAYASHIBARA, Chiba Institute of Technology

We have proposed an online imitation method of path-tracking behavior based on end-to-end learning of vision and action. In recent years, many studies of autonomous movement using end-to-end learning have been reported, but these studies have also observed deviations from the target path. One of the possible reasons for this is the lack of training data for returning to the path. In this paper, we perform end-to-end learning to follow a route generated by a map-based navigation system. The dataset was collected in two ways, one is to learn only the area around the route and the other is to learn the state away from the route, and the generated path-tracking behaviors were analyzed. In addition, we proposed a new method of collecting teacher data to reinforce the behavior of returning to the path, and verified the effectiveness of the method by experiments using a simulator.

Key Words: autonomous movement, navigation, end-to-end learning, dataset

1 緒言

我々は、入出力関係を直接学習する end-to-end 学習器により、人の操作ではなく、地図ベースのナビゲーションシステムを用いた経路追従行動を視覚に基づいてオンラインで模倣する手法を提案し、その有効性を実験により検証してきた [1][2].

地図ベースのナビゲーションとは、LiDAR や IMU、ホイールオドメトリなど複数のセンサから占有格子地図を作成し、その地図を用いて自己位置推定、経路計画、制御などの複数のタスクを達成することで目的となる場所へ自律的に移動する手法である [3].

近年、画像を入力とした end-to-end 学習による自律移動手法が注目されている。例えば、Muller らは、人が操作したコントローラ操作を教師データとして学習することで、オフロード環境で障害物を回避しながら走行できることを確認した [4]。また、Bojarski らは画像と人が操作した制御コマンドを end-to-end 学習することで、自動車を対象とした自律移動手法を提案した [5].

しかし、それらの研究では必ずしも学習した経路を追従できるわけではなく、目標とする経路から外れていく様子が確認されている。その要因の一つとして、目標経路へ復帰するための学習データが不足していることが考えられる。

そこで本稿では、カメラ画像を入力とした end-to-end 学習による自律移動手法において、常に経路周辺を学習する場合と、目標経路周辺だけでなく目標経路から離れた状態も学習する場合の二つの方法でデータセットを収集し、学習を行う [6]。そして、その学習によって生成された経路追従行動の解析を行い、離れた状態からの復帰行動を学習することが、有効であることを明らかにする。それに加え、目標経路への復帰行動を強化する教師データ収集法を新たに提案し、有効性を検証する。

2 地図ベースのナビゲーションを模倣する手法

地図ベースのナビゲーションをオンラインで模倣する手法を示す。本手法は、地図ベースのナビゲーションシステムの出力である角速度を教師データとして学習し、画像を入力とした学習器の出力によってロボットを制御する。手法は学習器の訓練を行う学

習フェーズと、訓練結果を検証するテストフェーズに分かれる。

2.1 学習フェーズ

図 1 に学習フェーズのシステムを示す。学習フェーズでは、ロボットは ROS のナビゲーションパッケージである navigation[7] を用いて自律移動を行う。それと同時に、画像と navigation から出力される目標角速度をペアにしたデータセットを収集し、end-to-end 学習を行う。地図ベースのナビゲーションは LiDAR とオドメトリを入力とするのに対し、学習器は画像を入力とする。本手法を用いて学習することで、LiDAR、オドメトリを使用した時と同じ行動を、画像のみで行うことが期待される。

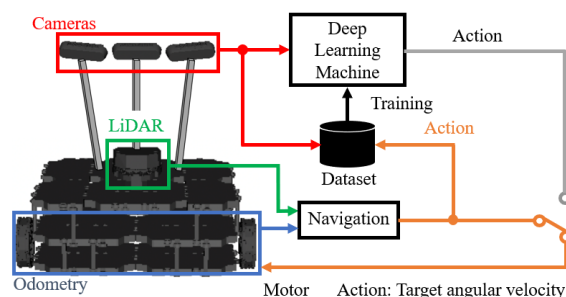


Fig.1: Learning phase

2.2 テストフェーズ

テストフェーズでは、図 2 のように学習フェーズで学習したモデルを用いてロボットを制御し、自律移動できるかテストする。64 x 48 にリサイズした画像を学習器に入力し、その推定結果をロボットの目標角速度としてロボットを自律移動させる。ロボットの角速度は学習器の出力を用いるが、並進速度は学習フェーズと同様に一定 (0.2m/s) とする。なお、テストフェーズでは 3 台のカメラのうち、中央のカメラのみ使用する。

