

# 制動力保持機構を持つブレーキシステムにおける 学習手法を用いた保持制御

佐藤 昭彦<sup>†</sup> 新井 義和<sup>†</sup> 猪股 俊光<sup>†</sup> 今井 信太郎<sup>†</sup> 曾我 正和<sup>††</sup> 千葉 和幸<sup>‡</sup>

岩手県立大学ソフトウェア情報学部<sup>†</sup> 岩手県立大学地域連携研究センター<sup>††</sup> 有限会社 中央車体<sup>‡</sup>

## 1. はじめに

日本の自動車保有台数は増加しており，都市部や主要道路などの渋滞も深刻化している．渋滞に遭遇したドライバーは停止や発進の度に頻繁にペダル操作を行わなくてはならず，疲労が蓄積するにつれ，運転ミスを起こしやすくなる．オートマチック車はブレーキペダルから足を離すとクリープ現象によって前進／後退するため，疲労によって意図せずブレーキペダルを開放した場合，追突などの危険性が指摘されている．

著者らはクリープ現象を必要としない停止時に自動的に制動力を保持するブレーキシステムを開発し，実車実験を行ってきた[1]．しかし，多様な道路状況の変化やドライバーの個性などに対して特定の評価基準だけでは制動力の保持／不保持の判断を行うことが困難であった．

本研究では，実際の走行状態に応じて適切な判断基準を獲得するための学習手法を提案する．

## 2. ブレーキシステム

### 2.1. 制動力保持機構

本システムは，ブレーキペダルを踏むことで生じた力を各タイヤのブレーキパッドへと伝える油圧回路中に電磁弁を設け，その弁を制御することによって，ブレーキペダルを開放した後も制動力が保持可能な構造となっている．

### 2.2. ブレーキシステムの構成

本システムの構成を図 1 に示す．車両の状態を検出するために，1)車速度センサ，2)第 1 ブレーキセンサ，3)第 2 ブレーキセンサ，4)アクセルセンサの各センサを設置する．また，アクチュエータとして油圧回路中に 5)電磁弁を設置し，状況に応じてこれをコントローラが開閉する．1)は走行速度を把握，2)，3)はブレーキペダルの踏み込みの強さを検出し，通常の停止と

緊急停止を区別するために用いる．緊急停止時には危険回避の妨げにならないよう制動力は保持しない．4)はアクセルペダルの踏み込みの有無を検出する．停止時に制動力を保持していた場合，この検出によって電磁弁を開放し，制動力を解除する．

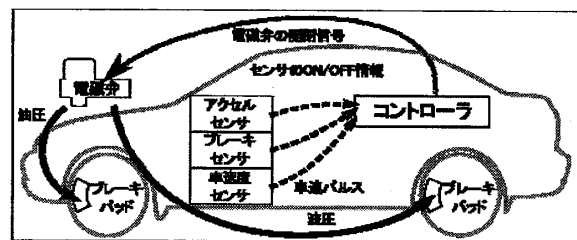


図 1 ブレーキシステムの構成

### 2.3. 従来の保持制御方式の問題点

一般に渋滞が軽度である場合，運転者はクリープ現象を積極的に活用して前方の車両に追従するため，制動力を保持することは望ましくない．このように，走行状況は時間とともに変化するので，システムはクリープ力を活用すべきか否かを判断する必要がある．そこで，停止時間に着目し，実時間でそれらを判断する“停車時間履歴方式”と“固定停車時間方式”の二つの手法を提案した[1]．しかし，停車時間履歴方式では，限られた環境における実験によって得られた閾値や係数に依存しているため，変化に富んだ走行状況に対応できない．一方，固定停車時間方式は，停止時間が一定時間を経過すると強制的に制動力を保持する方式であり，制動力が保持されるまでの時間遅れが発生する．

### 2.4. 学習手法による保持制御方式

走行状況は時間の経過，走行する場所，またはドライバーによってもその特性が変化する．したがって，制動力の保持／不保持の判断基準はその変化に応じて常に調整されるべきである．そこで，学習手法として Q-Learning[2]を導入した保持制御方式を提案する．

図 2 に示すように，学習エージェントに与える状態のパラメータとして前回の停止から今回の停止に至るまでの最高速度とその間の平均速

Retaining Control using Learning Method in Braking System with Braking Force Retaining Mechanism

<sup>†</sup> Akihiko SATO, <sup>†</sup> Yoshikazu ARAI

<sup>†</sup> Toshimitsu INOMATA, <sup>†</sup> Shintaro IMAI

<sup>††</sup> Masakazu SOGA, <sup>‡</sup> Kazuyuki CHIBA

Software and Information Science,

Iwate Prefectural University (<sup>†</sup>)

Regional Cooperative Research Center, Iwate Prefectural

University (<sup>††</sup>), Chuo Shatai, Inc. (<sup>‡</sup>)

