

走行経路情報を用いた車エージェントの運転行動学習の改善

米元聡¹ 松尾祐輝¹ 武田颯馬¹

概要: 現在、実用化されつつある公道での自動運転では、LiDAR やカメラから得られる実世界の計測情報と、リアルタイムに与えられる前提のウェイポイントと呼ばれる走行経路の情報を用いて、その走行経路との乖離がないよう運転行動を学習させることが多い。本研究では、走行経路の情報を行動学習には利用するが、観測情報として利用できない状況を想定した運転行動学習を走行シミュレーションにて実現した。具体的には、観測情報にフロントカメラ画像を用いる車エージェントを用いて、走行シミュレータ上の道路を指示した走行経路に沿って走行するよう深層強化学習により運転行動を学習させた。また、観測画像をもとに進路を推定する手法についても検討した。

キーワード: 自動運転シミュレーション, 深層強化学習, ウェイポイントナビゲーション

Simulation of Autonomous Driving Using Waypoints

SATOSHI YONEMOTO^{†1} YUKI MATSUO^{†1} SOMA TAKEDA^{†1}

Abstract: This paper presents a simulation of autonomous driving method based on deep reinforcement learning. We use only the front camera images as observation and discretized steering as actions for learning. In the learning process, the rendered routes that can be generated by using waypoints are introduced to drive stably. We also use the CNN-based methods to estimate the vehicle routing from the observed image. In several experiments, we show the learning results using waypoints and the estimation results of the vehicle routing.

Keywords: Self-Driving Simulation, Deep Reinforcement Learning and Waypoint Navigation

1. はじめに

現在、自動運転技術の実用化が進み、自動運転車が公道を走れるレベルに近づきつつある。また、大量の走行データの収集、リアルタイムで高精度な計測が可能な機器の導入、マップデータの高精度化・整備などにより、多くの企業、自治体が自動運転の技術の開発に取り組めるようになってきている。自動運転を実現するためには、車周辺で計測した実世界情報と受信したマップデータなどを即座に統合し、ルートに従って安全に走行させることが課題となる。

一方、シミュレーション環境を用いた運転行動学習の技術にも期待が高まっている。自動運転技術は、人工知能(AI)技術の進展とともに発展してきている。特に、観測情報をもとに行動を決定する強化学習によるアプローチが盛んに研究されている。その中でも深層強化学習は、ゲーム画面をもとにゲームを攻略する AI への適用から始まり、カメラ画像をもとに運転行動を学習する自動運転の研究にも応用されている。例えば、「End-to-End Learning」というアプローチでは、道路画像を入力、車の行動を出力とする形で学習を行うため、陽に特徴抽出を必要とせず、畳み込みニューラルネットワークの学習のみで行動学習が完結する。

自動運転に利用可能な走行シミュレータが数多く提案

されている[1][2]。例えば、CARLA[1]では、天候や街並みなどをリアリスティックに表現した道路シーンの生成が可能であり、公道を走る自動運転車の学習を想定した本格的なシミュレータとなっている。また、このシミュレータで生成できる走行経路情報を利用し、既知のルートについて走行を評価する環境も整っている。

自動運転のために、マップデータの1つであるウェイポイント(Waypoint)情報を利用してリアルタイムに走行経路を推定し、行動学習に利用することも検討されている[3][4][5]。ウェイポイント情報とは、走行すべき経路上の位置や走行速度などのデータのことであり、実機車両の車両モデルが既知で観測情報として利用可能であれば、車の軌道とその走行経路に当てはめる推定問題となる。しかし、車両の種類ごとにモデルの構築が必要なことや、常時ウェイポイント情報が得られると想定できない場合もあるため、ウェイポイント情報が事前に収集できる学習時だけ利用するアプローチも必要と考える。A.Amini らの研究[3]では、ルートが未知の周辺マップ画像と観測画像を DNN に入力として与え学習することで、進路の GMM モデルを用いた確率的方法で進路を推定している。また、ルートが既知の場合は、進路を決定論的方法により求めることができるよ

¹ 九州産業大学
Kyusyu Sangyo Univ.

