

移動ロボットにおける センサ形態と制御系の同時設計法

正員 杉浦 孔明*,**, 正員 川上 浩司***
正員 片井 修***

Simultaneous Design Method of the Sensory Morphology and Controller
of Mobile Robots

Komei Sugiura*,**, Member, Hiroshi Kawakami***, Member, Osamu Katai***, Member

This paper proposes a method that automatically designs the sensory morphology of a mobile robot. The proposed method employs two types of adaptations –ontogenetic and phylogenetic– to optimize the sensory morphology of the robot. In ontogenetic adaptation, reinforcement learning searches for the optimal policy which is highly dependent on the sensory morphology. In phylogenetic adaptation, a genetic algorithm is used to select morphologies with which the robot can learn tasks faster. Our proposed method is applied to the design of the sensory morphology of a line-following robot. We carried out simulation experiments to compare the design solution with a hand-coded robot. The results of the experiments revealed that our robot outperforms a hand-coded robot in terms of the line-following accuracy and learning speed, although our robot has fewer sensors than the hand-coded one. We also manufactured a physical robot using the design solution. The experimental results revealed that this physical robot uses its morphology effectively and outperforms the hand-coded robot.

キーワード：センサ進化, 生態学的バランス, 身体性, 学習と進化

Keywords: sensor evolution, ecological balance, embodiment, learning and evolution

1. はじめに

Uexküll が、生物の知覚世界と行動の密接な結びつきを指摘したのは、今から 100 年ほど前のことであった [14]。Uexküll は、「全ての種は客観的な世界像を共有しているわけではなく、自身のニッチに応じた主観的世界（環境世界）に基づいて行動している」ことを明らかにした。この種固

有の環境世界を生み出す感覚器は、生物の身体性や神経系、生態学的ニッチと密接に関わって進化してきた [5, 18]。

一方、ロボットのセンサは、アクチュエータや制御系とは別の観点から設計されることが多い。例えば、センサ形態の設計においては、設計者が物理量を計算しやすいような設計などが重視されてきた反面、振る舞いや制御系の学習則との関連は考慮されることが少なかった。ここで形態は、身体形状、アクチュエータ/センサの種類や配置などを含む [10]。またセンサ形態とは、形態のうちセンサに関するものとする。

しかしながら、環境の変化に適応的なロボットを構築するためには、形態・制御系・環境のバランス（生態学的バランス）を考慮することが重要である [11]。近年、形態と制御系をうまく組み合わせて、簡単な制御系で安定な制御や知的な振る舞いを実現する研究や、形態と制御系を同時設計することによって制御性能の向上を目指す研究が注目されてきている [19]。これらは、形態と制御系の時間的变化を積極的に利用して、より適応的なロボットの設計を目指すアプローチといえる。

ただし、これまでの研究では、形態のうち主に身体形状に

* (独) 情報通信研究機構 知識創成コミュニケーション研究センター
〒 619-0288 京都府相楽郡精華町光台 2-2-2

Knowledge Creating Communication Research Center, National Institute of Information and Communications Technology
2-2-2 Hikaridai, Seika, Soraku, Kyoto 619-0288

* (株) 国際電気通信基礎技術研究所 音声言語コミュニケーション研究所
〒 619-0288 京都府相楽郡精華町光台 2-2-2

ATR Spoken Language Communication Research Laboratories
2-2-2 Hikaridai, Seika, Soraku, Kyoto 619-0288

*** 京都大学 情報学研究科

〒 606-8501 京都市左京区吉田本町
Graduate School of Informatics, Kyoto University
Yoshidahonmachi, Sakyo, Kyoto 606-8501

限定して形態と制御系の相互作用を扱うことが多い [8, 12]。そのため、センサ形態と時間発展する制御系（学習器）の相互作用は、十分に検討されてきたとはいえない。

しかし、時間発展する制御系が利用する情報は、センサ形態に大きく依存する。例えば、障害物回避の学習では、センサの配置や数の変更が入力情報（ロボットにとっての環境世界）を変化させ、学習時間に影響を与える。また工学的には、実機での実現を考えたとき、形態のうちセンサ形態に着目することで、設計の自由度やコストの低さといった利点を得られる。例えば、自動車の衝突回避センサの配置を設計する場合、形状に関する自由度は低いが、センサ配置については自由度が高い。このようなセンサ配置設計の別の応用例としては、自律移動車椅子の衝突回避センサの配置 [20] などが挙げられる。

本研究の目的は、実機実現可能なセンサ形態と制御系の同時設計法を実現することである。この提案手法は、強化学習と遺伝的アルゴリズムを組み合わせることにより、学習に適したセンサ形態を持つロボットを設計する。本手法の特徴は、第一にセンサ形態と制御系を同時に設計できること、第二に学習能力に適した大きさの状態空間の構築が可能であること、第三に実機に適用可能であること、の3点である。本手法をライントレーサに適用し、学習に適したセンサ形態を設計させる。

本稿では、まず関連研究について述べ、次に本研究で提案する手法について説明する。その後、提案手法による設計実験について述べた後、得られた設計解を実機に実装して行なった比較実験について述べる。

