Deep Q-Networkを用いた自動運転車の ゆずりあいによる交通流の効率化

Efficiency of Traffic Flow with Mutual Concessions of Autonomous Cars Using Deep Q-Network

小川一太郎 *1 Ichitaro Ogawa	横山想一郎 *1 Soichiro Yokoyama	山下倫央 *1 Tomohisha Yamashita	川村秀憲 *1 Hidenori Kawamura	酒徳哲 *2 Akira Sakatoku
	柳原正 *2 Tadashi Yanagihara	大岸智彦 *2 Tomohiko Ogishi	田中英明 *2 Hideaki Tanaka	

*1北海道大学大学院 情報科学研究科

Graduate School of information Science and Technology, Hokkaido University.

*²KDDI 総合研究所 KDDI Research Inc.

In recent years, the development of autonomous operation technology has been actively carried out in a various research institutions and companies. Many experiments are conducted in public road to confirm whether one autonomous car can drive safely. However, autonomous operation with inter-vehicle communication has not been much researched. In this paper, we implement mutual concessions of autonomous cars with Deep Q-Network (DQN), which is a deep neural network structure used for estimation of Q-value of the Q-learning method. To verify the effectiveness of mutual concessions, we develop an experiment environment for verification of autonomous cars at the confluence in the roundabout. DQN is applied for the decision-making mechanism to decide velocity in the roundabout based on the position of others and the status of congestion. As a result of our experiment, we confirmed that the autonomous cars with DQN can realize high transfer efficiency in the roundabout.

1. はじめに

自動運転車が広く普及すると、人間が運転するために設計された現在の交通ルールが不適切な可能性がある。例えば合流時に優先車がいる場合、合流車は道を譲らなければならないが、 自動運転車が優先車を決めれば待ち時間は発生しなくなるだろう.本研究では深層学習の一つである Deep Q-Network を使用して、自動運転車両同士が経路を譲り合ってもここの移動が 阻害されることなく全体の効率を増加させることを目指す.

Deep Q-Network を導入できる自動運転システムアーキテ クチャの開発を行った.本研究では自動運転車のみが多数存在 している状況を想定して,交通流量を増加させることを目的 としたシミュレーションを行う.ラジオコントロールカーを使 用した物理シュミレーションを採用した.ラジオコントロール カーを使用した理由は、コンピュータ上の理想的な環境ではな くノイズなどの外乱が発生する状況で物理特性を再現可能で あることと、実車両よりも非常に安価であるとこがあげられ る.ラジオコントロールカーの制御を実現するため、行動をモ ジュール化することで層状に行動を決定するアーキテクチャを 作成した.

本論文ではラウンドアバウトを含んだコースを作成し, ラウ ンドアバウト内の走行を Deep Q-Network で調整することで 交通流量の増加方式を検討する.完成したシステムは日本の千 葉県,幕張メッセで開催された CEATEC2017 の KDDI ブー スにてビデオ中継で展示した.

2. 実験システム

ラジオコントロールカーを使用した交通流のシミュレーショ ンシステムを作成した.現在16台までのラジオコントロール カーが走行可能である.ラジオコントロールカーは実際の自動 車の1/12スケールであり図1に示している.16台の走行の ために6m四方のコースを作成した,図2に示している.全 体のシステム構成を説明する.



図 1: 1/12 スケール ラジオコントロールカー

2.1 ハードウェアの構成

ラジオコントロールカーにはそれぞれ RaspberryPi3 が取り 付けられており,それを使用して制御する.ラジオコントロー ルカーの天井部には赤外線 (IR)LED が取り付けられており, 位置認識と ID 認識に使用されている.上空に設置した赤外線 カメラでラジオコントロールカーを含めたコース全体を撮影す る.カメラ画像をイーサネットケーブル経由で PC に取り込 み,画像中の IRLED による白色領域の重心位置を割り出す. 検知した IRLED の座標に対してカメラ歪の補正,高さ補正を 行い実座標の IRLED の位置に変換する. IRLED の位置関係 からラジオコントロールカーの位置と方向,そして車 ID を求 め,透視変換により実座標に変換する.

制御用 PC では車の位置情報から行動を決定する.計算される行動はセクション (ソフトウェアの構成) で解説する.計算された行動を WiFi 経由でラジオコントロールカーに搭載されている raspberrypi3 に送信する. RaspberryPi3 はラジオコントロールカーのサーボモーターと ESC を介したモーターにつながっており, PWM を使用してステアリング制御とアクセル制御を行う. 走行時に制御用 PC で得られたデータをサー



図 2: 使用したコース



図 3: 装置の構成図

バー用 PC に送信し、MySQL 上に走行ログを保存する. 走行 データを使用してサーバー用 PC で NN の更新を行う. 装置の構成図を図 3 に示す.