うになっている。

2. 深層強化学習を用いた運転行動の学習

2.1 走行シミュレータ

本研究では、運転行動学習の問題に焦点を当てるため、できるだけ簡素化した走行シミュレータの開発を目指している。本研究では、道路を2本の曲線対による走行レーンのみで表現する。図1に走行シミュレータで作成した道路コースの例を示す。この例のように、道路は連結した曲線セグメントであり、曲線セグメントにベジェ曲線セグメントを用いている[6][7]。曲線セグメントを自由に連結し、制御点のパラメータを自由に変動させることで、様々な道路コースを表現できる。また、複数の道路を重ねることで、分岐(T字路、十字路など)も表現可能である。走行レーンを示す白線をフロントカメラから観測すると、観測画像は二値のエッジ画像となる(解像度は自由に変更できる)。

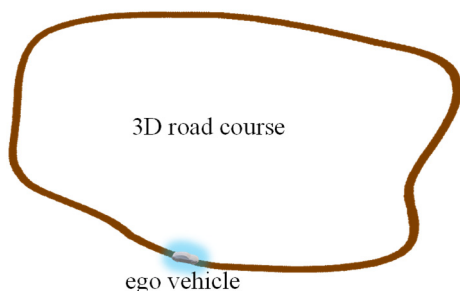


図1 道路コースの例

Figure 1 An example of 3D road course.

2.2 車両モデル

車エージェントの仕様について述べる。車両モデルの制御にはキネマティックモデルを利用する[6][7]。行動学習時は、左右にハンドルを切るか、直進するかの3種類のステアリング操作のみ行えるものとする。アクセルにより加速した後は、一定の速度で運転する。ブレーキ制御の行動の際には、減速によるブレーキ操作を行うことができるが、本稿では扱わないものとする。以降説明する実験では、車の全長45、車幅20とし、地面から20の高さで車の中心付近にフロントカメラを設置した。

2.3 PPO

深層強化学習は、強化学習に深層学習を組み合わせた手法であり、もともとゲーム画面をもとにAIコントローラがよりスコアを稼ぐ行動を学習していくというゲームAIの問題のために提案された方法[8]で、2人ゲーム[9]、ロボットの制御、自動車、船の自動運転などに応用されている。DQN法[8]が有名であるが、その後、改良が進み、様々な手法が提案されている。中でも、方策勾配法にもとづくPPO(Proximal Policy Optimization)[10]がよく利用されるようになってきている。本研究も運転行動学習にPPOを利用する。

なお、実装にはStable BaselinesのPPOを利用する[11]。

2.4 車エージェントの定義

基本となる車エージェントを以下のように定義する。

(1) 行動

車エージェントは、直進・左右回転の3種類のステアリング操作を行う。一定の速度になるまで加速し、その後定速で走行する。急カーブを想定する場合は、ブレーキ操作も追加する必要があるが、今回は操作の対象としないものとする。

(2) 観測

車エージェントのフロントカメラから見たレーン画像を観測画像として与える。画像サイズは最小で 64×64 とした。

(3) 報酬

基本的に走行できれば報酬 r として+1を与える。車中心がレーンからはみ出さず障害物に当たると-1を与える。

(4) エピソード終了条件

車エージェントが走行レーンを横切るか、障害物に当たった場合にエピソード終了もしくは目標のステップ数(走行距離)達成でもエピソード終了となる。

以上より、本研究で定義する車エージェントの行動空間は3つの離散値で、観測空間は最小サイズで $64 \times 64 \times 3$ となる。

2.5 フロントカメラ画像の生成

走行シミュレータにおいて、車エージェントのフロントカメラから3次元レーン曲線の投影像を得る方法について述べる。コース上の3次元のレーンを投影することにより、車エージェントから見た道路の投影像を生成することができる。車エージェントの観測画像にこの投影像を用いる。ワールド座標系を車両中心(地面が高さ0)に設定し、カメラ座標系を車エージェントの前方付近に設定する。カメラ座標系のパラメータであるピッチ角は 30° に設定する。レーン上の3次元の点から、ワールド座標からカメラ座標への変換、ビューポイント変換を経て、観測画像上の点を求めることができる。また、シミュレーション用に、車視点の鳥瞰画像(車両上方から見た画像)も生成することが可能である。