2. 関連研究

Fig. 1 に環境-学習エージェント間相互作用の概略図を示す。学習エージェントとは、学習と意思決定を行なう主体であり、生物やロボット、ソフトウェアエージェントを含む。エージェントの身体は、身体形状、センサ/アクチュエータとその配置などからなる形態と材質などから構成され、環境との相互作用のインターフェースとなる。また、制御系は認識と方策から構成される。認識はセンサ入力ベクトル s から観測ベクトル o への写像 R であり、方策 P は o から行動ベクトル a への写像である。 E は a から s への写像であり、環境の力学的特性の変化やエージェントの身体の変化に影響を受ける。

従来のエージェント学習の研究では、身体の設計と制御系の設計は独立に行なわれるのが一般的であった [19]。方策の自動設計、例えばロボットによる行動学習などにおいては、あらゆる身体に適用可能なユニバーサルな学習手法（強化学習、ニューラルネットなど）について議論することが多い。しかし、身体と制御系が共に設計可能であるような条件下においては、身体と制御系を同時に設計することで利点が得られる可能性がある。例えば、ロボットの形態の特徴を利用することで認識器を簡素化できる可能性などが指摘されている [19]。これは、身体の変化が、 a から s

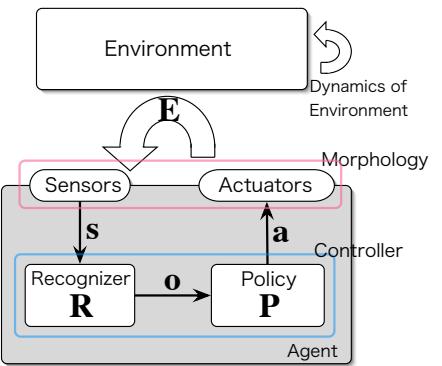


Fig. 1. Interaction between learning-agent and environment

への写像 E を、 R や P の学習に有効であるように改善せ得るためである。

形態と制御系を同時設計に関する先駆的研究としては、Sims のものが挙げられる [12]。Sims は、遺伝的アルゴリズムを用いて、エージェントの形態と制御系をソフトウェア上で自動設計する手法を提案した。Lipson らは、この手法とラピッドプロトタイピングを組み合わせることによって、ソフトウェア上で得られた形態を、ハードウェアとして実現することに成功している [8]。また、このような同時設計において、センサ数や種類、配置など、より多くの形態の属性を扱った研究も進められている [3, 9, 15]。

しかしこれらの研究は、形態と制御器がセットで時間発展する設計手法であって、形態に対して制御器が時間発展することはない。すなわち、設計された形態が学習に適しているかは明らかではない。これに対し、最近になって、Bongard らにより形態に対して制御器が時間発展する手法も提案されている [2]。彼らは、形態とニューラルネットコントローラが遺伝子型から発生するモデルを用いて、シミュレーション上で鎖状の球からなるロボットの設計を行なった。

これらの研究の流れと並行して、センサ進化 (sensor evolution) のアプローチが提案されている [4]。センサ進化は、センサ形態に着目して環境との相互作用をボトムアップ的に構築することにより、設計者が想定していなかった相互作用の発見を目指すアプローチである。これは、強化学習の系譜に連なる状態汎化 [13] の研究が、Fig. 1 の前処理に着目していたのとは異なる。すなわち「センサデータを前提とした上で、どのような状態縮約を前処理で行なうか」という問題のみではなく、「そもそもどのようなセンサデータを獲得すればよいか」という問題を含めてセンサレベルから設計解を探る枠組みである。このような問題意識から、センサ形態が学習に与える影響に関する研究も行なわれている [1, 6, 7]。

本研究はこれらの研究と関連するが、以下の点に独自性を有する。すなわち、1) 学習能力に適した大きさの状態空間の構築ができること、2) 実機に適用可能であること、の

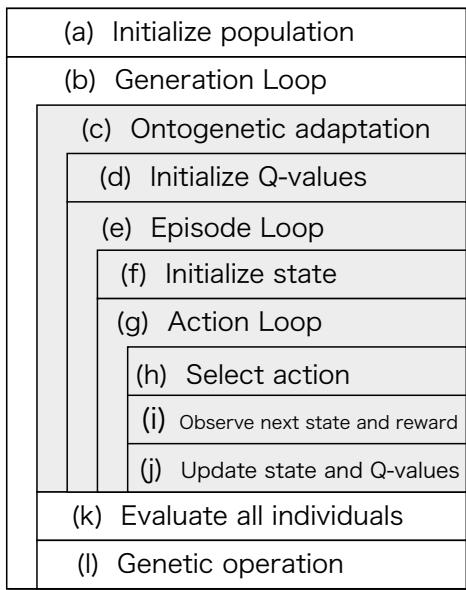


Fig. 2. Schematic diagram of proposed method

2点である。以下では提案手法の詳細について述べる。

3. センサ形態と制御系の同時設計法

3・1 提案手法の概略 Fig. 2 に、提案手法の概略を NS チャートで表現したものを示す。本手法では、2 種類の時間フェーズにおいて異なる機械学習手法を利用する。まず個体発生的適応フェーズでは、センサ形態は固定され、ロボットは強化学習を用いてタスクを学習する (Fig. 2 の網掛け部分)。次に、系統発生的適応フェーズにおいて、進化的手法を用いて形態を変更する (Fig. 2 の白色部分)。これらの詳細については、3・2 節で述べる。

