

令和 6 年 6 月 27 日現在

機関番号：12608

研究種目：基盤研究(B) (一般)

研究期間：2021～2023

課題番号：21H03517

研究課題名(和文) 安定・安全を指向する逆強化学習に基づく運転行動モデリング

研究課題名(英文) Driving behavior modeling based on stability and safety oriented inverse reinforcement learning

研究代表者

下坂 正倫 (Shimosaka, Masamichi)

東京工業大学・情報理工学院・准教授

研究者番号：40431796

交付決定額(研究期間全体)：(直接経費) 13,100,000円

研究成果の概要(和文)：近年、先進運転支援システムの開発が盛んである。本研究ではこれらのシステムの高度化を念頭に、逆強化学習(模倣学習・逆最適制御)に基づく運転行動モデリング(運転行動予測とシミュレーション生成)する手法を開発した。特に、自動車運転行動モデリングの応用の観点にたち、従来の技法で不足していた安定性と安全性を指向する技術を構築した。具体的には、確率的経路生成手法と経路候補に基づく重点サンプリングによる効率的な報酬場推定手法と、負例(交通事故に該当するデータ)を用いた効率的な学習法を開発した。また、負例に該当するデータを安全に収集するため、運転行動データ収集環境を構築し、実際に運転データ収集を行った。

研究成果の学術的意義や社会的意義

本研究で構築した技術の一つである、RRTに基づく経路生成法、及び、重点サンプリング法に基づく報酬場最適化手法は、逆強化学習に基づく運転行動モデリングの適用範囲を拡張するものである。また、負の事案を用いた手法は、熟練ドライバーの運転行動の表層的な模倣だけでなく、安全性という価値感を含めた模倣を目指す基盤となっている。構築した技術は交通事故を低減する新たな運転支援システムの開発に貢献する。例えば、事故要因分析や事故回避のシミュレーションへの応用が期待できる。

研究成果の概要(英文)：In recent years, the development of advanced driver assistance systems has been active. In this study, we developed a method for modeling driving behavior (driving behavior prediction and simulation generation) based on inverse reinforcement learning (imitation learning and inverse optimal control) with a view to the advancement of these systems. In particular, from the viewpoint of application to modeling of automobile driving behavior, we developed techniques oriented toward stability and safety, which have been lacking in conventional techniques. Specifically, we developed a probabilistic path generation method, an efficient reward estimation method using sampling based on path candidates, and an efficient learning method using negative examples (data corresponding to traffic accidents). In addition, to safely collect data corresponding to negative examples, we constructed a driving behavior data collection environment and actually collected driving data.

研究分野：知能ロボティクス

キーワード：逆強化学習 運転行動モデリング 経路生成 RRT 運転行動シミュレーション

1. 研究開始当初の背景

近年、先進運転支援システム(Advanced Driver Assistance Systems: ADAS) と呼ばれる、自動ブレーキや追従機能などの技術の開発が盛んである。これらは現在市場に出回る多くの車両に搭載され、事故を未然に防ぐための効果的なシステムとして注目を集めている。一方、ADASは幹線道路における活用が主であり、生活道路において十分に機能するシステムの実現には至っていない。この原因として、生活道路には信号なし交差点や歩行者の存在等の潜在的な危険因子が多く存在することが考えられる。この場合、人の運転特性を踏まえた行動のモデル化(運転行動の予測とシミュレーション生成)が重要であり、この課題において逆強化学習(模倣学習・逆最適制御)が注目されている。逆強化学習は、熟練者の一連の行動の観察から、エージェントの制御規範および制御則を獲得する枠組である。しかしながら、逆強化学習は、自動車運転行動モデリングの応用の観点、特に効率性と安全性の観点に立てば、技術は未完の状態であった。特に、与えられた報酬関数下で効率的に車両の最適経路を予測する手法が存在しないため、たとえ報酬関数が正確に付与されたとしても、熟練ドライバー運転行動の「模倣」自体不安定であった。また、「見たものすべて」を模倣する通常の逆強化学習では、負の事案を反面教師として学ぶという、模倣におけるもう一つの側面の実現が不可能である。これは、工学応用上重要な指標である、「安全性」の規範の獲得が難しいことを意味する。

2. 研究の目的

冒頭で述べた背景にもとづき、安定・安全を指向する逆強化学習ベースの運転行動モデリング技術の構築と、そのための評価基盤の構築を本研究の目的とする。具体的には以下の3つの課題に取り組む。

- ・課題1 安定かつ効率的な逆強化学習の基盤技術の構築
- ・課題2 安全性を指向する逆強化学習の基盤技術の構築
- ・課題3 逆強化学習ベース運転行動モデリングの評価基盤の構築

3. 研究の方法

課題1 従来の方策の最適化に基づく運転行動予測を置き換え、ロバストかつ効率的な「探索技術」に基づく逆強化学習を開発する。課題1における技術基盤は、課題2の構成要素として活用する。

課題2 熟練ドライバーの運転データ(正例)と共に、ヒヤリハット・事故発生時の運転データを反面教師として用いて行動規範を獲得する逆強化学習技術を開発する。負の事案を明示的に収集したデータを活用し、正の事例と負の事例を弁別しながら正の事例に基づく制御規範の獲得を行う方法論である。

