カオス進化を用いた強化学習の高速化と 運転支援表示システムへの適用

正員 山口 亨 (東京都立科学技術大学)
 非会員 高 秀 誠 (NTT 東 日 本)
 非会員 中 村 吉 寿 (宇 都 宮 大 学)
 正員 木 幡 直 樹 (高崎健康福祉大学)

Speed up of Reinforcement Learning using Chaotic Evolutionary Computation and its Application to Driver's Support Display System

Toru Yamaguchi, Member (Tokyo Metropolitan Institute of Technology), Makoto Takahide,
 Non-member (NTT EAST), Yoshimoto Nakamura, Non-member (Utsunomiya University), Naoki
 Kohata, Member (Takasaki University of Health and Welfare)

Recently, the research and development of the highly intelligent agent systems, such as ITS(Intelligent Transport Systems), have been performed with the progress of the information infrastructure. In such an intelligent agent system, a man-machine system is very important.

In this paper, we pay attention to the display function that becomes an interface between the human and the machine. This display function needs to present the suitable display for the user.

Therefore, we propose a method which is able to acquire the display knowledge with a few learning times by the interaction with the user. We realize this method based on Q-Learning, and we also use the "Chaotic Evolution method" to improve Q-Learning that requires many learning times.

In this paper, the proposed "Q-Learning with Chaotic Evolution" is applied to the driver's support display agent in the commuter car system, which is one of man-machine systems of ITS. The effectiveness is verified by simulation and experiment.

キーワード：ITS、人間-機械系システム、強化学習、Q-Learning、カオス、オンライン・高速な知識獲得と学習

1. はじめに

高度情報インフラストラクチャの整備が進む中、ITS(Intelligent Transport Systems)に代表されるような、高度な知的エージェントシステムの研究・開発が進められている。従来の知的エージェントシステムはハード的な側面が中心であった。しかし、現在、人間を中心とする柔軟な処理を行うシステムが注目されてきている。このソフト的・人間中心な知的システム、即ち人間-機械系システムでは、人間と機械とのコミュニケーションの機会が多く、特に人間(利用者)にシステムが様々な情報を表示し運転支援などのサポートを行う機能は非常に重要である⁽¹⁾⁽²⁾⁽³⁾⁽⁴⁾。

通常の情報表示システムでは、利用者に一方的に情報を表示しているものが大半である。更に、何らかの危険な状況が起こった時に、その直前に表示を行うため、利用者がいきなりの表示に驚きパニックになってしまったり、それ

のみに注意を向けてしまい他の状況に注意を払えなくなってしまうという欠点がある。

本稿ではこの欠点を改善するシステムとして対人間表示エージェントを提案する。この提案する対人間表示エージェントは利用者との対話を行う事で、直前に強く警告して人間を慌てさせるような堅い表示ではなく、余裕を持ってアニメーション等で弱めに知らせることで、人間が落ち着いて、その表示の意図を判断し対応することが可能な柔らかなパターン表示を獲得・表示する機能を持つ。このような、音楽で言えばバックグラウンドミュージックに相当する人間に負担をかけない柔らかな表示を、本稿では「ソフトタッチ(な表示)」と呼ぶ。パターン表示を取り上げたのは、単に静止画像ではなくそれを連続表示することで、より理解しやすい表示が行えるためである。このパターン表示の知識獲得においては、オンラインで、かつ少ない学習回数で表示知識を学習し獲得する必要がある。本稿では、ソフト

タッチで理解が容易な表示を、少ない学習回数で高速に獲得することを目的とする。

運転支援表示システムにおいて、新たな画像表示パターンを獲得するのに適用可能な代表的手法としては、遺伝的アルゴリズム (GA: Genetic Algorithm) のような進化型計算手法や強化学習の一手法である Q-Learning⁽⁵⁾ などが考えられる。しかし、従来の GA 的な手法は、基本的には確率的操作を実施するためランダム性が高く、望みのパターンへ行き着くまでに、理論上は全てのパターンをたどらねばならない可能性がある。また、Q-Learning も基本的に状態遷移先が確率的に選択されていき、それに対し報酬を与えていく手法であり、ランダム性が高いため、多数回の試行を経る必要があり、学習に手間がかかる。