3・2 設計手法

3・2・1 個体発生的適応 個体発生的時間スケールにおいて、ロボットは Q 学習 [13] を用いて行動を学習する。ただし、ロボットが観測できる情報はセンサ値のみとする。そのため、センサ形態の違いにより状態空間の次元や状態の意味が変化する。したがって、一定期間内にタスクを学習するためには、学習に有利なセンサ形態を有している必要がある。

ロボットには、 N_o 回の試行機会 (エピソード) を与える。各エピソードでの学習の流れを以下に示す。ただし、括弧付きアルファベットで示された記号は、Fig. 2 の対応する部分を表す。

- (1) 基本状態 (位置など) を初期化する (f)
- (2) 行動ステップ t が最大行動ステップ数 T に達するまで、以下の 3. から 5. を繰り返す (g)
- (3) 行動ステップ t における行動を選択する (h)
ロボットが t において、状態 s_t にあるとする。 s_t は、各センサの離散値入力のみに依存する。このとき、ロボットは行動価値関数を参照し、 ε -greedy 戦

略に基づいて行動 a_t を選択する。

(4) 次状態 s_{t+1} と報酬 r_{t+1} を観測する (i)

(5) Q 値および状態を更新する (j)

Q 学習の学習則として、以下に示す一般的な更新則を用いる。

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_{t+1} + \gamma \max_a Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)]$$

ここに、 α は学習率、 γ は割引率である。

$s_{t+1} \leftarrow s_t$ とする。

3・2・2 系統発生的適応 系統発生的適応では、遺伝的アルゴリズム (GA) を用いることによって、ロボットの形態を設計する。各 GA 個体には、センサ形態がバイナリコーディングされる。すなわち、センサ基板の特定の位置にセンサを配置するかしないかを 2 値で表現する。設計の流れは以下の通りである。ただし、Fig. 2 の対応する部分をアルファベットで示した。

(1) 初期世代をランダムに生成する (a)

(2) 世代数 n が最大世代数 N_p に達するまで、3. から 5. を繰り返す

(3) n 世代目の各個体が個体発生的適応 (c) を行なう

(4) 全個体の学習が終了した後、その結果に応じて評価値 ϕ を与える (k)

(5) 遺伝的操作を加えて $n + 1$ 世代を生成する (l)

親個体がトーナメント選択により選ばれた後、遺伝的操作 (1 点交叉と点突然変異) が加えられて子個体が生成される。このとき、エリート数 10% のエリート保存戦略も併用して用いる。なお、前に述べた通り、コントローラに関しては、子個体に親子体の学習結果を引き継がせない。

4. 提案手法の適用実験

以降では、提案手法の有効性を検証するために行なう 2 種類の実験について述べる。実験の概要は以下の通りである。

• 実験 A : センサ形態の設計実験

提案手法を用いてライントレースの学習に適したセンサ形態を自動設計させる。

• 実験 B : 人手によるセンサ形態との比較実験

センサ形態が学習に与える影響に関して、人手によるセンサ形態と実験 A で得られた設計解を比較する。

4・1 タスク環境 本研究では、提案手法の適用対象としてライントレーサを採用する。ライントレーサとは、床に引かれたラインに沿って移動しゴールを目指すロボットのことである。用いるコースを Fig. 3 に示す。コースは $2\text{ m} \times 5\text{ m}$ の長方形で、4 分の 1 ずつの地点にチェックポイント p_1, p_2, p_3, p_4 を設けてある。なお、 p_4 をスタートおよびゴールとする。

タスクとしてライントレースを用いる理由として、以下の 3 点が挙げられる。

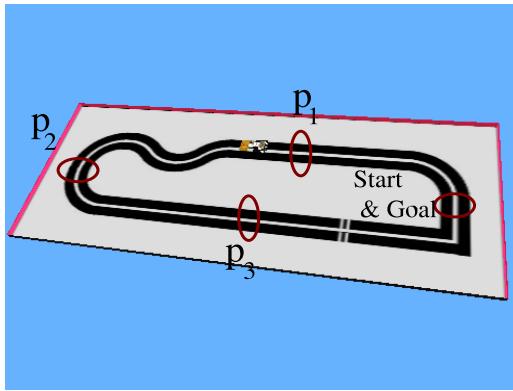


Fig. 3. Task Environment

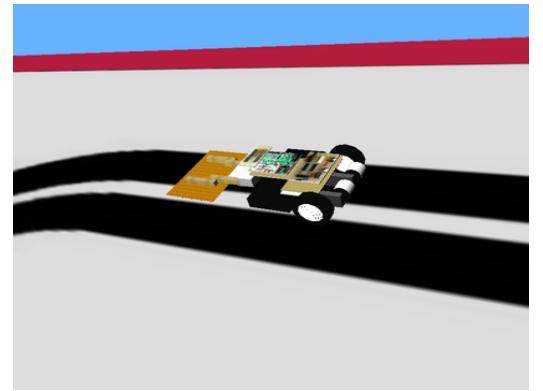


Fig. 4. Line-following robot (Simulation)

- 不規則なセンサ配置による性能の改善可能性

大会で用いられるロボットの多くでは、設計者が制御系(以下コントローラ)を設計しやすいように、センサ配置が単純かつ規則的である[17]。これに対し、ロボットが学習可能なコントローラを有する場合には、センサ配置が規則的である必要はない。さらに、規則的でないセンサ配置により性能を改善できる可能性がある。