課題3 課題1, 2で構築する安定・安全を指向する運転行動モデリング技術の有効性検証を実施する。そのため、開発する技術の評価のための運転データを収集する。特に、課題2で重要になるヒヤリハット・事故発生運転データ(負の例)の収集は危険を伴う。安全性・件数を担保するため、シミュレーションベースの運転行動計測プラットフォームを構築し、シミュレータ上でも運転データ収集を行う。

4. 研究成果

得られた研究成果を課題1~課題3の順に示す。

課題1: 逆強化学習の構成要素の一つである、与えられた報酬場の下での最適経路生成手法を開発した。本研究で扱うのは、古典的な逆強化学習で議論されてきたような離散的状態空間での大域的最適な経路生成ではなく、高次元連続状態空間中の局所最適性の経路生成手法である。従来の経路生成、特に、近年深層学習で議論される方策最適化は経路生成では、安定的な経路の品質保証は困難であったが、本研究では、車両の力学的特性が既知であることを活用し、探索空間全体を確率的・網羅的に探索するRRTを用いる。RRTはロボット工学分野では古くから活用が進んできた手法であるが、逆強化学習との組み合わせについ

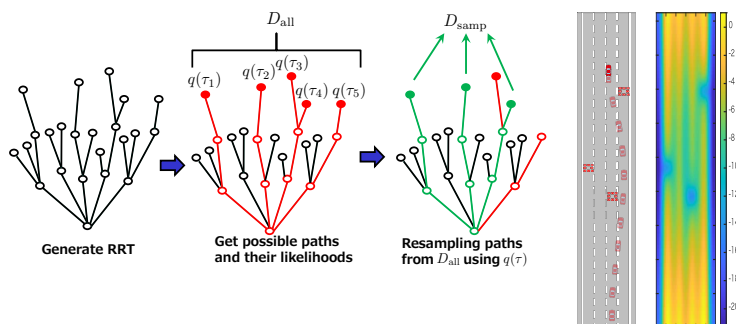


図1: テンプレート RRT に基づく

重点サンプリング法および経路生成、報酬場推定の様子

て議論の余地があった。

本成果は、自動車の力学的特性の一つである非ホロノミック運動に適したテンプレートベース探索手法であり、従来の枠組みと比べ、効率的に安定的な経路生成が可能である。また、本成果では、生成した経路の候補の利活用を工夫する、重点サンプリング法に基づく報酬場更新を行う。これにより、学習の効率化を達成する。様々な運転行動の予測問題に本手法を適用し、有効性を検証した。図1は、生成された木を利用した重点サンプリングの様子および、障害物回避・レーン変更の経路生成および、復元された報酬場を可視化したものである。

一連の成果を国際会議 IEEE IV2022 で発表した。また、テンプレートベースの RRT 手法を無信号交差点における停止行動に応用した事例を日本ロボット学会 RSJ2022 に、また、終端条件の整合性を考慮した RRT 経路生成手法として、日本機械学会 Robomech 2023 にて発表を行なった。

課題2：既存の逆強化学習とは対比的に、本課題で開発する逆強化学習は、負の事案を反面教師として学習する。負の事案を盛り込む逆強化学習は、これまでにいくつか提案されてきたが、システムに提示する正負の行動単位が経路全体であり、報酬場の学習が安定せず、収束に至らない点が課題となっていた。これは、正例と判別が不能な負例の局在的な経路が含まれることを意味し、判別不能なデータに対して正負のラベルを与えることで、学習が不安定化する。この問題を受け、本研究では、局所経路上の正負ラベル（例えば衝突行動の直前数秒程度のみ負ラベル）を付与し、学習を安定化する枠組みを開発した。具体的には、本成果は、最大エントロピー強化学習の枠組みを活用し、正負の事案の弁別を制約条件に加味するものである。これは、ラグランジュ未定乗数法で求解し大域的最適解を与え、効率的な報酬場の獲得が可能である。提案する枠組みを、無信号交差点における停止行動、障害物回避行動に適用し提案手法の有効性を検証した。評価では、理想的な運転行動の教示、および負の事案（例えば障害物に突入する）を与え、障害物に関する回避行動を励起するための報酬場を効率的に学習できるか否かを評価した。図2左は、提案手法の学習の過程を表現しており、既存手法に比して安定してモデルが獲得できることを示している。図2中は獲得した報酬場を示しており、図2右は生成された経路を表す。得られた成果は、IEEE IV2024 にて口頭発表（ポスター発表を含む501件中20件のみ口頭発表として選ばれている）で発表している。

