



**Title:** Hierarchical Reward Model of Deep Reinforcement Learning for Enhancing Cooperative Behavior in Automated Driving (自動運転における協調行動を強化するための深層強化学習を用いた階層型報酬モデルの提案)

**Authors:** Kenji Matsuda, Tenta Suzuki, Tomohiro Harada, Johei Matsuoka, Mao Tobisawa, Jyunya Hoshino, Yuuki Itoh, Kaito Kumagae, Toshinori Kagawa, and Kiyohiko Hattori  
(松田 賢治 (東京工科大学大学院生), 鈴木 天太 (東京工科大学大学院生), 原田 智広 (埼玉大学 准教授), 松岡 丈平 (東京工科大学 講師), 飛澤 真大 (東京工科大学大学院生), 星野 順哉 (東京工科大学大学院生), 伊藤 優希 (東京工科大学大学院生), 熊谷 海斗 (東京工科大学大学院生), 加川 敏規 (電力中央研究所 主任研究員) 服部 聖彦 (東京電機大学 教授))

**Journal:** Journal of Advanced Computational Intelligence and Intelligent Informatics

**掲載年月:** 2024 年 3 月

**研究概要:** 車線を想定しない自動運転専用道路環境において、強化学習を用いたゼロベースでの運転制御の最適化に着目する。強化学習を用いた行動獲得に焦点を当てた既存研究では、未学習環境での効率的な運転制御を獲得した一方で、スループットや衝突率に課題があった。本研究では一方通行・車線なしの専用道路において、効率的かつ安全性の高い運転制御の実現を目的とする。具体的には強化学習の報酬に(1)個別報酬と(2)共通報酬の2つを用いる階層型報酬モデルを提案した。報酬モデルの違いによる車両制御の比較を図1に表す。シミュレーションを用いて評価した結果、スループットの増加と、衝突率の低減を達成した。加えて、追い越し行動の創発による効率的な制御が実現した



Combination of Individual and Common Rewards

図1 報酬モデルがもたらす

**研究背景:** 近年、AI技術やセンサの性能向上に伴い、自動運転の研究が盛んになっている。特に自動運転車の走行制御にディープラーニングを用いた機械学習手法が用いられることが多い。それらの自動運転に関わる研究は、既存の一般的な道路インフラの上で人間が運転する車両との共存を前提としている。一方で、自動運転車の普及が100%に近づくと、人間が運転する車両が存在しない状況を積極的に活用、最適化された制御が可能になる。そこで本研究では、安全性と道路利用の効率化を両立させる自動運転の制御方法を強化学習を用いて車自身に発見させることを考える。

**研究成果:** 道路全体における車両のスループット増加と、衝突率の低減を達成した。加えて、追い越し行動の創発による効率的な制御が実現した。

**社会への影響:** 自動運転の普及に伴ってヒューマンエラーによる事故の減少といった安全性の向上が考えられる。また、車両同士のコミュニケーションによって、交通流が最適化され渋滞の緩和、24時間稼働が可能になることによる物流業界の効率化といった影響があることが考えられる。

#### 専門用語:

**強化学習:** エージェントが環境と相互作用しながら、試行錯誤を通じて報酬を最大化する行動方針を学習する機械学習の一分野である。エージェントは、状態、行動、報酬を繰り返し観察し、最適な行動を選択する能力を向上させる。

**AI:** 人間の知能を模倣するように設計されたシステムやアルゴリズムを指し、自然言語処理、画像認識、意思決定などのタスクを自動化可能である。

**ディープラーニング:** 多層のニューラルネットワークを使用してデータから特徴を自動的に抽出し、高度なパターン認識を可能にする機械学習の一手法である。画像認識や音声認識など、多くの分野で高い性能を発揮する。

**報酬**：強化学習においてエージェントが特定の行動を取った結果として得られるフィードバックの値で、行動の良し悪しを評価する。エージェントは、報酬を最大化するように行動を調整し、最適な方針を学ぶ。