- 人手による設計との比較が可能

ライントレースは多くの大会が開かれ、かつ多くの専用キットが販売されているため、人手によるセンサ形態のサンプルが多く存在している。そのため、提案手法によるセンサ形態の設計解との比較が可能である。

- 代表的な「身体が相互作用に意味を与える」[16]タスク

ライントレースは、部分観測マルコフ決定過程(POMDP)[13]としてモデル化できる。このタスクでは、コース上の位置により最適な行動が決定されるが、ロボットは床センサのみを用いて環境を観測するため、位置を直接観測することはできない。したがって、効果的なセンサ形態の使用により、(タスク遂行に必要なだけの)マルコフ性を伴う状況識別ができることが望ましい。ライントレースタスクを用いて、環境との相互作用のチャネルであるセンサ形態を自動設計することで、身体と状況識別との関係を検討することができる。

4・2 設計対象 設計対象として、Fig. 4 に示すようなロボットを選んだ。このロボットは、マイコンカーラリーにおいて標準的に用いられている日立インターメディックス製のマイコンカーを元にしたものである。ただしこのロボットには、シミュレーションで扱いやすいように、2輪型へと改造を加えた。ロボットの前部のセンサボードには、床センサを格子状に取り付けることができる。床センサは赤外発光LEDと受光素子から構成され、床面の色に応じた値をコントローラに渡す。センサ値はしきい値を設けて2値化する。また、コースとタイヤの摩擦に存在する不完全性をモデル化するために、モータ出力に0.5%の白色雑音

をスリップノイズとして加える。

本研究では、床センサの配置および個数に対して提案する設計システムを適用し、自律的にセンサ形態を獲得させる。簡単のため、床センサはロボットの底面に格子点状に配置されるようにした。センサは最大32個まで取り付けることができる。

4・3 パラメータ設定 実験で用いたパラメータの設定をTable 1に示す。これらのパラメータは実機での実験を念頭において設定した。特に実機では、バッテリの消耗や部品の摩耗の影響があるため、有限の時間しか試行できない。そのため、1エピソードでの最大行動ステップ数 T は、実機実験の1試行とほぼ同程度にした。1ステップは64 msecであるため、2000ステップは約2分である。このように有限の試行時間が与えられることにより、状態空間の大きさも制限される。よって、センサはタスクに必要なもののみに絞られるようになることが予測される。

ロボットがステップ t で選択する行動 a_t は、5種類の行動(直進、左折大、左折小、右折大、右折小)のいずれかとする。5種類の行動に対応する左右のモータ角速度[rad/s]は、 $\{(15, 15), (6, 13), (2, 8), (13, 6), (8, 2)\}$ とした。行動は、予備実験を行なったうえで典型的な結果が得られるパラメータを選択した。ステップ t における報酬 r_t は、 $r_t = -\beta x_t^2$ とする。ここで β は減衰率を表す定数であり、 x_t は、 t におけるトレースラインの中央線とセンサボードの中心との距離を表す。報酬の設計に際しては、ラインから離れるほど r_t は小さくなるように $\beta = 1.75$ とし、センサボードがライン外に出た場合は $r_t = -100$ とする。なお報酬は、仮想的な報酬観測センサをロボットに取り付けて得られた値から計算する。報酬観測センサは、マイコンカーのデフォルトのセンサ配置(Fig. 7左図)と同様になるよう横一列に配置されるため、厳密には、 x_t は報酬観測センサ列がなす直線上におけるボード中心とライン中央線との距離である。

個体の評価値 ϕ は、以下の ϕ_i の平均とする。

$$\phi_i = \frac{1}{N_c} \sum_{j=1}^{N_c} \left\{ 1 - \frac{t_{ij}}{T} \right\} \quad (1)$$

ここで、 t_{ij} は、 i 番目のエピソードにおいて、 j 番目とそ

Table 1. Parameter settings

	Parameters	Value
Ontogenetic adaptation	Number of action steps T	2000
	Number of episodes N_o	100
	Learning rate α	0.8
	Discount rate γ	0.999
	ε	0.01
Phylogenetic adaptation	Gene length	32
	Max. number of generation N_p	50
	Population size P	50
	Cross-over rate p_c	1.0
	Mutation rate p_m	0.03

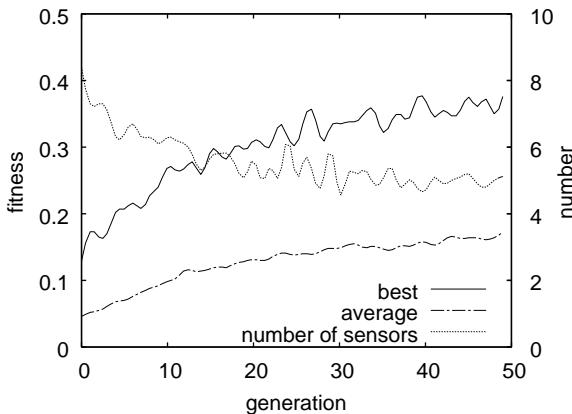


Fig. 5. Variation of fitness and the number of sensors

の直前のチェックポイント間に要したステップ数である。また、 N_c はチェックポイントの数 ($N_c = 4$ で固定)、 T は最大行動ステップ数を表す。ただし、コースアウトによりチェックポイント j に到達できなかった場合は、 $t_{ij} = T$ とする。 ϕ_i の大きさは、多くのチェックポイントを早く回ったことを意味する。