2.2 ソフトウェアの構成

ラジオコントロールカーの走路をパスで表し bez 曲線で定 義している.ラジオコントロールカーは自身が走行しているパ スを参照でき,現在のパスが接続している次のパスも参照でき るため経路選択を行いながら走行することができる.ラジオコ ントロールカーの位置と方向を得た後に,現在とその次のパス のうち最もラジオコントロールカーとの距離が近い点を走路 中の現在位置として得る.ラジオコントロールカーの進行方向 は現在位置から少し先の走路上のガイド点に向かって PID 制 御を行うことで決定される.ラジオコントロールカー同士の協 調行動は走行ラインに沿いながら速度を変化させることで実 現している.速度は包括関係のモジュールで決められている. 各モジュールを通して目標速度が求められた後,PID 制御で 速度を調整する.以上の構成図を図4に示した.

2.3 Deep Q-Network による動作

以上のラジオコントロールカーの走行システムを使用して, Deep Q-Network を用いた行動制御を行う. Deep Q-Network を使用せずとも走行が可能なラジオコントロールカーの機能 を、部分的に Deep Q-Network で置き換えることで実現する. 各機能をモジュール化して置き換えることによって,部分的な 判断を Deep Q-Network で行うことができ,ほかの機能に干 渉しなくなる. 今回はラウンドアバウト走行中の車の速度決定



図 4: モーター値決定までの計算過程

モジュールのうちの一つとして Deep Q-Network を用いるこ とで Deep Q-Network による推論は制御用 PC で行っている. 学習はサーバー用 PC で行う. 制御用 PC で集められた走行 データはサーバー用 PC 中の MySQL に保存され,サーバー 用 PC がそのデータを使用して学習を行う.サーバー用 PC で 作成された NN のモデルを制御用 PC に送信し,更新するこ とによって学習が完了となる.

2.4 Deep Q-Network

2.4.1 学習アルゴリズム

制御用 PC で得た走行データを、車ごとにサーバに送信す る.サーバでは MySQL でデータを保存できるようになって おり、システムの1ステップごとに「時刻、実験番号、車番号、 データ番号、ログデータ」を保存する.DQNの学習はサーバ 中のデータを使って行い、NNのモデルをアップデートするこ とで実現する.今回の研究ではラウンドアバウト内の車に対 して、合流車がいる場合に速度調整をするという形で導入し た.入力、出力に使用しているログデータ、報酬計算は表1に 表した.NNのモデルは 10-10-10-3 である.活性化関数は

表 1: DQN が扱うデータ			
種類	内容		
	制御車の合流地点までの距離		
	制御車の速度		
	合流車の合流地点までの距離		
	合流車の速度		
	先行車と制御車の距離		
入力	制御車の速度		
	ラウンドアバウト内の車台数		
	合流車の待ち台数		
	コース全体の車台数		
	車全体の平均速度		
	高速 0.65m/s		
出力	低速 0.50m/s		
	停止 0.00m/s		
報酬	全体の速度平均による加点		
	停止が続いた時の減点		
	緊急停止時の減点		

ReLU, 最適化関数は RMSProp.



図 5: 実験結果 総走行距離

3. Deep Q-Network による走行

3.1 比較実験

公道の走行を再現した走行と,自動運転車のみの想定の走行 を比較した.使用したコースは図2のようにラウンドアバウト と,その周辺の道路で構成されている.自動運転車のみの走行 では道路ごとの優先順位を考慮せず,安全である限り前進を続 ける走行である.この時,自動運転車のみの走行で,ラウンド アバウト内の速度調整をDQNを使用して行った.使用した学 習データは227,358 個で,ランダムサンプリングで100 デー タを抜き出し,10万回の学習を行った.比較の走行は12 台同 時走行で3回ずつ行った.1回の走行で5分間走行を行い,後 半4分間のデータのみを使用している.結果は表5に示した. 2つの結果を比較すると自動運転車のみの走行のほうが交通流 量が多くなっていた.

4. おわりに

本研究では、自動運転車のためのシミュレータを、ラジオコ ントロールカーを使用して作成した. DQN を使用して自動運 転車の速度制御を行うことで交通流量の増加が見込めるか検証 を行った.特にラウンドアバウトの合流時の速度調整を行うこ とで、交通流の増加が見込める結果となった.今回使用したシ ミュレータは一車線の走行が前提となり、走行ルートにある程 度の制限がある.そのため今後の課題は、車線変更が行えるよ うに拡張し走行経路を変更できるように拡張することであると 考えている.

参考文献

- Satoshi Suzuki. and Keiichi Abe., Topological Structural Analysis of Digitized Binary Images by Border Following, CVGIP 30 1, pp 32-46 (1985)
- [2] Zhengyou Zhang., A flexible new technique for camera calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330-1334, 2000.
- [3] Volodymyr Mnih. Koray Kavukcoglu. David Silver. et al. : Human-level control through deep reinforcement learning, Nature, Vol. 518, issue 7540, pp. 529-533, (2015)
- [4] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama, Jpn. J. Ind. Appl. Math. 11, 203 (1994).

[5] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama, Phys. Rev. E 51, 1035 (1995).