一方、我々は、ニューラルネットワーク上のカオス的想起⁽⁶⁾を基盤として、カオス現象に基づく知識の発想・進化の研究を継続的に実施してきており⁽⁴⁾⁽⁷⁾⁽⁸⁾、その中で、パターンの想起範囲を容易に決定することが可能であり、新たな知識をオンラインで獲得できるカオス進化が提案されている⁽⁸⁾。我々のカオス進化の手法では、カオスの持つ構造的な複雑さのために、記憶してある既存のパターンにひきずられる形で新たなパターンが想起されるため、想起されるパターンの範囲を限定することが可能である。そのため、想起範囲の限定を適切に行うことにより、従来の GA 的な手法と比較して、望みのパターンを効率良く獲得可能な利点を持っている。

したがって、このカオス進化の手法を Q-Learning における学習過程に導入することができれば、Q-Learning における状態遷移の範囲をあらかじめ絞り込んでおくことが可能となり、学習過程の効率化が見込める。このため、運転支援表示システムにおける連続表示画像の獲得に、カオス進化 (+ Q-Learning) を用いることは、他の手法との比較においても有効性が期待される。

従来、Q-Learning とニューラルネットワークは全く別々の手法として研究され、両者の結び付きに目を向けるものはいなかったが、本研究では、Q-Learning における一連の状態遷移をニューラルネットワーク上に表現するというアイデアに基づき、Q-Learning の一連の状態遷移に対してカオスの想起の適用を可能にしたことで、カオス進化と Q-Learning の融合という技術課題を解決した。

本稿では、カオス進化により Q-Learning の状態遷移候補を絞り込み、その上で Q-Learning を実施する一連の処理を行うことを、カオス進化と Q-Learning の融合と位置付け、この提案する手法を「カオス進化型 Q-Learning」と呼ぶ。また、この一連の処理により Q-Learning の処理を効率化することを、強化学習の高速化と呼ぶ。

提案の手法を実際のコンピューターカーにおける運転支援表示知識獲得に応用し、実験及びシミュレーションを行うことで本手法の有効性を検証する。

2. ITS と対人間表示エージェント

図1 ITSにおけるコンピューターカーシステム

Fig. 1. Commuter Car System in ITS

図2 対人間表示エージェント

Fig. 2. Display Agent to human

近年、ITS では道路の交通渋滞・交通公害の軽減だけではなく、人間の安全性・快適性も重視され、高齢者にとっても快適なシステムが求められている。その1つが、図1に示すコンピューターカーシステムである。

コンピューターカーシステムにおける重要なシステムの一つとして、運転補助等を行う表示システムが挙げられる。この表示システムは一つの知的エージェントであり、図1下段に示すように、モバイルパッドとして、コンピューターカーシステム以外の、例えば、Web-TV や e-コマース等に活用できる情報端末として扱うことも将来的に可能であると考えている。本稿ではこの情報表示システムであるモバイルパッドを対人間表示エージェントとして扱う。

この対人間表示エージェントは図2中央に示す様に、利用者にとってより理解しやすく、かつソフトタッチでパニックにならないような情報表示を行うために、利用者との対話によりオンラインで、かつ少ない回数で学習する事で表示知識を獲得する機能が必要である。本稿での表示知識獲得とは、表示を行う際のルールを獲得するという意味である。

本稿では、このような表示知識獲得を、カオス進化と Q-Learning を融合した、カオス進化型 Q-Learning の手法により実現する。次章において、このカオス進化型 Q-Learning について説明する。

図3 カオス進化型 Q-Learning

Fig. 3. Unit Chaotic retrieval to Q-Learning

3. カオス進化型 Q-Learning

対人間表示エージェントにおける表示知識の学習法として、本稿では、カオス進化型 Q-Learning を提案する。本稿での進化とは、良好な知識を獲得することで、より良い知識を構築していくという意味での進化である。この手法は、Q-Learning の欠点である過多となってしまう学習回数を軽減できる手法である。以下に、提案するカオス進化型 Q-Learning とそれを構成する際に用いたカオス進化、Q-Learning の手法についてそれぞれ説明する。