4・4 結 果

4・4・1 設計実験の結果 (A) まず、設計解を定量的に検討するために、評価値 ϕ の変化について調べた。Fig. 5 は、各世代における ϕ と、その世代の最良個体のセンサ個数の変化を示したものである。ただし図において、実線は世代における最大の ϕ を、破線は平均の ϕ を、点線は最良個体のセンサ個数をそれぞれ表している。また、横軸は世代数、左縦軸は評価値 ϕ 、右縦軸はセンサ数を表す。実験を 10 回行なって、その平均値をプロットした。

次に、定性的な結果として、提案手法によって設計されたセンサ形態の例を Fig. 6 に示す。Fig. 6 に示した例は、最終世代において ϕ がおよそ 0.35 の個体群から、センサ数が 4, 5, 6 個の個体を選んだものである。

4・4・2 人手によるセンサ形態との比較結果 (B)

実験 A において得られた設計解と、人手によるセンサ形態を用いて学習過程の比較を行なった結果について述べる。比較対象のセンサ形態を Fig. 7 に示す。Fig. 7 左図がマイコンカーのデフォルトのセンサ配置 (M1) であり、右図が提

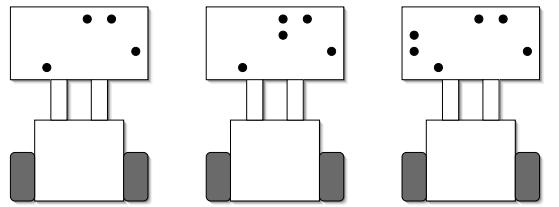


Fig. 6. Examples of automatically designed sensory morphology: four sensors (left), five sensors (middle), six sensors (right). Black dots represent sensors.

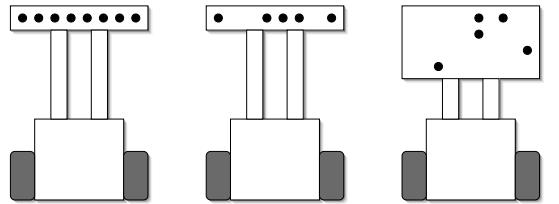
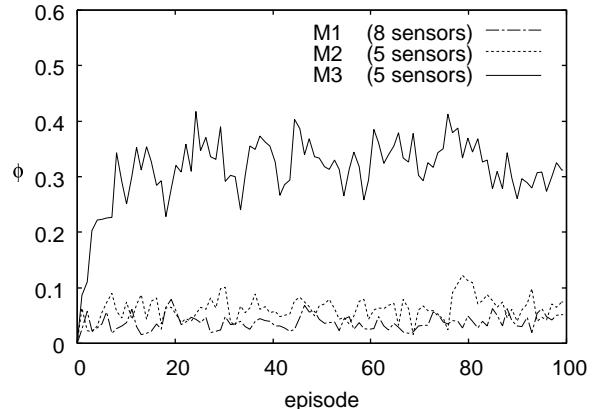


Fig. 7. Sensory morphology of robots: M1 (left), M2 (middle), M3 (right)

Fig. 8. Variation of fitness ϕ_i

案手法による設計解 (M3) である。この 2 種類に加え、設計解と状態空間の次元が等しい形態の M2 (Fig. 7 中図) を用意し、三者で比較を行なった。

Fig. 8 にエピソード (横軸) に対する評価関数 ϕ_i (縦軸) の変化を示す。図において破線、点線、実線は、それぞれ M1, M2, M3 の結果を表す。

Fig. 8 より、提案手法の設計解 M3 が、全てのエピソードにおいて高い値を示していることがわかる。このことを確認するために、エピソード数を増やして同様の実験を行なったところ、Fig. 8 と同じ傾向が見られた。

4・5 考 察 提案手法により得られたセンサ形態に関して、センサ配置とセンサ数の点から考察を行なう。

4・5・1 センサ配置 Fig. 6 に示すように、獲得されたセンサ形態は以下のようないくつかの特徴を持つ。

- 配置が左右非対称である

今回のタスクは、コースを半時計回りするようにしたため、S 字カーブを除けば左カーブが多い。そのため左カーブ得意とするセンサ形態の評価値が高くなっ

たと考えられる。

- 前後に分散して配置している

横一列に並んだ配置は、ラインの前後関係を読み取ることができない。そのため、ロボットがカーブの上にいるのか、あるいは直線の上にいるのか、をロボット自身が判断することが難しい。前後にセンサが分散することで、ラインの前後関係から直線やカーブといったラインの形状をロボットが読み取ることが出来るようになったと考えられる。

4・5・2 センサ数 Fig. 5より、最良個体においては、センサ数が減少するにしたがって評価値 ϕ が増加していることがわかる。この結果から、ライントレース環境においては、5個前後のセンサを有していることが学習に有利であるといえる。これは、学習器を用いることによって、マイコンカーラリーで標準的に用いられるセンサ個数(8個)を、減らすことができることを示唆している。

