

遺伝的プログラミングと強化学習による自動運転モデルの比較

山村 光平[†] 中川 将輝[†] 渡辺 秀行^{††} 片桐 滋[†] 大崎 美穂[†]
[†] 同志社大学 ^{††} ATR

1. はじめに

近年高齢者ドライバーによる交通事故が増加している。この問題を解決するにあたり自動運転技術が注目されている。本研究では車線復帰課題を取り上げ、遺伝的プログラミング(GP: Genetic Programming)と強化学習の一つである決定論的方策勾配(DPG: Deterministic Policy Gradient)を用いたActor-Critic法(以下AC法)とによる運転制御モデルの獲得及び異なるモデルサイズ毎の汎化性能について比較を行った。

2. 遺伝的プログラミング

GPは生物の進化の仕組みを基に作られた進化論的手法の1つである遺伝的アルゴリズム(GA: Genetic Algorithms)の遺伝子型(運転モデル)を木構造として扱えるように拡張したものであり、環境における適合度が最適となるように学習(進化)させる。また、本研究におけるGPのモデルサイズは木構造のノード数で定義する。

3. 強化学習

強化学習ではエージェント(運転モデル)が環境との相互作用を通して方策関数と価値関数の最適化を目指す。方策関数はエージェントがどのような行動を実行するかを決定し、価値関数は状態や行動の良さを定量的に評価する。なお、採用した我々のAC法では方策関数と価値関数を3層パーセプトロンによって表現する。また、本研究における強化学習のモデルサイズは、3層パーセプトロンの隠れ層のノード数とする。

4. 評価実験

本実験の環境は運転シミュレータTORCSにて行った。課題は、時速50kmで直線道路を中央線に沿って走行し、1秒間左もしくは右に3°ステアリングを傾けることによって車線逸脱してしまった車両を中央線に復帰させる自動運転モデルを獲得することであった。評価は、表1と2に示した5つのサイズのモデルにおいて汎化性能を比較して行った。汎化性能は表3の180通りの運転条件で検証した。

表1. GPのモデルサイズ

Model ID	1	2	3	4	5
#Node	5	15	31	39	51

表2. AC法のモデルサイズ

Model ID	1	2	3	4	5
#Node(Actor)	6	30	60	90	120
#Node(Critic)	7	35	70	105	140

表3. 評価のテストケース(太字は学習時の条件)

Angle[°]	±1, ±2, ±3 , ±4, ±5, ±6
Time[sec]	0.1, 0.2, ..., 1.0 , ..., 1.5

図1と2にそれぞれの最大モデルサイズによる動作の軌跡を示す。実験から、最大サイズ以外の小さなモデルでは、GPとAC法のいずれもが、全ての運転条件において中央線に復帰させ得ることがわかった。しかし、GPの最大サイズのモデルは、[Angle=±5, Time=1.3~1.5]と[Angle=±6, Time=1.1~1.5]との学習時よりも厳しい運転条件では復帰することができなくなった。その一方でAC法は、運転軌跡に振動は含むものの(例えば、図2橙曲線)、モデルサイズが大きとも全ての運転条件において中央線へ復帰させ得ることが確認できた。

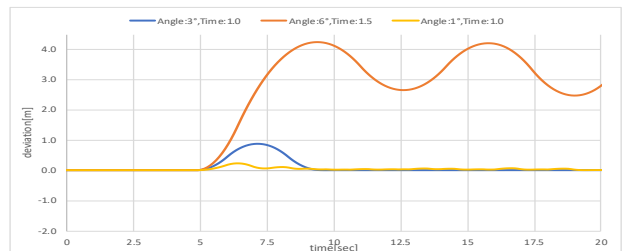


図1. GP (Model ID = 5)による自動運転軌跡。縦軸(deviation)は自動車と中央線の距離(m)、横軸は走行時間(sec)。

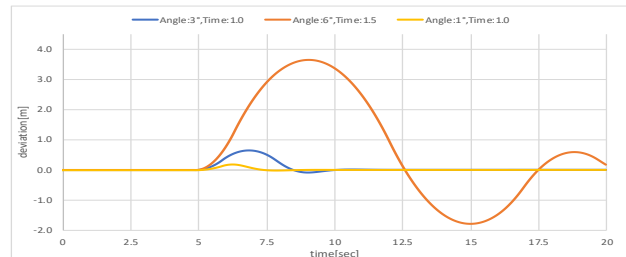


図2. 強化学習 (Model ID = 5)による自動運転軌跡。

5. おわりに

自動運転モデル獲得の課題を用い、GPとAC法との学習の汎化能力の比較を行った。実験の結果、モデルサイズが大きい時、GPよりもAC法の方がより高い汎化能力を持つことが示唆された。

謝辞 本研究の一部は、科研費(18H03266)の支援を受けた。

参考文献

[1] T. P. Lillicrap, et al. Proc. ICLR, May, 2016.
 [2] B. Wymann, et al. Torcs: The open racing car simulator, 2015.
 [3] 中川, 他. 信学会総合大会, ISS-P-023, Mar, 2018.
 [4] 八上. 同志社大学修士論文, Jan, 2018.