(3.1) カオス進化型 Q-Learning カオス進化型 Q-Learning とは、カオス進化の手法を Q-Learning の遷移過程に応用する事で、次の状態の遷移 (行動順) を考えられる有効な範囲内に決定し、学習回数が多い Q-Learning の欠点を改善することができる手法である。

図3上部に示す様に、通常の Q-Learning では、全ての状態遷移先へのパターンを試す必要があり、爆発的にパターン数が増加する事があった。これにより、結果の善し悪しを機械的に判断し自動的に報酬を与えるような学習シミュレーションでは、確かに良好な結果の得られる学習法であると言えるが、結果の判断や報酬を与えることを人間が行うとなると、学習回数が多くなればなるほど利用者のストレスが溜り真の意味で良い学習法とは言えない。

しかし、カオス進化と融合させた場合、その非線形性を調節することで任意の範囲内での状態遷移が可能となる。つまり、非線形性を弱くする事で初期の有効な状態遷移から大きく逸脱しない状態遷移先をある程度決定することができ、それに沿った学習を行う事ができるという事である。そのため図3下部に示す様に、試行数を削減し、学習の効率を上げる事が可能である。その際、非線形性を調節することができるパラメータ d_2 は 3.5 とした。このパラメータに関しては(3.2)で詳しく説明する。

以下に提案する「カオス進化型 Q-Learning」を構築する際に用いたカオス進化と Q-Learning について詳しく説明する。

(3.2) カオス進化 本稿でのカオス進化の手法とは、ニューラルネットワーク上のカオスの想起により、既存の

図4 非線形抵抗の特性変動とエネルギー曲面上でのカオスの遍歴

Fig. 4. Characteristic variation of non-linear resistance and chaotic wandering on energy curved surface

知識からカオスの非線形性に基づき新しく良好な知識を想起し、知能を進化していくものである⁽⁸⁾。このカオスの想起の手法として、カオスの最急降下法⁽⁶⁾(Chaotic Steepest Descent method, 以下 CSD 法と略す)を用いている。この方法は、ニューラルネットワークのエネルギー曲面での運動を記述する力学方程式の散逸項に図4のような周期的に変動する非線形抵抗を導入することにより、カオス的にエネルギー極小解の間を遍歴させる方法である。

以下、実際にカオス進化型 Q-Learning において、各状態に相当する画像パターンを記憶させ、カオスの想起を実施するのに用いたホップフィールド型ニューラルネットワークを例にして簡単に CSD 法を紹介する。CSD 法は(1)式によって表される。

$$\text{CSD 法} \quad m\ddot{u}_i + f(\dot{u}_i, \omega t) = \varepsilon \sum_j \omega_{ij} a_j \quad \dots (1)$$

ここで、 u_i 、 a_i はそれぞれ i 番目のユニットの内部状態および出力、 ω_{ij} はユニット j から i への結合の重みであり、 $f(\cdot)$ は非線形抵抗を表す関数、 m, ε は正の係数、 ω は非線形抵抗の特性を変動させるための角速度である。更に、 i と j は、それぞれ $[1, l]$ である。 l はホップフィールド型連想メモリの大きさであり、図3においては、各層のノード数である。ただし、系のエネルギー E と、 $a_i \in [0, 1]$ 、 ω_{ij} は次式で表されるものとする。

$$E = \sum_{i,j} \omega_{ij} a_i a_j \quad \dots (2)$$

$$a_i = 1 / (1 + \exp(-u_i)) \quad \dots (3)$$

$$\omega_{ij} = \sum_k (2a_i^k - 1)(2a_j^k - 1) \quad \dots (4)$$

ここで、 a_i^k は k 番目の記憶させるパターンのユニット i の値である。系の持つ時定数に対し、比較的遅い周期で正抵抗及び負抵抗の値を変動させると、正抵抗が大きくなる位相で状態は安定化し極小に漸近し、逆に負抵抗が大きくなる位相で不安定化し極小から飛び出そうとする(図4)。この変動を適当に行うことによって、カオス的にエネルギー曲面の極小から極小へ遍歴させることができる。

$$f(\dot{u}_i, \omega t) = [d_0 \sin(\omega t) + d_1] \dot{u}_i + d_2 \dot{u}_i^2 \text{sgn}(\dot{u}_i) \quad (5)$$

