

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4670007号
(P4670007)

(45) 発行日 平成23年4月13日(2011.4.13)

(24) 登録日 平成23年1月28日(2011.1.28)

(51) Int.Cl. F 1
G 0 6 F 17/50 (2006.01)
 G 0 6 F 17/50 6 0 4 D
 G 0 6 F 17/50 6 8 0 Z

請求項の数 8 (全 16 頁)

<p>(21) 出願番号 特願2005-223343 (P2005-223343) (22) 出願日 平成17年8月1日(2005.8.1) (65) 公開番号 特開2007-41723 (P2007-41723A) (43) 公開日 平成19年2月15日(2007.2.15) 審査請求日 平成20年3月27日(2008.3.27)</p> <p>特許法第30条第1項適用 2005年3月17日 社団法人計測自動制御学会発行の「第32回 知能システムシンポジウム資料」に発表</p> <p>(出願人による申告)平成17年度独立行政法人情報通信研究機構、研究テーマ「人間情報コミュニケーションの研究開発」に関する委託研究、産業活力再生特別措置法第30条の適用を受ける特許出願</p>	<p>(73) 特許権者 393031586 株式会社国際電気通信基礎技術研究所 京都府相楽郡精華町光台二丁目2番地2</p> <p>(74) 代理人 100067828 弁理士 小谷 悦司</p> <p>(74) 代理人 100096150 弁理士 伊藤 孝夫</p> <p>(74) 代理人 100109438 弁理士 大月 伸介</p> <p>(72) 発明者 杉浦 孔明 京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内</p> <p>(72) 発明者 下原 勝憲 京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内 最終頁に続く</p>
---	--

(54) 【発明の名称】 センサ設計装置、センサ設計方法、センサ設計プログラム及びロボット

(57) 【特許請求の範囲】

【請求項1】

センサ、アクチュエータ及びコントローラを有し、コントローラにより、センサの出力に基づき行動学習しつつアクチュエータが制御されるロボットの前記センサの形態を設計するセンサ設計装置であって、

前記センサの形態を特定するための複数の遺伝子型を作成し、各遺伝子型により特定されるセンサの形態を有し、かつ共通のアクチュエータ及び共通のコントローラを有する複数のロボットを仮想的に作成する初期世代作成手段と、

前記初期世代作成手段により作成された複数のロボットに仮想的に学習を行わせ、学習結果を基に各ロボットの適応度を算出する学習手段と、

前記学習手段により算出された各ロボットの適応度を基に親個体となる複数のロボットを選択し、選択した複数のロボットの遺伝子型から遺伝的アルゴリズムに基づき次世代の遺伝子型を作成し、各次世代の遺伝子型により特定されるセンサの形態を有し、かつ共通のアクチュエータ及び共通のコントローラを有する複数のロボットを仮想的に作成する次世代作成手段とを備え、

前記学習手段は、前記次世代作成手段により作成されたロボットに仮想的に再度学習を行わせ、学習結果を基に各ロボットの適応度を算出し、

前記次世代作成手段及び前記学習手段による処理を所定数繰り返すことによりセンサの形態を決定することを特徴とするセンサ設計装置。

【請求項2】

前記学習手段は、Q学習を用いてロボットに学習を行わせることを特徴とする請求項1に記載のセンサ設計装置。

【請求項3】

前記次世代作成手段は、

前記学習手段により算出された各ロボットの適応度を基に、前記複数のロボットの中から学習性能の高い所定数のロボットを親個体として選択するとともに、残りのロボットからトーナメント選択により同数のロボットを親個体として選択する選択手段と、

前記選択手段により選択された親個体の遺伝子型から遺伝的アルゴリズムに基づき次世代の遺伝子型を作成し、各次世代の遺伝子型により特定されるセンサの形態を有し、かつ共通のアクチュエータ及び共通のコントローラを有する複数のロボットを次世代ロボットとして仮想的に作成する作成手段とを備え、

前記学習手段は、前記作成手段により作成された次世代ロボットに仮想的に再度学習を行わせ、学習結果を基に各ロボットの適応度を算出することを特徴とする請求項1又は2に記載のセンサ設計装置。

【請求項4】

前記遺伝子型は、センサの位置、個数、分解能、センシング間隔の少なくとも一つを特定することを特徴とする請求項1～3のいずれかに記載のセンサ設計装置。

【請求項5】

前記次世代作成手段は、親個体となるロボットの遺伝子型から交叉及び突然変異の少なくとも一方を用いて次世代の遺伝子型を作成し、各遺伝子型により特定されるセンサの形態を有し、かつ共通のアクチュエータ及び共通のコントローラを有する複数のロボットを仮想的に作成することを特徴とする請求項1～4のいずれかに記載のセンサ設計装置。

【請求項6】

初期世代作成手段、学習手段及び次世代作成手段を備えるセンサ設計装置を用いて、センサ、アクチュエータ及びコントローラを有し、コントローラにより、センサの出力に基づき行動学習しつつアクチュエータが制御されるロボットの前記センサの形態を設計するセンサ設計方法であって、

前記初期世代作成手段が、前記センサの形態を特定するための複数の遺伝子型を作成し、各遺伝子型により特定されるセンサの形態を有し、かつ共通のアクチュエータ及び共通のコントローラを有する複数のロボットを仮想的に作成する第1のステップと、

前記学習手段が、前記初期世代作成手段により作成された複数のロボットに仮想的に学習を行わせ、学習結果を基に各ロボットの適応度を算出する第2のステップと、

前記次世代作成手段が、前記学習手段により算出された各ロボットの適応度を基に親個体となる複数のロボットを選択し、選択した複数のロボットの遺伝子型から遺伝的アルゴリズムに基づき次世代の遺伝子型を作成し、各次世代の遺伝子型により特定されるセンサの形態を有し、かつ共通のアクチュエータ及び共通のコントローラを有する複数のロボットを仮想的に作成する第3のステップと、

前記学習手段が、前記第3のステップにおいて作成されたロボットに仮想的に再度学習を行わせ、学習結果を基に各ロボットの適応度を算出する第4のステップとを含み、

前記第3及び第4のステップによる処理を所定数繰り返すことによりセンサの形態を決定することを特徴とするセンサ設計方法。

【請求項7】

センサ、アクチュエータ及びコントローラを有し、コントローラにより、センサの出力に基づき行動学習しつつアクチュエータが制御されるロボットの前記センサの形態を設計するためのセンサ設計プログラムであって、