本実験では、センサ数が状態空間の次元を決定する。一般的に、次元が少ない方が学習の収束が早いが、学習収束後のパフォーマンスは次元が多い場合よりも低くなると考えられる。しかし 4・4・2 節の結果から、学習が収束した後であっても、M1 に比べてセンサ数が少ない M3 のパフォーマンスが高い、ということがわかる。これは、少ない状態空間を効果的に用いれば、収束速度だけでなく、学習結果も改善できることを示している。

5. 実機による検証実験

シミュレーションでは、物理環境を完全には再現できないため、シミュレーションで得られた設計解が実環境でも同様に働くとは限らない。文献 [8]においても、シミュレーションと実環境で同一の評価値が得られているわけではない。これは、実環境での摩擦などの影響によるものと考えられている。そのため、本手法で得た設計解を実機に実装して、シミュレーションと実機の相違について検討する。また、設計解の有効性を検証するため、人手により形態とコントローラを設計した場合との比較を行う。

5・1 設 定 本実験では、3種類のロボット M1C, M2, M3 (Fig. 7 参照) にライントレースタスクを行なわせる。各タスク試行は、10周するかコースアウトすると終了する。各試行ごとに、コースの状態を初期化して5回実験を行なう。

センサ形態とコントローラに関する比較を Table 2 にまとめた。M1C のセンサ形態は、マイコンカーのデフォルトのものを用いた。また、M1C のコントローラは、付属のサンプルプログラムを改良したものを用いた。すなわち、M1C のセンサ形態は Fig. 7 の M1 に等しいが、コントローラは人手により設計されたものである。この M1C を人手による設計の代表とする。

M2 のセンサ形態は、M1C と同様にデフォルトのものを用いた。M2 のコントローラは、4節の実験(B)での学習結果を利用する。すなわち M2 は、状態に対して最大の Q

Table 2. Comparison of manufactured robots

	Sensory morphology	Controller
M1C	Human-designed	Hand-coded
M2	Human-designed	Learned in simulation
M3	Designed by proposed method	Learned in simulation

値を持つ行動を常に選択する。そのため、本実験ではコントローラを学習させることはない。

M3 のセンサ形態は、本手法による設計解のうち、センサ数 5 個のものを実装した (Fig. 9 参照)。Fig. 10 に実験で用いたコースを示す。M3 のコントローラは、M2 と同様に実験(B)で得られた学習結果を用いたものである。

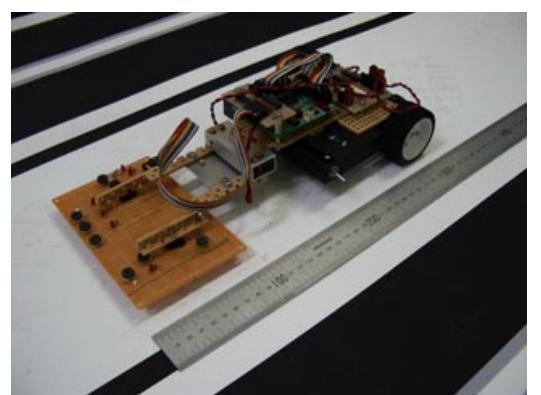


Fig. 9. Manufactured line-following robot M3

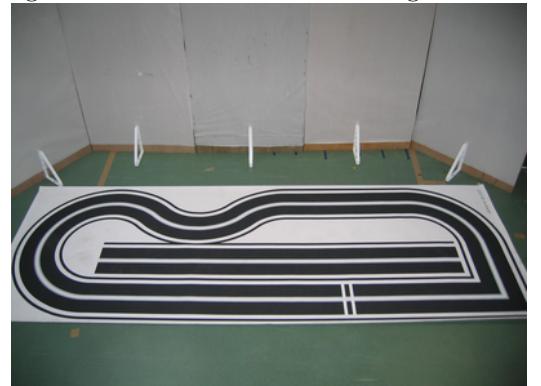


Fig. 10. Course used in the experiment

5・2 結 果 初めに、走行性能を検討するため、三者の平均ラップタイムを比較する。Fig. 11 は、5回の実験における三者の平均ラップタイムを標準偏差とともに示したものである。Fig. 11 より、平均ラップタイムを比較すると M3 が最も速いことがわかる。

次に、コースアウトするまでのラップ数(滞在ラップ数)を調べた。その結果、M3 の平均滞在ラップ数は 8.4 周であるのに対し、M1C と M2 の平均滞在ラップ数はそれぞれ 4.4 周と 3.0 周であった。すなわち M3 は、M1C や M2 に比べてコースアウトすることが少ない。