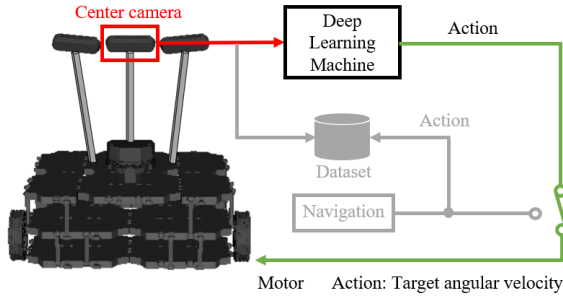


Fig.2: Test phase

2.3 学習器について

ネットワークの構成と学習方法について述べる。本研究は、画像を入力、角速度を出力として end-to-end 学習を行う。ネットワークは、図 3 に示すように、入力層 1、畳み込み層 3、全結合層 2、出力層 1 の全 7 層で構成されている。また学習をする際、1 ステップにつきデータセットの中からランダムに 8 個のデータを抽出し、その 8 個のデータで学習を行っている。

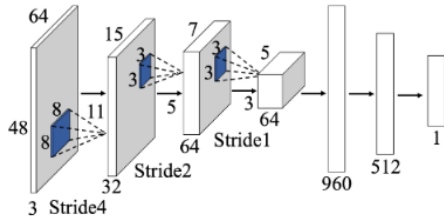


Fig.3: Structure of the network

3 地図ベースのナビゲーションを模倣する実験

経路周辺だけでなく、経路から離れた状態を学習することが経路追従行動を模倣する上で有効であることを明らかにするためにシミュレータを用いた実験を行う。実験は、経路周辺を学習した場合と、経路から離れた状態も学習する場合の二つのデータセット収集方法で行い、学習器の出力で経路を追従できるか調査する。

3.1 実験装置

実験には Gazebo[8] というシミュレータを用いた。環境は図 4 に示すような、千葉工業大学津田沼キャンパス 2 号館 3 階を模したモデルである。ロボットは、Turtlebot3[9] にカメラを 3 つ取り付けたモデルを用いた。図 5 にロボットが追従する経路を示す。ロボットはこの経路を追従するように学習を行う。

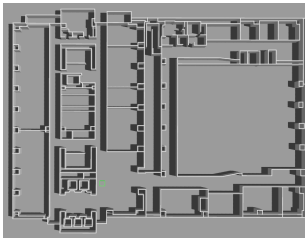


Fig.4: Experiment model

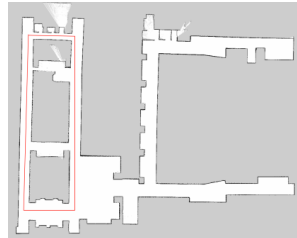


Fig.5: Target path

3.2 比較するデータセット収集方法

比較する 2 つのデータセットの収集方法を述べる。

3.2.1 経路周辺の状態のみを学習する手法

経路周辺の状態のみを学習する手法を図 6 に示す。ロボットは常に navigation の出力で経路を追従するように行動する。そして、その行動を常にデータセットとして収集する。本稿では、このデータセット収集方法を Path following method とする。