前記センサの形態を特定するための複数の遺伝子型を作成し、各遺伝子型により特定されるセンサの形態を有し、かつ共通のアクチュエータ及び共通のコントローラを有する複数のロボットを仮想的に作成する初期世代作成手段と、

前記初期世代作成手段により作成された複数のロボットに仮想的に学習を行わせ、学習結果を基に各ロボットの適応度を算出する学習手段と、

10

20

30

40

50

前記学習手段により算出された各ロボットの適応度を基に親個体となる複数のロボットを選択し、選択した複数のロボットの遺伝子型から遺伝的アルゴリズムに基づき次世代の遺伝子型を作成し、各次世代の遺伝子型により特定されるセンサの形態を有し、かつ共通のアクチュエータ及び共通のコントローラを有する複数のロボットを仮想的に作成する次世代作成手段としてコンピュータを機能させ、

前記学習手段は、前記次世代作成手段により作成されたロボットに仮想的に再度学習を行わせ、学習結果を基に各ロボットの適応度を算出し、

前記次世代作成手段及び前記学習手段による処理を所定数繰り返すことによりセンサの形態を決定することを特徴とするセンサ設計プログラム。

【請求項 8】

請求項 1 ~ 5 のいずれかに記載のセンサ設計装置により設計されたセンサと、前記アクチュエータと、前記コントローラとを有することを特徴とするロボット。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ロボットの行動学習に使用されるセンサの形態を設計するセンサ設計装置、センサ設計方法及びセンサ設計プログラム、並びに前記センサ設計装置により設計されたセンサを有するロボットに関するものである。

【背景技術】

【0002】

環境の変化に適応的なロボットを構築するためには、ロボットの形態、制御系及び環境のバランスを考慮しなければならない。近年、形態と制御系とをうまく組み合わせてロボットに適応的な振る舞いを獲得させる研究が注目されている。例えば、非特許文献 1 では、エージェントの形態と制御系とをソフトウェア上で進化させる手法と迅速成形技術と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している。

【非特許文献 1】リプソン エイチ (Lipson H.) 他、「ロボット形態の自動設計及び製造」(Automatic design and manufacture of robotic lifeforms)、ネイチャー (Nature)、2000年、Vol. 406、No. 6799、p. 974 - p. 978

【発明の開示】

【発明が解決しようとする課題】

【0003】

しかしながら、上記のような手法では、個体発生的なスパンにおける適応、すなわち学習にとって有利なセンサの形態を自動的に設計することはできない。

【0004】

本発明の目的は、ロボットの行動学習に有利なセンサの形態を自動的に設計することができるセンサ設計装置、センサ設計方法、センサ設計プログラム及びロボットを提供することである。

【課題を解決するための手段】

【0005】

本発明に係るセンサ設計装置は、センサ、アクチュエータ及びコントローラを有し、コントローラにより、センサの出力に基づき行動学習しつつアクチュエータが制御されるロボットのセンサの形態を設計するセンサ設計装置であって、センサの形態を特定するための複数の遺伝子型を作成し、各遺伝子型により特定されるセンサの形態を有し、かつ共通のアクチュエータ及び共通のコントローラを有する複数のロボットを仮想的に作成する初期世代作成手段と、初期世代作成手段により作成された複数のロボットに仮想的に学習を行わせ、学習結果を基に各ロボットの適応度を算出する学習手段と、学習手段により算出された各ロボットの適応度を基に親個体となる複数のロボットを選択し、選択した複数のロボットの遺伝子型から遺伝的アルゴリズムに基づき次世代の遺伝子型を作成し、各次世代の遺伝子型により特定されるセンサの形態を有し、かつ共通のアクチュエータ及び共通

10

20

30

40

50

のコントローラを有する複数のロボットを仮想的に作成する次世代作成手段とを備え、学習手段は、次世代作成手段により作成されたロボットに仮想的に再度学習を行わせ、学習結果を基に各ロボットの適応度を算出し、次世代作成手段及び学習手段による処理を所定数繰り返すことによりセンサの形態を決定するものである。

【0006】

本発明に係るセンサ設計装置では、センサの形態を特定するための複数の遺伝子型が作成され、各遺伝子型により特定されるセンサの形態を有し、かつ共通のアクチュエータ及び共通のコントローラを有する複数のロボットが仮想的に作成され、作成された複数のロボットに学習を行わせ、学習した各ロボットの学習結果を基に各ロボットの適応度が算出される。

10

【0007】

次に、算出された各ロボットの適応度を基に親個体となる複数のロボットが選択され、選択された複数のロボットの遺伝子型から遺伝的アルゴリズムに基づき次世代の遺伝子型が作成され、各次世代の遺伝子型により特定されるセンサの形態を有し、かつ共通のアクチュエータ及び共通のコントローラを有する複数のロボットが仮想的に作成され、作成された複数のロボットに再度学習を行わせ、学習結果を基に各ロボットの適応度が算出され、これらの次世代作成処理及びその学習処理が所定数繰り返されることによりセンサの形態が決定される。

【0008】

したがって、学習結果に基づいて、適応度の高いロボットすなわち学習性能が高いロボットが、死滅することなく、親個体のロボットとして選択され、センサの形態を特定する遺伝子型を進化させることができるので、ロボットの行動学習に有利なセンサの形態を自動的に設計することができる。

20

【0009】

学習手段は、Q学習により複数のロボットに学習を行わせることが好ましい。この場合、種々の遺伝子型すなわち種々のセンサの形態を有するロボットを効率的に学習させることができ、好適な形態のセンサを高速に設計することができる。

【0010】

次世代作成手段は、学習手段により算出された各ロボットの適応度を基に、複数のロボットの中から学習性能の高い所定数のロボットを親個体として選択するとともに、残りのロボットからトーナメント選択により同数のロボットを親個体として選択する選択手段と、選択手段により選択された親個体の遺伝子型から遺伝的アルゴリズムに基づき次世代の遺伝子型を作成し、各次世代の遺伝子型により特定されるセンサの形態を有し、かつ共通のアクチュエータ及び共通のコントローラを有する複数のロボットを次世代ロボットとして仮想的に作成する作成手段とを備え、学習手段は、作成手段により作成された次世代ロボットに仮想的に再度学習を行わせ、学習結果を基に各ロボットの適応度を算出することが好ましい。

30

【0011】

この場合、エリート戦略により適応度の高いロボットがそのまま親個体として選択されるので、学習性能が高いロボットが偶然選択されずに死滅することを防止することができる。この結果、学習性能が高いロボットを死滅させることなく、順次学習させることができるので、学習結果に基づいて最適なセンサの形態を設計することができる。