3. ウェイポイント情報の利用

3.1 ウェイポイント情報とは

公道で実現する自動運転車の技術には、目的地までのルートを示したマップ情報が用いられる。その1つであるウェイポイント情報は、目的地までの走行経路の情報であり、車周辺の進路や目先の道路情報を知ることができる。利用にあたっては、実際の車両位置と照合する必要があり、周辺のウェイポイントをもとにした経路推定が行われる。ウェイポイント情報をリアルタイムに観測情報として利用できれば、制御の問題として解くことで経路に沿った走行を

実現できるが、本研究では、ウェイポイント情報を観測情報として想定せず、フロントカメラ画像のみを観測情報とする。一般に、ウェイポイント情報は、通過すべき地点の系列であり、ウェイポイントの1点は通過すべき位置、走行速度などから構成される。本研究では、ウェイポイント情報を2次元位置の系列とみなし、速度は利用しないものとする。

3.2 ウェイポイント画像の生成

ウェイポイント情報をもとに本走行シミュレータ上に走行経路を生成し、コースマップ上の領域として表現する。図2に、このウェイポイント情報をコースマップ上に重ねた例を示す。マゼンタ色の部分が生成した走行経路である。また、車エージェントから見た走行経路を図中央に示す。図中央左が観測画像、図中央右がウェイポイントの車エージェント視点の鳥瞰画像である。本研究ではこれを「ウェイポイント画像」と呼ぶこととする。

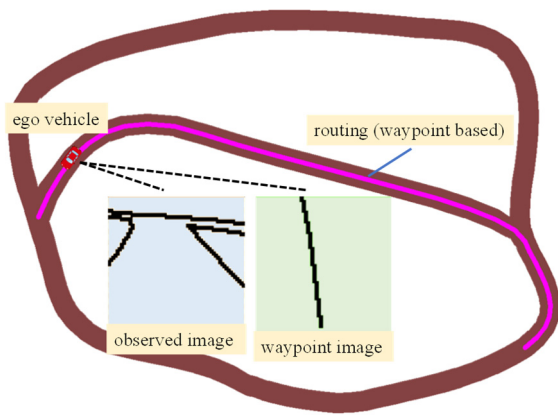


図2 ウェイポイント情報と走行経路

Figure 2 Waypoints and the routing.

3.3 ウェイポイント画像の生成

ウェイポイント画像をコースマップ上に重ねることで、現在の車エージェントがその走行経路上にあるかについておおよそ知ることができる。すなわち、走行経路の領域内に位置すれば、向きは多少異なっても正しい経路に従っていると判断できる。シミュレーションによる学習時は、ウェイポイント情報が既知の道路コースを用意できるため、ウェイポイント情報を運転行動の修正に積極的に利用することができる。

次に、ウェイポイント画像を報酬計算に利用する方法について述べる。ウェイポイント情報を利用する運転行動学習の場合は、以下の式(1)に示す報酬 r に変更する。

$$r = r_1 + (1 - r_1) \times r_{wp} \quad \dots (1)$$

この報酬 r は、単に走行可能であれば r_1 のみ与えられ、走行経路上にある確率 r_{wp} が1に近ければ $r = 1$ となる報酬を意味する。図3にウェイポイント画像から走行経路上にある確率 r_{wp} を求める方法を示す。なお、車中心が走行経路上に