5・3 考 察 以下では、2つのポイントに注目して、本手法の設計解の有効性を検討する。

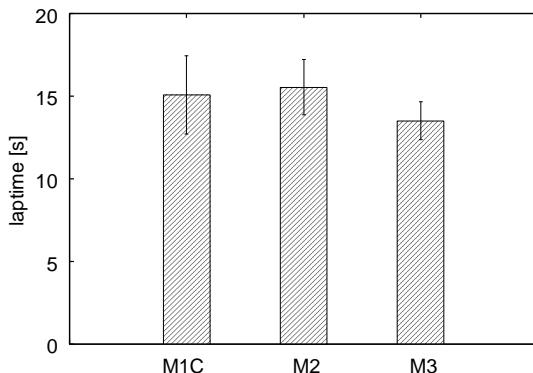


Fig. 11. Average Lap time

5・3・1 バッテリ消耗の影響 4節の設計実験では、バッテリの消耗の影響を考慮しなかった。しかし実際には、実機によるライントレースでは、バッテリーの消耗のため、ラップを重ねるごとにラップタイムが大きくなる傾向がある。M1C のラップタイムを例に挙げると、10周後にはおよそ 10% 大きくなる。このことは、1週目と 10週目では、たとえコントローラの出力値が同じであっても、モータのトルクが同じであるとは限らないことを示している。そのため、周を重ねるほど同じコントローラでの走破が困難になる。実際に、M1C や M2 の平均滞在ラップ数はそれぞれ 5周以下であった。

しかしながら、M3 の平均滞在ラップ数は、他の二者に比べて顕著に大きい。すなわち、M3 は M1C や M2 に比べてコースアウトの少ないロボットであるといえる。これは、4節の実験 (B) で、バッテリの消耗に対して頑健なコントローラが学習されていたことを示している。

以上の点は、提案手法によって、設計者が想定していなかった特徴に関して有効なコントローラが獲得された例として興味深い。

5・3・2 直角コーナーの走破 次に、ライントレースにおいて難易度が高いとされる直角コーナーの走破を例として、三者がどのようにセンサ配置を利用しているかを述べる。

M1C は直角コーナーの直前に存在するクロスラインを合図に用いる方法を採用している。すなわち、直角コーナーの直前のクロスラインを読み取ると、直角コーナー用のモードに移行し、直角コーナー専用の制御則を用いて直角コーナーをクリアする。このように、人手によるコントローラでは、「クロスラインの後には直角コーナーが存在する」というコースに関する設計者の知識を利用できるため、通常のカーブとは異なる制御則を使って直角コーナーを走破できる。

しかし、M2 や M3 はコースに関する設計者の知識を利用できない。そのため、直角コーナーを通常のカーブと同じ制御則を用いて走破しなければならない。横一列に並んだセンサを持つ M2 は、直角コーナーでコースアウトすることが多い。この理由は、主に M2 のセンサが前後に分散

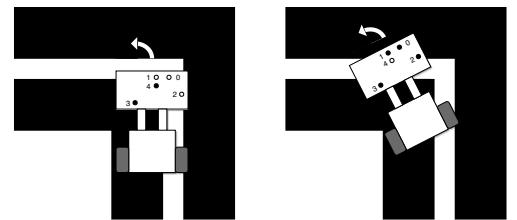


Fig. 12. Effective use of morphology at a perpendicular corner

していないためラインの前後関係を捉えにくることによる。これに対し M3 では、直角コーナーにさしかかると Fig. 12 左図に示すようにセンサが反応する。これは左カーブと同じセンサ状態であるので、M3 は少し左に曲がる。ところが、実際にはコースは直角であるため、Fig. 12 右図に示すようにコースから外れそうになる。ここで 4 番目のセンサがセンターラインを検出し、マイコンカーは大きく左折するため、直角コーナーを走破することができる。

6. おわりに

環境の変化に適応的なロボットを設計するためには、形態・制御系・環境の相互作用を積極的に利用しなければならない。特に、学習器を有するロボットにとって、センサ形態は環境との相互作用のチャネルであるため、センサ形態の変更が環境の解釈に及ぼす影響は大きい。このことから、本研究の提案手法は、センサ形態を変更することによって、より適した解釈系を獲得するアプローチであると考えられる。本論文では、このアプローチが学習の探索空間を改善し得る点について議論した。

提案手法は、強化学習と遺伝的アルゴリズムを組み合わせることにより、学習に適したセンサ形態を持つロボットを設計する。本手法をライントレーサに対して適用し、得られた形態と人手による形態の学習結果の比較を行なった。実験の結果、本手法で設計されたロボットは、センサ数が少ないとても関わらず、学習速度とパフォーマンスの両方において優れることを示した。さらに、本手法によって得られた設計解を実機に実装して、既存手法と比較した。その結果、本手法で設計されたロボットは、形態を有効に利用することによって、速くかつ正確にライントレースを行なうことを示した。今後の研究の課題としては、センサの種類の拡張や、センサ形態が学習に与える影響の定量化などが挙げられる。

謝 辞

本研究は、日本学術振興会科学研究費補助金 18·2972 による研究助成を受け実施したものである。
(平成 19 年 10 月 30 日受付、平成 20 年 3 月 26 日再受付)

文 献

- (1) Balakrishnan, K. and Honavar, V.: On Sensor Evolution in Robotics, *Proceedings of the First International Conference*