更に、CSD 法に外部入力を与え続けると、外部入力に対して距離的に近い範囲にある記憶パターンを動的に想起することができる。かつ、その範囲をネットワークにおける力学系の非線形性を変化させる式 (5) の 1 つのパラメータ d_2 で制御することができる。 d_2 が大きい場合非線形性は弱く、小さい場合非線形性は強い。

また、ある程度常識を逸脱しない知識を、あらかじめホップフィールドに少数 (k 個) 記憶させておき、初期記憶を構築する。

本稿では、全てのカオス的想起において、それぞれ 5 つの初期記憶を用意した。この初期ネットワークにより発想させる知識の決定を行うことができる。

この手法を用いる事により、新たな表示知識を発想し獲得する事が出来る。又、CSD 法の大きな特徴である非線形性抵抗の特性を変動させる事が容易に行える事により、任意の範囲内での知識の発想が可能となり、Q-Learning の行動遷移の想起範囲を任意に指定する事が出来る。その際、 ω_{ij} を、図 3 を高度化した各ノード間を繋ぐホップフィールドの重みとし、 $l \times l$ の大きさに設定する。そして、カオス的想起を行い、ある隣り合ったレイヤのそれぞれ一つのノードが活性化した時、Hebb 学習則⁽⁹⁾を用いてそのノード間が結合する。

〈3・3〉 Q-Learning 利用者の直感に合った表示知識を獲得するには、一連のパターン表示全体を学習する必要がある。本稿では、結果に応じた報酬を与える事によって適切な問題の解決方法を獲得する事の出来る強化学習を取り上げる。強化学習法は、試行錯誤的な経験から適切な行動を適応的に学習する手法である。本稿では、この強化学習の代表的なアルゴリズムとして Q-Learning⁽⁵⁾を用いている。

Q-Learning のアルゴリズムを以下に示す。

時刻 t において状態 x_t であるときにある行動選択法にしたがって行動 a_t をとった結果、状態 x_{t+1} に遷移し報酬 r_t が得られたとき、Q 値は式 (6) で示す更新幅で更新される。

$$\Delta Q(x_t, a_t) = \alpha(r_t + \gamma \max_b Q(x_{t+1}, b) - Q(x_t, a_t)) \quad (6)$$

α は学習率、 γ は割引率を示す。これは、次の状態での最大の Q 値を割引率 γ だけ割り引いた値と、そこで直接得られた報酬 r_t との和に、元の Q 値を近づける様に更新する事を意味する。Q-Learning では、多数の試行後に Q 値が収束した時、各状態において最大の Q 値を持つルールが、各ルールの中では最適となる事が証明されている。

図 3 上部において、それぞれのノードが状態を表し、一番左の層から順にノード間を遷移していく。各状態間のネットワークに Q 値を記憶する事でパターン学習を行う。又報酬は状態遷移の逆伝搬でそれぞれのネットワークに与えられる。

今回は、式 (7) に示す Boltzman 分布に基づく選択が行われている。

図 5 選択したシチュエーション

Fig. 5. Situation

$$P(a|x) = \frac{\exp(Q(x,a)/T)}{\sum_{b \in \text{actions}} \exp(Q(x,b)/T)} \dots\dots\dots (7)$$

ここで、actions は遷移先の総数であり、 T は温度定数と呼ばれる。 T を徐々に 0 に近づけることで、学習された結果に遷移先を依存させていくことが可能である。本稿においては、この温度定数 T は 0.98 とした。

本稿では、状態 x をパターン表示の画像とし、また、状態遷移 a を連続で行う事により、パターン表示を実現している。提案するカオス進化的 Q-Learning では、この Q-Learning における一連の状態遷移をニューラルネットワーク上に表現することで、カオス的想起の適用を可能にし、カオス進化と Q-Learning の融合を図っている。