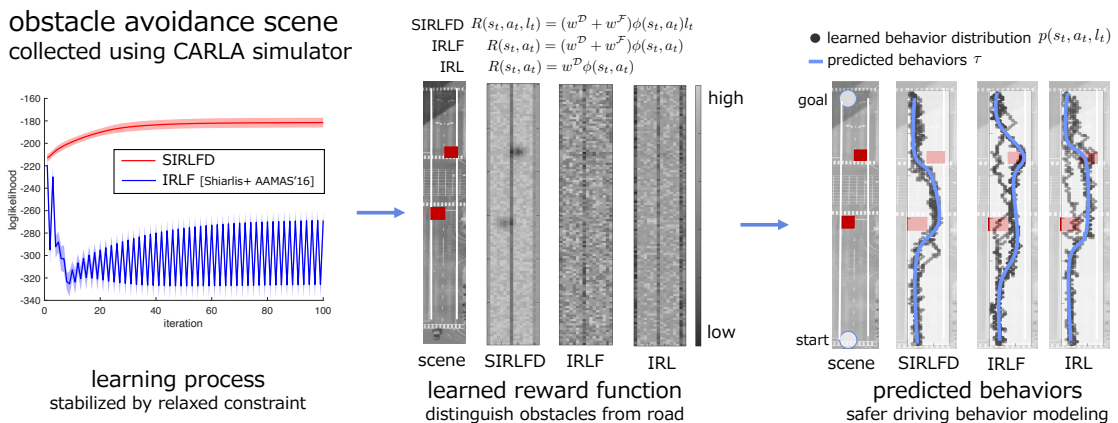


図2：正例・負例から学習する逆強化学習の対数尤度の推移，報酬場推定の結果，経路生成の様子

課題3：課題2では負の事案（交通事故）に該当する運転データを収集するため、実車両データではなく運転行動をシミュレータ上で収集する必要がある。本研究課題に取り組むに当たり、既存のオープンソースソフトウェアの利活用に着目した。LGSVL, Autoware, Summit, VISTA, AirSim, Gazebo, Udacity, CARLA といったシステムを調査し、当初 Autoware での開発を進めた一方、プロジェクト後半から CARLA を利用したデータ収集環境を構築した。運転データ収集システムでは、GPU を搭載した Linux ワークステーションおよびレーシングゲーム用コントローラ（ハンドル・アクセル・ブレーキ）を利用し、データを収集する。これにより、被験者および周辺環境に危険性を与えることなく、所望の負の事案データの収集が可能となった。CARLA が提供する仮想都市環境において、想定する運転状況、例えば障害物を動的に設置し、様々な条件の運転データを収集した。



図3：構築した運転行動データ収集用シミュレータ

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 0件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 下坂正倫	4. 巻 39
2. 論文標題 逆強化学習を用いた生活道路中の運転行動モデリング	5. 発行年 2021年
3. 雑誌名 日本ロボット学会誌	6. 最初と最後の頁 631 ~ 636
掲載論文のDOI（デジタルオブジェクト識別子） 10.7210/jrsj.39.631	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計7件（うち招待講演 0件 / うち国際学会 3件）

1. 発表者名 Shaoyu Yang, Hiroshi Yoshitake, Motoki Shino, and Masamichi Shimosaka.
2. 発表標題 Smooth and Stopping Interval Aware Driving Behavior Prediction at Un-signalized Intersection with Inverse Reinforcement Learning on Sequential MDPs.
3. 学会等名 2021 IEEE Intelligent Vehicles Symposium (IV) (国際学会)
4. 発表年 2021年

1. 発表者名 平川 優伎, 下坂 正倫
2. 発表標題 特徴量探索とパラメータ最適化の交互実行によるコンパクトな運転行動推定モデル
3. 学会等名 第22回 計測自動制御学会 システムインテグレーション部門講演会
4. 発表年 2021年

1. 発表者名 Shinpei Hosoma, Masato Sugasaki, Hiroaki Arie, and Masamichi Shimosaka
2. 発表標題 RRT-based maximum entropy inverse reinforcement learning for robust and efficient driving behavior prediction
3. 学会等名 2022 IEEE Intelligent Vehicles Symposium (IV) (国際学会)
4. 発表年 2022年

1. 発表者名 Minglu Zhao and Masamichi Shimosaka.
2. 発表標題 Inverse Reinforcement Learning with Failed Demonstrations towards Stable Driving Behavior Modeling
3. 学会等名 2024 IEEE Intelligent Vehicles Symposium (IV) (国際学会)
4. 発表年 2024年

1. 発表者名 Shaoyu Yang, Minglu Zhao, and Masamichi Shimosaka
2. 発表標題 Improved Template-Based RRT for Efficient Kinodynamic Motion Planning
3. 学会等名 日本機械学会ロボティクス・メカトロニクス講演会2023
4. 発表年 2023年

1. 発表者名 Shaoyu Yang, and Masamichi Shimosaka
2. 発表標題 Sequential 2D Continuous Kinodynamic RRT for Driving Behavior at Un-signalized Intersections with Stop Lines
3. 学会等名 第40回日本ロボット学会学術講演会
4. 発表年 2022年

1. 発表者名 趙 Minglu (JIS第二水準漢字の制限のためアルファベット表記), 羊 少宇, 下坂 正倫
2. 発表標題 危険行動の時間的局所性に着目した負例導入逆強化学習の安定化
3. 学会等名 日本機械学会ロボティクス・メカトロニクス講演会2023
4. 発表年 2023年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分担 者	小竹 元基 (Shino Motoki) (10345085)	東京工業大学・工学院・教授 (12608)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------