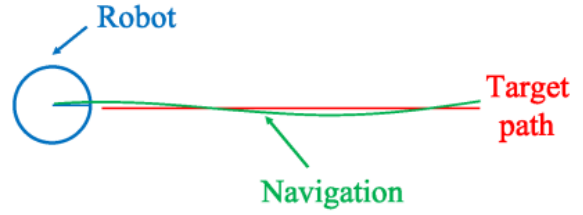


Fig.6: Path following method

3.2.2 経路から離れた状態も学習する手法

経路から離れた状態も学習する手法を図 7 に示す。経路周辺と経路から離れたときで、ロボットを制御するモジュールを切り替えている。経路周辺では、学習器の出力を用いている。ロボットが経路から離れた場合、navigation で経路へ復帰させる。データセットに追加するデータは常に navigation の出力である。このように経路周辺で学習器の出力を用いてロボットを制御することで、Path following method に比べて、経路から離れた状態のデータを多く収集できることが期待できる。このデータセット収集方法を本稿では Using learning machine output method とする。

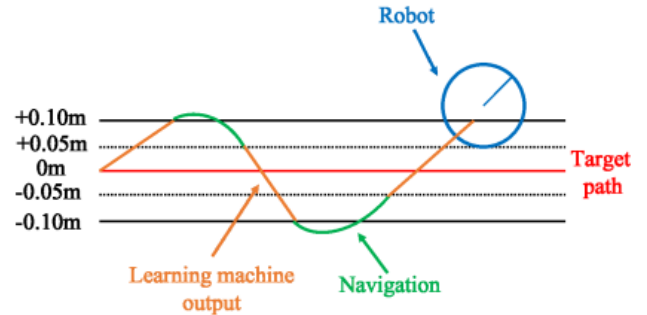


Fig.7: Using learning machine output method

3.3 実験方法

実験は学習フェーズの後、テストフェーズに移行し、コースを 1 周できるか確認する。学習フェーズでは 2 章で説明した、地図ベースのナビゲーションを模倣する手法を用いる。データセットは 0.2 秒周期で取得し、10000 エピソード学習行った。壁に衝突することなくコースを 1 周することができれば経路追従に成功したと判断し、この実験を二つのデータセット収集方法でそれぞれ 50 回ずつ行った。

3.4 実験結果

実験の結果を表 1 に示す。Path following method で学習したモデルでロボットを制御した場合、一度もコースを 1 周することができなかった。それに対し、Using learning machine output method では、50 回中 29 回コースを 1 周することに成功した。

ここで、データセットに含まれる目標経路からの距離で分類した学習データの割合のグラフを図 8 に示す。これは、それぞれのデータセット収集方法で 50 回ずつ実験を行い、平均した結果である。Using learning machine output method の方が Path following method に比べて目標経路から離れた状態のデータセットを多く収集していることを確認した。これらのことから、目標経路から離れた状態のデータセットを多く収集することで経路を追従できる確率が高まることが分かった。

Table 1: Number of successes

Experiment	Number of successes
Path following method	0/50
Using learning machine output method	29/50