40

【0012】

遺伝子型は、センサの位置、個数、分解能、センシング間隔の少なくとも一つを特定することが好ましい。この場合、センサの位置、個数、分解能、センシング間隔等の形態を自動的に設計することができる。

【0013】

次世代作成手段は、親個体となるロボットの遺伝子型から交叉及び突然変異の少なくとも

50

も一方を用いて次世代の遺伝子型を作成し、各遺伝子型により特定されるセンサの形態を有し、かつ共通のアクチュエータ及び共通のコントローラを有する複数のロボットを仮想的に作成することが好ましい。この場合、広範な探索空間から最適な遺伝子型を探索することができ、最適な形態のセンサを効率的に設計することができる。

【0014】

本発明に係るセンサ設計方法は、初期世代作成手段、学習手段及び次世代作成手段を備えるセンサ設計装置を用いて、センサ、アクチュエータ及びコントローラを有し、コントローラにより、センサの出力に基づき行動学習しつつアクチュエータが制御されるロボットのセンサの形態を設計するセンサ設計方法であって、初期世代作成手段が、センサの形態を特定するための複数の遺伝子型を作成し、各遺伝子型により特定されるセンサの形態を有し、かつ共通のアクチュエータ及び共通のコントローラを有する複数のロボットを仮想的に作成する第1のステップと、学習手段が、初期世代作成手段により作成された複数のロボットに仮想的に学習を行わせ、学習結果を基に各ロボットの適応度を算出する第2のステップと、次世代作成手段が、学習手段により算出された各ロボットの適応度を基に親個体となる複数のロボットを選択し、選択した複数のロボットの遺伝子型から遺伝的アルゴリズムに基づき次世代の遺伝子型を作成し、各次世代の遺伝子型により特定されるセンサの形態を有し、かつ共通のアクチュエータ及び共通のコントローラを有する複数のロボットを仮想的に作成する第3のステップと、学習手段が、第3のステップにおいて作成されたロボットに仮想的に再度学習を行わせ、学習結果を基に各ロボットの適応度を算出する第4のステップとを含み、第3及び第4のステップによる処理を所定数繰り返すことによりセンサの形態を決定するものである。

【0015】

本発明に係るセンサ設計プログラムは、センサ、アクチュエータ及びコントローラを有し、コントローラにより、センサの出力に基づき行動学習しつつアクチュエータが制御されるロボットのセンサの形態を設計するためのセンサ設計プログラムであって、センサの形態を特定するための複数の遺伝子型を作成し、各遺伝子型により特定されるセンサの形態を有し、かつ共通のアクチュエータ及び共通のコントローラを有する複数のロボットを仮想的に作成する初期世代作成手段と、初期世代作成手段により作成された複数のロボットに仮想的に学習を行わせ、学習結果を基に各ロボットの適応度を算出する学習手段と、学習手段により算出された各ロボットの適応度を基に親個体となる複数のロボットを選択し、選択した複数のロボットの遺伝子型から遺伝的アルゴリズムに基づき次世代の遺伝子型を作成し、各次世代の遺伝子型により特定されるセンサの形態を有し、かつ共通のアクチュエータ及び共通のコントローラを有する複数のロボットを仮想的に作成する次世代作成手段としてコンピュータを機能させ、学習手段は、次世代作成手段により作成されたロボットに仮想的に再度学習を行わせ、学習結果を基に各ロボットの適応度を算出し、次世代作成手段及び学習手段による処理を所定数繰り返すことによりセンサの形態を決定するものである。

【0016】

本発明に係るロボットは、上記いずれかに記載のセンサ設計装置により設計されたセンサと、アクチュエータと、コントローラとを有するものである。

【発明の効果】

【0017】

本発明によれば、学習結果に基づいて、適応度の高いロボットすなわち学習性能が高いロボットが、死滅することなく、親個体のロボットとして選択され、センサの形態を特定する遺伝子型を進化させることができるので、ロボットの行動学習に有利なセンサの形態を自動的に設計することができる。

【発明を実施するための最良の形態】

【0018】

以下、本発明の一実施の形態によるセンサ設計装置について図面を参照しながら説明する。図1は、本発明の一実施の形態によるセンサ設計装置の構成を示すブロック図である

。

【 0 0 1 9 】

図 1 に示すセンサ設計装置は、入力部 1、初期世代作成部 2、行動学習部 3、選択部 4、次世代作成部 5 及び出力部 6 を備える。センサ設計装置は、ROM (リードオンリメモリ)、CPU (中央演算処理装置)、RAM (ランダムアクセスメモリ)、外部記憶装置、記録媒体駆動装置、入力装置及び表示装置等を備える通常のコンピュータを用いて、後述するセンサ設計処理を実行するためのセンサ設計プログラムを CPU 等で実行することにより上記の各機能を実現することができる。

【 0 0 2 0 】

本実施の形態では、上記コンピュータと、シミュレータとして、Cyberbotics 社製 Webots を用い、センサやアクチュエータの特性を定義して後述するロボットを仮想的に作成してセンサの形態をシミュレーションしている。なお、センサ設計装置の構成は、上記の例に特に限定されず、上記各機能の一部又は全てを専用のハードウェア回路により実現する等の種々の変更が可能である。

【 0 0 2 1 】

入力部 1 は、ロボットを構成するセンサ、アクチュエータ及びコントローラ (エージェント) 等を定義するためのモデルデータ等をユーザが入力するために使用される。

【 0 0 2 2 】

初期世代作成部 2 は、入力部 1 から入力されたモデルデータ等を用いて、センサの形態を特定するための遺伝子型がランダムになるように複数の遺伝子型を作成し、各遺伝子型により特定されるセンサの形態を有する複数の初期世代ロボットを仮想的に作成し、作成した初期世代ロボットのデータを行動学習部 3 へ出力する。なお、初期世代ロボットのセンサ以外の構成要素であるアクチュエータ及びコントローラ等は、すべてのロボットについて共通であり、後述する次世代ロボットも同様である。

【 0 0 2 3 】

ここで、センサの形態は、物理的な形態のみではなく、センサの特性及び使用状態等をも含み、本実施の形態に用いられる遺伝子型は、センサの位置、個数、分解能、センシング間隔、及びセンサ値の制御系に対する重みを表す結合強度等のパラメータを特定するコード等を特定する情報であり、「1」又は「0」を用いて表現される。例えば、1 のパラメータが 4 bit でデコードされ、4 種類のパラメータから遺伝子型が表される場合、遺伝子型の長さは、256 bit となる。