あれば1とすることも可能である。

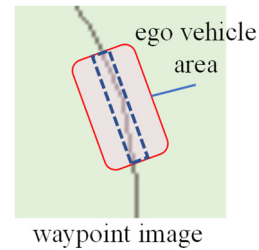


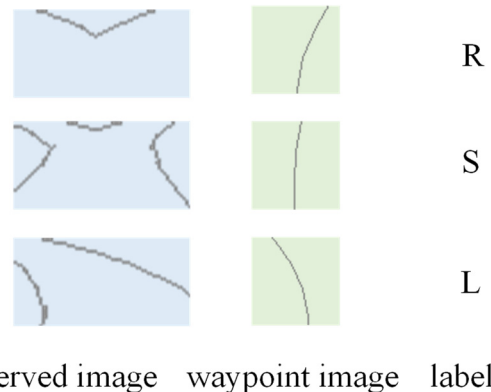
図3 走行経路上にある確率の計算

Figure 3 Probability calculation for the route occupancy.

4. CNN を用いた進路の学習

4.1 ウェイポイント画像の解析

本シミュレータでは、ウェイポイント情報を表す走行経路を曲線で表現する。上述したとおり、ウェイポイント画像とは、ウェイポイント情報である走行地点の系列をつないだ曲線群を道路上へ描画し、それを車エージェント視点の鳥瞰画像へ変換することにより取得できる画像のことである。図4に、車エージェントの観測画像とウェイポイント画像の例をいくつか示す。この図の例では、上段の観測画像の場合、分岐で右に走行する必要があることを示している。同様に中段は直進、下段は左にカーブを意味する。



observed image waypoint image label

図4 観測画像とウェイポイント画像の例

Figure 4 Observed images and waypoint images.

ウェイポイント情報が利用可能な場合、行動学習時に正確な進路を意味するウェイポイント画像を安定に得ることができる。進路を示す曲線または直線は、車エージェント視点の鳥瞰画像の中央付近に位置する。ウェイポイント情報が不確かなルートの走行では、観測画像のみから進路を予測する必要がある。学習により、進路を想起することができればよいが、非常に難しい問題である。そこで本研究では、進路を「left」、「straight」、「right」の3種類の方向に絞り、予測することを検討した。今後、「この先、左に急カーブ」のように推定する進路の精度を上げることも可能と考える。なお、右方向、左方向といった「進路」を画

像より推定することは、観測画像からその時刻における行動を決定することに直結するものではない。

次に、ウェイポイント画像の解析の手順を図 5 に示す。進路のラベルを $\{S, L, R\}$ とする。このラベルが、対応する観測画像の教師データとなる。最終的には、ウェイポイント画像上の進路を曲線または直線として表現する。ウェイポイント画像を解析し、図のように x 座標ごとに走行経路上の画素の平均座標を求め、その左右への偏りを判定することでおよその判定が可能である。今回の実装ではこの判定方法を用いている。以降で説明予定の、学習に用いるウェイポイント画像（走行経路）は、安定に走行している状況で収集したものであるため、画像のほぼ中央の曲線または直線として出現する。

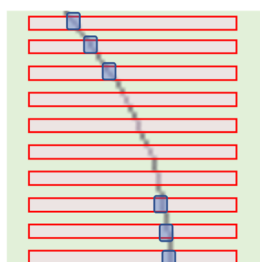


図 5 進路の判定

Figure 5 The predictive routing.

4.2 CNN による学習

観測画像と対応するウェイポイント画像をもとに進路の学習を行う。学習結果をウェイポイント情報の与えられない場合の進路として利用することを想定する。本研究では、 64×64 のウェイポイント画像および 64×64 の観測画像の一部 (64×32 の画像) を学習データに用いる。このように観測画像から車前方の半分の高さの画像を切り出して用いる。この切り出した 64×32 の観測画像 n 枚および教師データとなる進路のラベル $\{t_1, \dots, t_n\}$, $t_i \in \{S, L, R\}$ をもとに学習する。繰り返しになるが、進路のラベルは、上述したウェイポイント画像の解析により事前に求めておく。