4. シミュレーションによる対人間表示エージェントの表示知識獲得

提案するカオス進化的 Q-Learning を用いて、被験者 4 人に対し、表示知識獲得のシミュレーションを行った。

〈4・1〉 シチュエーションとシミュレーション条件 今回のシミュレーションでは、対人間表示エージェントをコンピューターカーシステムに搭載し、実際に路上を走行している状態を想定して行った。対人間表示エージェントはコンピューターカーを運転している被験者に、危険になりうる可能性のある歩行者の存在を知らせる情報表示を行う。この歩行者の情報提示を行う表示知識を、提案するカオス進化的 Q-Learning の手法を用いて学習・獲得していく。

〈4・1・1〉 シチュエーション 今回のシミュレーションにおいて、図 5 にあるように、3 つのシチュエーションを選択した。1 つは、路上駐車をしている車が運転者の死角となり、車の奥から歩行者が路上に飛び出して来るというシチュエーションである。2 つ目は、前方の大型車により視界が無く、その奥の横断歩道を歩行者が歩いているというシチュエーション、3 つ目は、向かって右側の舗道から歩行者が飛び出して来るシチュエーションである。その 3 つのシチュエーション毎に、表示知識を学習していく。この 3 つのシチュエーションによる歩行者の情報表示を応用

図6 シミュレーションのシステム構成
Fig. 6. Simulation composition

図7 シミュレーション
Fig. 7. Simulation

することで他の場面に対しても対処することができると考える。

今回は、この3つのシチュエーションのうち、1つ目のシチュエーションのみを実際の路上でビデオ撮影したもので、また、3つ全てを3次元コンピュータグラフィックス(3DCG)で再現したものでシミュレーションを行った。

〈4・2〉シミュレーションの流れとシステム構成 シミュレーションのシステム構成を図6に示す。コンピューターには対人間表示エージェントとしてのノート型パソコン(以下PCと略す)が搭載してある。PCでは、事前に提案するカオス進化的型Q-Learningを用いた表示知識の学習・獲得を行う。そして被験者が理解しやすく、かつソフトタッチなコンピュータグラフィック(以下CGと略す)によるパターン表示を行う。このパターン表示は、実画像もしくは3DCGの背景に図7左に示すように、スーパーインポーズで表示される。

表示するCGは、コンピューターカーを運転中に歩行者の存在を運転手に教えるものとし、人型のCGを5回連続で表示することでパターン表示を行う。その際、カオス進化的型Q-Learningでは、第〈4・3〉章で獲得した人型のCGを図7右にあるQ-Learningのノード(状態)とする。このノードを図のように10個並べたものを一つの層とし、それを5層構築し、左の層から順に状態遷移を行うことでパターン表示を行う。即ち、一度の状態遷移において、その遷移先は10個となる。

更に、あらかじめ常識の範囲を大きく逸脱しないであろう状態遷移順の知識を5つカオス進化に記憶させることで、状態遷移の想起範囲を限定する。この記憶させる状態遷移

図8 カオス進化により新たに獲得された画像
Fig.8. New acquisition Images using Chaotic Evolution

の知識は、必ずしもソフトタッチな表示ではないがある程度は、誰にでも理解してもらえると考えられる表示を行えるものとする。また、その際の初期知識は第〈4・3〉章で被験者が最初に選択した画像を5つ連続で表示する状態遷移順とする。

そして、一連の状態遷移を行い、その結果、被験者が理解しやすく、かつソフトタッチな表示が行われた、と思えば報酬を与える。この学習を数回繰り返し、同じパターン表示が5回連続で行われ、かつ、そのパターン表示がある程度見やすく、かつソフトタッチであると被験者が感じる事ができた時に学習を終了する。