【 0 0 2 4 】

行動学習部 3 は、初期世代ロボットのデータを用いて、初期世代ロボットに仮想的に強化学習を行わせ、その学習結果に基づく適応度を算出して各ロボットのデータとともに選択部 4 へ出力する。例えば、行動学習部 3 は、Q 学習を用いてロボットに学習を行わせ、ロボットのコントローラに以下の処理を仮想的に実行させる。

1. センサを用いて環境の状態 s_t を観測する。
2. 行動選択戦略に従って行動 a_t を実行する。
3. 状態に応じて報酬 r_t を受け取る。
4. センサを用いて状態遷移後の状態 s_{t+1} を観測する。
5. 下記式 (1) に従って Q 値を更新する。

【 0 0 2 5 】

$$Q(s_t, a_t) = Q(s_t, a_t) + \{ r_t + \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t) \} \dots (1)$$

ここで、 α は学習率 ($0 < \alpha < 1$)、 γ は割引率 ($0 < \gamma < 1$) である。

6. 時間ステップ t を $t+1$ に進めて手順 1 に戻る。

【 0 0 2 6 】

選択部 4 は、エリート戦略に従い、適応度を基に上位のロボットを親個体として所定数選択するとともに、トーナメント選択に従い、選択しなかった残りのロボットから同数のロボットを親個体としてさらに選択してそれらのデータを次世代作成部 5 へ出力する。

【 0 0 2 7 】

次世代作成部 5 は、選択部 4 により選択された親個体の遺伝子型から遺伝的アルゴリズムに基づき次世代の遺伝子型を作成し、各遺伝子型により特定されるセンサの形態を有する複数の次世代ロボットを仮想的に作成し、作成した次世代ロボットのデータを行動学習部 3 へ出力する。例えば、次世代作成部 5 は、選択部 4 により選択された親個体の遺伝子型から交叉及び突然変異により次世代の遺伝子型を作成する。

【 0 0 2 8 】

なお、選択部 4 及び次世代作成部 5 による選択操作及び遺伝子操作は、上記の例に特に限定されず、適応度を基に下位（例えば、10%）のロボットを取り除き、残ったものからトーナメント選択、ルーレット選択、期待値選択、又はランキング選択等により同数の親を選択して遺伝子操作を加える等の種々の変更が可能である。また、交叉についても、1点交叉、多点交叉、又は一様交叉等の種々のものを用いることができる。

【 0 0 2 9 】

次世代作成部 5 から出力される次世代ロボットのデータを受け取った行動学習部 3 は、そのデータを用いて、次世代ロボットに仮想的に強化学習、例えば、Q 学習を行わせ、その学習結果に基づく適応度を各ロボットのデータとともに選択部 4 へ出力する。

【 0 0 3 0 】

上記の選択部 4、次世代作成部 5 及び行動学習部 3 による処理を所定世代繰り返すことにより最終世代のロボットが仮想的に作成され、選択部 4 は、最終世代ロボットの中から適応度の最も高いロボットを選択し、最良個体としてそのデータを出力部 6 へ出力する。

【 0 0 3 1 】

出力部 6 は、最良個体のデータを表示又は印刷等して最終的に決定されたセンサの形態をユーザに知らせる。なお、出力部 6 が出力するセンサの形態としては、遺伝子型、表示型等のいずれを用いてもよい。

【 0 0 3 2 】

本実施の形態では、初期世代作成部 2 が初期世代作成手段の一例に相当し、行動学習部 3 が学習手段の一例に相当し、選択部 4 及び次世代作成部 5 が次世代作成手段の一例に相当し、選択部 4 が選択手段の一例に相当し、次世代作成部 5 が作成手段の一例に相当する。

【 0 0 3 3 】

次に、上記のように構成されたセンサ設計装置によるセンサ設計処理について説明する。図 2 は、図 1 に示すセンサ設計装置によるセンサ設計処理を説明するためのフローチャートである。

【 0 0 3 4 】

まず、ユーザが、入力部 1 を用いて、ロボットを構成するセンサ、アクチュエータ及びコントローラ等を定義するためのモデルデータ等を入力すると、ステップ S 1 において、初期世代作成部 2 は、入力されたモデルデータ等を用いて、センサの形態を特定するための遺伝子型がランダムになるように複数の遺伝子型を作成する。

【 0 0 3 5 】

ここで、本センサ設計装置によりセンサの形態が設計されるロボットの一例について説明する。図 3 は、図 1 に示すセンサ設計装置の設計対象となるロボットの一例を示す底面模式図であり、図 4 は、図 3 に示すロボットのタスク環境を示す模式図であり、図 5 は、図 4 に示すコース上のロボットの状態を示す模式図である。

【 0 0 3 6 】

図 3 に示すロボット 10 は、赤外発光 LED 及び受光素子から構成されるセンサ 11、センサ 11 が固定される固定台 12、固定台 12 を保持する本体部 13、及び本体部 13 に回転可能に支持される 2 個の車輪 14 を備える。ロボット 10 は、床に引かれたラインに沿って移動してゴールを目指すロボットであり、ライントレーサと呼ばれる。

【 0 0 3 7 】

図 4 に示すコースは、2 m × 5 m の長方形であり、4 分の 1 ずつの地点にチェックボイ

10

20

30

40

50

ント P 1 ~ P 4 が設けられ、図 5 に示すように、ロボット 1 0 は、ライン L I に沿って移動してゴールを目指す。ここで、ロボット 1 0 のセンサ 1 1 は、コース上に設けられたライン L I を検出するために床面の色に応じた値を本体部 1 3 内のコントローラ（図示省略）へ出力し、コントローラは、本体部 1 3 内のモータ及び駆動回路（図示省略）を制御することにより、センサ 1 1 の出力に応じてライン L I 上をトレースするように車輪 1 4 を駆動させる。

【 0 0 3 8 】

センサ 1 1 は、予め規定された 8×4 のマトリックス状の 3 2 個の配置可能位置の中から任意の個数及び位置を選択して配置されることができ、本センサ設計装置は、ロボットの行動学習に有利なセンサ 1 1 の位置及び個数を自動的に設計する。この場合、ステップ S 1 において、上記のライントレーサ及びコースをモデル化するためのモデルデータ等が入力され、初期世代作成部 2 は、入力されたモデルデータ等を用いて、センサ 1 1 の位置及び個数を特定するための 3 2 b i t の遺伝子型がコード化され、5 0 個の遺伝子型をランダムに作成する。

10

【 0 0 3 9 】

次に、ステップ S 2 において、初期世代作成部 2 は、作成した各遺伝子型により特定されるセンサの形態を有する複数の初期世代ロボットを仮想的に作成し、作成した初期世代ロボットのデータを行動学習部 3 へ出力する。このとき、行動学習部 3 は、時間ステップ t 等を初期化する。