実験には、ウェイポイント情報である走行経路を既知とした道路コースを走行させ取得した 1,715 枚を用いた。観測画像は学習時に 32×32 にリサイズし、正規化する。進路のラベルは、対応するウェイポイント画像を解析し求める。CNN の実装には Tensorflow, Keras ライブラリを用い、計算時に GPU を用いた。エポック数は 50 とした。利用した CNN のモデルは、畳み込み層 $\times 2$, MAX プーリング層 $\times 1$, 畳み込み層 $\times 2$, MAX プーリング層 $\times 1$, 全結合層 $\times 2$ を連結する構成とした。図 6 に学習時の分類精度・誤差の推移を示す。学習データに対しては、分類精度 (accuracy) は 98.95%, 誤差 (loss) は 0.013 であった。

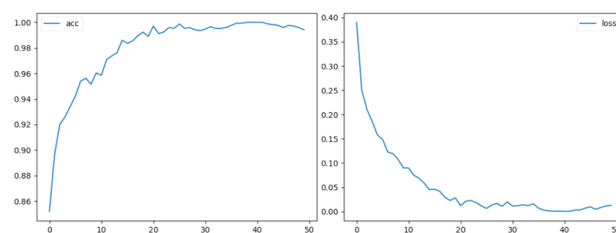


図 6 分類精度・誤差の推移

Figure 6 Accuracy and loss .

4.3 推定した進路の利用

推定した進路が精度のよいものであれば進路を逆投影することにより、走行経路の推定情報として活用できる。

A.Amini らの研究[3]では、ルートが未知の周辺マップ画像と観測画像を入力として与えることで、3 種類の進路についての確率的な判定が可能となっている。本研究では、およその方向しか求めていないが、ウェイポイント情報のない状況下で進路を判断する場合に利用可能である。

5. 実験結果

5.1 概要

本研究では、車エージェントの実装に強化学習向けのエージェント開発のフレームワークである OpenAI Gym[12] を利用した。この車エージェントを含む環境に Stable Baselines の PPO を組み込んだ。また、走行シミュレータ（道路コース生成、フロントカメラから見た 3 次元レーン曲線の投影像の生成）については独自に開発したものを用いた[6][7]。

5.2 運転行動の学習

まず、生成した道路コースについて行動学習した結果について述べる。観測画像には、レーン画像の解析結果を用いることも可能であるが[13][14][15]、今回は投影像をそのまま観測画像として用いている。学習には図 2 に示す短い道路コースを用いる。ただし、走行の難易度を上げるため、道路コースの道幅を広めに設定している。これにより、学習した走行にふらつきが発生することが予想される。また、分岐点を含んでいることから、車エージェントにとって難しい判断が必要と予想する。図 7 に 20 万ステップ学習した結果を示す[16]。横軸はエピソード数、縦軸はそのときの報酬である。グラフは、(a)ウェイポイント情報を利用する場合 (WP), (b)利用しない場合 (NO-WP) の結果を示している。不安定な報酬の推移になっているものの、ともに 2 万回程度で目標ステップ数 (コース 1 周半) に達していた。(a) の場合は、報酬関数が 1.0 以下となるため、予想通り平均報酬は下がった。また、(b) の場合は、15 万回程度で報酬は安定したが、これは、(b) の報酬が「長時間走行可能か」のみを評価しているため、走行レーン内を大きくふらついていても達成できる。一方、(a) の場合については、走行経路上を完全に走行するまで学習回数を要し、平均報酬の推

移が安定しないことも確認できた。そこで、テスト走行において、(a)(b)の場合についてそれぞれ走行経路上にあるかを目標ステップ数の間、計測し確認した。それぞれ 90.8%, 25.7%であった。このことから「走行経路を安定に走行する」という目的は、(a)の場合のみ、達成できていることを確認できた。また、(a)の場合のみ、設定した分岐点で指定した方向に間違えることなく走行できることも確認した。



図7 エピソード報酬の推移

Figure 7 Transition of the episode rewards .