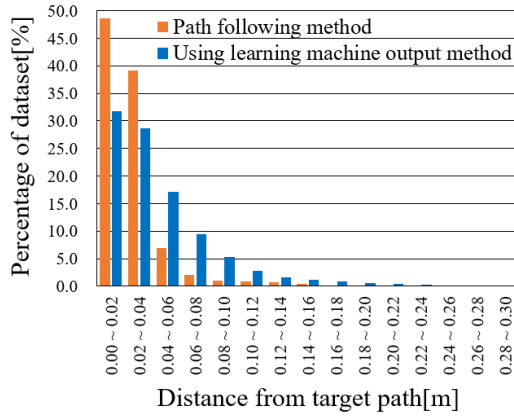


Fig.8: Percentage of the dataset in the experiment

4 生成された経路追従行動の解析

4.1 解析方法

生成された経路追従行動を比較するため、実験で生成した学習モデルでロボットを制御し、その経路追従行動を調査する。ロボットの初期位置を、図9のような場所に指定してロボットを走行させる。ロボットの進行方向を基準にして右側を正の値とし、左側を負の値とした。調査する場所は、曲がり角付近とした。この曲がり角というのは、実験において、経路追従行動に失敗する回数が特に多かった場所である。この場所を対象に、84カ所からロボットを走行させ、経路に戻るることができるのか調査した。また、一つの座標につき5方向（目標経路を追従するように制御した時のロボットの向きを基準として-10度～10度を±5度ずつ傾けた方向）をロボットの初期姿勢とした。つまり、一つの学習モデルにつき420カ所の初期位置からロボットを走行させ、経路へ復帰できるのか調査した。ここで、図10に示すように、壁に衝突することなく目標経路に近づいて、経路周辺（0.1m以内）を4秒間走行することができれば経路復帰に成功したと判断した。逆に、壁に衝突した場合を経路復帰に失敗したと判断した。

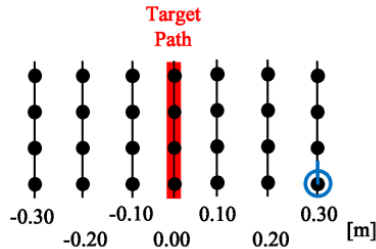


Fig.9: Initial position of the robot during analysis

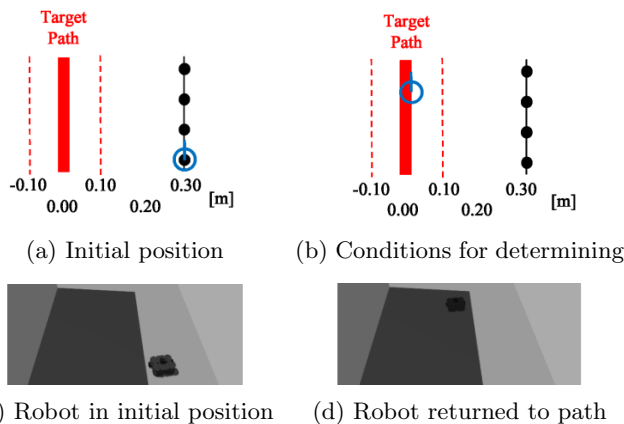
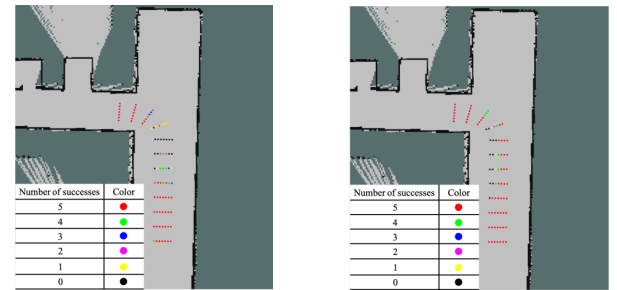