【 0 0 4 0 】

20

例えば、上記のライントレーサの場合、ステップ S 2 において、初期世代作成部 2 は、5 0 個の遺伝子型の各々により特定されるセンサの位置及び配置を有する 5 0 個の初期世代ロボットを仮想的に作成し、初期世代ロボットの個体数（集団数）は 5 0 となる。

【 0 0 4 1 】

次に、ステップ S 3 において、行動学習部 3 は、初期世代ロボットのデータを用いてモデル化した仮想ロボットのコントローラにセンサを用いて環境の状態 s_t を観測させる。例えば、環境の状態 s_t は、各センサの出力から構成され、センサの出力が 2 値で、4 つのセンサを有するロボットの場合、状態数は $2 \times 2 \times 2 \times 2 = 1 6$ になる。なお、センサの出力値は、上記の例に特に限定されず、多値を用いてもよい。

【 0 0 4 2 】

30

次に、ステップ S 4 において、行動学習部 3 は、行動選択戦略として グリーディ戦略に従った行動 a_t をコントローラに実行させる。ここで、グリーディ戦略とは、の確率でランダムな行動を選択し、それ以外の場合は Q 値が最大の行動を取る戦略である。

【 0 0 4 3 】

例えば、上記のライントレーサの場合、コントローラが取り得る行動 a_t は、直進 a_{t0} 、左に曲がる a_{t1} 、右に曲がる a_{t2} 、大きく左に曲がる a_{t3} 、大きく右に曲がる a_{t4} の 5 種類あり、ロボットの左右の車輪の角速度 (rad/s) を ω_L, ω_R とすると、 a_{t0} の場合は $(\omega_L, \omega_R) = (15, 15)$ 、 a_{t1} の場合は $(\omega_L, \omega_R) = (13, 6)$ 、 a_{t2} の場合は $(\omega_L, \omega_R) = (6, 13)$ 、 a_{t3} の場合は $(\omega_L, \omega_R) = (8, 2)$ 、 a_{t4} の場合は $(\omega_L, \omega_R) = (2, 8)$ になるように車輪の角速度が制御され、 $\Delta = 0.01$ を用いることができる。

40

【 0 0 4 4 】

次に、ステップ S 5 において、行動学習部 3 は、状態に応じて報酬 r_t をコントローラに与える。例えば、上記のライントレーサの場合、ロボットの中心がラインから外れるほど罰を大きくした報酬を用いることができ、図 3 に示す 8 列のセンサをロボットの下から見て左側から第 0 列、第 1 列、...、第 7 列とすると、観測した状態が、(1) 第 0 列のセンサ又は第 7 列のセンサがライン上にある場合は $r_t = -20.0$ 、(2) 第 1 列のセンサ又は第 6 列のセンサがライン上にある場合は $r_t = -10.0$ 、(3) 第 2 列のセンサ又は第 5 列のセンサがライン上にある場合は $r_t = 0.0$ 、(4) 第 3 列のセンサ又は第 4 列のセンサがライン上にある場合は $r_t = 1.0$ 、(5) 全てのセンサがライン外にあ

50

る場合は $r_t = -100.0$ となる報酬 r_t を用いることができる。なお、上記の条件 (1) ~ (4) は、(4)、(3)、(2)、(1) の順に優先して使用される。

【0045】

次に、ステップ S6 において、行動学習部 3 は、上記式 (1) に従って Q 値をコントローラに更新させる。上記式 (1) により、ロボットが採った行動 a_t により報酬 r_t を得て $Q(s_t, a_t)$ が更新され、その報酬が正であったならば、 $Q(s_t, a_t)$ が増加するため、次回同じ状態になった場合、その行動を採る可能性が高くなる。例えば、上記のライントレーサの場合、上記式 (1) において、学習率 $= 0.8$ 、割引率 $= 0.99$ を用いることができる。

【0046】

次に、ステップ S7 において、行動学習部 3 は、時間ステップ t が最大ステップ数に達して 1 試行が終了したか否かを判断し、最大ステップ数に達していない場合は、時間ステップ t を $t+1$ に進め、ステップ S3 に戻って以降の処理を継続し、最大ステップ数に達した場合はステップ S8 へ処理を移行する。例えば、上記のライントレーサの場合、最大ステップ数 $= 2000$ を用いることができ、また、ライントレーサが最大ステップ数以内にゴールに到達した場合も 1 試行が終了するため、ステップ S8 へ処理を移行する。

【0047】

次に、ステップ S8 において、行動学習部 3 は、試行回数がエピソード数に達して学習が終了したか否かを判断し、学習が終了していない場合は、学習条件等を初期化してステップ S3 に戻って次の試行を継続し、学習が終了した場合はステップ S9 へ処理を移行する。例えば、上記のライントレーサの場合、エピソード数 $= 100$ を用いることができ、また、試行機会 (エピソード) 毎にロボットの位置及び向きを初期化して次の試行を継続する。

【0048】

学習が終了した場合、ステップ S9 において、行動学習部 3 は、学習結果に基づく適応度を算出し、学習を行った各ロボットのデータとともに対応する応答度を選択部 4 へ出力する。例えば、上記のライントレーサの場合、1 個体に付き 100 回の試行を行い、下記の式 (2) により適応度 ϕ を算出する。

【0049】

【数 1】

$$\phi = \frac{1}{N} \sum_{i=1}^N H_i \quad \dots \quad (2)$$

$$H_i = \begin{cases} 0.25 \times \left(1.0 - \frac{t_1}{T_{\max}/4}\right) & (P_1 \text{ に到達}) \\ 0.50 \times \left(1.0 - \frac{t_2}{(T_{\max}/4) \times 2}\right) & (P_2 \text{ に到達}) \\ 0.75 \times \left(1.0 - \frac{t_3}{(T_{\max}/4) \times 3}\right) & (P_3 \text{ に到達}) \\ 1.0 - \frac{t_4}{T_{\max}} & (P_4 \text{ に到達}) \end{cases}$$

【0050】

ここで、 N はエピソード数(100回)、 H_i は*i*回目の試行の達成度、 t_i は図4に示すチェックポイント P_i への到達時間(sec)、 T_{max} は最大試行時間(128sec)である。上式より、学習が進んだ個体(ロボット)ほど各試行における達成度が高くなり、適応度も高くなる。すなわち、適応度の高い個体は、学習しやすいセンサの形態を有しているといえることができる。なお、適応度の計算方法は、上記の例に特に限定されず、種々の変更が可能である。