なお、今回のシミュレーションでは、学習回数を比較するために、通常のQ-Learningによる学習も行い、被験者毎の学習回数を記録した。

次に、Q-Learningの各状態(画像)を獲得する方法について説明する。

〈4・3〉カオス進化による表示知識獲得 〈3・2〉で述べたように、カオス進化の手法は新たな知識の獲得を可能としている。本稿では、提案するカオス進化的型Q-Learningによる表示知識の学習を行う前に、カオス進化を用いて様々な画像、すなわち、Q-Learningの各状態を獲得する。

具体的には、ある既存の初期画像を元とし、その初期画像の、'色'、'各部位の大きさ'を様々に変化させることで、新たな画像を獲得していく。'色'はRGB値を、'大きさ'は初期画像の大きさに対し0.5倍から1.5倍の倍率で変化させる。その際、カオス進化の非線形性はある程度強くすることで、初期画像から大きく異なった画像を獲得することが可能である。この時の非線形性を調節するパラメータ d_2 は0.3とした。この値は、ハミング距離やユークリッド距離の変動量を指標として決定した。

図8は、実際にカオス進化により既存の初期画像から新たに獲得された画像群である。初期画像は、図8の一番左上のものである。

これらの画像の中から、利用者(被験者)が最も理解しやすい・ソフトタッチな表示であると思ったものを一つ選び、その画像を元に、今度は非線形性の弱いカオス進化を行い、

図9 実画像の背景で行った結果
Fig.9. Real background simulation result

図10 シチュエーション1における学習結果
Fig.10. 3DCG simulation result of case in situation No.1 using Q-Learning with Chaotic Evolution

Q-Learning の各状態となる画像を獲得する。この時の非線形性を調節するパラメータ d_2 は 4.0 とした。この値も、ハミング距離やユークリッド距離の変動量を指標として決定した。

次に、カオス進化によって獲得した画像を用いて、カオス進化型 Q-Learning による一連のパターン表示を学習するシミュレーションについて説明する。

〈4・4〉 シミュレーション結果 以下にそれぞれのシチュエーションにおいてパターン表示を学習した結果の画像と、それぞれの学習回数の平均、被験者毎の表示の評価結果を挙げる。

〈4・4・1〉 背景が実画像の場合 図9は背景に録画された実画像を用いてシミュレーションを行った結果の一例である。先の図5における、1番目のシチュエーションのみを取り上げて学習のシミュレーションを行った。

この図9にあるような、一連のパターン表示が獲得できた。

次に、背景に3DCGを用いた場合のシミュレーション結果を挙げる。

〈4・4・2〉 背景がコンピュータグラフィックの場合 以下に、背景に3DCGを用いた場合の学習結果の一例を挙げる。それぞれ図10がシチュエーションの1番目、図11が2番目、図12が3番目の学習した結果の一連のパターン表示である。

それぞれのシチュエーションにおいて、これらのパター

図11 シチュエーション2における学習結果
Fig.11. 3DCG simulation result of case in situation No.2 using Q-Learning with Chaotic Evolution

図12 シチュエーション3における学習結果
Fig.12. 3DCG simulation result of case in situation No.3 using Q-Learning with Chaotic Evolution

表1 シミュレーションの平均学習回数

Table 1. Average of learn times.

背景が実画像		
シチュエーション	通常の Q-Learning	カオス進化型 Q-Learning
路上駐車	87.3	25.8
背景が3DCG		
路上駐車	87.0	21.0
横断歩道が見えない	126.7	20.8
右側から歩行者	55.0	18.8
全平均		
———	89.0	21.6

ン表示を獲得することができた。次に、提案するカオス進化型 Q-Learning と通常の Q-Learning の学習回数を比較する。

〈4・4・3〉 各学習回数の比較 表1はそれぞれのシチュエーションにおける、通常の Q-Learning を用いた場合と、提案するカオス進化型 Q-Learning を用いた場合の被験者4人の平均学習回数を比較したものである。

この表1から、通常の Q-Learning に対してカオス進化型 Q-Learning を用いた学習の方が、それぞれのシチュエーション毎では最大6分の1に減少している。また、全体平均でも約4分の1に減少している。よって、カオス進化型