5.3 CNN を用いた進路の学習

図2のコースについて、ウェイポイント上の進行方向に正確に走行させることで獲得した未学習データすなわち観測画像およびウェイポイント画像をもとに、推定精度の検証を行った。実験にはウェイポイント129点を用いた。進行方向の連続する2点からその方向を求め、車エージェントの向きとした。学習時と同様に、ウェイポイント画像から対応するラベルを解析により求め、真値とした(目視で正確なラベルであることも確認した)。学習したCNNで進路を判定したところ、分類精度は90.7%であった。この結果より、進路が正確に求められたことがわかる。学習に用いていない、運転に失敗したシーンの観測画像に対する評価は今後検討する。

図4上段の分岐の例のように、左右どちらの進路を割り当てるかは、学習時の与え方に依存する。本手法は、過去(学習データ)から進路を学ぶ手法であるが、この例のように、そもそもAI自身で判断できない場合は進路について明確な指示が必要である。

6. おわりに

本稿では、深層強化学習を用いた運転行動学習について述べた。深層強化学習アルゴリズムの1つであるPPOを用い、ステアリング操作の行動を学習させた。観測情報は走行シミュレータを用いて生成されるフロントカメラ画像のみとした。ウェイポイント情報が利用可能な場合は、その走行経路上にあるかどうかについて評価する報酬を設定した。実験の結果、ウェイポイント情報を利用する場合は、1)走行経路を完全に通過できるようになるまで学習に時間を要すること、2)走行経路に沿って学習するため走行が安

定すること、3)分岐点において、確実に指定した方向を通過するよう学習できることを確認できた。

また、CNNを用いて進路を学習する手法についても検討した。観測画像とウェイポイント画像を用いて、進路のラベル3種類について学習し、観測画像をもとに進路を推定できるようになった。

今後の課題は、1)直角に曲がるレーンや、交差点、ラウンドアバウトなど、より複雑な道路コースに対する学習実験と評価、2)、レーンチェンジへの適用、3)推定した進路にもとづく行動の予測、4)より精度の高い進路の推定などである。

参考文献

- [1] Dosovitskiy, Alexey, et al. CARLA: An open urban driving simulator. In: Conference on robot learning, PMLR, pp.1-16, 2017.
- [2] Shah, Shital, et al. Airsim: High-fidelity visual and physical simulation for autonomous vehicles, In: Field and service robotics. Springer, Cham, pp.621-635, 2018.
- [3] A. Amini, et al. Variational end-to-end navigation and localization, International Conference on Robotics and Automation(ICRA). IEEE, 2019.
- [4] J. Xu, and Tao Wu. End-to-end autonomous driving based on image plane waypoint prediction, International Symposium on Control Engineering and Robotics (ISICER). IEEE, 2022.
- [5] J. Xu, et al. Trajectory Prediction for Autonomous Driving with Topometric Map, International Conference on Robotics and Automation (ICRA). IEEE, 2022.
- [6] 米元聡, 國武佑哉, 道路俯瞰画像を用いた車エージェントの運転行動学習, 火の国シンポジウム, 2020年.
- [7] 米元聡, 菅河凌太, 深層強化学習を用いた運転行動学習のための走行シミュレータの検討, 火の国シンポジウム, 2021年.
- [8] V. Mnih, et al. Playing atari with deep reinforcement learning, 2013, arXiv:1312.5602.
- [9] D. Silver, et al. Mastering the game of Go with deep neural networks and tree search, 2016, nature, 529(7587), 484.
- [10] Schulman, John, et al. Proximal policy optimization algorithms, 2017, arXiv:1707.06347.
- [11] Stable Baselines, <https://github.com/hill-a/stable-baselines>
- [12] 米元聡, 鶴山侑生, 新田龍一, 深層強化学習を用いた自動駐車シミュレーション, 火の国シンポジウム, 2022年.
- [13] Luca Venturi, Krisstof Korda, Hands-On Vision and Behavior for Self-Driving Cars, Packt Press, 2020.
- [14] Udacity Self-Driving Car Nanodegree Project, Finding lane lines, <https://in.udacity.com/nanodegree/>.
- [15] OpenAI <https://github.com/openai/baselines>.
- [16] Tensorboard <https://www.tensorflow.org/tensorboard>.