【0051】

次に、ステップS10において、選択部4は、行動学習部3から出力されるロボットのデータを基に学習が終了したロボットが最終世代ロボットであるか否かを判断する。最終世代ロボットである場合、選択部4は、適応度の最も高いロボットを選択し、最良個体としてそのデータを出力部6へ出力し、出力部6は、最良個体のセンサの遺伝子型を表示する。例えば、上記のライトレーサの場合、最終世代数として50が用いられ、第50世代目のロボットの学習が終了した後、適応度の最も高いロボットのセンサの位置及び個数を表す遺伝子型が表示される。

10

【0052】

一方、最終世代ロボットでない場合は、ステップS11において、選択部4は、エリート戦略に従い、適応度が上位のロボットを所定数選択し、親個体としてそのデータを次世代作成部5へ出力する。例えば、上記のライトレーサの場合、エリート数が5に設定され、50個のロボットのうち適応度が上位5位以内の5個のロボットが親個体として選択され、無条件で次世代に残される。

20

【0053】

次に、ステップS12において、選択部4は、トーナメント選択に従い、選択しなかった残りのロボットから所定数の個体をランダムに選択し、その中で適応度の最も高い個体を選択し、この過程を集団数が得られるまで繰り返すことにより、親個体を選択してそのデータを次世代作成部5へ出力する。例えば、上記のライトレーサの場合、選択されなかった45個のロボットからトーナメント選択により45個のロボットが親個体として選択される。

【0054】

次に、ステップS13において、次世代作成部5は、選択部4により選択された親個体の遺伝子型から交叉及び突然変異により次世代の遺伝子型を作成し、各遺伝子型により特定されるセンサの形態を有する複数の次世代ロボットを仮想的に作成し、作成した次世代ロボットのデータを行動学習部3へ出力する。例えば、上記のライトレーサの場合、親個体として選択された50個のロボットから、突然変異率=0.03及び交叉率=1.0で遺伝子操作が行われ、50個の次世代ロボットが作成される。

30

【0055】

上記ステップS13の処理後、ステップS3に戻って次世代のロボットが学習を行い、遺伝的アルゴリズムによりさらに次世代のロボットが順次作成され、最終世代のロボットの学習が終了するまでステップS3～S13の処理が繰り返される。

【0056】

上記の処理により、本実施の形態では、センサの形態を特定するための複数の遺伝子型が作成され、各遺伝子型により特定されるセンサの形態を有する複数の初期世代ロボットが仮想的に作成され、複数の初期世代ロボットにQ学習を行わせ、学習した各ロボットの学習結果を基に各ロボットの適応度が算出される。

40

【0057】

次に、算出された各ロボットの適応度を基に、複数のロボットの中から学習性能の高い所定数のロボットがそのまま親個体として選択されるとともに、残りのロボットからトーナメント選択により同数のロボットが親個体として選択され、選択された親個体の遺伝子型から交叉及び突然変異により次世代の遺伝子型が作成され、各次世代の遺伝子型により特定されるセンサの形態を有する複数のロボットが次世代ロボットとして仮想的に作成され、複数の次世代ロボットに再度Q学習を行わせ、学習結果を基に各ロボットの適応度が

50

算出され、これらの次世代作成処理及びその学習処理が最終世代まで繰り返されることによりセンサの形態が決定される。

【0058】

したがって、学習結果に基づいて、適応度の高いロボットすなわち学習性能が高いロボットが、死滅することなく、親個体のロボットとして選択され、センサの形態を特定する遺伝子型を進化させることができるので、ロボットの行動学習に有利なセンサの形態を自動的に設計することができ、ロボットの学習能力に適した状態空間を構築するために学習結果に基づいてセンサの形態を自律的に設計することができる。

【0059】

また、設計時に仮想的に行われる学習により自動設計されたセンサの形態に最適な学習アルゴリズムをも獲得することができる。さらに、ロボットの形態のうちセンサに特化してその形態を自動設計しているので、物理世界と情報世界のインターフェースであるセンサをボトムアップ的に構築することができるとともに、アクチュエータが固定されるため、ハードウェア上で実現しやすいという利点もある。

【0060】

次に、上記センサ設定装置によるセンサの設計結果について、図3に示すラインレーサを例に具体的に説明する。図6は、図1に示すセンサ設定装置により設計されたラインレーサの各世代における適応度と、各世代における最良個体のセンサ個数との変化を示す図である。図中、実線は最大適応度を、破線は最良個体のセンサ個数を、一点鎖線は実験を10回行ったときの平均適応度をそれぞれ示している。

【0061】

図6から、最良個体においては、センサ個数が減少するに従って適応度が増加していることがわかる。例えば、第5世代（センサ数7前後）の適応度が約0.2であるのに対し、第50世代（センサ数5前後）の適応度は約0.4に上昇している。この結果、図4に示すラインレース環境の場合、5個前後のセンサが学習に有利であり、学習器を用いることにより、マイコンカーラリーで標準的に用いられているセンサ個数（8個）を削減することができることがわかった。

【0062】

図7は、図1に示すセンサ設計装置により設計されたセンサ形態の代表例を示す図であり、図中の黒丸がセンサを表している。例えば、センサが4個の場合、図7の（a）に示すセンサ配置が設計され、センサが5個の場合、図7の（b）に示すセンサ配置が設計され、センサが6個の場合、図7の（c）に示すセンサ配置が設計された。

【0063】

次に、図1に示すセンサ設計装置により設計されたセンサ形態と、人手により設計したセンサ形態とを用いて学習結果の比較を行った。図8は、図1に示すセンサ設計装置により設計されたセンサ形態及び人手により設計されたセンサ形態の例を示す図であり、図中の黒丸がセンサを表している。

【0064】

図8の（c）に示すセンサ形態は、図1に示すセンサ設計装置により設計されたセンサ形態であり、5個のセンサが平面的に所定間隔で配置されている。図8の（a）に示すセンサ形態は、マイコンカーラリーにおいて標準的に用いられているセンサ形態であり、8個のセンサが一直線上に等間隔で配置されている。図8の（b）に示すセンサ形態は、図8の（c）に示すセンサ形態と状態空間の次元が等しいマイコンカーのセンサ形態であり、5個のセンサが一直線上に所定間隔で配置されている。

【0065】

図9は、図8に示す各センサ形態のエピソード数に対する適応度の変化を示す図である。図9に示すエピソード数に対する適応度の変化は、ゴールに到達する早さの変化を表しており、実線は図8の（c）に示すセンサ形態の変化を、一点鎖線は図8の（a）に示すセンサ形態の変化を、破線は図8の（b）に示すセンサ形態の変化をそれぞれ示している。