図 13 各表示の見やすさの評価値

Fig. 13. Evaluation value of easiness to see

Q-Learning による学習の高速性を示すことができた。次に、獲得されたパターン表示の見やすさ・ソフトタッチな表示であるかどうかの評価を行う。

〈4・4・4〉被験者の評価 4人の被験者に、システム開発者が任意で作った一連のパターン表示、単に注意を促すだけの人型のCGを連続表示したもの、通常のQ-Learningのみを用いて学習したパターン表示、そして、今回提案するカオス進化型Q-Learningにより学習され獲得したパターン表示の見やすさの評価値を各シチュエーション毎に5段階でつけてもらい、それを平均シグラフとしたものが図13である。その際、評価値の基準は、1:見にくい、理解しにくい、ソフトタッチな表示ではない。2:少々見にくい、少々理解しにくい。3:可も不可もない表示、基準的。4:理解しやすいがソフトタッチとは言いがたい。5:見やすい、理解しやすい、ソフトタッチな表示である。と設定した。

図13にあるように、被験者にとって最も見やすい・理解しやすい・ソフトタッチな表示と感じられる一連のパターン表示を獲得したのは、通常のQ-Learningを用いた学習であった。しかし、提案するカオス進化型Q-Learningを用いた学習結果であっても、必ずしも100%よい結果が得られているわけではないが、被験者にとってある程度は見やすい・ソフトタッチな表示知識を獲得している。これにより、先に述べた学習の高速性を考えれば提案のカオス進化型Q-Learningは、十分に有効な学習手段であると言える。

また、各被験者が選択したパターンにおける類似性を考察すると、色は黄色～橙色の暖色系の色が選択されやすく、画像の形状は手足のバランスがとれたもので、5段階表示される画像の大きさの変化としては、徐々に大きくなるものが選択される、という類似性が見られた。

5. 実機コンピューターカー実験による対人間表示エージェントの検証

提案するカオス進化型Q-Learningにより、注意情報表示知識を学習することのできる対人間表示エージェントを、実際のコンピューターカーシステムに搭載した実験について以下に説明する。

〈5・1〉システム構成 図14に実機実験のシステム構成を示す。実験には、2台のコンピューターカーと、外界認識用の地上局パーソナルコンピュータ2台、コンピューター

図 14 実機実験のシステム構成

Fig. 14. System of real machine experiment

図 15 実機実験風景とその表示結果

Fig. 15. Real machine experiment result

カーに搭載する外界認識と対人間表示エージェントを兼ねるノート型パーソナルコンピュータが2台設置されている。それぞれのパーソナルコンピュータはネットワーク化されている。地上局のパソコンや路上駐車のパソコンは認識された歩行者情報を、後方の走行車に送信することができる。

今回の実験では、1台のコンピューターカーは路上駐車をしており、その影から、飛び出して来る歩行者を、路上駐車の後ろを走行しているもう1台のコンピューターカーに情報表示を行うというシチュエーションを想定している。その際の、歩行者の認識やその情報の伝達手法等に関しては参考文献⁽¹⁰⁾⁽¹¹⁾を参照されたい。

〈5・2〉実験結果 図15が実際に実験を行っている様子であり、その際に運転手に表示された画像である。

図15にあるように、実際に屋外でコンピューターカーを用いた実験を行った。使用したコンピューターカーは光岡社製の電気自動車EV-70である。図15のついたてから歩行者が飛び出して来るようになっている。

図15のように、ついたてから飛び出して来る歩行者に合わせて、ソフトな情報表示が対人間表示エージェントにより運転手に行われた。この歩行者の情報表示により、運転手は前もって物影から飛び出して来る歩行者を知ることができ、更に、そのみに注意を向けるのではなく、他の情報にも気を配ることができるので、安全に運転することができる。

6. 終 り に

本稿では高度な知的システムにおける、よりユーザに分かり易く、ソフトタッチで危険の少ない表示を行う対人間表示エージェントの表示知識の獲得をオンラインで高速に行う事を目的とした。この目的を実現するために、「カオス進化型 Q-Learning」を提案した。この手法は従来の Q-Learning のパターン学習とカオス進化のオンラインでの高速な知能獲得機能を融合した物である。この手法を ITS におけるコンピューターカーシステムの運転支援対人間表示エージェントの知能進化に応用し、シミュレーション及び、実機実験を行った。シミュレーションの結果から、通常の Q-Learning に比べ提案の手法では学習回数が約 4 分の 1 に減少し、また、実機実験からソフトタッチな表示を行っていることを確認したことで、本手法の有効性を示すことができた。