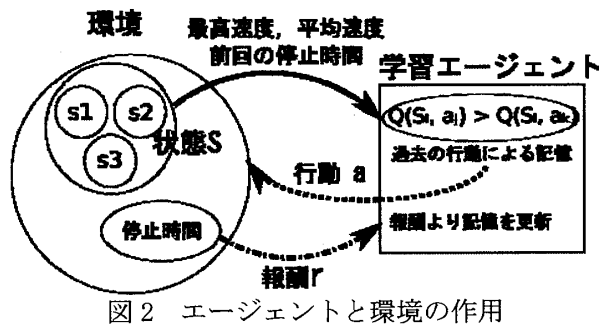
度, 前回停止した際の停止時間の 3 つを考える. また, 学習の収束を促すため各速度は 5[km/h] ごとに離散化し, 停止時間は 10[sec] を閾値として, それ以上およびそれ以下の 2 段階に離散化する.

エージェントは与えられた状態から, 過去の記憶を元に制動力の保持/不保持の 2 つの行動の中から行動価値  $Q(s_{t+1}, a_{t+1})$  の高い行動を選択する.

行動選択の結果得られる報酬  $r_{t+1}$  はシステムの意図に合致する行動を選択した際に正, 逆の場合は負となるよう, 次のように定義する.

$$r_{t+1} = \begin{cases} -1 & (\text{「保持する」を選択し, } T_s < 10 \text{ [sec]}) \\ 1 & (\text{「保持する」を選択し, } T_s \geq 10 \text{ [sec]}) \\ 1 & (\text{「保持しない」を選択し, } T_s < 10 \text{ [sec]}) \\ -1 & (\text{「保持しない」を選択し, } T_s \geq 10 \text{ [sec]}) \end{cases}$$

ただし,  $T_s$  は直前の停止時間である.



### 3. 提案手法の実験方法

提案手法ではオンラインで学習を行うことが可能であるが, 提案手法の評価のために必要な車速信号を, 車両に計測器を装着し 0.5 [sec] ごとに取得した.

学習は得られた車速情報に基づいてエピソードを生成し, オフラインで行う. すなわち, 車速情報から各走行区間の最高速度, 平均速度, 直前の停止時間を算出し, これらを状態とする. この状態に対して行動を選択する. その次の走行区間が始まった時点で停止時間を確定し, それに基づいて報酬を算出する. さらに, 次の走行区間が終了し, 次の状態が確定した際に行動価値関数の更新を行う. 以上の手順を反復することによって, 学習が成立する.

本学習はオフラインで行われているが, 車速情報は実際の車両を用いて取得されており, 行動選択後に遷移する状態がそこで選択された行動に依存しないことから, オンラインで学習し

たのと同じ効果を得ることが可能である.

### 4. 実験結果

図 3 に制動力保持実験の結果を示す. Q-Learning では, 学習の過程において, 現在の状態に対して最も行動価値関数の値の大きい行動を選択することが基本となる. しかし, 局所解に陥らないために一定の割合  $\epsilon$  でランダムに行動を選択する  $\epsilon$ -グリーディ方策を採用. 同図は, 提案手法において  $\epsilon$  を最も報酬が得られる 0.01 に設定した際の, 学習を行ったエピソード数に対する行動最適化度の変化を示している. ここで, 行動最適化度とは, 全てのエピソードに対して正の報酬が得られた割合である. 結果として, およそ 3500 エピソードを終了した時点でほぼ学習が収束しており, およそ 70[%] の割合で正しい行動が選択されていることが確認できる.

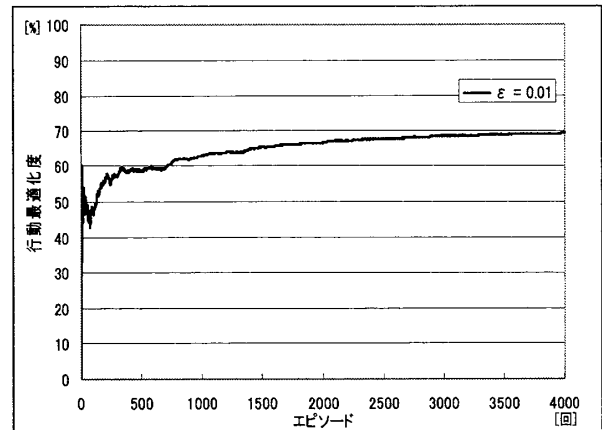


図 3 行動最適化度の変化

### 5. まとめ

本研究では, 制動力を保持するか否かの判断に強化学習を導入した. また, 実車による走行を通して得た多様な走行状況に対して一定の割合で正しい行動選択が可能であることを確認した.

今後の課題としては, 提案されてきた各手法を組み合わせ, より状況に応じた行動選択ができるような手法を確立していくことである.

### 参考文献

- [1] 石下, 新井, 猪股, 曾我, 千葉: “制動力保持機構を持つブレーキ制御システムのための自動起動手実車試験による評価”, 情報処理学会第 70 回全国大会, 4ZA-3, 2008
- [2] Richard S. Sutton, Andrew G. Barto: “強化学習”, 森北出版, 2000