。

10

20

30

40

50

【 0 0 6 6 】

図 9 から、図 1 に示すセンサ設計装置により設計されたセンサ形態、すなわち図 8 の (c) に示すセンサ形態は、図 8 の (a) 及び (b) に示す人手によるセンサ形態より全てのエピソードにおいて高い適応度を示していることがわかる。なお、エピソード数を増やしても、この傾向は同様であった。また、図 8 の (c) に示すセンサ形態は、状態空間が大きい図 8 の (a) に示すセンサ形態に対しても、収束速度とパフォーマンスとの面で勝っており、本実施の形態のように物理世界を適切に観測することで、より効果的な学習を行えることがわかった。

【 0 0 6 7 】

次に、図 1 に示すセンサ設計装置により設計された図 8 の (c) に示すセンサ形態の配置及び個数について考察する。まず、図 8 の (c) に示すセンサ形態は、センサの配置に関して以下の特徴を有する。

【 0 0 6 8 】

(1) センサの配置が左右非対称である。これは、今回のタスクは、コースを半時計回りに回るように設定しているため、S 字カーブを除けば、左カーブが多く、左カーブを得意とするセンサ形態の適応度が高くなったためと考えられる。

【 0 0 6 9 】

(2) センサが前後に分散して配置されている。これは、横一列に並んだ配置では、ラインの前後関係を読み取ることができず、ラインレーサがカーブ上にいるのか、又は直線上にいるのかを判断することができないが、センサが前後に分散されることにより、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読み取ることができるためであると考えられる。

【 0 0 7 0 】

次に、センサの個数について、自動設計されたセンサ形態は、5 個前後のセンサを持つものが多く、センサの個数は状態空間の次元を決定するため、学習速度と取得できる情報量に影響を与える。一般的に、次元が少ない方が学習の収束が早い、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。しかしながら、上記の結果では、学習が収束した後であっても、センサ数が少ない方が、高いパフォーマンスを示した。この結果、少ない状態空間を効果的に用いることができれば、収束速度だけでなく、学習効果も改善できることがわかった。

【 0 0 7 1 】

このように、本センサ設計装置を用いることにより、タスク環境とロボットの学習能力とに適した状態空間を構成させることができ、ロボットにより適応的な行動を実行させることができる。

【 0 0 7 2 】

次に、上記のように設計されたセンサの形態を有するラインレーサ (図 8 の (c) に示すラインレーサ) を実際に作成してラインレースを行い、人手で設計されたラインレーサ (図 8 の (a) 及び (b) に示すラインレーサ) と比較した。なお、コントローラとして、図 8 の (a) に示すラインレーサは付属のサンプルプログラムを改良したハンドコーディングによるものを用い、図 8 の (b) 及び (c) に示すラインレーサは設計時のシミュレーションの学習結果を用い、設計時に得られた Q 値に基づいて行動を決定し、センサが観測した状態に対して最大の Q 値を持つ行動を選択するものを用いた。

【 0 0 7 3 】

上記の各ラインレーサに 5 回試行させ、1 回の試行はコースを 10 周するか、又はコースアウトした場合に終了させ、走行性能を比較するために平均ラップタイム (秒) を測定し、ロバスト性を比較するために平均滞在ラップ数を測定した。図 8 の (c) に示すラインレーサの平均ラップタイムは 13 . 5 秒、平均滞在ラップ数は 8 . 4 であり、図 8 の (a) に示すラインレーサの平均ラップタイムは 15 . 6 秒、平均滞在ラップ数は 4 . 4 であり、図 8 の (b) に示すラインレーサの平均ラップタイムは 16 . 0 秒、平均滞在ラップ数は 3 . 0 であり、図 1 に示すセンサ設計装置により設計されたセンサ形態を

10

20

30

40

50

有するライントレーサは、走行性能及びロバスト性ともに優れていた。

【0074】

次に、ラインレースにおいて難易度が高いとされる直角コーナーの攻略を例に、上記の各ライントレーサがどのようにセンサ配置を利用しているかについて説明する。

【0075】

まず、図8の(a)に示すライントレーサは、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おり、直角コーナーの直前のクロスラインを読み取ると、直角コーナー突破用のモードに移行し、直角コーナーを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することができるため、通常のカーブとは異なる制御則を用いて直角コーナーを突破することができる。

10

【0076】

図10は、図8の(c)に示すライントレーサの直角コーナーにおけるセンサ形態の利用方法を説明するための模式図である。図8の(c)に示すライントレーサは、コースに関する設計者の知識を利用できないため、直角コーナーを通常のカーブと同じ制御則を用いてクリアしなければならない。このため、直角コーナーにさしかかると、図10の(a)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ため、ライントレーサは、少し左に曲がる。しかしながら、実際にはコースは直角であるため、図10の(b)に示すように、コースから外れそうになる。このとき、4番目のセンサがセンタラインを検出し、ライントレーサは、大きく左折するため、直角コーナーをクリアすることができる。

20

【0077】

一方、図8の(b)に示すライントレーサは、横一列に並んだセンサを有するため、直角コーナーでコースアウトすることが多い。これは、センサが前後に分散していないため、コース上の前後関係を学習できなかったことによると考えられる。

【0078】

上記のように、図1に示すセンサ設計装置により設計されたライントレーサは、物理世界を適切に観測することにより、人手による設計に比べて、走行性能及びロバスト性に優れるとともに、学習性能にも優れることがわかった。

【0079】

なお、上記の説明では、ライントレーサを一例に説明したが、本発明が適用されるロボットは、この例に特に限定されず、センサを用いるロボットであれば、種々のロボットに適用可能である。