Fig.10: Analysis of the path-tracking behavior

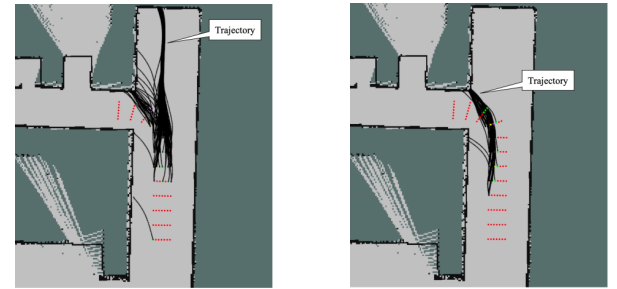
4.2 解析結果

解析の結果の一部を図11、図12に示す。図12は、ロボットが走行し始めた場所と経路復帰に成功した回数を示している。1つの座標につきロボットの方向を変えながら5回経路に戻るかを調査しており、経路復帰に成功した回数を色で表している。この結果から、2つの手法ともに曲がり角に入る前で経路復帰に失敗する傾向があることが確認できた。図12は経路復帰に失敗した時のロボットの軌跡を表している。図12(a)に示すように、Path following methodでは曲がり角で、曲がることなく直進する傾向があったのに対し、図12(b)に示すように、Using learning machine output methodでは、曲がり角で曲がる動作がより多く確認できた。また、3章で示した実験で生成した50個の学習モデルのうち5個を解析し、その平均した結果を、図13に示す。この図は、目標経路からの距離ごとに経路復帰に成功した割合を表している。Path following methodでは経路周辺が経路復帰できる確率が高いが、経路から離れると経路に復帰する確率は低くなっていく傾向が見られた。それに対し、Using learning machine output methodは、進行方向を基準とした経路の右側からロボットを走行させた場合、Path following methodと比べて経路復帰できる確率が高いことが分かった。これらのことから、経路周辺の状態だけでなく経路から離れた状態も学習することで、経路から離れても経路復帰できる確率が高まることが確認できた。



(a) Path following method (b) Using learning machine output method

Fig.11: Visualization of the number of path returns



(a) Path following method (b) Using learning machine output method

Fig.12: Robot trajectory during analysis

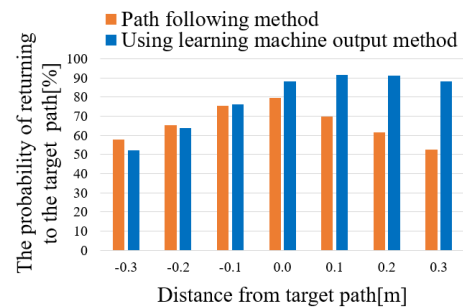


Fig.13: Percentage able to return to the pathway

5 経路復帰行動を強化する教師データ収集法の検討

経路から離れた状態も学習することで、経路を追従できる確率が高まることが明らかになった。よって、経路復帰行動を強化するためのデータセット収集方法を提案し、実験を行う。

5.1 提案するデータセット収集法

提案する手法では、経路からの距離が 0.05m 未満であれば 1 つ、0.05m 以上 0.10m 未満であれば 2 つ、0.1m 以上であれば 3 つ同じデータをデータセットに加える。よって、経路から離れた状態からの復帰行動を多く学習することになる。本稿では、このデータセット収集法を Enhanced return action method とする。

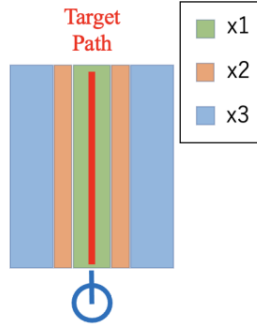


Fig.14: Enhanced return action method

5.2 シミュレータを用いた実験

提案するデータセット収集法の有効性を検証するため、シミュレータを用いた実験を行う。

5.2.1 実験装置

実験装置は 3.1 章で説明したものと同様である。

5.2.2 実験方法

実験方法は 3.3 章で説明したものと同様である。

5.2.3 実験結果

実験の結果を表 2 に示す。提案するデータセット収集法の方が経路を追従できる確率が高いことが分かった。ここで、データセットに含まれる目標経路からの距離で分類した学習データの割合のグラフを図 15 に示す。Enhanced return action method は、より経路から離れた状態のデータを多く収集していることが確認された。

Table 2: Number of successes

Experiment	Number of successes
Using learning machine output method	29/50
Enhanced return action method	42/50