(平成 12 年 11 月 28 日受付, 同 13 年 4 月 19 日再受付)

文 献

- (1) 山口, 中村, 高秀: インテリジェントトランスポートシステムにおける対人間表示の進化と学習, 計測自動制御学会 Dynamical System Theory シンポジウム 講演論文集, pp213-216(1999)
- (2) 高秀, 木幡, 山口: インテリジェントスペースとその「人間-機械」系システムへの応用, 電気学会 産業システム情報化研究会 研究会資料, IIS-99-4, pp.19-24 (1999)
- (3) 高秀, 木幡, 山口: 対人間知能の進化・学習, 電気学会 電子・情報・システム部門大会 講演論文集, pp319-322 (1999)
- (4) M. Takahide, K. Hirayama, T. Yamaguchi, H. Hashimoto: Intelligent Space and its Application to Man-machine System, Proc. of IEEE International Conference on Systems, Man and Cybernetics (SMC'99), Tokyo, Japan (1999)
- (5) C. J. C. H. Watkins, "Q-Leaning," *Machine Learning*, vol. 8, pp. 279-292 (1992)
- (6) 谷淳: カオスの最急降下法を適用したニューラルネットにおける学習および記憶想起の動特性について, 電子情報通信学会論文誌 A, Vol.J74-A, No.8, pp.1208-1215 (1991)
- (7) 佐藤, 牛田, 山口, 井村, 高木: ファジィ連想記憶システムにおけるカオスの記憶探索, 日本ファジィ学会誌, Vol.6, No.4, pp.765-774 (1994)
- (8) N. Kohata, T. Yamaguchi, T. Baba, H. Hashimoto: Chaotic Evolutionary Parallel Computation on Intelligent Agents, *Journal of Robotics and Mechatronics*, Vol.10, No.5, pp.424-430(1998)
- (9) 萩原, 山口 共著: ニューラルネットワークとファジィ信号処理, コロナ社, pp.51 (1998)
- (10) 平山, 新田, 山口: 情報の再構成と感覚拡張, 平成 11 年 電気学会 電子・情報・システム部門大会 講演論文集, pp.315-318 (1999)
- (11) T. Yamaguchi, K. Hirayama, M. Takahide, H. Hashimoto: Intelligent Space and its Application to Augmented Sensing, IEEE/IEEJ/JSAI Conference on Intelligent Transportation Systems(ITSC'99), Tokyo, Japan, pp.901-906(1999)

山 口 亨 (正員) 1981 年千葉大学大学院電気工学専攻修

士課程修了。工学博士。1981 年 (株) 東芝入社。1989 年 4 月から技術研究組合国際ファジィ工学研究所へ出向。1992 年 4 年より (株) 東芝システム・ソフトウェア技術研究所。同年、電気学会論文発表賞受賞。ファジィとニューラルネットに関する研究に従事。1993 年 4 月より宇都宮大学情報工学科助教授。2000 年 4 月より東京都立科

学技術大学教授。計測自動制御学会、電気学会、日本ファジィ学会、日本神経回路網学会各会員、日本ロボット学会正会員。

高 秀 誠 (非会員) 1998 年宇都宮大学情報工学科卒業。

2000 年同大学大学院博士前期課程情報工学専攻修了。2000 年 NTT 東日本入社。人工知能学会会員。

中 村 吉 寿 (非会員) 2000 年宇都宮大学情報工学科卒業。

木 幡 直 樹 (正員) 1997 年宇都宮大学大学院博士前期課程情

報工学専攻修了。2000 年同大学大学院博士後期課程生産・情報工学専攻修了。2001 年 4 月より高崎健康福祉大学講師。電気学会、日本ファジィ学会、電子情報通信学会各会員。