30

【図面の簡単な説明】

【0080】

【図1】本発明の一実施の形態によるセンサ設計装置の構成を示すブロック図である。

【図2】図1に示すセンサ設計装置によるセンサ設計処理を説明するためのフローチャートである。

【図3】図1に示すセンサ設計装置の設計対象となるロボットの一例を示す底面模式図である。

40

【図4】図3に示すロボットのタスク環境を示す模式図である。

【図5】図4に示すコース上のロボットの状態を示す模式図である。

【図6】図1に示すセンサ設定装置により設計されたライントレーサの各世代における適応度と、各世代における最良個体のセンサ個数との変化を示す図である。

【図7】図1に示すセンサ設計装置により設計されたセンサ形態の代表例を示す図である。

【図8】図1に示すセンサ設計装置により設計されたセンサ形態及び人手により設計されたセンサ形態の例を示す図である。

【図9】図8に示す各センサ形態のエピソード数に対する適応度の変化を示す図である。

【図10】図8の(c)に示すライントレーサの直角コーナーにおけるセンサ形態の利用

50

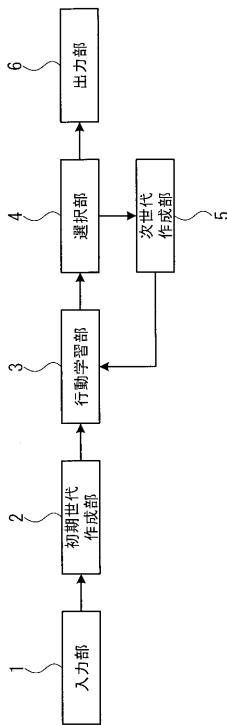
方法を説明するための模式図である。

【符号の説明】

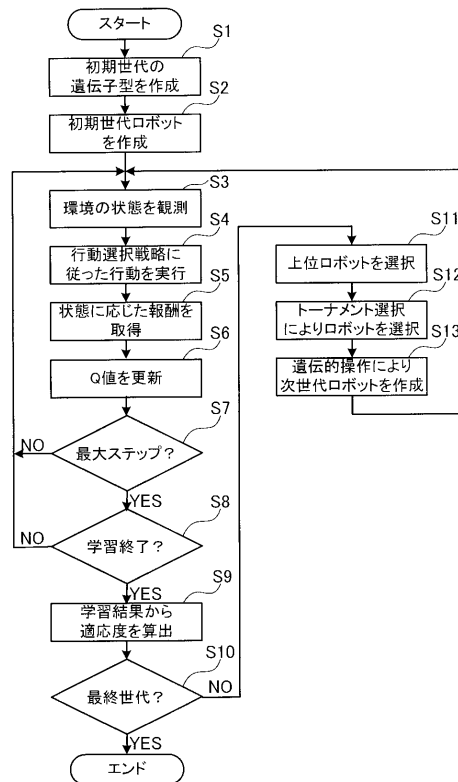
【0081】

- 1 入力部
- 2 初期世代作成部
- 3 行動学習部
- 4 選択部
- 5 次世代作成部
- 6 出力部

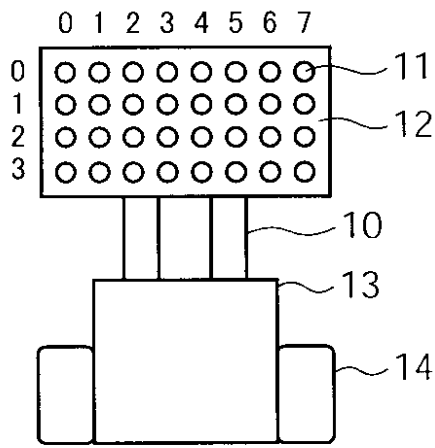
【図1】



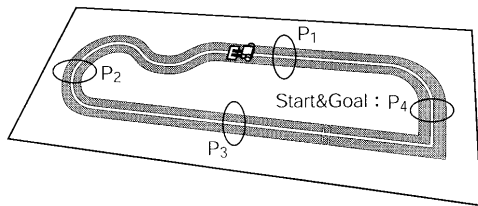
【図2】



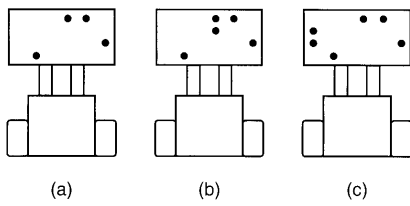
【図3】



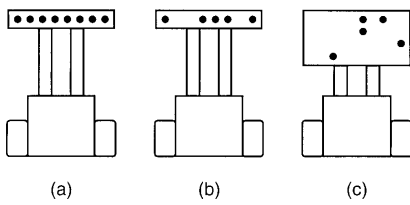
【図4】



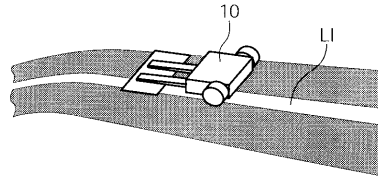
【図7】



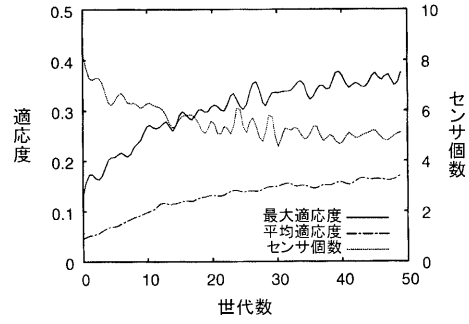
【図8】



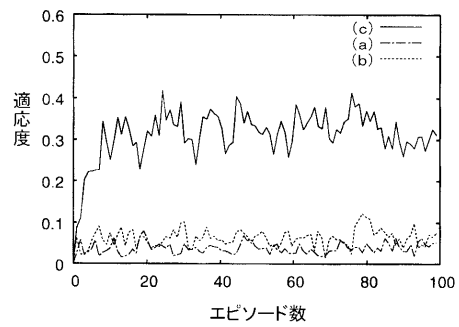
【図5】



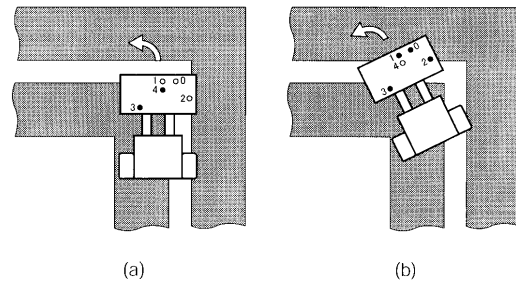
【図6】



【図9】



【図10】



フロントページの続き

(72)発明者 劉 健勤

京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内

審査官 松永 隆志

(56)参考文献 特開2003-326482(JP,A)

特表2003-531425(JP,A)

LICHTENSTEIGER L, PFEIFER R (Univ. Zurich, Zuerich, CHE) , An Optimal Sensor Morphology Improves Adaptability of Neural Network Controllers , Lect Notes Comput Sci , ドイツ , 2002年, Vol.2415 , Page.850-855

亀井 圭史, 遺伝的アルゴリズムによる移動ロボットの強化学習パラメータ最適化 , 電子情報通信学会技術研究報告, 日本, 2005年 3月22日, Vol.104 No.759, 119~124頁

(58)調査した分野(Int.Cl., DB名)

G06F 17/50