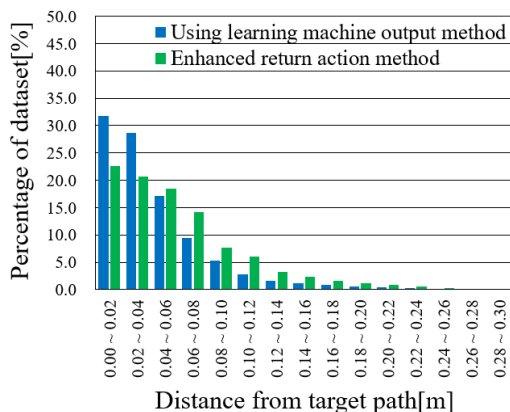


Fig.15: Percentage of the dataset in the experiment

5.3 生成された経路追従行動の解析

4 章で行った解析と同様に、本章で提案する手法によって生成した経路追従行動を調査する。4.1 章で示した方法と同様に、実験により生成した学習モデルを用いて解析を行う。5.2.3 章で示した実験で生成した 50 個の学習モデルのうち 5 個を解析し、その結果を平均したものを図 16 に示す。この図は、目標経路からの距離ごとに経路復帰に成功した割合を表している。この結果から、目標経路の周辺全てにおいて、提案する手法は、既存の手法である Using learning machine output method と比べて経路復帰に成功する確率が高いことが確認できた。

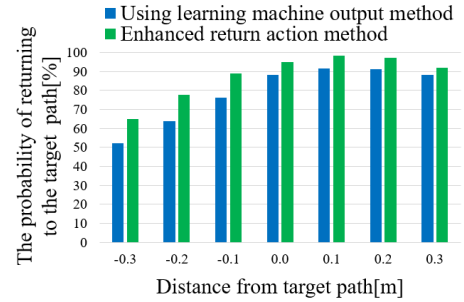


Fig.16: Percentage able to return to the pathway

6 結言

本稿では、従来から提案する経路追従行動を模倣する手法において、目標経路から離れた状態も学習することが有効であることを明らかにした後、経路復帰行動を強化するための教師データ収集方法を新たに提案し、シミュレータを用いた実験により、手法の有効性を検証した。また、学習器の出力を用いて、経路から離れた位置から経路に復帰できるかを解析し、経路から離れた状態も学習したほうが経路追従できる確率が高くなることを確認した。

参考文献

- [1] 岡田真也, 清岡優祐, 上田隆一, 林原靖男: “視覚と行動の end-to-end 学習により経路追従行動 をオンラインで模倣する手法の提案”, 計測自動制御学会 SI 部門講演会 SICE-SI2020 予稿集, pp.1148-1152(2020).
- [2] 岡田真也, 清岡優祐, 春山健太, 上田隆一, 林原靖男: “視覚と行動の end-to-end 学習により経路追従行動 をオンラインで模倣する手法の提案 -経路追従行動の修正のためにデータセットを動的に追加する手法の検討-”, 計測自動制御学会 SI 部門講演会 SICE-SI2021 予稿集, pp.1066-1080(2021).
- [3] W. Schwarting, J. Alonso-Mora, and D. Rus. Planning and decision making for autonomous vehicles . Annu. Rev, Vol. 1, No. 1, pp. 188-210, May 2018.
- [4] U. Muller, J. Ben, E. Cosatto, B. Flepp, and Y. Cun. "Off-road obstacle avoidance through end-to-end learning ". Advances in neural information processing systems, Vol. 18, 2005.
- [5] Bojarski, Mariusz, and et al. "End to end learning for self-driving cars". arXiv:1604.08316, 2016.
- [6] 清岡優祐, 岡田真也, 岩井一輝, 上田隆一, 林原靖男: “視覚と行動の end-to-end 学習により経路追従行動 をオンラインで模倣する手法の提案 -データセットと生成された経路追従行動の解析-”. 計測自動 制御学会 SI 部門講演会 SICE-SI2021 予稿集, pp.1071-1075(2021).
- [7] ros-planning, navigation リポジトリ (最終閲覧日 :2023 年 2 月 6 日). <https://github.com/ros-planning/navigation>.
- [8] Nathan Koenig and Andrew Howard. "Design and use paradigms for gazebo, an open-source multi-robot simulator ". In 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)(IEEE Cat. No. 04CH37566), Vol. 3, pp. 2149-2154. IEEE, 2004.
- [9] Erico Guizzo and Ackerman Evan. "The Turtlebot3 Teacher [Resources Hands On] ". IEEE Spectrum, Vol. 54, pp. 19-20, August 2017.