- on Genetic Programming*, Stanford University, CA., pp. 455–460 (1996).
- (2) Bongard, J. C. and Pfeifer, R.: Evolving Complete Agents using Artificial Ontogeny, *Morpho-functional Machines: The New Species (Designing Embodied Intelligence)* (Hara, F. and Pfeifer, R., eds.), Springer-Verlag, pp. 237–258 (2003).
- (3) Bugajska, M. and Schultz, A.: Coevolution of Form and Function in the Design of Micro Air Vehicles, *4th NASA / DoD Workshop on Evolvable Hardware (EH 2002)*, IEEE Computer Society, pp. 154–166 (2002).
- (4) Dautenhahn, K., Polani, D. and Uthmann, T.: Special Issue on Sensor Evolution, *Artificial Life*, Vol. 7, No. 2, pp. 95–98 (2001).
- (5) Dusenberry, D. B.: *Sensory ecology: how organisms acquire and respond to information*, W. H. Freeman & Co. (1992).
- (6) Jung, T., Dauscher, P. and Uthmann, T.: Evolution and Learning: Evolving Sensors in a Simple MDP Environment, *Adaptive Behavior*, Vol. 11, No. 3, pp. 159–177 (2003).
- (7) Lichtensteiger, L. and Pfeifer, R.: An Optimal Sensor Morphology Improves Adaptability of Neural Network Controllers, *ICANN* (Dorronsoro, J. R., ed.), Springer, pp. 850–855 (2002).
- (8) Lipson, H. and Pollack, J. B.: Automatic design and manufacture of robotic lifeforms, *Nature*, Vol. 406, No. 6799, pp. 974–978 (2000).
- (9) Lund, H. H., Hallam, J. and Lee, W.-P.: Evolving Robot Morphology, *Proceedings of IEEE 4th International Conference on Evolutionary Computation*, IEEE Press, p. Invited paper (1997).
- (10) Pfeifer, R. and Bongard, J.: *How the Body Shapes the Way We Think*, Mit Press, Cambridge, MA (2006).
- (11) Pfeifer, R. and Scheier, C.: *Understanding Intelligence*, MIT Press, Cambridge, MA. (1999). (石黒章夫, 小林宏, 細田耕監訳: 知の創成 – 身体性認知科学への招待, 共立出版, (2001)).
- (12) Sims, K.: Evolving 3D Morphology and Behavior by Competition, *Artificial Life IV: Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems* (Brooks, R. and Maes, P., eds.), Cambridge, MA., MIT Press, pp. 28–39 (1994).
- (13) Sutton, R. S. and Barto, A. G.: *Reinforcement Learning: An Introduction*, MIT Press (1998). (三上貞芳, 皆川雅章共訳: 強化學習, 森北出版, 2000).
- (14) von Uexküll, J.: A stroll through the worlds of animals and men, *Instinctive Behavior: The Development of a Modern Concept* (Shiller, C. H., ed.), International University Press, pp. 5–80 (1957).
- (15) 朝井勇次, 有田隆也 : ブロック型ロボットを用いた構造と行動の共進化の試み, 情報処理学会論文誌 , Vol. 43, No. SIG 10, pp. 110–118 (2002).
- (16) 國吉康夫, ベルトウーズリュク : 身体性に基づく相互作用の創発に向けて, 日本ロボット学会誌 , Vol. 17, No. 1, pp. 29–33 (1999).
- (17) 滝田好宏 : SSM を用いた起動誘導車両の高速化 , 日本機械学会論文集 (C 編) , Vol. 65, No. 630, pp. 622–629 (1999).
- (18) 富永佳也 (編) : 昆虫の脳を探る , 共立出版 (1995).
- (19) 細田耕 : 形態が学習にもたらすもの , 学習が形態にもたらすもの , 日本ロボット学会誌 , Vol. 22, No. 2, pp. 186–189 (2004).
- (20) 矢入郁子 , 永合宣明 , 矢入健久 , 猪木誠二 : ピーカル操縦者意図認識のためのセンサ配置の考察 , 第 16 回人工知能学会全国大会オンライン資料 (2003). 2E2-02.

川 上 浩 司 (正員) 1987 年京都大学工学部精密工学科卒業。1989 年同大学大学院工学研究科修士課程修了。同年、岡山大学工学部情報工学科助手 , 1998 年 4 月京都大学大学院情報学研究科助教授 , 現在に至る。博士 (工学)。生態学的・創発的システム設計, 知的情報処理の研究に従事。計測自動制御学会論文賞 (1991, 2003 年度) 受賞。計測自動制御学会, 日本人工知能学会, 日本機械学会, 日本知能情報ファジィ学会などの会員。

片 井 修 (正員) 1969 年京都大学工学部機械工学科卒業。修士・博士課程を経て , 1974 年京都大学工学部精密工学科助手。1983 年同助教授 , 1994 年同教授。1998 年京都大学大学院情報学研究科の創設に参画し , 同年同研究科システム科学専攻教授。その間 , 1980-81 年フランス国 INRIA 客員研究員 , 工学博士。自然システムと人工システムの共生の在り方について興味を持つ。計測自動制御学会論文賞 (1989, 1991, 2003 年度) , 同著述賞 (1992 年度) など受賞。

杉 浦 孔 明 (正員) 2002 年京都大学工学部電気電子工学科卒業。2004 年同大学大学院情報学研究科修士課程修了 , 2007 年同大学大学院情報学研究科博士後期課程修了。2004 年より (株) 国際電気通信基礎技術研究所研修研究員 , 2007 年同連携研究員を経て , 2008 年より現職。その間 , 2006-2008 年日本学術振興会特別研究員。博士 (情報学)。機械学習 , 言語獲得 , センサ進化の研究に従事。計測自動制御学会 , 日本人工知能学会 , 日本ロボット学会などの会員。

